

FACULDADE DE DIREITO DE VITÓRIA
PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM DIREITOS E
GARANTIAS FUNDAMENTAIS
DOUTORADO EM DIREITO

RONALDO FÉLIX MOREIRA JÚNIOR

RACISMO ALGORÍTMICO E FERRAMENTAS DE POLICIAMENTO PREDITIVO:
PROTEÇÃO DE DIREITOS FUNDAMENTAIS OU UM RETORNO A CESARE
LOMBROSO?

VITÓRIA
2025

RONALDO FÉLIX MOREIRA JÚNIOR

**RACISMO ALGORÍTMICO E FERRAMENTAS DE POLICIAMENTO PREDITIVO:
PROTEÇÃO DE DIREITOS FUNDAMENTAIS OU UM RETORNO A CESARE
LOMBROSO?**

Tese apresentada ao Programa de Pós-Graduação Stricto Sensu em Direitos e Garantias Fundamentais da Faculdade de Direito de Vitória, como requisito para obtenção de título de Doutor em Direito.
Orientador: Professor Doutor Nelson Camatta Moreira.

VITÓRIA
2025

RONALDO FÉLIX MOREIRA JÚNIOR

**RACISMO ALGORÍTMICO E FERRAMENTAS DE POLICIAMENTO PREDITIVO:
PROTEÇÃO DE DIREITOS FUNDAMENTAIS OU UM RETORNO A CESARE
LOMBROSO?**

Tese apresentada ao Programa de Pós-Graduação Stricto Sensu em Direitos e Garantias Fundamentais da Faculdade de Direito de Vitória, como requisito para obtenção de título de Doutor em Direito.

Orientador: Professor Doutor Nelson Camatta Moreira.

Aprovada em 12 de março de 2025

COMISSÃO EXAMINADORA

Prof. Dr. Nelson Camatta Moreira
Faculdade de Direito de Vitória
Orientador

Prof. Dr. Adriano Sant'Ana Pedra
Faculdade de Direito de Vitória

Prof. Dr. Raphael Boldt de Carvalho
Faculdade de Direito de Vitória

Prof. Dra. Angela Araujo da Silveira
Espindola
Universidade Federal de Santa Maria

Prof. Dr. David Sanchez Rubio
Universidad de Sevilla

Aos meus pais, aos professores e a todos que, de alguma forma, tornaram esta tese possível.

RESUMO

Este trabalho investiga a relação existente entre inteligência artificial (IA), racismo algorítmico e colonialismo digital, com foco na análise de ferramentas de policiamento preditivo. O estudo demonstra como os algoritmos treinados com dados derivados de experiências humanas (como abordagens policiais) refletem preconceitos e estereótipos, reforçando um racismo já presente durante a persecução penal. Com efeito, sistemas de policiamento preditivo (como o PredPol e Palantir, apresentados no decorrer da tese), exemplificam essas práticas ao basear suas previsões delitivas em dados enviesados, o que, conseqüentemente, possui o potencial de aumentar a vigilância e até mesmo a violência praticada pelo Estado contra grupos já tradicionalmente marginalizados. O presente estudo se insere na linha de pesquisa “Democracia, Cidadania e Direitos Fundamentais”, do Programa de Pós-Graduação em Direitos e Garantias Fundamentais da Faculdade de Direito de Vitória, e tem como escopo demonstrar que a adoção por parte de países em desenvolvimento dessas tecnologias, sob um discurso de neutralidade, pode ser encarada como um reflexo do colonialismo digital e um retorno aos preceitos da criminologia lombrosiana do séc. XIX. Ao examinar essas questões a partir de uma metodologia predominantemente qualitativa e bibliográfica, a pesquisa também apresenta a hipótese de que é viável (e possível) a criação de ferramentas não violadoras de garantias, até mesmo no campo da persecução penal, partindo da premissa de que a IA pode ser desenvolvida e aplicada de forma a observar direitos humanos e fundamentais. Nesse sentido, em conclusão, como forma de comprovar a hipótese fundante, a tese apresenta projetos concretos de desenvolvimento de inteligências artificiais a partir de uma comunicação com o conceito de direitos humanos instituintes, destacando a importância de uma abordagem regulatória estatal que evite a reprodução de preconceitos históricos e culturais nesses sistemas tecnológicos.

Palavras-chave: colonialismo digital; direitos humanos instituintes; inteligência artificial; policiamento preditivo; racismo algorítmico.

ABSTRACT

This work investigates the relationship between artificial intelligence (AI), algorithmic racism, and digital colonialism, focusing on the analysis of predictive policing tools. The study demonstrates how algorithms trained with data derived from human experiences (such as police interactions) reflect stereotypes, reinforcing existing racism within criminal prosecution processes. Indeed, predictive policing systems (such as PredPol and Palantir, discussed throughout the thesis) exemplify these practices by basing their crime predictions on biased data, which can potentially increase surveillance and even state violence against historically marginalized groups. This study falls within the study area "Democracy, Citizenship, and Fundamental Rights" of the Graduate Program in Rights and Fundamental Guarantees at the Faculdade de Direito de Vitória. Its aim is to demonstrate that the adoption of these technologies by developing countries under a rhetoric neutrality can be seen as a reflection of digital colonialism and a return to the precepts of 19th-century Lombrosian criminology. By examining these issues through a predominantly qualitative and bibliographic methodology, the research also hypothesizes that the creation of non-violating tools respecting fundamental guarantees is feasible, particularly in the field of criminal prosecution, with the premise that AI can be developed and applied in a way that upholds human and fundamental rights. To support this hypothesis, the thesis concludes by presenting concrete AI development projects aligned with the concept of foundational human rights, emphasizing the need for a state regulatory framework that prevents the perpetuation of historical and cultural biases in technological systems.

Keywords: digital colonialism; foundational human rights; artificial intelligence; predictive policing; algorithmic racism.

LISTA DE ABREVIATURAS E SIGLAS

API – Interface de Programação de Aplicações

ARPANET – Advanced Research Projects Agency Network

CENIA – Centro Nacional de Inteligencia Artificial de Chile

CEPAL – Comissão Econômica para América Latina e Caribe

ChatGPT – Chat Generative Pre-Trained Transformer

COMPAS – Correctional Offender Management Profiling for Alternative Sanctions

CTIA – Comissão Temporária sobre Inteligência Artificial no Brasil

D2D – Device to Device

GenAI – Generative Artificial Intelligence

GPT – Generative Pre-Trained Transformer

HCI – Infraestrutura Hiperconvergente

IA – Inteligência Artificial

ICQ – "I Seek You"

IHC – Interação Humano-Computador

ILIA – Índice Latinoamericano de Inteligência Artificial

IoT – Internet of Things

LGPD – Lei Geral de Proteção de Dados

LSI-R – Level of Service Inventory – Revised

mIRC – Internet Relay Chat

MIT – Massachusetts Institute of Technology

ML – Machine Learning

PBIA – Plano Brasileiro de Inteligência Artificial

PL – Projeto de Lei

PNAS – Proceedings of the National Academy of Sciences

PP – Policiamento Preditivo

PPAs – Predictive Policing Algorithms

PredPol – Predictive Policing

RAM – Random-Acess Memory

SIA – Sistema Nacional de Regulação e Governança de Inteligência Artificial

SSL – Strategic Subject Lists

TIC – Tecnologias de Informação e Comunicação

WWW – World Wide Web

SUMÁRIO

1	INTRODUÇÃO.....	10
2	A VIOLAÇÃO DE DIREITOS FUNDAMENTAIS NO CONTEXTO DAS TECNOLOGIAS DIGITAIS: RACISMO DIGITAL E O MITO DA NEUTRALIDADE ALGORÍTMICA.....	17
2.1	DEFINIÇÕES NECESSÁRIAS PARA COMPREENDER A RELAÇÃO ENTRE TECNOLOGIA E DIREITOS HUMANOS.....	19
2.1.1	Tecnologia e seus diferentes conceitos.....	20
2.1.2	A múltipla utilização do termo ciberespaço.....	24
2.1.3	O que são algoritmos e inteligências artificiais.....	29
2.1.4	Uma definição de policiamento preditivo.....	34
2.2	TÉCNICA, TECNOLOGIA E IDEOLOGIA: OS BANCOS DE DADOS QUE ALIMENTAM A VIOLAÇÃO DE DIREITOS.....	37
2.2.1	As relações humanas e a influência dos aparatos tecnológicos...	37
2.2.2	Direitos humanos e fundamentais a partir de uma teoria crítica...	40
2.2.3	Tecnologia e a possibilidade de violação a direitos.....	43
2.2.4	Ausência de neutralidade na concepção de ferramentas tecnológicas.....	46
2.3	O DIREITO E A INTELIGÊNCIA ARTIFICIAL.....	50
3	DO DETERMINISMO BIOLÓGICO AO DETERMINISMO ALGORÍTMICO: A HERANÇA DO POSITIVISMO CRIMINOLÓGICO.....	54
3.1	O POSITIVISMO CRIMINOLÓGICO E A SUPOSTA SUPERAÇÃO DA ESCOLA CLÁSSICA.....	55
3.2	O PENSAMENTO LOMBROSIANO E AS INFLUÊNCIAS DE FERRI E GAROFALO.....	57
3.3	O BRASIL E A INFLUÊNCIA DO POSITIVISMO CRIMINOLÓGICO....	62
3.3.1	Influências para o surgimento do pensamento criminológico no Brasil.....	64

3.4	A CRIMINOLOGIA CRÍTICA E O DESLOCAMENTO DA QUESTÃO CRIMINAL PARA OS PROCESSOS DE CRIMINALIZAÇÃO.....	70
3.4.1	A crítica ao direito penal no Brasil e um novo pensamento criminológico.....	71
3.4.2	A questão racial no pensamento criminológico crítico brasileiro..	73
4	RACISMO ALGORÍTMICO E SELETIVIDADE PENAL: UMA ANÁLISE A PARTIR DE PROJETOS CONCRETOS DE POLICIAMENTO PREDITIVO.....	76
4.1	RACISMO ALGORÍTMICO E UMA POLÍTICA CRIMINAL CIBERLOMBROSIANA.....	77
4.2	O RACISMO, O COLONIALISMO E O DIGITAL.....	81
4.2.1	A Tecnologia e vigilância histórica sob a perspectiva de Fanon e Browne.....	85
4.3	O PODER DE PUNIR E A SELETIVIDADE PENAL.....	88
4.4	AS NOVAS FERRAMENTAS DE SELETIVIDADE PENAL.....	94
4.4.1	A polêmica quanto à utilização da ferramenta PredPol.....	97
4.4.2	Nova ferramenta e antigos problemas: o Palantir.....	101
4.4.3	O Brasil e a tecnologia na segurança: o projeto Smart Sampa.....	103
4.4.4	Outro breve exemplo brasileiro: a Guarda Civil de Vitória....	108
4.5	O POLICIAMENTO PREDITIVO COMO A APROXIMAÇÃO DE UMA DISTOPIA.....	109
5	A POSSÍVEL CRIAÇÃO DE TECNOLOGIAS NÃO VIOLADORA DE DIREITOS: COMO EVITAR O RETORNO A CESARE LOMBROSO...	113
5.1	REGULAMENTAÇÃO DA IA PELO MUNDO E SEUS PRINCIPAIS ASPECTOS.....	114
5.1.1	A União Europeia e sua regulamentação sobre IA.....	116
5.1.2	A regulamentação da IA no Brasil.....	119
5.1.3	A regulamentação da IA na América Latina.....	123
5.2	A MITIGAÇÃO DO RACISMO ALGORÍTMICO E UMA POSSÍVEL RELAÇÃO COM DIREITOS HUMANOS INSTITUINTES.....	125

5.2.1	É possível erradicar vieses ocultos (ou não) em inteligências artificiais?.....	127
5.2.2	O desenvolvimento de IA e sua conexão com direitos humanos instituintes.....	132
5.2.3	Falsas tecnologias garantidoras: inclusões abstratas e exclusões concretas.....	135
5.2.4	Direitos humanos instituintes para além de garantias impostas “de cima para baixo”.....	137
5.3	NOVAS FERRAMENTAS NA BUSCA PELA EFETIVAÇÃO DE DIREITOS.....	139
5.3.1	Projetos do Sul Global.....	140
5.3.2	Um novo rumo para a inteligência artificial.....	143
6	CONCLUSÃO.....	145
	REFERÊNCIAS.....	149

1 INTRODUÇÃO

O avanço no estudo da inteligência artificial (IA) permitiu a criação de ferramentas que não apenas otimizam os mais diversos procedimentos, mas que também transformam a maneira como o ser humano interage com o mundo ao seu redor (o que se mostra cada vez mais presente em atividades rotineiras, como selecionar uma música ou um filme em um aplicativo).

No âmbito profissional, não se pode negar que desde a automação de tarefas repetitivas até na análise preditiva de grandes volumes de dados, essas tecnologias têm encontrado aplicação em praticamente todos os campos do saber. No contexto da saúde, por exemplo, a IA pode auxiliar na interpretação de exames e na personalização de tratamentos (conforme as necessidades do paciente). No campo jurídico, pode facilitar a constante e necessária revisão de documentos e a análise de bancos de jurisprudência.

No entanto, esse crescimento também suscita algumas preocupações. Até mesmo os próprios desenvolvedores de IA têm manifestado receios quanto ao rumo que essas tecnologias podem tomar.

Alguns desses alertas são bem radicais e beiram o lúdico. Eliezer Yudkowsky, pesquisador do *Machine Intelligence Research Institute*, em 2023, pediu a todos os laboratórios de IA para cessarem o treinamento de seus sistemas mais poderosos que o GPT-4, o que, certamente, não aconteceu (Nakagawa, 2023).

Yudkowsky alertou que os laboratórios mais avançados podem, talvez de forma inconsciente, acabar possibilitando o desenvolvimento de uma inteligência artificial sobre-humanamente inteligente. Nas condições atuais, o pesquisador compreende que isso pode representar um risco extremo para toda a humanidade. Em seu entendimento, sem uma preparação adequada, novos conhecimentos científicos e precisão no desenvolvimento, o cenário mais provável seria catastrófico. Ele enfatiza que, no momento, a humanidade não está preparada e não saberia como alcançar o nível exato de segurança no desenvolvimento de uma IA (Nakagawa, 2023).

Previsões catastróficas, como as de Yudkowsky, podem parecer temas para obras cinematográficas ou literárias, como *Neuromancer*, de William Gibson, mas elas têm gerado discussões reais sobre os perigos da inteligência artificial. Contudo, os aspectos distópicos dessas tecnologias já estão presentes na realidade de maneira mais sutil, mas não menos significativa.

É possível mencionar, dentre as consequências contemporâneas do uso de IAs, o próprio tratamento inadequado de dados pessoais. Hoje, com algoritmos processando quantidades massivas de informações, a privacidade dos usuários é constantemente ameaçada (ainda mais com a presença de inúmeros de setores que não são claros em relação às suas políticas de uso e armazenamento de informações individuais coletadas).

Outro ponto bem discutido diz respeito à responsabilidade por danos provocados por IAs. Tendo em vista que essas tecnologias se integram (nos dias de hoje) a áreas críticas como transporte, saúde e finanças, é natural que surjam incertezas sobre quem deve ser responsabilizado em caso de falhas ou de acidentes.

Não obstante, também se fala no impacto da IA no mercado de trabalho a partir de uma cada vez maior automação de tarefas manuais e intelectuais, o que está transformando rapidamente diversos setores, com eventuais substituições de empregos humanos por soluções mais rápidas e econômicas.

Quanto ao tema central do trabalho, a questão do viés algorítmico expõe um dos problemas mais preocupantes da IA na atualidade. Como os algoritmos são treinados com dados que são, em verdade, informações derivadas de experiências humanas, eles podem manifestar preconceitos previamente existentes (seja na figura do programador ou na sociedade na qual ele está inserido). Isso é particularmente notável em sistemas de reconhecimento facial, que muitas vezes apresentam resultados menos precisos para pessoas de etnias não brancas, ou em algoritmos de recrutamento, que podem favorecer perfis semelhantes aos padrões estabelecidos a partir de uma herança racista e colonial.

Dentro dessa discussão a respeito do viés algorítmico, é preciso mencionar que a inteligência artificial em programas policiamento preditivo é uma das aplicações mais controversas desse campo tecnológico. A partir da análise de dados pregressos (e muitas vezes derivados de relatos policiais) sobre criminalidade, esses algoritmos tentam prever onde e quando crimes podem ocorrer, permitindo que as polícias direcionem suas ações de maneira (supostamente) mais eficiente. No entanto, essa prática levanta preocupações, já que esses sistemas tendem a reproduzir preconceitos presentes nos dados que foram usados para treiná-los, frequentemente associando determinados bairros ou grupos populacionais a maiores índices de criminalidade. Os possíveis resultados são múltiplos, como a eventual vigilância desproporcional de comunidades marginalizadas e até mesmo a desigualdade no

tratamento policial (com um potencial aumento da violência aplicada), dentro de um contexto de racismo algorítmico.

Os pontos mencionados demonstram que os problemas contemporâneos relacionados à inteligência artificial são, em verdade, problemas humanos que muitas vezes são mascarados por uma falsa noção de neutralidade tecnológica.

O presente trabalho tem, assim, seu objeto definido nas discussões crescentes a respeito de ferramentas de policiamento preditivo e do racismo algorítmico, mas também busca abordar a questão a partir de um problema mais amplo: o colonialismo digital.

Conforme demonstrado no decorrer da tese, o colonialismo digital se refere a uma nova modalidade de dominação e exploração, na qual grandes corporações tecnológicas e países desenvolvidos exercem o controle sobre os recursos digitais e a infraestrutura tecnológica de países menos desenvolvidos. Assim como no colonialismo tradicional, em que potências estrangeiras extraíam recursos naturais e impunham seu controle político e econômico sobre territórios, o colonialismo digital envolve a exploração de dados, a imposição de padrões tecnológicos e a dependência de países em desenvolvimento de plataformas, tecnologias e serviços controlados por grandes corporações de tecnologia. O trabalho busca explicar, assim, que o colonialismo digital tem se tornado presente de diferentes formas, como na importação acrítica de tecnologias de policiamento, sob uma justificativa de eficiência e neutralidade.

Não se busca, na presente tese, verificar ou não a existência do racismo algorítmico em tais ferramentas. Conforme demonstrado no decorrer da pesquisa, há extensa base teórica que já se aprofundou no assunto, confirmando um claro desbalanceamento em relação às prisões realizadas a partir desses sistemas em detrimento de pessoas de grupos tradicionalmente marginalizados. O que se pretende mostrar é que a adoção dessas ferramentas tem o potencial de estigmatizar ainda mais populações que já estão no campo da clientela da persecução penal, o que pode culminar em ainda mais violência contra esses grupos e na manutenção de uma já existente naturalização da desigualdade. Trata-se da criação de uma política criminal “ciberlombrosiana” que busca resgatar, do séc. XIX para os dias atuais, a figura do criminoso nato.

Com efeito, tendo em vista a constante violação a direitos por meio das tecnologias algorítmicas, a presente tese tem como escopo demonstrar a possibilidade de estigmatização de certos grupos pelo uso dessas ferramentas e, a partir daí, discutir se é possível evitar a violação de direitos por meio dessas

tecnologias.

Busca-se, dessa forma, responder à seguinte indagação: o uso de ferramentas de policiamento preditivo pode reforçar o racismo e a violência e quais alternativas éticas podem ser adotadas para que a tecnologias de IA sejam utilizadas de forma não violadora?

Para responder a essa questão é preciso que haja o devido o aprofundamento nos seguintes temas: 1) o racismo algorítmico nas ferramentas de policiamento preditivo como um dos reflexos do colonialismo digital; 2) a potencialização da criminalização secundária a partir da adoção dessas novas tecnologias; 3) a necessidade de regulação e até mesmo a criação de ferramentas não violadoras a partir de uma aproximação com a noção de direitos humanos instituintes.

É importante destacar que o trabalho não pretende fazer uma análise exaustiva a respeito do tema colonialidade, apenas demonstrar a existência de um colonialismo digital e que uma de suas consequências pode ser o aumento da violência e vigilância contra as populações marginalizadas. A hipótese central do trabalho é de que essas ferramentas tecnológicas, ao serem importadas de países desenvolvidos (ou criadas a partir de uma lógica similar), carregam consigo os mesmos problemas que enfrentam em seus contextos de origem (como o racismo). O fato de países em desenvolvimento (como o Brasil) adotarem essas ferramentas, muitas vezes sob um discurso de neutralidade tecnológica, pode ser encarado como uma extensão do colonialismo digital capaz de aprofundar ainda mais o problema da seletividade penal já presente na realidade brasileira.

Especificamente, a hipótese deste estudo é que a utilização do policiamento preditivo reforça a estigmatização e a violência contra determinados grupos sociais. No entanto, também se explora a possibilidade de que a inteligência artificial possa ser empregada de maneira a evitar violações de direitos humanos e fundamentais, desde que desenvolvida e implementada sob uma perspectiva crítica.

Assim, quanto à metodologia, trata-se de pesquisa qualitativa de cunho bibliográfico a partir dos estudos existentes sobre os temas relevantes ao trabalho, como recepção de teorias do positivismo criminológico, racismo algorítmico e colonialismo digital.

A partir da leitura de Gustin, Dias e Nicácio (2020, p. 73), é possível afirmar que a tese se insere na vertente de pesquisa jurídico-social, tendo em vista que ela trata do impacto das tecnologias de policiamento preditivo e do racismo algorítmico na sociedade contemporânea.

Nesse contexto, uma análise interdisciplinar (que conecta direito e tecnologia) com foco nos efeitos sociais das normas e práticas jurídicas, é uma característica central da vertente jurídico-social, que busca entender o direito em seu contexto real, para além da mera formalidade normativa.

Para a análise dos conceitos criminológicos (e da própria criminologia lombrosiana), selecionou-se importantes autores no campo da criminologia nacional e latino-americana, como Evandro Piza e Raúl Zaffaroni. Autores como Faustino, Lippold e Silva, por outro lado, foram selecionados pelo peso de seus trabalhos sobre colonialismo digital e racismo algorítmico. No que diz respeito a Fanon e Browne, suas obras sobre a vigilância dos corpos negros justificou sua inclusão. Ademais, no capítulo final, a obra de Sánchez Rubio foi utilizada devido à sua importante definição de direitos humanos instituintes. Este trabalho busca estabelecer a comunicação entre as ideias desses distintos autores, no intuito de promover um diálogo entre suas perspectivas teóricas.

O trabalho busca apresentar projetos concretos de implementação de ferramentas de policiamento preditivo no estrangeiro e no Brasil, como PredPol; Palantir; e Smart Sampa (os projetos foram escolhidos a partir da repercussão gerada após a implementação ou divulgação de tais sistemas). Essas análises buscam auxiliar na identificação das práticas algorítmicas que podem reforçar o racismo e a violência estatal. Importa ressaltar que, no que pese a pesquisa qualitativa poder possuir conteúdo altamente descritivo, ela pode lançar mão também de dados quantitativos incorporados nas análises, apesar de preponderar sempre o exame rigoroso da natureza, do alcance e interpretações possíveis do fenômeno objeto do estudo (Mezzaroba; Monteiro, 2009, p. 110).

Frisa-se que a análise de casos concretos presente na pesquisa não se limita aos sistemas de policiamento preditivo (explicados no segundo capítulo), mas também em relação a eventuais projetos ligados ao desenvolvimento de inteligência artificial a partir de uma ótica de direitos humanos, conforme demonstrado no último capítulo.

Dessa forma, o trabalho é segmentado em quatro momentos. O primeiro capítulo possui o intuito de apresentar o contexto de violação de garantias a partir do

uso de tecnologias (digitais ou não) e desconstruir a ideia de neutralidade tecnológica, destacando o crescimento e o impacto das ferramentas algorítmicas na sociedade contemporânea. Neste capítulo, são apresentados conceitos essenciais, como tecnologia, algoritmos, inteligência artificial, policiamento preditivo e seus desdobramentos em relação a tais garantias fundamentais.

O segundo capítulo apresenta as principais ideias do positivismo criminológico, examinando como essas ideias foram importadas e adotadas no Brasil, influenciando a visão do sistema penal e a abordagem das políticas de segurança pública. Dessa maneira, o capítulo estabelece um vínculo entre o positivismo criminológico e a seletividade penal contemporânea, mostrando como a utilização de ferramentas de policiamento preditivo pode representar um retrocesso na proteção a direitos fundamentais e um retorno a essas concepções criminológicas já ultrapassadas.

No terceiro capítulo, por sua vez, a pesquisa aprofunda o problema proposto, explorando as definições de racismo e colonialismo digital, com ênfase nas implicações dessas noções no contexto brasileiro (já gravemente marcado pela seletividade penal). A análise é ilustrada por meio de casos concretos de sistemas de policiamento preditivo, cujas implementações geraram controvérsias e discussões entre profissionais e acadêmicos de diferentes áreas.

O quarto e último capítulo se propõe levantar questões sobre a viabilidade da regulação dessas ferramentas para mitigar os efeitos do racismo digital. Este capítulo também examina propostas e projetos desenvolvidos por países do Sul Global que buscam utilizar a inteligência artificial de maneira não violadora de direitos, havendo um necessário diálogo com o conceito de direitos humanos instituintes. O objetivo é fomentar uma discussão sobre a necessidade (e possibilidade) de se opor a essa forma contemporânea de colonialismo sem que, para isso, seja necessário abandonar as novas tecnologias.

A presente tese, que se insere na linha de pesquisa “Democracia, Cidadania e Direitos Fundamentais”, do Programa de Pós-Graduação em Direitos e Garantias Fundamentais da Faculdade de Direito de Vitória, foi desenvolvida a partir das discussões e leituras realizadas no âmbito do Grupo de Pesquisa Teoria Crítica do Constitucionalismo. Autores como David Sanchez Rúbio e Franz Fanon (que, conforme destacado, compõem parte da base teórica do trabalho) são alguns dos autores recorrentes nas discussões do grupo. Os temas “racismo algorítmico” e “colonialismo digital” (e obras relacionadas, tais como “Racismo algorítmico:

inteligência artificial e discriminação nas redes digitais”, de Tarcízio Silva e “Colonialismo digital: por uma crítica hacker-fanoniana”, de Deivison Faustino e Walter Lippold) surgiram no Grupo como um aprofundamento natural do tema “direitos humanos” na medida em que as discussões se deparavam com o inevitável fenômeno da ascensão das inteligências artificiais e seus impactos na vida cotidiana.

A tese aqui apresentada se mostra, portanto, como resultado desses anos de pesquisa, estudos e troca de ideias entre as áreas das ciências criminais, direito digital e teoria crítica dos direitos humanos.

2 A VIOLAÇÃO DE DIREITOS FUNDAMENTAIS NO CONTEXTO DAS TECNOLOGIAS DIGITAIS: RACISMO DIGITAL E O MITO DA NEUTRALIDADE ALGORÍTMICA

“Governos do Mundo Industrial, vocês gigantes aborrecidos de carne e aço, eu venho do Ciberespaço, o novo lar da Mente”. É com essas palavras que John Perry Barlow iniciou seu texto: *A Declaration of the Independence of Cyberspace* (Uma Declaração de Independência do Ciberespaço), escrito em 1996, com o intuito de defender toda forma de manifestação e utilização da rede mundial de computadores da ingerência estatal (Venturini; Cappi, 2018).

O documento mencionado mostra que não é tão recente a alegação da existência de uma eventual neutralidade do mundo virtual ou das tecnologias digitais. Tais argumentos são uma marca de um “tecnolibertarianismo¹”, conforme menciona Tarcízio Silva (2022, p.18).

Vale destacar que a segunda metade da década de 1990 foi marcada por um importante debate a respeito dos direitos de liberdade de expressão no cenário norte-americano e estadunidense. A legislação de telecomunicações dos Estados Unidos da América: *Communications Decency Act*, ou Lei de Decência das Comunicações, buscou regular o conteúdo da internet, proibindo a transmissão de material considerado "obsceno" ou "indecente" em redes públicas de computadores.

No entanto, a lei gerou um grande debate sobre a liberdade de expressão e em 1997 a Suprema Corte dos Estados Unidos declarou que algumas das disposições do ato violavam a Primeira Emenda da Constituição do país. A existência dessa norma fez com que surgissem inúmeros defensores de uma ausência de controle do Estado sobre essa temática, o que culminou na mencionada declaração que foi publicada em Davos, na Suíça, e considerada um documento defensor de ideais libertários, embasados nas premissas de teóricos como Thomas Jefferson, John Stuart Mill, Alexis de Tocqueville e Louis Brandeis, autor da obra clássica *The Right to Privacy* (Zulmar, 2021).

¹ “Tecnolibertarianismo” ou “ciberlibertarianismo” pode ser entendido como um movimento que defende a impossibilidade de sujeição do ciberespaço aos mesmos modelos regulatórios que limitam os estados nacionais. O virtual, na perspectiva dos tecnolibertários, só pode ser regimentado por meio de uma autorregulação a partir da ação conjunta dos próprios usuários (Freitas, 2018, p. 135).

É possível afirmar que o documento citado é uma das primeiras manifestações em apoio à autorregulamentação da rede mundial de computadores. Trata-se da crença em uma suposta neutralidade tecnológica capaz de permitir o pleno funcionamento do ciberespaço que poderia ser prejudicado por meio do controle do Estado.

Para Silva (2022, p. 18), a declaração é uma reação da elite tecnológica estadunidense que oferecia uma concepção determinista da internet. Essa visão “tecnolibertária” carece de uma necessária análise a respeito de questões de gênero e raça, bem como ignora a existência de novas formas de colonialismos. Nas palavras de Barlow, autor do texto original, eventuais problemas existentes no campo tecnológico seriam solucionados naturalmente e internamente:

Vocês afirmam que existem problemas entre nós que vocês precisam solucionar. Vocês usam essa alegação como desculpa para invadir nossos recintos. Muitos desses problemas não existem. Onde existirem conflitos reais, onde existirem erros, iremos identificá-los e resolvê-los por nossos próprios meios. Estamos formando nosso próprio Contrato Social. Essa maneira de governar surgirá de acordo com as condições do nosso mundo, não do seu. Nosso mundo é diferente (Barlow, citado por Venturini e Cappi, 2018).

Muito embora a presente tese não tenha, dentre seus objetivos, o intuito de pesquisar o tema “liberdade de expressão” na *web* ou de analisar os meandros da mencionada declaração, a menção a essa temática é importante, pois a própria existência desse documento reforça uma concepção comum (que afeta não apenas o campo jurídico) de que as tecnologias digitais não são apenas neutras, mas capazes de impedir e solucionar conflitos sem interferências externas, principalmente aquelas realizadas pelo Estado.

Lawrence Lessig (2009, p. 34) menciona que foi muito disseminada a ideia de que o ciberespaço teria o condão para a criação de uma sociedade definida “de baixo para cima”, dotada de ordem por si própria, livre de ações governamentais e atuações políticas. Ocorre que, conforme demonstrado no decorrer da pesquisa, existem apontamentos por parte da comunidade científica e por ativistas contemporâneos de que até mesmo os processos de construção das novas tecnologias possuem vieses, inclusive de categorizações feitas segundo critérios raciais.

Dessa forma, o primeiro capítulo do trabalho tem como escopo demonstrar a inexistência de neutralidade na projeção algorítmica, de modo que se torna possível

discutir não apenas questões como racismo na estrutura dos novos artefatos digitais, mas também discutir a existência de elementos da seletividade penal no contexto de determinadas ferramentas, em especial as tecnologias de policiamento preditivo.

No intuito de facilitar a compreensão do tema e alcançar o objetivo proposto de forma coerente, o capítulo é segmentado da seguinte maneira:

A primeira parte se concentra na apresentação de conceitos e definições necessárias para o entendimento do tema abordado na tese. Discute-se nesse momento conceitos ligados diretamente às tecnologias digitais, entre eles: 1) as diferentes definições de tecnologia; 2) a definição do termo ciberespaço; 3) a apresentação dos termos algoritmos e inteligências artificiais; e 4) a definição de policiamento preditivo.

A segunda parte do capítulo busca apresentar temas não ligados diretamente às tecnologias de informação e comunicação, mas que são igualmente relevantes para a pesquisa. Nesse momento, discute-se questões como "técnica e ideologia", além do conceito de direitos humanos utilizado no trabalho, sendo demonstrado como os aparatos tecnológicos têm sido tradicionalmente utilizados no contexto de sua violação.

2.1 DEFINIÇÕES NECESSÁRIAS PARA COMPREENDER A RELAÇÃO ENTRE TECNOLOGIA E DIREITOS HUMANOS

A discussão quanto à forma ou formas pelas quais as novas tecnologias afetam determinados direitos e garantias depende, preliminarmente, da compreensão do que é uma tecnologia e a razão pela qual se utiliza o termo "novas tecnologias". Outros termos relevantes para o presente trabalho também merecem ser explicados, como ciberespaço e inteligência artificial. Entretanto, a discussão inicial deve se manter no que é tecnologia.

A expressão mencionada normalmente vem associada a outro termo: técnica. É importante ressaltar, contudo, que as expressões possuem suas próprias características.

Técnica e tecnologia costumam ser expressões associadas a uma certa origem: a utilização, pelo homem, de objetos que foram transformados em ferramentas complexas durante o processo de surgimento da sociedade. Técnica e tecnologia possuem uma origem comum na palavra grega *techné*, que significa mais

um modo de alteração do mundo do que sua compreensão (Veraszto; et. al., 2008, p. 61-62).

Pode-se afirmar, assim, que enquanto técnica possui diferentes significados (como fabricar, produzir e construir), tecnologia parte da junção entre techné e logus (razão). Trata-se, dessa forma, do próprio estudo da técnica (Veraszto; et. al., 2008, p. 62).

2.1.1 Tecnologia e seus diferentes conceitos

É importante frisar que o termo tecnologia tem sido utilizado, com o passar do tempo, em diferentes contextos. Fala-se hoje em diferentes tipos de tecnologia, como tecnologia da educação, tecnologias algorítmicas e até mesmo de forma geral: “novas tecnologias”. Para Lévy (1993), a própria linguagem pode ser encarada como uma das primeiras tecnologias desenvolvidas.

Veraszto menciona que o ser humano (como é entendido hoje) surgiu quando seu pensamento foi capaz de se aliar à capacidade de transformação e ressignificação, o que é exemplificado pelo autor a partir da cena inicial do filme “2001, Uma Odisseia no Espaço”, dirigido em 1968 por Stanley Kubrick² (2008, p. 63). Não é possível, portanto, falar em ser humano sem associá-lo a seus instrumentos, ainda que rudimentares. Pode-se dizer que houve um processo de aprendizagem quanto à possibilidade de criação de novas extensões para ação dos membros de forma mais eficiente (Veraszto; et. al., 2008, p. 63).

Tecnologia é, assim, um termo que abrange mais do que apenas a criação de uma ferramenta (seja pela criação em si de um novo objeto ou pela ressignificação de um objeto já existente – como o osso utilizado como tacape pelos hominídeos primitivos). É preciso ressaltar que o termo também não se refere necessariamente a objetos ou produtos individualmente considerados.

Compreender corretamente o que se pode considerar como tecnologia é vital para a compreensão dos demais temas que estão inseridos na tese. Nos termos de Veraszto (2008, p. 67), há uma concepção equivocada e uma pré-suposição de autonomia e de neutralidade da tecnologia que:

² No prelúdio “Aurora do homem”, após o encontro com um monólito negro na África pré-histórica, um hominídeo percebe ser capaz de utilizar um osso como uma ferramenta e como uma arma para caçar ou dominar grupos rivais (Kubrick, 1968).

[...] tem favorecido uma imagem da evolução tecnológica que mantém o dilema, errôneo, da eficiência interna x interferência externa, dando lugar a uma determinada concepção da avaliação tecnológica que sustenta, por sua vez, um modelo tecnocrático de ordem política.

Dessa maneira, de forma a não cometer equívocos, é importante que sejam apresentadas as principais concepções do termo tecnologia, quais sejam: a concepção 1) intelectualista; 2) utilitarista; 3) artefactual; e 4) de sociossistema.

Em relação ao primeiro conceito (intelectualista), tecnologia pode ser considerada um conhecimento prático derivado do conhecimento teórico científico (Veraszto; et. al., 2008, p. 67). Em outras palavras, trata-se de uma verdadeira aplicação do conhecimento científico por meio de procedimentos práticos, como é o caso da transformação de matéria-prima em bens de consumo.

O conceito de tecnologia também pode ter cunho utilitarista, de uma maneira que técnica pode ser considerada um termo sinônimo. Nesse sentido, a tecnologia deve ser “[...] precisa e eficiente (mais ou menos como a técnica grega) e não deve preocupar-se em constatar teorias [...]” (Veraszto; et. al., 2008, p. 67). Pode-se dizer que se trata, portanto, da utilização do conhecimento para a melhoria da qualidade de vida humana.

Entretanto, a concepção mais recorrente no cotidiano é uma vertente instrumental ou artefactual do termo. Tecnologia, nesse ponto de vista, são as ferramentas ou artefatos que foram construídos para diferentes fins. Esse conceito traz uma série de problemas (e simplificações), haja vista que sua sustentação implica em gerar uma ausência de diferenciação entre, por exemplo, os utensílios de pedra da antiguidade e os contemporâneos itens tecnológicos (Veraszto; et. al., 2008, p. 68).

Por mais que não seja o conceito mais apropriado, a presente tese utiliza o termo “tecnologia” em diferentes momentos como sinônimo de artefato tecnológico, haja vista que o termo também é utilizado de forma similar por autores que compõem o arcabouço teórico da tese.

Por fim, o significado mais interessante à pesquisa é o conceito de tecnologia como sociossistema. Essa abordagem permite que seja estabelecida uma relação entre tecnologia, técnica, demanda social, produção tecnológica, política e até mesmo econômica. Nesse sentido, tecnologia pode ser considerada um processo aberto cujo curso é determinado por meio de interações de diferentes grupos sociais relevantes.

Assim, quando um artefato tecnológico se converte em produto (por meio de uma técnica), há sempre uma carga política concreta (Veraszto; et. al., 2008, p. 73).

Torna-se possível dizer que mudanças tecnológicas podem causar alterações nos meandros culturais e organizacionais de uma sociedade. Esse conceito se comunica com os estudos de Jürgen Habermas (a respeito da relação entre técnica, ciência e ideologia) e Frantz Fanon (quanto ao uso da tecnologia pelas populações colonizadas), ambos explorados no decorrer do trabalho.

Vale ressaltar que outros termos relacionados à expressão “tecnologia” merecem destaque no presente trabalho: 1) a concepção de neutralidade tecnológica; e 2) a concepção de determinismo tecnológico.

Como ilustrado na parte inicial do capítulo, a Declaração de Independência do Ciberespaço reforça cristalinamente ambas as concepções. Quanto à primeira, parece lógico acreditar que a tecnologia é sempre alheia a qualquer interesse particular, não apenas na sua criação e desenvolvimento, como também no produto que é concebido na etapa final.

Pensar a tecnologia, ou mesmo a ciência, como elementos neutros³ não é algo especificamente recente. De maneira rasa, é possível afirmar que o pensamento iluminista, no intuito de se contrapor à religião, compreendia que a ciência, como ferramenta para se alcançar a verdade, não admitiria vieses.

O iluminismo, reforçado pelo positivismo do final do século XVIII (com base no pensamento de Francis Bacon e René Descartes), potencializou a ideia de neutralidade científica, como se a natureza fosse permanentemente surda e inflexível quanto aos desejos humanos (Veraszto; et. al., 2008, p. 69).

Ocorre que, se é impossível dissociar a ideia de tecnologia da técnica, também não é possível desconectá-la da política (Veraszto; et. al., 2008, p. 70):

Se considerada como independente de qualquer sistema político ou social, a tecnologia pode ser transferida de um país a outro sem dificuldade alguma. Essa visão reducionista da tecnologia impede sua análise crítica e ignora as intenções ou interesses sociais, econômicos e políticos daqueles que a idealizam, financiam e controlam. Sabemos que a tecnologia não é neutra; um artefato aparentemente inócuo pode estar carregado de interesses políticos (e/ou outros).

³ Entendamos neutralidade como uma postura de ausência de favorecimento a qualquer assunto, pessoa, grupo, característica ou ideia.

No intuito de trazer exemplos que reforcem a temática proposta pela tese e corroborar com a informação citada, é possível mencionar tecnologias (mais uma vez utilizadas aqui no sentido de artefatos concebidos por determinadas técnicas) que historicamente foram utilizadas como ferramentas de segregação e violação de direitos.

Por exemplo, em Long Island, Nova Iorque, é possível notar viadutos de alturas incomuns, mais baixos que em diversos outros lugares do país. Muito embora possa parecer uma simples escolha estética (ou quem sabe até mesmo um erro de cálculo), há certamente uma escolha política oculta nessas plataformas de concreto (Winner, 2017, p. 198-199).

Os viadutos da região foram projetados por Robert Moses, empreiteiro de estradas e outras obras de Nova Iorque, responsável por construir durante quase cinquenta anos (décadas de 1920 até 1970) pontes que desencorajassem a presença de ônibus nas avenidas de parques, que normalmente eram utilizados por pessoas pobres e/ou negras. Conforme mencionado pelo biógrafo de Moses, a “estética” seria relevante para encorajar a presença de automóveis (que até então eram utilizados por pessoas brancas de classe alta ou da classe média burguesa). Portanto, veículos com mais de três metros de altura não podiam passar por esses viadutos, o que evidenciava uma verdadeira vontade política de segregação (Winner, 2017, p. 199).

Compreender essas ocorrências torna possível criticar também a ideia de determinismo tecnológico, ou seja, criticar a concepção determinista das relações entre tecnologia e sociedade, como se os progressos científico e tecnológico seguissem um caminho determinado, inalterado por fatores políticos, econômicos ou sociais (Veraszto; et. al., 2008, p. 70).

O que se pode afirmar, portanto, é que as tecnologias em uma sociedade podem determinar como vivem (ou até morrem) os indivíduos que nela habitam. Afirmar isso é diferente de dizer que o modo de vida de uma população (ou parte dela) é determinada inexoravelmente pelo imutável desenvolvimento dos artefatos tecnológicos.

Com base no que até o momento foi exposto, é possível que se entenda tecnologia não apenas como artefatos. Da mesma maneira, não é possível afirmar que tecnologia é sinônimo de processo de produção, como as máquinas e instrumentos necessários para a fabricação de objetos. Para Veraszto (2008, p. 78):

Na medida em que muda padrões, a tecnologia também cria novas rotas de desenvolvimento. Portanto, trabalhar com tecnologia é trabalhar com algo dinâmico. O que hoje é ponta, amanhã é obsoleto, exigindo novos procedimentos, conceitos e atitudes para inovar. A tecnologia faz parte do acervo cultural de um povo, por isso existe na forma de conhecimento acumulado, e por essa mesma razão está em contínua produção. A tecnologia em si constitui-se, portanto, como uma forma de conhecimento e todas as tecnologias são produtos de todas as formas de conhecimento humano produzidas ao longo da história.

A definição mencionada pelo autor é interessante ao presente estudo, de modo que é possível concluir que tecnologia se trata de um agrupamento de conhecimentos diferentes ligados ao desenvolvimento de diferentes instrumentos por meio de técnicas específicas. Tais conhecimentos versam sobre uma necessidade de se solucionar problemas da vida, mas de uma forma que não se pode deixar de levar em consideração a existência de interesses e vieses na própria concepção desses saberes.

2.1.2 A múltipla utilização do termo ciberespaço

As discussões a respeito de inteligências artificiais não estão necessariamente ligadas à rede mundial de computadores, mas entender o termo “ciberespaço” é importante para o trabalho, tendo em vista que falar de novas tecnologias nos leva a discutir, eventualmente, o termo, já que inúmeras ferramentas de IA são utilizadas a partir da rede mundial de computadores.

A expressão foi cunhada por William Gibson, escritor de ficção científica, no início da década de 1980. Em seu livro "Neuromancer", publicado em 1984, O ciberespaço pode ser compreendido como uma espécie de consciência artificial, de cunho global, marcada pela completa interação entre humano e máquina. A ideia parte de um espaço virtualizado (não material), onde as pessoas poderiam se conectar e interagir com o mundo digital. Conforme é descrito na obra:

Ciberespaço. Uma alucinação consensual vivenciada diariamente por bilhões de operadores autorizados, em todas as nações, por crianças que estão aprendendo conceitos matemáticos... uma representação gráfica de dados abstraídos dos bancos de todos os computadores do sistema humano. Uma complexidade impensável. Linhas de luz alinhadas no não espaço da mente, aglomerados e constelações de dados. Como luzes da cidade, se afastando... (Gibson, 2016, p. 77)

Desde então, o termo ciberespaço se tornou amplamente utilizado para descrever a internet e o ambiente digital em geral, vinculando-o a outros conceitos como *World Wide Web*, as redes sociais, a nuvem, a inteligência artificial e outras tecnologias. O ciberespaço é visto como uma nova realidade, em que as pessoas podem criar, compartilhar e acessar informações, fazer compras, jogar e se conectar com outras pessoas de todo o mundo.

Tendo em vista que os mencionados termos (nuvem, redes sociais, etc.) hoje dificilmente podem ser dissociados, é compreensível a utilização da expressão ciberespaço para todos eles. Como consequência, o termo também é utilizado como um sinônimo de internet, muito embora essa última seja uma ferramenta que almejava diferentes finalidades, a depender da época em que era utilizada e desenvolvida.

A internet surgiu na década de 1960, como uma iniciativa do Departamento de Defesa dos Estados Unidos para criar uma rede de computadores capaz de continuar funcionando mesmo em caso de guerra ou de falhas em uma parte da rede. Essa rede, chamada de ARPANET (*Advanced Research Projects Agency Network*), foi a base da internet tal como a conhecemos hoje (Corrêa, 2013, p. 17).

A conexão em rede permitida pela ARPANET otimizava, em grande escala, a utilização do tempo nos computadores daqueles anos (bem diferente dos contemporâneos computadores pessoais). No sistema original, havia um computador central (chamado de *mainframe*), responsável por armazenar e processar todos os dados compartilhados entre os terminais disponíveis (Corrêa, 2013, p. 18).

Na década de 1990, a popularidade da internet cresceu rapidamente com a introdução da rede mundial de computadores – *World Wide Web* (WWW) e dos navegadores *web*. É a partir dessa ferramenta que a internet (ou o ciberespaço) passou a se aproximar dos conceitos hoje conhecidos (como páginas para vendas, mecanismos de busca e a utilização de imagens). A WWW tornou mais fácil para as pessoas acessarem informações e recursos na internet, e a popularidade da web explodiu, muito embora sua popularização no Brasil ainda tenha levado mais uma década (Vieira, 2014, p. 2-3).

Como mencionado, foi apenas com a *World Wide Web* que a concepção contemporânea do ciberespaço ganhou forma, aproximando-se vagamente do que havia escrito Gibson em sua obra *Neuromancer*. Os navegadores (ou *browsers*) foram fundamentais para que a internet fosse associada a um imaginário de local ou “mundo virtual” no qual as pessoas pudessem se relacionar (algo que se tornou ainda mais

notável com a chamada internet das coisas, conforme demonstrado adiante). Até a década de 2010, a navegação pela rede mundial podia ser feita por pelo menos dez navegadores (Vieira, 2014, p. 3-4):

- Mosaic, 1993: Primeiro navegador a rodar na plataforma Windows, o que determinou a abertura da WWW ao público;
- Netscape, 1994: Modernizou o browser com a inclusão de navegação por abas, bloqueio de pop ups, e histórico de visitas, entre outros. Foi líder até 2002, perdendo seu posto para o Internet Explorer, sistema operacional Windows Microsoft;
- Internet Explorer 1.0, 1995: Início da guerra dos navegadores Web; período de quatro anos de 1995 a 1999;
- Internet Explorer 3.0, 1996: Lançado em agosto pela Microsoft; no mesmo ano, é lançado a versão 3 do Netscape e a primeira do Opera;
- Internet Explorer 4.0, 1997: Incluída no sistema operacional “Windows 98”, deu origem a ação antitruste movida pelo Departamento de Justiça dos Estados Unidos;
- Mozilla, 1998: A Netscape anuncia a liberação do código-fonte de seu navegador para alterações por qualquer colaborador e cria a comunidade Mozilla, que em 2004 lançaria o Firefox;
- Opera, 2000: Criado na Noruega, foi a primeira alternativa “leve” para os usuários;
- Safari, 2003: Criado pela Apple, que anuncia seu navegador próprio como o navegador padrão a partir do sistema operacional Mac OS X v10.3;
- Firefox, 2004: Versão simplificada do Mozilla, gratuito e com navegação por abas, se tornou o maior rival da Microsoft;
- Chrome, 2008: Lançamento do Google, com a promessa de ser mais rápido, seguro e estável que os concorrentes.

Até os primeiros anos do século XXI, a imagem do ciberespaço como um verdadeiro “espaço virtual” onde interações são possíveis ainda era algo especialmente restrito às obras de ficção científica. A conexão entre indivíduos era possível por meio de softwares como o mIRC⁴. Em regra, a rede ainda era formada quase que exclusivamente por páginas acessadas por usuários de forma que quase não havia interação entre eles.

A existência de uma maior interação entre usuários da WWW pode ser identificada a partir de dois momentos: 1) a criação das redes sociais digitais; 2) o surgimento da chamada “internet das coisas”.

O termo “rede” merece ser rapidamente discutido. Apesar da expressão ser comumente e originalmente utilizada para designar um aparato de pesca, foi com a

⁴ O mIRC é considerado um *Internet Relay Chat* (o que pode ser traduzido como bate-papo de retransmissão de internet) utilizado para comunicação em diferentes nichos, como jogos e trabalhos profissionais. A comunicação é possível por usuários individuais ou em grupo. Trata-se de uma das primeiras ferramentas de comunicação na rede mundial de computadores, tendo sido lançado no ano de 1995 (mIRC, 1995-2023).

abstração de tal conceito que seu significado foi amplificado (e utilizado nas mais diferentes áreas do conhecimento) para representar diversos sistemas ou pontos interligados dentro de uma interface de gestão em espaço e tempo (Zenha, 2018, p. 21).

Uma rede social digital (ou online) pode ser definida, conforme Zenha (2018, p. 24), como um ambiente digital organizado por uma interface virtual própria organizada a partir de perfis de seres humanos classificados conforme afinidades e características semelhantes.

Por meio dessas redes, tornou-se possível uma maior interação entre diferentes usuários. Se em aplicações anteriores era necessário que a pessoa tivesse acesso a uma espécie de endereço ou número virtual (como no caso do ICQ⁵), ou a interação era restrita a salas de bate-papo, com as novas redes sociais virtuais, tornou-se possível interagir de diferentes maneiras (fóruns, chats individuais e salas de bate-papos eram muitas vezes integrados). Tornou-se possível também a criação de conteúdo por parte dos usuários.

A chamada “internet das coisas”, por sua vez, pode ser considerada uma nova faceta da comumente chamada “era digital”. É possível imaginar, por meio de sua infraestrutura, a possibilidade de realização de serviços mais avançados e uma maior conexão entre pessoas e artefatos tecnológicos.

A expressão “internet das coisas” foi proposta por um estudioso do MIT (Instituto de Tecnologia de Massachusetts), Kevin Ashton, no final da década de 1990. Haveria, conforme entendia, uma maior necessidade (no século vindouro) de conexão com a rede mundial de computadores, o que deveria ser feito por formas variadas (*device-to-device*, ou D2D), não apenas pelo uso dos computadores pessoais (Magrani, 2018, p. 45).

É possível pensar em inúmeras vantagens da conexão à internet a partir de aparelhos diversos (como veículos, brinquedos e até mesmo eletrodomésticos), principalmente em solução de problemas práticos da vida cotidiana.

Se a ideia clássica da ficção sobre ciberespaço tratava da existência de um espaço virtual de interação que pudesse ser conectado de qualquer lugar. Na obra de

⁵ Aplicação criada em 1996 pela empresa *Mirabilis*, sua sigla é uma derivação da expressão “*I seek you*”. Após sua instalação, ele atribuía ao usuário um código de identificação chamado de *Unique Identification Number* (Carniel, 2021).

ficção distópica (*Neuromancer*), essa conexão é feita por meio de aparelhos chamados decks.

Na mencionada obra, os decks são dispositivos eletrônicos portáteis que permitem aos usuários o ingresso no um ambiente virtual (ciberespaço). Esses dispositivos são usados para acessar redes de computadores e sistemas de segurança, bem como para realizar outras atividades ilegais no mundo dos bits. Os decks são descritos como pequenos o suficiente para caber no bolso, sendo operados por meio de interfaces neurais, que permitem que o usuário se conecte diretamente ao seu sistema nervoso. Isso permite que a experiência de estar no ciberespaço seja tão real quanto a vida e permite que os usuários realizem tarefas complexas em um espaço imaterial (Gibson, 2016, p. 317).

Embora haja conjecturas e até mesmo estudos a respeito, não há ainda instrumentos de conexão neural. No entanto, a conexão à rede mundial de computadores pode ser feita por uma série de aparatos portáteis, como celulares e tablets.

A utilização dessa multiplicidade de dispositivos pode, em verdade, facilitar as tarefas cotidianas e profissionais dos indivíduos. Há maior comodidade no acesso a certos tipos de aplicações e informações por meio de aparelhos celulares ou ferramentas similares. É preciso mencionar, contudo, que o crescimento das formas de conexão pode levar (e já leva) a uma fragilidade no que diz respeito à privacidade e a segurança das informações pessoais dos usuários, favorecendo inúmeras violações a direitos.

Conforme demonstrado no decorrer do trabalho, a constante violação a direitos fundamentais por meio de artefatos tecnológicos ocorre devido ao fato de que a sociedade contemporânea é altamente dependente dessas ferramentas. Elas, por outro lado, são, em sua própria concepção, vulneráveis a interferências de todo tipo.

Magrani (2018, p. 922-93) menciona uma análise feita pela HP Security Research que apontou para uma falha de segurança em aproximadamente 70% dos dispositivos móveis existentes. Essas falhas normalmente estão ligadas a problemas com privacidade, autorizações de uso insuficientes, falta de criptografia no tratamento dos dados, interfaces com pouca segurança ou mesmo inadequação em softwares de proteção.

Não é, entretanto, a possibilidade de um ataque hacker o maior problema dos instrumentos móveis de conexão. Muitos dispositivos são criados com a exata

finalidade de coleta de dados, muitas vezes de forma dissimulada e para fins desconhecidos por parte do titular.

Em linhas gerais, a existência de 1) uma rede que permite diversos modelos de interação entre os usuários, bem como a criação de conteúdo por parte desses mesmos usuários (como em perfis de Instagram) e 2) de meios diversos de conexão a essa rede devido a existência de inúmeros aparatos que permitem esse acesso de quase qualquer lugar, faz com que a realidade se aproxime minimamente do conceito de ciberespaço nascido na ficção, tanto que a expressão vem sendo utilizada para tratar de diferentes aspectos relacionados à rede mundial de computadores e à inteligência artificial.

2.1.3 O que são algoritmos e inteligências artificiais

É imprescindível para a compreensão do tema proposto pela tese a discussão quanto aos conceitos de algoritmo e inteligência(s) artificial(ais). A falsa noção de neutralidade tecnológica parte, na maioria das vezes, da tentativa de “culpar” os algoritmos ou a inteligência artificial (que, conforme demonstrado no decorrer do trabalho, não são inteligentes e, tampouco, artificiais⁶) por determinados resultados, como se o processo de “escolha” algorítmica fosse algo completamente independente da vontade e das escolhas humanas.

Não é possível dissociar os termos trabalhados nesse tópico da ação do ser humano. Um exemplo claro de como esse contato é fundamental para os resultados obtidos com o uso de uma ferramenta reside no caso do polêmico Chatbot Tay, desenvolvido pela Microsoft.

A ferramenta citada foi desenvolvida para realizar interações e aprender com jovens com idade entre 18 e 24 anos. A inteligência artificial acabou sendo influenciada por internautas por meio do Twitter e, com pouco tempo, tornou-se agressiva e preconceituosa (Rodrigues, 2016).

Segundo o executivo da Microsoft responsável, a ferramenta estava preparada e protegida contra abusos e erros, mas devido a uma vulnerabilidade específica no

⁶ Para o professor da USP Ernane Costa, há uma definição quase mítica quanto ao que seria uma inteligência artificial. Para o professor, é preciso definir antes o que é inteligência e o que é artificialidade, tendo em vista que é complexo para a ciência contemporânea responder o que é uma inteligência sem incorporar toda complexidade carregada pelo termo, como sua relação com a ideia de poder (Ramos; Costa, 2022).

sistema, um ataque coordenado alterou de forma completa o propósito do programa em menos de 24 horas (Rodrigues, 2016).

A forma como a interação e interferência humana está relacionada a eventuais danos e violações a direitos que podem ser realizados por esses dispositivos é um tema para os próximos tópicos. No momento, é preciso que, primeiro, seja possível compreender algumas conceituações.

A primeira delas, algoritmo, pode ser compreendida, de forma simples, como uma sequência específica e precisa de diretrizes que, a partir do momento em que é executada por uma determinada pessoa, produz um resultado previsto (Castilho; Silva; Weingartner, 2020, p. 13).

O objetivo de um algoritmo é a resolução de um determinado problema. Trata-se de um método registrado em sequência, um verdadeiro “passo-a-passo” de manipulação de dados para que o referido obstáculo seja vencido. Dessa maneira, é possível afirmar que uma fórmula matemática e até mesmo as instruções de uma receita de pizza ou bolo podem ser consideradas exemplos de algoritmo.

Valendo-se do exemplo da receita de um bolo, é possível afirmar que ela não será necessariamente lida e seguida pela mesma pessoa que a escreveu, mas é preciso que ela esteja escrita e detalhada de uma forma clara o suficiente para que outros possam compreender e executar a tarefa descrita.

Pode-se afirmar, portanto, que os algoritmos possuem um formato padrão que deve atentar aos seguintes detalhes (Castilho; Silva; Weingartner, 2020, p. 14):

- as instruções serão escritas uma em cada linha;
- as instruções serão executadas uma a uma, da primeira até a última linha,
- nesta ordem, a menos que o próprio algoritmo tenha instruções que alterem este comportamento;
- em cada linha, uma instrução faz somente uma coisa;
- tudo o que está implícito deverá ser explicitado.

Algoritmos podem, assim, ser executados não apenas por seres humanos, mas também por programas de computador. Contudo, as instruções devem ser precisas, categorizadas e organizadas, sem uso de frases ou comandos divergentes, imprecisos ou ambíguos.

Um programa de computador pode ser considerado um algoritmo elaborado de uma maneira tão detalhada de forma que até mesmo uma máquina é capaz de “entender” as instruções para que uma diretriz seja executada com a finalidade

desejada. A utilização de algoritmos por softwares possui finalidades variadas, como criptografia, funcionamento de redes neurais, e processamento geral de grandes volumes de dados, como é feito em redes sociais digitais na classificação de conteúdos, na formação de um perfil de usuário para recomendações personalizadas e, conforme apontado na presente tese, no processamento de dados em ferramentas de policiamento preditivo.

Em resumo, tendo em vista que algoritmos são utilizados com base em regras e diretrizes completas e complexas, ele fará uso dessas regras já estabelecidas para a execução de tarefas, como, por exemplo, a análise de todo conteúdo indexado por um buscador (como Google) para entregar o resultado que corresponda ao critério de pesquisa de um usuário.

Conforme demonstrado por Brasil (2022), algoritmos e softwares são extremamente eficientes no processamento de grandes volumes de informação em tempo recorde, de modo que essa categorização seria impraticável e até mesmo impossível de ser realizada por seres humanos.

Com efeito, Brasil (2022) menciona que os algoritmos são essenciais para questões como automação de tarefas (executando determinadas diretrizes de maneira eficiente, como a análise de uma quantidade exacerbada de dados em busca de padrões ou na criação de modelos matemáticos e estatísticos) e previsão de tendências (a análise desse grande volume de dados permite a realização de previsões relacionadas ao comportamento de determinados grupos).

Um ponto essencial para a compreensão do problema tratado na tese reside no fato de que as instruções e os dados fornecidos para o processamento realizado pelas máquinas são criações humanas, de modo que o tipo de informação fornecida poderá afetar drasticamente a forma como as previsões e a criação de padrões são feitas.

Discorrer sobre algoritmos é algo que eventualmente levará ao debate acerca de outro termo: inteligência artificial. Este último termo, por sua vez, possui também diferentes conceitos, mas, conforme já mencionado, inteligências artificiais não são, de fato, inteligências, nem mesmo artificiais.

Alguns dos conceitos do termo IA (Inteligência Artificial) incluem: “[...] a ciência de fazer máquinas que fazem coisas que requerem inteligência, caso feitas pelo homem” ou “[...] a ciência e a engenharia de criação de máquinas inteligentes”

(Andersen, 2018, p. 8). Independente da definição, é possível afirmar que inteligência artificial se refere muito mais a um campo de estudo do que a uma ferramenta.

Também é importante ressaltar que a maior parte dos sistemas legais também não trazem uma definição do que seria uma inteligência artificial. O presente trabalho busca, portanto, trazer uma definição com base em estudos no campo filosófico e científico.

De acordo com o *Oxford Computer Science Explanatory Vocabulary*, a IA é a parte do campo da tecnologia da informação que lida com a criação de programas capazes de resolver problemas que requerem inteligência humana (Eszteri, 2015, p. 58). Não obstante, também é necessário mencionar as quatro abordagens trazidas por Russel e Norvig (citados por Eszteri, 2015, p. 58) em termos do conceito filosófico do desenvolvimento de uma inteligência artificial:

- 1) System thinking as human: This trend considers such systems AI that model the functions of human mind and cognition.
- 2) System acting as human: This approach is linked to the British mathematician Alan Matheson Turing, who claimed in his famous Turing-test that the criteria and purpose of AI is human-like acting.
- 3) Rationally thinking system: This viewpoint considers the purpose of AI in developing more rational or perfect systems than human cognition.
- 4) Rationally acting system: The approach of modern information technology sciences. It does not aim its purpose to create systems that think or imitate human-like behavior just to behave rational (for example to clearly diagnose diseases, predict natural disasters etc.)⁷.

Os mencionados autores apontam para a conjectura de quatro diferentes sistemas de inteligências artificiais, de modo que o conhecimento científico contemporâneo é capaz (e pretende) apenas de se aproximar de um deles. Os três primeiros sistemas, conforme apontado, são aqueles que: 1) pensam como humanos, funcionando como a mente humana e é capaz de cognição; 2) agem como humanos, mas sem a habilidade de cognição; 3) pensam racionalmente, de modo que uma AI

⁷ 1) Pensamento sistêmico como humano: Essa tendência considera sistemas que modelam as funções da mente humana e da cognição.

2) Sistema que age como humano: Essa abordagem está relacionada ao matemático britânico Alan Matheson Turing, que afirmou em seu famoso teste de Turing que o critério e o propósito da Inteligência Artificial é agir de forma semelhante a humana.

3) Sistema que pensa racionalmente: Essa visão considera o propósito da Inteligência Artificial em desenvolver sistemas mais racionais ou perfeitos do que a cognição humana.

4) Sistema que age racionalmente: A abordagem das ciências modernas de tecnologia da informação. Seu objetivo não é criar sistemas que pensem ou imitem o comportamento humano, apenas para se comportarem de forma racional (por exemplo, para diagnosticar claramente doenças, prever desastres naturais etc.) (tradução livre).

poderia, em tese, desenvolver um sistema de pensamento ainda mais racional que o ser humano. A abordagem existente e pretendida, contudo, diz respeito apenas aos sistemas que agem racionalmente.

A ideia desse sistema não é a imitação do comportamento humano, mas um agir de forma coerente e racional, como em situações de diagnóstico de doenças e antecipação de desastres naturais.

Em conformidade com a definição acima, Dániel Eszteri (2015, p. 58-59) ainda menciona o conceito de John R. Sarle a respeito de Inteligência Artificial fraca e forte. Enquanto a AI fraca age como se fosse inteligente, apesar de não haver qualquer evidência de que haja uma mente individual, a AI forte é considerada um sistema que realmente pensa e possui sua própria consciência, algo que o conhecimento científico hoje não é capaz de alcançar, ficando reservado a obras ficcionais.

Portanto, quando falamos em inteligência artificial, da forma como existe nos dias de hoje, como campo de estudo científico, falamos em inteligência artificial fraca, capaz de obedecer a comandos conforme uma programação antecipada, ainda que o sistema seja capaz de “aprender” por meio do que é conhecido como aprendizado de máquina ou *machine learning*⁸. Nesse último caso, o que se entende como aprendizado nada mais é que uma programação que permite que o sistema seja capaz de reconhecer padrões com base em informações já existentes em um banco de dados.

Finda a classificação, é importante mencionar que não apenas o estudo, mas também os usos de inteligências artificiais, saíram da pura especulação científica e, atualmente, acabou por se tornar parte da vida cotidiana. Os usos de ferramentas relacionadas a IA variam. Enquanto alguns programas tentam prever como o mercado afeta o funcionamento das bolsas de valores, outros mostram qual o caminho mais barato e mais curto de carro para a uma localidade próxima e outros ainda controlam os adversários em jogos de computador que devem ser superados pelo jogador humano (Eszteri, 2015, p. 58).

A verdade é que a inteligência artificial se tornou uma parte importante da vida do ser humano comum. No espaço virtual, os usuários geralmente se comunicam com softwares que são controlados por essas funções artificiais.

⁸ O aprendizado de máquina (ou *machine learning*) é uma das áreas mais estudadas no campo da inteligência artificial, permitindo a identificação de critérios com base em casos e experimentações prévias, tal como ocorre com a inteligência humana (Santos, et. al., 2019).

O perigo em relação à existência das inteligências artificiais está longe do que apontam os apocalípticos filmes ou obras de ficção científica (apesar do alerta de desenvolvedores como Yudkowsky), mas isso não significa que não haja riscos em relação ao seu uso e criação. Muito embora hoje sejam discutidos muitos problemas relacionados à IA, como a criação de textos, imagens e toda discussão quanto à propriedade intelectual e eventual diminuição de empregos humanos, a presente tese apresenta um risco muito mais específico, como o de um indivíduo negro ser identificado como suspeito a partir de uma análise algorítmica de reconhecimento facial feita por uma inteligência artificial.

2.1.4 Uma definição de policiamento preditivo

A previsão de crimes por meio de ferramentas de aprendizado de máquina é algo naturalmente complexo, tendo em vista as inúmeras tarefas envolvidas, como mineração de dados com base em dados históricos de crimes. Não se pode, assim, afirmar que todas as ferramentas de policiamento preditivo funcionam da mesma maneira.

O presente tópico tem o simples escopo de apresentar informações gerais sobre esse tipo de ferramenta, tendo em vista que as ferramentas específicas (como PredPol) merecem destaque em um capítulo próprio.

É crucial destacar que o policiamento preditivo vai além das ferramentas básicas de mapeamento *online* que monitoram a ocorrência de crimes. As metodologias estatísticas utilizadas incluem a teoria da repetição e a análise de pontos de acesso a crimes, que se baseiam na premissa de que, uma vez que um crime violento ou de propriedade ocorra em um local específico, há uma alta probabilidade de que ocorra novamente na mesma área. Além disso, o modelo de terreno de risco enfoca a análise geográfica, identificando fatores de risco e características dos locais afetados por crimes, como iluminação pública inadequada e possíveis rotas de fuga (Aguirre; Bradan; Muggah, 2019, p. 3).

Em um estudo realizado pelo Instituto Igarapé, conduzido por Aguirre, Badran e Muggah (2019, p. 3), os pesquisadores afirmam que as metodologias estatísticas utilizadas por essas ferramentas incluem *near-repeat theory*, *crime hot-spot analysis* e *risk terrain model*.

O que se pode afirmar sobre esses métodos, é que a Teoria da repetição (*Near-repeat theory*) propõe a ideia de que crimes (em especial aqueles violentos e ligados à propriedade) possuem a tendência de ocorrerem perto de locais onde crimes parecidos já aconteceram (se um telefone celular é roubado em uma rua de um bairro, há, segundo a ideia apresentada, uma maior chance de que outros aparelhos sejam roubados no mesmo local. A análise de Pontos de acesso a crimes (*Hot-spot analysis*) envolve a identificação de pontos onde crimes podem ocorrer com maior frequência a partir de dados pregressos. Por sua vez, o Modelo de terreno de risco (*Risk terrain model*) tem o escopo de identificar e analisar fatores geográficos que podem contribuir para a ocorrência de crimes (como pontos de fuga).

Não obstante, também se sustenta que o policiamento preditivo é baseado em várias teorias estabelecidas sobre comportamento criminal e oportunidades de crime, que explicam a concentração e repetição de crimes, assim como por que os crimes ocorrem em alguns lugares e não em outros.

Dentre essas teorias, a Teoria das atividades rotineiras (*Routine activity theory*) afirma que o crime depende de vários fatores, incluindo a motivação dos infratores, alvos adequados e a ausência de guardiões capazes. a Teoria da escolha racional (*Rational Choice Theory*) destaca que os criminosos tomam decisões racionais baseadas em oportunidades e custos estimados, como a possibilidade de serem presos e punidos. já a Teoria dos padrões de crime (*Crime Pattern Theory*) explica por que, quando e onde os crimes ocorrem, focando nas interseções e semelhanças entre vítimas e perpetradores.

Com base em uma combinação de tais teorias, os *softwares* de policiamento existentes desenvolvidos e aplicados por departamentos de polícia pelo mundo podem ser classificados em duas categorias: 1) sistemas que predizem a localização do crime; e 2) sistemas que preveem os prováveis autores e vítimas de crimes (Aguirre; Bradan; Muggah, 2019, p. 4).

O primeiro método utiliza relatórios policiais pregressos, chamadas de emergência, previsões meteorológicas e informações sobre grandes eventos públicos para calcular a probabilidade de crimes em determinados locais e momentos. Já o segundo método, que é considerado pelos autores como mais controverso, analisa dados de prisões, registros criminais, perfis em redes sociais, e informações demográficas como localização, raça, idade e gênero. Este método gera uma lista de indivíduos de alto risco, considerados como potenciais envolvidos em crimes futuros.

Dentre os sistemas mais utilizados, provavelmente o mais conhecido é o PredPol, uma plataforma de previsão de crimes baseada em localização que aplica a mencionada teoria da repetição e processa dados de eventos criminais passados. Foi utilizado em diversas áreas, incluindo Los Angeles, Atlanta e Kent (no Reino Unido).

Outro sistema (também controverso) é o Palantir, que inicialmente envolvia ferramentas de diagnóstico para o exército dos Estados Unidos durante operações no Afeganistão. Sua solução aplica reconhecimento de padrões, combinando dados de crimes, dados de prisões e análise de redes sociais (Aguirre; Bradan; Muggah, 2019, p. 4).

Esses sistemas são analisados com o devido aprofundamento em capítulo específico. O que se pode afirmar, também em linhas gerais, é que essas ferramentas serviram como base para outras em várias partes do mundo. Uma das primeiras ferramentas de previsão de crimes desenvolvidas na América Latina foi criada no Chile pelo Centro de Análise e Modelagem em Segurança da Universidade do Chile. Atualmente, a previsão de crimes é comum em países como Dinamarca, Itália e Japão, e novos projetos estão em andamento na Argentina, Brasil, Colômbia e Uruguai. Na Colômbia também foi criado um projeto na cidade de Bogotá, envolvendo a Universidade Nacional e a empresa privada Quantil, está estabelecendo uma plataforma de previsão de crimes que busca evitar a reprodução de vieses e perfis discriminatórios (Aguirre; Bradan; Muggah, 2019, p. 4).

Com consonância com as informações analisadas, é possível conceituar essas ferramentas, de modo geral, como sistemas que utilizam dados históricos e teorias criminológicas para prever onde e quando crimes podem ocorrer, bem como quem são os possíveis autores e vítimas, de modo a otimizar a alocação de recursos de segurança pública.

É importante destacar que mesmo com teorias bem delimitadas como as citadas, as ferramentas de policiamento preditivo ainda estão atreladas em casos de racismo algorítmico. Isso ocorre porque esses algoritmos dependem de dados que podem refletir preconceitos e discriminações sistêmicas (conforme demonstrado adiante).

2.2 TÉCNICA, TECNOLOGIA E IDEOLOGIA: OS BANCOS DE DADOS QUE ALIMENTAM A VIOLAÇÃO DE DIREITOS

Até o presente momento, a tese procurou estabelecer conceitos importantes ligados ao estudo das tecnologias de informação e comunicação. O entendimento sobre esses termos é fundamental para uma compreensão de um complexo fenômeno estabelecido como objeto principal da tese: a existência do racismo algorítmico e sua possível contribuição para a violação de direitos de indivíduos pertencentes a grupos periféricos.

Assim, esses termos de grande relevância (tecnologia, ciberespaço, algoritmos e inteligência artificial) foram preliminarmente desenvolvidos para que seja possível, no presente tópico, apontar como os atuais artefatos tecnológicos não apenas possibilitam a inserção do racismo no ambiente virtual, como também tornam possível a replicação da lógica racista em “ações não humanas”, por meio dos bancos de dados que alimentam determinados instrumentos algorítmicos.

Para que seja possível demonstrar essa ocorrência, é necessário que sejam aprofundados dois pontos: 1) a forma como as relações humanas foram alteradas a partir do constante e crescente uso das ferramentas mencionadas; 2) o que são direitos fundamentais.

Demonstrar-se-á que muito se cogitou a respeito do aspecto positivo da interação entre homem e máquina, até mesmo sobre o provável papel que as tecnologias poderiam ter no processo de busca por efetivação de direitos e garantias fundamentais. O que será apontado, contudo, é que novas tecnologias são desenvolvidas a todo momento com base em já conhecidas técnicas de exclusão e dominação.

2.2.1 As relações humanas e a influência dos aparatos tecnológicos

É imperativo destacar que a ideia de que as relações humanas são moldadas pelo uso crescente das tecnologias de comunicação e informação não é algo recente. Tem-se, como um dos exemplos mais evidentes, a metáfora da “aldeia global”, uma expressão criada por Marshall McLuhan ainda na década de 1960.

O autor canadense explorou o que se compreende como tribalismo (tal como visto em diversos grupos humanos desde a antiguidade) e extrapolou sua essência

para criar o conceito de uma “aldeia” na qual o uso das tecnologias de comunicação e informação exerceriam o papel de encurtar as distâncias, não apenas no aspecto geográfico, mas também cultural. Dessa forma, sustentou que o uso dessas novas ferramentas levaria os indivíduos a viverem em uma espécie de aldeia planetária na qual os conhecimentos são substituídos de forma célere e o papel que antes era exercido pelo líder tribal passaria a ser ocupado por uma rede de fios e de fluxos formados pelos novos aparatos tecnológicos (McLuhan, 1969, p. 91).

É importante destacar que o autor já previa o enfraquecimento do poder exercido pelos Estados nacionais a partir do nascimento de uma nova força global. Conforme mencionou (McLuhan, 1969, p. 44), eventualmente, os tradicionais grupamentos cívicos, estatais ou até mesmo nacionais tornar-se-iam impraticáveis.

Não apenas McLuhan previa essa espécie de evolução conjunta entre homem e máquina. Em uma conjectura baseada no que hoje podemos compreender como “internet das coisas” (como já foi mencionado anteriormente nesse mesmo capítulo), Philip Howard acreditava no surgimento de um período cunhado de *pax technica* (Magrani, 2018, p. 91).

No entendimento do autor, os sistemas de governo soberanos da contemporaneidade eventualmente seriam substituídos por “sociotecnocracias” embasadas na ampla utilização de dispositivos móveis. Dessa maneira, não seria mais necessário que as pessoas expressassem suas crenças e valores, tendo em vista que os dados comportamentais fariam esse trabalho automaticamente (Magrani, 2018, p. 91).

Quanto à *pax technica* de Howard, não parece exagerado afirmar que ela se faz cada vez mais presente na medida em que cresce a automatização de dispositivos que muitas vezes tratam os dados pessoais de seus titulares sem o devido consentimento.

Pode-se afirmar que a aposta de McLuhan em uma profunda mudança nas culturas e valores, diante do surgimento de instrumentos globais, tornou-se, ao menos em parte, algo real. Por exemplo, é possível mencionar o surgimento de certas tecnologias e o tempo que levou para que elas se tornassem ferramentas cotidianas. O uso da rede mundial de computadores, muito embora seja mais recente que ferramentas como rádio e televisão, já está intimamente acoplado ao dia a dia da maioria das pessoas.

Portanto, a rede mundial de computadores é, dentre os instrumentos de comunicação criados nas décadas recentes, o que exerce maior influência na vida cotidiana e profissional dos indivíduos.

A existência de uma ferramenta tecnológica ligada à comunicação eletrônica e a sistemas de informação faz com que se torne cada vez mais possível dissociar a proximidade espacial do desempenho de certas funções, como trabalho, compras e até entretenimento (Castells, 2002, p. 493). Dessa forma, para que um indivíduo realize suas tarefas rotineiras, em relação a grande parte das ações, não é necessário que ele se locomova a determinado local. Ele pode comprar produtos por uma página da internet, realizar seu trabalho e se comunicar com outros indivíduos que possuam gostos, valores ou ideias similares.

Essa dissociação entre os locais é algo que se tornou tão acentuado que Castells (2002, p. 483) menciona que não foram poucos os futurologistas que predisseram o fim da sociedade tal como é hoje conhecida. Entretanto, o mencionado autor sustenta a possibilidade de existência de uma nova sociedade que é baseada no conhecimento e organizada ao redor de redes e fluxos. Consoante menciona, a sociedade está inteiramente construída em volta dessas diferentes formas de fluxos, que são entendidos como:

[...] sequências intencionais, repetitivas e programáveis de intercâmbio e interação entre posições fisicamente desarticuladas, mantidas por atores sociais nas estruturas econômica, política e simbólica da sociedade (Castells, 2002, p. 501).

Dessa forma, é possível apontar a existência de diferentes fluxos, sejam eles de capital; informação; tecnologia; ou mesmo interação social. Eles possibilitam que esses elementos se movam de um ponto a outro e, graças aos avanços tecnológicos mais recentes, muitas vezes essa transição ocorre de forma instantânea.

Entre os efeitos da existência desses fluxos, é possível afirmar que esse processo favorece o surgimento de novos mercados que tenham algumas características peculiares como: 1) a produção de bens imateriais; 2) a possibilidade de negociação com clientes em diferentes locais do globo; e 3) a valorização dos dados pessoais de usuários. Esse tipo de mercado pode ser muito bem ilustrado pelas redes sociais digitais mais tradicionais, como o Facebook; o Instagram e o aplicativo de mensagens WhatsApp.

É preciso destacar que tanto governos nacionais e locais, além de empresas e até mesmo movimentos sociais, estão contemporaneamente conectados por meio de uma rede de informações que é capaz de impactar radicalmente a forma como estas realizam suas atividades típicas. Isso acaba refletindo nos mais diversos campos, como a economia, cultura e até política (Lopes; Moreira Júnior, 2019, p. 34).

Em aliança às já conhecidas tecnologias de informação e comunicação (e com o surgimento do *Big Data*⁹), percebe-se contemporaneamente o aumento de ferramentas que se valem de algoritmos e processos de *machine learning*. Essas novas possibilidades tecnológicas favorecem o surgimento das chamadas tecnologias preditivas.

Se os algoritmos permitem a antecipação de certos fatos, não é difícil imaginar as diversas implicações desse tipo de tecnologia em diferentes ambientes, tais como no mercado financeiro, no desenvolvimento de conteúdo cultural (sejam filmes, séries ou músicas) e até mesmo no campo da segurança pública. Muitas vezes tais tecnologias são tratadas como ferramentas neutras, alheias à ingerência humana, um verdadeiro resultado apenas da inovação científica. Certamente esse tipo de argumentação possui incontáveis falhas.

2.2.2 Direitos humanos e fundamentais a partir de uma teoria crítica

O presente trabalho, conforme já apontado, busca demonstrar a eventual existência de uma relação entre o fenômeno do racismo algorítmico e a perpetuação de um estado de violação de direitos fundamentais (e um retorno às bases da criminologia positivista e racista do séc. XIX). É certo que a existência de um racismo estrutural (digital ou não) é, por si, uma violação aos mencionados direitos e garantias. Antes que o trabalho comece a abordar a explicação e análise dos fenômenos citados, é preciso que o tema “direitos fundamentais e direitos humanos” seja bem definido.

Usualmente, os termos direitos humanos e direitos fundamentais são vistos como figuras sinônimas. Uma definição jurídica tradicional consiste em afirmar que

⁹ Trata-se também de um importante conceito para o presente trabalho. O termo pode ser caracterizado como um conjunto de dados muito grande e complexo para serem analisados até mesmo por um *software* de processamento de dados tradicional. Graças ao uso cada vez maior da rede mundial de computadores pela sociedade e com a rápida melhoria na capacidade de processamento de computadores, ocorreram avanços significativos no setor de Inteligência Artificial nos últimos dez anos (Andersen, 2018, p. 8).

direitos humanos são direitos inerentes à condição humana e anteriores até mesmo a uma eventual afirmação por meio do direito positivo. São direitos que transcendem os direitos fundamentais, tendo em vista que seu conteúdo é dotado de uma ordem de princípios universais. Afirma-se, assim, que os direitos humanos, incorporados pela Constituição de um Estado, acabam por ganhar a condição de direito fundamental (Alvarenga, 2019).

Esses termos também costumam remeter a conhecidos instrumentos normativos. O primeiro deles é a Declaração de Direitos do Homem e do Cidadão (1789) que inclui, em sua introdução, o seguinte parágrafo:

Os representantes do povo francês, constituídos em Assembleia Nacional, considerando que a ignorância, o esquecimento ou o desprezo dos direitos do homem são as únicas causas das desgraças públicas e da corrupção dos Governos, resolveram expor em declaração solene os Direitos naturais, inalienáveis e sagrados do Homem, a fim de que esta declaração, constantemente presente em todos os membros do corpo social, lhes lembre sem cessar os seus direitos e os seus deveres; a fim de que os actos do Poder legislativo e do Poder executivo, a instituição política, sejam por isso mais respeitados; a fim de que as reclamações dos cidadãos, doravante fundadas em princípios simples e incontestáveis, se dirijam sempre à conservação da Constituição e à felicidade geral (Declaração de Direitos do Homem e do Cidadão, 1789).

O segundo texto comumente lembrado é a Declaração Universal dos Direitos Humanos (1948) que, de forma similar, estabelece em seu preâmbulo os seguintes dizeres:

Agora portanto a Assembleia Geral proclama a presente Declaração Universal dos Direitos Humanos como o ideal comum a ser atingido por todos os povos e todas as nações, com o objetivo de que cada indivíduo e cada órgão da sociedade tendo sempre em mente esta Declaração, esforce-se, por meio do ensino e da educação, por promover o respeito a esses direitos e liberdades, e, pela adoção de medidas progressivas de caráter nacional e internacional, por assegurar o seu reconhecimento e a sua observância universais e efetivos, tanto entre os povos dos próprios Países-Membros quanto entre os povos dos territórios sob sua jurisdição (Organização das Nações Unidas, 1948).

Os citados textos possuem inúmeras especificidades e inúmeras similaridades. Entre essas últimas, é possível e necessário destacar as seguintes: 1) a elaboração após a ascensão burguesa por grupos formados a partir de uma elite; 2) a afirmação de que se trata de um texto universal, capaz de alcançar todos os grupos em todos os tempos em todo o mundo.

A definição utilizada na presente tese parte de uma análise cuidadosa das chamadas teorias críticas dos direitos humanos. Para Sánchez Rubio (2010, p. 58), é possível afirmar que direitos humanos é um conceito complexo que implica em elementos normativos e não normativos que vão além de processos de lutas gerados e institucionalizados pela burguesia no caminhar que levou ao surgimento da modernidade.

Nesse sentido, é possível afirmar que uma das principais objeções feitas no campo da teoria crítica dos direitos humanos ao estudo tradicional do tema é que este aborda essas garantias de forma abstrata e descontextualizada, sem levar em conta as relações de poder que influenciam sua concepção e efetivação na prática.

É necessário frisar, assim, que os direitos humanos não são simplesmente princípios universais a serem aplicados de forma igual para todos, mas são moldados por interesses políticos e econômicos dominantes em uma determinada sociedade. Não é incomum, por exemplo, que a discussão quanto aos direitos humanos acabe por justificar a intervenção militar em países estrangeiros, a fim de promover os interesses geopolíticos de uma nação, ou para legitimar a desigualdade econômica em nome da liberdade individual.

Na perspectiva de Herrera Flores (2009, p. 173), após o século XVIII, ocorreram inúmeros processos de resistência vinculados aos efeitos provocados pela ascensão do modo de produção capitalista. A partir do entendimento do autor, é necessário estabelecer que o conceito de direitos humanos abordado na presente tese não possui conceituação semelhante à concepção adotada até a primeira metade do século passado.

A definição tradicional, ao pretender estender a ideia de direitos humanos a todo e qualquer agrupamento humano, acaba por equivocar-se na consideração da humanidade como um coletivo homogêneo que buscaria a maximização de seus interesses individuais (Moreira; Moreira Júnior, 2017, p. 286).

Com efeito, o conceito adotado na tese não analisa os direitos humanos como uma benesse global homogênea, mas como um verdadeiro resultado de processos de luta contra as violações advindas do modo de produção capitalista e pela busca de dignidade por grupos marginalizados.

Na perspectiva de David Sánchez Rubio (2022, p. 54-55), o direito estatal (considerado legítimo) não é capaz de promover plenamente as garantias constitucionais (um tema mais bem elaborado no terceiro capítulo). O autor espanhol

sustenta que as estruturas legais existentes e institucionalmente estabelecidas acabam por refletir e reforçar as desigualdades e relações de poder presentes na sociedade.

Dessa maneira, o direito institucional está estritamente vinculado a sistemas legais e políticos historicamente gerenciados pelos interesses econômicos, políticos e sociais dominantes. O resultado é a criação de normas e práticas incapazes de garantir integralmente a igualdade e dignidade.

Essa análise crítica é relevante para a tese, uma vez que Sánchez Rubio (2022) traz ao debate a ideia de direitos humanos instituintes, argumentando que é necessário um enfoque mais amplo, capaz de incorporar as perspectivas e vozes de pessoas afetadas por uma série de violações.

É nesse sentido que surgem projetos como o “DecolonizAI” (destacado no último capítulo da pesquisa), que emergem com o propósito de empregar a inteligência artificial para questionar os algoritmos e conjuntos de dados capazes de perpetuar desigualdades e injustiças sociais. A meta desses projetos é promover o conhecimento a respeito do desenvolvimento de IA, para que os cidadãos possam pensar na criação de suas ferramentas conforme suas necessidades, de modo a promover e efetivar direitos (como no caso de ferramentas que analisam redes sociais digitais e identificam a prática de discurso de ódio, racismo e outras violações).

Em consonância com Sánchez Rubio, é possível mencionar as ideias de Frantz Fanon quanto à possibilidade de uma abordagem anticolonial na utilização da tecnologia. Conforme destacado adiante, Fanon (2008, p. 22) investigou as interações de autoridade e subjugação que são intrínsecas ao colonialismo (que se manifesta atualmente até mesmo pela seara digital). É crucial, portanto, ressaltar a relevância de questionar as estruturas contemporâneas, abrangendo tanto as técnicas quanto os instrumentos que mantêm a sistemática mitigação a direitos.

2.2.3 Tecnologia e a possibilidade de violação a direitos

Com base na trabalhada ideia de que ferramentas podem ser pensadas e criadas (conscientemente ou não) para a violação de direitos, o presente tópico busca introduzir o conceito de policiamento preditivo. Esse modelo tecnológico está ligado à utilização da inovação das tecnologias algorítmicas no intuito de identificar áreas

consideradas de maior risco para a ocorrência de crimes e indivíduos potencialmente capazes de cometer delitos.

Os modelos de PP (Policiamento Preditivo) se baseiam em algoritmos que analisam diversos tipos de dados, tais como: históricos de criminalidade, dados demográficos e outros fatores considerados aptos na tarefa de apuração da probabilidade de um crime ocorrer em uma determinada área ou ser praticado por uma determinada pessoa. O uso desse tipo de ferramenta permite que a polícia utilize essas informações para direcionar suas atividades de policiamento, sejam elas ostensivas ou de vigilância.

Muito embora o policiamento preditivo possa ser visto como uma abordagem inovadora para prevenir a criminalidade sob uma possível argumentação de neutralidade tecnológica, ele também apresenta alguns pontos críticos justamente porque não é possível atestar qualquer tipo de imparcialidade. Pelo contrário, conforme demonstrado no decorrer do trabalho, esse tipo de ferramenta é fundado em pré-conceitos que replicam digitalmente atitudes racistas (em especial aquelas praticadas por órgãos de Estado) e perpetuam a seletividade penal.

Quanto à existência (ou não) de uma “imparcialidade tecnológica” na utilização desse tipo de ferramenta preditiva, é preciso mencionar duas questões cruciais: 1) a dificuldade de compreensão de como verdadeiramente ocorre o tratamento de dados e alimentação do banco de dados das tecnologias aqui estudadas; e 2) a tradicional incorporação de subjetividade na criação de qualquer tipo de técnica (tema abordado no tópico seguinte).

Quanto ao primeiro ponto, Lucena (2019, p. 2) menciona que a dificuldade de compreender o funcionamento dos algoritmos e estruturação do banco de dados ocorre porque as empresas não divulgam esses elementos em sua totalidade. As corporações alegam que agem na garantia legal de proteção de segredos comerciais, no intuito de impedir que essas informações venham a ser utilizadas por empresas rivais, mas garantem que há respeito a cláusulas contratuais e com os dados pessoais coletados dos indivíduos.

Para Lucena (2019, p. 2), trata-se de uma argumentação frágil, haja vista o histórico de casos concretos de empresas que ultrapassam as cláusulas de proteção

e tratamento de dados na intenção de obter maiores lucros ou maior controle de mercado, como ocorreu no caso Cambridge Analytica¹⁰.

É verdade que os próprios Estados têm buscado (ao menos nos anos recentes) uma maior proteção dos dados de seus cidadãos, como no caso da *European General Data Protection Regulation* na União Europeia ou até da Lei Geral de Proteção de Dados (LGPD) no Brasil. Ocorre que mesmo com a edição de legislações específicas sobre o tema, não há qualquer garantia de que informações pessoais (incluindo informações sensíveis¹¹) não possam ser utilizadas para violar direitos fundamentais ou perpetuar sua já existente violação. Conforme menciona Lucena (2019, p. 2-3):

Há, mesmo com a criação de leis e agências reguladoras, certo uso e aplicação de dados pessoais que nos assombra frente às potencialidades de seu mau uso, e que apresenta em diversos casos, difíceis dificuldades de análise, compreensão e transparência à sociedade civil. São os atos em que o Estado, seja como o verdadeiro coletor dos dados, ou em parceria com empresas privadas, aplica, muitas vezes de forma irregular e descumprindo tratados internacionais, a interpretação de dados coletados por si, ou por intermédio de terceiros, na realização de políticas públicas.

Essa informação é de extrema relevância, tendo em vista que, no Brasil, a própria legislação de proteção de dados pessoais (Lei n. 13.709/18) é cristalina ao afirmar que não há sua aplicabilidade em situações ligadas a tratamentos de dados (sejam eles quais forem) em tarefas vinculadas a atividades de investigação e repressão penal¹².

Dessa maneira, mesmo por meio de uma análise preliminar, é possível afirmar que há uma preocupação quanto à análise de dados por meio de ferramentas de

¹⁰ O caso se trata de um escândalo político ocorrido em 2018 envolvendo a empresa britânica de consultoria política Cambridge Analytica e o uso indevido de dados de usuários do *Facebook*. A empresa teria coletado dados pessoais de milhões de usuários da rede social sem o consentimento explícito e usado esses dados para criar perfis psicológicos detalhados e personalizados no intuito de influenciar a opinião pública em favor de campanhas políticas e eleitorais, como a campanha presidencial de Donald Trump nos Estados Unidos e o referendo do Brexit no Reino Unido. O caso levantou preocupações sobre a privacidade de dados e o uso indevido de informações pessoais (Fornasier; Beck, 2020, p. 184).

¹¹ Segundo a LGPD, dados pessoais sensíveis são dados que se relacionam a certas características da personalidade da pessoa humana, mas também estão ligados a certas escolhas pessoais. Para Pinheiro (2020, p. 20), são exemplos de dados pessoais sensíveis: “[...] origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente a saúde ou a vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural”.

¹² O art. 4º da LGPD menciona expressamente as situações nas quais a lei não terá aplicação. Entre as hipóteses é interessante mencionar o que consta no inciso I, alínea “b”, que inclui as atividades ligadas a questões como: segurança pública; defesa nacional; segurança de Estado; ou atividades de investigação e repressão de infrações penais (Brasil, 2018).

policiamento preditivo perpetuar uma série de violações, como discriminação racial e social. Conforme abordado no decorrer do trabalho, os dados históricos refletem as práticas de policiamento passadas, que muitas vezes foram baseadas em preconceitos raciais. O uso, consoante apontam os estudos, têm levado a uma concentração excessiva de policiamento em áreas e pessoas que já estão sob vigilância excessiva.

Não obstante, outra crítica que pode ser inicialmente levantada em relação ao policiamento preditivo reside no fato de ele pode criar um clima de vigilância constante e de invasão da privacidade. O uso de tecnologias de vigilâncias, como câmeras de reconhecimento facial e outros sensores, podem levar a uma percepção que órgãos estatais (e, em alguns casos, até mesmo privados) estão constantemente observando as pessoas e prejudicando suas liberdades civis.

Com base no que foi afirmado até o momento, o uso de tecnologias algorítmicas, em especial as de policiamento preditivo, gera algumas preocupações: 1) o fato de que mesmo com a criação de legislações e agências reguladoras para tratar de dados pessoais, o potencial mau uso de certas tecnologias podem dificultar a fiscalização e punição de eventuais violações a direitos; e 2) a dificuldade em se saber se determinados órgãos estatais e entidades privadas estão realmente preocupados com a eventual violação desses direitos, principalmente quando se trata de informações relacionadas a indivíduos pertencentes a populações periféricas e marginalizadas.

Importa dizer que o tema tem seu debate ampliado no segundo capítulo, no qual são analisados modelos reais de PP utilizados nos cenários nacional e internacional.

2.2.4 Ausência de neutralidade na concepção de ferramentas tecnológicas

O tópico anterior, além de introduzir o objeto central da pesquisa (policiamento preditivo), foi responsável por apontar um questionamento indispensável ao presente trabalho: a suposta existência de uma “imparcialidade tecnológica” na criação e utilização de ferramentas preditivas.

Assim, antes que seja discutida noções sobre racismo digital ou algorítmico, é preciso que a própria noção sobre “imparcialidade” nas ferramentas tecnológicas seja discutida. Para tanto, a obra “Técnica e ciência como ‘ideologia’”, de Jürgen

Habermas, pode servir de ponto de partida para uma reflexão a respeito da responsabilidade dos desenvolvedores e usuários dos mais diversos produtos tecnológicos, em especial quanto à forma como os algoritmos são treinados e utilizados, de modo a garantir que tais artefatos não sejam usados e nem desenvolvidos para justificar ou mesmo perpetuar relações de poder pré-existentes.

Certamente não é possível identificar temas como algoritmos ou inteligência artificial no clássico texto do autor alemão, mas o termo “racionalidade” é abordado como um conceito introduzido por Max Weber com o intuito de determinar a forma de uma atividade econômica do capitalismo; das relações de direito privado da burguesia e da chamada dominação burocrática (Habermas, 1980, p. 313).

É possível estabelecer, a partir do texto, que a ideia de racionalidade (ou racionalização) significa a existência de uma amplificação de setores sociais que são submetidos a padrões de uma decisão racional vinculada a uma institucionalização da ciência e da técnica. Dessa maneira, pode-se afirmar que a técnica (e ciência) são elementos capazes de adentrar em diversos campos da sociedade, orientando e até mesmo determinando seu funcionamento (o que, inclusive, corrobora com o pensamento de Veraszto a respeito dos conceitos de tecnologia trazidos na primeira parte do capítulo atual).

A primeira crítica a essa noção de racionalização é apontada por Habermas a partir das ideias de Herbert Marcuse, que estaria convencido de que esse processo de estabelecimento de uma “racionalização” é responsável não por disseminar na sociedade uma racionalidade (no sentido tradicional do termo), mas por disseminar um modelo de dominação política, conforme se pode perceber a partir o trecho a seguir:

O método científico que levou à dominação cada vez mais eficaz da natureza passou assim a fornecer tanto os conceitos puros, como os instrumentos para a dominação cada vez mais eficaz do homem pelo homem através da dominação da natureza [...]. Hoje a dominação se perpetua e se estende não apenas através da tecnologia, mas enquanto tecnologia, e esta garante a formidável legitimação do poder político em expansão que absorve todas as esferas da cultura (Habermas, 1980, p. 315).

Tomazella (2011, p. 1) compreende da mesma maneira ao afirmar que o conceito de racionalidade proposto por Max Weber é baseado na ideia de planejamento sistemático da vida através de processos burocráticos e do direito privado burguês. Isso implica a utilização de técnicas e meios de comunicação para

alcançar objetivos previamente estabelecidos. Essa forma de racionalidade é, assim, a base para a inserção do poder nos diversos aspectos da sociedade, como nas estruturas urbanas, na divisão social do trabalho e no processo industrial.

Tomazella (2011, p. 2) corrobora com a visão de Marcuse de que a continuidade do domínio político da perspectiva tecnicista é inevitável. Ele argumenta que o desenvolvimento do sistema sob a lógica da técnica tem como objetivo interno a sua própria manutenção e expansão, o que claramente faz com que as forças produtivas, representadas pelos trabalhadores, gastem a maior parte de suas vidas realizando tarefas repetitivas e monótonas (que não lhes oferecem nenhum sentido ou propósito como seres humanos). Isso os impede de desfrutar dos benefícios proporcionados pelo próprio desenvolvimento técnico que eles mantêm em movimento, criando um ciclo vicioso em que as pessoas são subordinadas ao sistema e a sua lógica tecnológica.

Com base nos textos apresentados, é possível concluir que a técnica e tecnologia são verdadeiramente responsáveis pela possibilidade de privação do homem de sua liberdade e de sua dominação por outros indivíduos.

Mais uma vez, a discussão no texto de Habermas não trata de tecnologias de comunicação e informação e a técnica mencionada diz respeito a um aparato técnico responsável por amplificar as comodidades da vida e aumentar a produtividade do trabalho (Habermas, 1980, p. 315), mas a utilização da tecnologia como instrumento de controle desde sua concepção é algo que está presente no texto e pode ser ilustrada a partir dos fenômenos estudados nesta tese.

No intuito de demonstrar a subjugação humana a partir da técnica, a crítica feita por Marcuse pode ser confirmada a partir da obra *"Code 2.0"*, de Lawrence Lessig. A discussão central do autor estadunidense não é, de forma alguma, discutir técnica e tecnologia da mesma maneira que a obra habermasiana, mas ao afirmar que os protocolos de um código não são neutros, mas criados por seres humanos, é possível inferir que esses códigos refletem os valores e interesses de seus criadores.

Ao tratar do tema "regulação do ciberespaço", Lessig (2006, p. 20) menciona a necessidade de se observar para além da regulação tradicionalmente concebida, aquela feita por meio de normas e legislações. Regulação deve ser compreendida em um conceito amplo, tendo em vista que no cenário digital as normas são regidas preliminarmente por meio do código e protocolos que geram as operações dos sistemas de computação.

De modo a exemplificar o que foi mencionado, é possível compreender a visão de Lessig no fato de que os criadores (ou arquitetos) de determinado código são aqueles que (antes mesmo de uma lei) possuem o poder de decidir que tipo de comportamentos são ou não permitidos dentro de um sistema.

No contexto de plataformas digitais, por exemplo, são os criadores do algoritmo os responsáveis por decidir que tipos de discursos são permitidos ou proibidos, bem como definir como a moderação é realizada e até mesmo as imagens e mensagens que aparecem na tela inicial dos usuários.

O fato de que a primeira regulamentação está nas mãos dos desenvolvedores da tecnologia explica como fenômenos (tais como a seletividade penal e o próprio racismo estrutural) estão inseridos no próprio design das tecnologias de policiamento preditivo, soterrando o mito de uma neutralidade tecnológica.

Assim, será a programação que determinará como os seres humanos, dentro de um determinado sistema, ou a partir dela, interagirão. Nos termos de Lessig (2006, p. 6):

We can build, or architect, or code cyberspace to protect values that we believe are fundamental. Or we can build, or architect, or code cyberspace to allow those values to disappear. There is no middle ground. There is no choice that does not include some kind of building. Code is never found; it is only ever made, and only ever made by us¹³.

Com essas palavras, Lessig enfatiza que não se pode evitar a responsabilidade de projetar a tecnologia de forma consciente e intencional. Um código que define um artefato tecnológico (como a internet, por exemplo) não pode simplesmente ser encontrado, mas construído.

Retornando a Habermas e Marcuse, a junção entre técnica (e por que não dizer “código”?) e dominação, racionalidade e opressão não pode ser vista de outra forma a não ser pelo fato de que ciência e tecnologia exigem um verdadeiro projeto de mundo determinado pelo interesse de uma classe e por uma situação histórica (Habermas, 1980, p. 316).

Habermas, contudo, deixa claro que os aparatos tecnológicos são indiferentes a fins políticos (1980, p. 319), de modo que um certo computador pode servir a uma

¹³ Podemos construir, arquitetar ou codificar o ciberespaço para proteger valores que acreditamos serem fundamentais. Ou podemos construir, arquitetar ou codificar o ciberespaço para permitir que esses valores desapareçam. Não há meio-termo. Não há escolha que não envolva algum tipo de construção. O código nunca é encontrado; ele só é feito, e feito sempre apenas por nós (tradução livre).

administração de Estado capitalista ou socialista, tal como uma pesquisa nuclear pode ser eficiente tanto no aspecto marcial ou pacifista de uma sociedade. Mais uma vez, isso não quer dizer que as tecnologias sejam neutras, haja vista que a própria concepção de um artefato está carregada de intencionalidades e seu “código” é, em verdade, uma forma de regulação e controle. Contudo, é possível afirmar, a partir da perspectiva de Habermas, que diferentes intencionalidades podem alterar até mesmo a finalidade de uma ferramenta.

Essa possibilidade de alteração de finalidades é algo mais bem abordado no terceiro capítulo, mas é importante adiantar que, colaborando com a ideia apresentada, é possível apresentar as ideias de Frantz Fanon quanto ao uso da tecnologia na luta contra o colonialismo e na mobilização das massas colonizadas (Faustino; Lippold, 2023, p. 185).

2.3 O DIREITO E A INTELIGÊNCIA ARTIFICIAL

É importante salientar que apesar do presente trabalho ter como objeto as tecnologias de policiamento preditivo, não é incomum no campo jurídico o uso de ferramentas algorítmicas e inteligências artificiais, seja no campo da segurança pública ou até mesmo nos usos cotidianos do Poder Judiciário. O presente tópico tem o intuito de brevemente demonstrar as ferramentas de IA no âmbito jurídico (antes que o próximo capítulo aprofunde o tema central).

Em 2018, o Supremo Tribunal Federal já havia anunciado a criação de um projeto de IA nomeado VICTOR, fruto de uma pesquisa sobre construção de redes neurais para rastrear em processos temas de repercussão geral, agrupar os processos relacionados por tema e até indicar ao Ministro Presidente da Corte os processos para fins de devolução de recurso à sua origem (Maia filho; Junquilha, 2018, p. 221). Nota-se que a ferramenta VICTOR (ou qualquer outra IA produzida até o momento) não é utilizada no ato de tomar decisões ou mesmo em qualquer momento do julgamento de mérito dos processos, muito embora em outras regiões do globo isso possa ser uma realidade.

Por exemplo, no sistema criminal estadunidense as ferramentas de inteligência artificial já são utilizadas em *evidence-based practices* e em *risk assessment practices*. Com o uso dessas ferramentas, os julgadores são informados sobre eventuais riscos de indivíduos submetidos ao sistema penal em relação a questões

como dosimetria de pena, regimes de cumprimento ou possibilidade de liberdade condicional e outros benefícios (Sulocki, 2020, p. 671).

Em um primeiro momento pode parecer que esse tipo de ferramenta auxiliará não apenas julgadores, mas toda uma cadeia de profissionais relacionados ao sistema de justiça criminal, como agentes penitenciários e psicólogos. Sulocki (2020, p. 672) explica que esses profissionais precisam apenas acessar um score que é obtido na avaliação de cada indivíduo, sendo tais ferramentas baseadas em métodos atuariais de avaliação dos riscos, no intuito de fornecer orientação a partir das análises probabilísticas de reincidência do acusado e sua periculosidade social.

O governo norte-americano tem até argumentado que a utilização de inteligência artificial em decisões de progressão de regime e liberdade condicional ajuda a reduzir o viés humano e aumentar a objetividade nas decisões. Em tese, a IA pode analisar grandes quantidades de dados, incluindo o histórico criminal do indivíduo, o tempo de serviço, a saúde mental e outros fatores relevantes para determinar o risco de reincidência. Essas informações até podem auxiliar os julgadores a tomarem decisões com maior precisão.

Entretanto, conforme demonstrado, essa ferramenta, em sua essência, possui um escopo diverso, tendo em vista que a IA pode perpetuar o viés humano, pois as decisões baseadas em dados históricos podem refletir preconceitos existentes na sociedade. Além disso, há preocupações em relação à transparência e responsabilidade da IA, já que os algoritmos utilizados podem ser complexos e difíceis de compreender, o que pode dificultar a revisão das decisões tomadas.

Sulocki (2020, p. 672) explica essa tendência nefasta com base nas políticas tradicionalmente adotadas pelo país no controle de sua criminalidade, como as políticas de “guerra às drogas” criadas nos idos de 1970 que buscavam, na realidade, controlar penalmente a juventude negra e pobre.

Esse resultado (aprofundamento da discriminação e seletividade do sistema penal, afastando-se da possibilidade do devido processo legal e um eventual julgamento justo) pode ser confirmado por meio da análise dos dados prisionais que se sucederam após a utilização de tais ferramentas.

Conforme pesquisa realizada pela ProPublica (agência não governamental de jornalismo), os dados sobre a utilização da ferramenta COMPAS eram preocupantes. Em análise de mais de dez mil condenações em Broward Conty, os acusados negros eram classificados como indivíduos com um maior risco de reincidência, a ponto de

até mesmo aqueles que não reincidiram no período de dois anos costumavam ser classificados equivocadamente como alto risco de reincidência. Da mesma forma, acusados brancos que reincidiram em um intervalo de dois anos também foram classificados como indivíduos de baixo risco de reincidência em uma proporção quase duas vezes maior que os acusados negros (Sulocki, 2020, p. 679).

É importante mencionar, portanto, os critérios utilizados no cálculo atuarial de probabilidade de reincidência. O que se pode afirmar é que o banco de dados é alimentado por informações de caráter subjetivo que carregam todos os pré-conceitos de seus elaboradores.

No caso específico da COMPAS, os dados são sobre acontecimentos pretéritos ligado a muitos indivíduos, como descumprimentos e detenções anteriores (desde meras abordagens policiais até prisões provisórias). Apesar desses primeiros elementos terem caráter objetivo, a ferramenta também era baseada em critérios subjetivos e privados, como:

[...] “personalidade distorcida” (personality disorder), “personalidade criminosa” (criminal personality) ou ainda “isolamento social” (social isolation), e de cunho privado, como se o indivíduo é solteiro ou casado (marital status/lived with intimate partner for yers), ou possui casa própria (own residence/residential stability) (Sulocki, 2020, p. 679).

O que se percebe é, portanto, que o uso de critérios subjetivos como “personalidade distorcida” ou “personalidade criminosa” em ferramentas utilizadas pelo poder judiciário para avaliar o risco de reincidência de indivíduos pode levar a práticas discriminatórias e seletivas na atuação do Estado. Esses termos são vagos e podem refletir preconceitos e estereótipos culturais sobre certos grupos sociais, como minorias étnicas, lembrando novamente a figura do “criminoso nato” ou “homem delinquente” lombrosiano (Lombroso, 2001, p. 247-248).

Além disso, o critério de “isolamento social” pode ser interpretado como uma forma de punir indivíduos que não se encaixam nos padrões sociais dominantes, como aqueles que vivem em comunidades de baixa renda ou têm dificuldades para se relacionar socialmente. Isso também pode levar a uma criminalização das pessoas mais marginalizadas e vulneráveis da sociedade, ao invés de se concentrar nas causas estruturais da criminalidade.

Os critérios privados também possuem forte caráter discriminatório, tendo em vista que podem ser usados para discriminar grupos específicos, como mulheres

solteiras ou minorias raciais que têm menos acesso a recursos financeiros. Isso pode levar a uma seleção desigual na aplicação da norma, na qual indivíduos mais privilegiados social e economicamente são mais propensos a serem considerados de baixo risco, enquanto indivíduos mais marginalizados são mais propensos a serem considerados de alto risco e receber punições mais duras.

Com base nas informações mencionadas, é possível estabelecer uma clara relação entre o uso de critérios subjetivos em ferramentas de avaliação de risco de reincidência e a ideia inicial do tópico sobre técnica e ideologia.

Dessa maneira, o uso de critérios subjetivos em ferramentas de avaliação de risco de reincidência pode ser entendido como uma forma de dominação ideológica que se perpetua por meio da tecnologia. Isso torna possível a introdução dos próximos conceitos relevantes para a presente tese: a existência do racismo digital e algorítmico.

3 DO DETERMINISMO BIOLÓGICO AO DETERMINISMO ALGORÍTMICO: A HERANÇA DO POSITIVISMO CRIMINOLÓGICO

Por que um retorno a Lombroso? Sustenta-se, na presente pesquisa, que o racismo algorítmico, em especial no contexto das ferramentas de policiamento preditivo, tem a capacidade de perpetuar a ideia de que comportamentos criminosos são praticados (e podem ser previstos e classificados) por grupos específicos, cujas características pessoais são bem definidas e levadas em consideração na própria alimentação do banco de dados desses instrumentos.

Demonstra-se, no decorrer de todo o texto, que os algoritmos de policiamento preditivo podem reforçar preconceitos já existentes em dados de treinamento, o que tem resultado em latente discriminação e até mesmo no aumento da violência e vigilância contra grupos marginalizados: uma ressurreição de uma abordagem determinista e reducionista, mas impulsionada pelo desenvolvimento tecnológico.

A criminologia como ciência tem sua base no positivismo, responsável pela formulação de seu método. Conforme destaca Evandro Piza (1998, p. 139), a partir das lições de Massimo Pavarini, o positivismo possui características como: 1) a interpretação mecanicista da sociedade; 2) o pressuposto de que o crime e o criminoso eram realidades naturais independentes de processos de definição; 3) a crença em uma possível resolução científica da questão criminal baseada na defesa social; 4) a afirmação da neutralidade do saber; 5) o consenso social como realidade natural; e 6) a legitimação do aparelho repressivo como algo natural.

Pode-se afirmar que há similaridades no discurso positivista criminológico e no discurso de neutralidade das ferramentas tecnológicas. Trata-se de uma crença errônea na capacidade da ciência de ser completamente imparcial e livre de vieses. No positivismo criminológico, sustenta-se que o crime e o criminoso podem ser compreendidos e tratados como fenômenos naturais, passíveis de análise objetiva e desprovidos de influências sociais ou culturais. Da mesma forma, o policiamento preditivo se baseia na premissa de que algoritmos e dados podem prever atividades criminosas de forma neutra, sem reconhecer os vieses embutidos nos dados históricos e no próprio processo de modelagem.

O presente capítulo busca apontar esse “retorno a Lombroso” não apenas como uma referência ao chamado “criminoso nato”, mas também demonstrando que há uma íntima relação entre a criminologia positivista e a própria seletividade penal,

presente na contemporaneidade. Para deixar isso ainda mais claro, o capítulo, além de explicar a criminologia positivista do sec. XIX, mostra sua inserção no cenário brasileiro até a ruptura existente a partir da criminologia crítica, que vem demonstrar a latente seletividade no sistema punitivo brasileiro.

3,1 O POSITIVISMO CRIMINOLÓGICO E A SUPOSTA SUPERAÇÃO DA ESCOLA CLÁSSICA

A criminologia positivista, em sua fase inicial como disciplina independente, engloba teorias criadas entre o fim do século XIX e o início do século XX, dentro do contexto do positivismo naturalista. Isso inclui a Escola Sociológica francesa, representada por Gabriel Tarde, e a Escola Social na Alemanha, com Franz von Liszt, mas principalmente a Escola Positiva na Itália, liderada por Cesare Lombroso, Enrico Ferri e Raffaele Garofalo (Duarte, 1998, p. 129-130).

A inovação trazida pela criminologia positivista, em termos de abordagem para o problema da criminalidade, residia na tentativa de identificar características antropológicas de criminalidade no indivíduo preso, ou seja, buscar sinais físicos e comportamentais que pudessem indicar tendências criminosas.

Conforme ressaltado por Evandro Piza (1998, p. 130) a Escola Positiva Italiana é frequentemente retratada em oposição à Escola Clássica. A fundação dessa escola é geralmente creditada a Cesare Lombroso em 1876, com a publicação de sua obra "O Homem Delinquente", pela qual ele começou a explorar a ideia do "tipo criminal". A partir desse ponto, outros autores, especialmente Enrico Ferri, discípulo de Lombroso, continuaram a desenvolver e refinar essa hipótese inicial, sugerindo que o crime poderia ser resultado de fatores atávicos¹⁴. A escola de Lombroso, portanto, entrou em conflito com as ideias da Escola Clássica, que incluía o contratualismo de Beccaria, o utilitarismo de Bentham e as construções jurídicas de Carrara.

Nesse sentido, a Escola Positiva Italiana, pode ser vista como uma adaptação do materialismo médico para o estudo da criminalidade, uma abordagem que começou na Inglaterra, se espalhou para a França no início do século XIX com a

¹⁴ Fatores atávicos são características herdadas que podem ressurgir em indivíduos de gerações posteriores. No contexto da criminologia positivista, Cesare Lombroso acreditava que certos comportamentos criminosos eram resultado da "reversão" a traços primitivos ou ancestrais (Lombroso, 2001, p. 266).

reforma penitenciária e, mais tarde, chegou à Itália sob a influência do positivismo (Duarte, 1998, p. 130).

Em termos de disputa teórica, a oposição entre a Escola Clássica e a Escola Positiva girava em torno de questões fundamentais como: em que se baseia a responsabilidade penal do criminoso, o que é crime e qual o seu conceito, se o criminoso é um indivíduo normal ou uma anomalia dentro da espécie humana, e quais são os conceitos e os efeitos da pena.

No contexto da Escola Clássica, o delito era visto a partir de uma visão rigorosamente jurídica. Beccaria, por exemplo, parte de uma base ideológica contratualista, alegando que a origem das penas reside no contrato social e na necessidade de sua defesa contra agressões perpetradas por particulares (Anitua, 2008, p. 161).

A Escola Clássica, conseqüentemente, enfatizava a clara existência de uma responsabilidade moral dos indivíduos, de modo que se afirmava que o crime era uma escolha racional feita após o cálculo dos benefícios e custos, não havendo distinção biológica entre os indivíduos. Segundo Beccaria (1999, p. 30):

A segunda conseqüência é que, se cada membro em particular está ligado à sociedade, essa sociedade está igualmente ligada a todos os seus membros por um contrato que, por natureza, obriga as duas partes. Essa obrigação, que desce do trono até a choupana e liga igualmente o mais poderoso ao mais desgraçado dos homens, nada mais é do que o interesse de todos, em observar pactos úteis à maioria. A violação, de um só pacto, gera a autorização da anarquia.

Para além do livre-arbítrio como elemento essencial na prática de um delito, os expoentes da Escola Clássica também afirmavam a necessidade de uma proporcionalidade entre os delitos praticados e as penas aplicadas. Beccaria (1999, p. 38) argumenta que, para uma ação ser considerada um delito e merecer punição, ela deve claramente violar limites bem definidos. Quando esses limites não são claros, resulta-se em uma moralidade confusa que contraria as próprias leis, que acaba por criar um sistema legal inconsistente e contraditório. Há críticas por parte do autor a uma multiplicidade de leis que acabam punindo severamente indivíduos por infrações que não são bem delineadas.

Percebe-se que a Escola Clássica era apoiada em um método metafísico, embasado em “[...] deduções apriorísticas de ideias abstratas, enquanto as demais, inspiradas na filosofia naturalística, assentar-se-iam no método experimental” (Duarte,

1998, p. 131). Tais ideias eram rechaçadas pelos autores positivistas, dos quais consta Cesare Lombroso, que teve grande influência até mesmo em território brasileiro.

No século XIX, o positivismo criminológico começou a ganhar destaque em detrimento da Escola Clássica devido a uma mudança no entendimento sobre ciência, a natureza do crime e do criminoso. Em relação a Lombroso, na publicação de sua obra "*L'uomo Delinquente*", de 1876, seu método positivo faz ampla utilização da estatística, defendendo a existência do tipo criminal, ou "criminoso nato", revelado a partir de seus sinais externos e estigmas que levariam, invariavelmente, à prática delituosa. Lombroso ainda sugeriu uma progressão das formas de vida mais simples, como animais e plantas, até a complexidade da vida humana. Dentro dessa visão, acreditava que existiam diferentes níveis de evolução, levando à noção de raças superiores e inferiores. Propôs, assim, uma hierarquia entre os seres vivos, baseada em suas capacidades orgânicas (Duarte, 1998, 136).

É importante mencionar, conforme demonstrado no tópico seguinte, que o pensamento lombrosiano foi se alterando com o passar do tempo, em especial, por conta dos estudos de seus sucessores: Ferri e Garofalo.

3.2 O PENSAMENTO LOMBROSIANO E AS INFLUÊNCIAS DE FERRI E GAROFALO

Cesare Lombroso, médico alienista nascido em 1836, é considerado um dos fundadores da criminologia, tendo desviado o objeto de estudo criminal do delito (como era na Escola Clássica) para o delinquente, o que fez com o auxílio de seus saberes no campo da anatomia, psiquiatria e fisiologia (Anitua, 2008, p. 302-303).

Segundo sustentava, o criminoso (ou delinquente) não era diferente do louco ou do insano moral, tendo em vista que suas características são psicossomaticamente reconhecíveis e são causadas pelo mencionado atavismo. O criminólogo usava os presídios e instituições psiquiátricas como seus laboratórios de pesquisa. Nessas instituições, ele conduzia estudos detalhados em prisioneiros e pacientes, observando e medindo características físicas e comportamentais. Lombroso acreditava que poderia identificar traços anatômicos específicos que predispussem indivíduos ao crime, como a forma do crânio ou outras anomalias físicas. Ele chegou até a publicar artigos que confirmavam que um suposto atraso evolutivo era a chave para a

compreensão da criminalidade como um comportamento atípico (Anitua, 2008, p. 3004).

Muito embora a ideia de tipo criminal possa levar ao entendimento sobre diferentes níveis de desenvolvimento dos seres vivos, as ideias lombrosianas estavam muito mais próximas dos modelos propostos por Cuvier e Gobineau do que do evolucionismo tratado por Charles Darwin e Spencer, tendo em vista que o criminólogo italiano defendia a ideia de degeneração das chamadas “raças humanas” (Duarte, 1998, p. 136).

Isso significa que Lombroso sustentava a adaptação de teorias raciais pessimistas, que afirmavam a decadência ou degeneração das principais raças, e a teoria da recapitulação embrionária, para explicar a etiologia do crime. Ele criou uma hierarquia de explicações conforme o modelo positivista de complexidade usado na zoologia. Conforme demonstra Evandro Piza (1998, p. 136), seu discurso reflete uma representação do poder e da hierarquia de diferenças presente na sociedade capitalista, que ele acreditava ser tendente a reproduzir essas desigualdades.

Lombroso fazia comparações entre diferentes grupos marginalizados. Primeiramente, ele equiparava os prisioneiros e internados psiquiátricos, que sofriam com a degradação do sistema penal, às classes pobres dos países centrais que também enfrentavam a degradação do sistema capitalista. Em segundo lugar, Lombroso relacionava esses indivíduos aos “selvagens” (povos que eram forçados a se integrar compulsoriamente na sociedade dominante e cuja diversidade era constantemente negada). Por último, ele comparava esses grupos às crianças, que eram submetidas a novas formas de disciplina tanto dentro quanto fora da família na sociedade industrial. Para o autor italiano, também se associava à criminalidade fenômenos como a prostituição e o homossexualismo (Duarte, 1998, p. 136-137).

Dessa forma, o que se pode considerar como original no pensamento do criminólogo é a elaboração de uma explicação coerente com a negação de qualquer comportamento que fosse minimamente destoante dos padrões culturais e estéticos da burguesia branca europeia. Diversidade equivaleria, para o autor, à degeneração (Duarte, 1998, p. 137).

Lombroso, contudo, não é o único expoente dessa corrente de pensamento. Anitua (2008, p. 308) afirma que a “famosa trindade do positivismo” (*Scuola Positivista*) era formada também por Enrico Ferri e Raffaele Garofalo. Diferentes de autores como Max Nordau, Hans Kurella e Luíz María Drago, que eram considerados

discípulos e partidários de Lombroso, os mencionados italianos apresentaram, em seus estudos, originalidade e até mesmo divergências em relação ao precursor do pensamento criminológico positivista.

Diferentemente de Lombroso, Ferri era advogado, jornalista e político. Como autor positivista, rechaçava a noção de livre-arbítrio e compatibilizava o naturalismo lombrosiano com a noção de defesa social. Considerava o delito o sintoma de uma personalidade perigosa individual, de modo que a pena serviria para transformar tal personalidade (Anitua, 2008, p. 310-311). A negação do livre-arbítrio, conforme apontado, é um elemento muito claro em sua obra *“La Scuola Criminale Positiva”*, conforme pode ser visto em:

Recisamente neghiamo il supposto di un libero arbitrio, anzitutto perchè scopriamo l'origine naturale dell'illusione comune ond'esso è affermato, dipendente dalla sola ignoranza o incoscienza dello cause fisiche o fisio—psicologiche, onde ogni nostra deliberazione è preceduta e determinata; tanto è vero, che quando di un atto umano si conoscono o si sentono dall'agente stesso, in precedenza, i votivi determinanti e prepotentemente decisivi, scompare l'idea che quell'atto sai libero (Ferri, 1885)¹⁵.

De forma diversa à proposição exclusivamente biológica da delinquência, Ferri propôs uma classificação dos delinquentes baseada em fatores biológicos, psicológicos e até mesmo sociais que influenciariam o comportamento criminoso, identificando, no processo, cinco categorias principais de delinquentes: 1) o delinquente nato, que possui predisposições genéticas para o crime, não podendo ser ressocializado de qualquer forma; 2) o delinquente louco, cujo comportamento criminoso resulta de transtornos mentais; 3) o delinquente habitual, que comete crimes devido a hábitos e condições sociais desfavoráveis; 4) o delinquente ocasional, que é impulsionado por circunstâncias temporárias ou situações específicas; e 5) o delinquente passional, cujos crimes são motivados por emoções intensas e súbitas, como ciúmes ou vingança. Dessa maneira, a conduta humana criminosa era sempre produto de um organismo fisiopsíquico, bem como de uma atmosfera física e até social que envolve o delito (Anitua, 2008, p. 312).

¹⁵ "Negamos precisamente o suposto livre-arbítrio, antes de tudo porque descobrimos a origem natural da ilusão comum que o afirma, dependente apenas da ignorância ou inconsciência das causas físicas ou físico-psicológicas, de onde cada nossa deliberação é precedida e determinada; tanto é verdade que, quando de um ato humano se conhecem ou se sentem pelo próprio agente, previamente, os motivos determinantes e preponderantemente decisivos, desaparece a ideia de que esse ato seja livre" (Tradução nossa).

Com os estudos de Enrico Ferri, a criminologia transcendeu os limites do campo médico, abrangendo também a sociologia. Uma vez que Ferri argumentava que o comportamento criminoso não poderia ser explicado apenas por fatores biológicos, mas também por influências sociais e ambientais, inaugurou-se uma corrente de pensamento na qual o contexto social e as interações humanas também são essenciais para a compreensão do crime e da delinquência.

É preciso destacar que a noção de defesa social sustentada por Ferri engloba reformas claramente antiliberais. Conforme menciona Anitua (2008, p. 312), algumas garantias processuais, marcas do processo penal contemporâneo, como a presunção de inocência e o princípio do *in dubio pro reo*, não encontrariam lugar no seu pensamento, uma vez que era necessário, em inúmeros casos, eliminar o determinismo individual e aproximar o processo penal na mão de especialistas.

Esse determinismo é relevante aos estudos do autor mencionado, tendo em vista que a categoria de responsabilidade moral, presente na Escola Clássica, passaria a ser substituída por uma categoria “positiva” de responsabilidade social. Conforme já destacado, Ferri não levava em consideração a existência de uma suposta livre vontade, mas um comportamento humano e individual que fosse concreto e causado, de forma que a reação da sociedade deve ser a de impedir a reiteração dessa conduta (Anitua, 2008, p. 313).

É interessante mencionar, por fim, uma semelhança entre os atuais instrumentos de policiamento preditivo e as ideias propostas pelo autor italiano devido ao fato de que a defesa social deveria não apenas atuar após a prática do crime, mas também a partir de “medidas pré-delituais” (Anitua, 2008, p. 313), de modo que o delinquente deveria não apenas sofrer os rigores de uma punição no cometimento do crime, mas deveria ser identificado e até mesmo excluído antes que pudesse vir a praticar seus atos de delinquência.

O terceiro expoente da Escola Positivista mencionado no trabalho é Raffaele Garofalo, responsável pela autoria da obra “Criminologia”, publicada no ano de 1885, na qual trabalha a ideia de “delito natural”. Nota-se que a noção de delito natural não diz respeito a um crime originário, que seria detestável em todas as sociedades e em todos os momentos históricos, conforme destacado pelo próprio autor:

Nós não procuraremos indagar se todos os atos que o nosso tempo e a nossa sociedade consideram criminosos tiveram ou deixaram de ter em todos os tempos e em todos os lugares a mesma significação. Sabem todos que os

costumes de muitos povos não só toleravam o homicídio por vingança de sangue, mas o impunham como dever sagrado ao filho assassinado; que o duelo, umas vezes sujeito a gravíssimas penas, era, outras, considerado pela própria lei uma das principais formas de julgamento; que a blasfêmia, a heresia, o sacrilégio e a feitiçaria, tidos outrora em conta de negros malefícios, desapareceram dos códigos atuais em nações cultas; que a pilhagem de navios naufragados era autorizada em alguns países; que a depredação e a pirataria foram, durante séculos, os principais meios de existência de povos hoje pacificados; enfim, que, fora da raça europeia, existem sociedades meio civilizadas que toleram o infanticídio e a venda de crianças, que honram a prostituição e elevaram a altura de instituição cívica o adultério. São por demais conhecidos estes e análogos fatos (Garofalo, 1997, p. 10).

O que o autor entende como delito natural, conforme explicado por Anitua (2008, p. 314), seria a violação aos sentimentos básicos e naturais da piedade e da probidade, capazes de serem desenvolvidas apenas pela “raça europeia”. Essas violações são aquelas comumente praticadas pelas sociedades degeneradas.

Dessa maneira, a existência do delito natural pressupõe a existência de um delinquente natural, que seria o indivíduo que carece de tais sentimentos, eivados do que ele chama de “periculosidade” (Garofalo, 1997, p. 10-11; Anitua, 2008, p. 314). Contra o delinquente natural não há forma de ressocialização possível, sendo necessária, conforme o autor (Anitua, 2008, p. 314), sua inoculação por meio de deportação, expulsão ou morte (para aqueles sem piedade, ou verdadeiramente irrecuperáveis).

As ideias do autor serviram para que outros autores positivistas extrapolassem o “estritamente penal” para tratar de uma questão chamada “estado perigoso sem delito”, o que se tratava de uma medida de controle e defesa social. Esse estado perigoso dizia respeito a condutas que, muito embora não fossem propriamente crimes em alguns lugares, eram condutas consideradas degeneradas (*mala vita*) e que deveriam ser repreendidas, como a mendicância, a prostituição e o jogo. Importa dizer, conforme explica Anitua (2008, p. 315), que a lista incluía também religiões não oficiais e todo tipo de ação ou cultura que não se adequasse ao estilo de vida burguês ou à “imitação para os proletários disciplinados”.

Podia-se falar em uma verdadeira difusão de uma marginalização social (até mesmo exagerada) para a criação de classes de indivíduos tratadas sempre como perigosas, consideradas o oposto da civilidade branca europeia.

É de relevância para o presente trabalho ressaltar que as obras de Garofalo (assim como as dos demais autores mencionados nesse capítulo) chegaram a ser

rapidamente traduzidas para outros idiomas, o que ocorreu não apenas nos países europeus, mas também nas Américas, o que serviu para a disseminação do pensamento criminológico positivista em países como o Brasil, conforme destacado devidamente no tópico seguinte.

3.3 O BRASIL E A INFLUÊNCIA DO POSITIVISMO CRIMINOLÓGICO

Antes que se possa falar na recepção de um pensamento considerado científico, é preciso ressaltar que, no período, a força do argumento do método era algo secundário, tendo em vista que não se questionava o discurso da inferioridade racial. Conforme leciona Evandro Piza (1998, p. 168):

Tome-se, por exemplo, a equivalência entre o tipo racial e o tipo criminal de Lombroso. De maneira simples, sua pergunta era: Por que são criminosos “os criminosos” (ou mais precisamente os encarcerados)? A resposta será dada a partir da análise empírica nas prisões italianas. Porém, em sua obra, ele não se preocupa em perguntar: Por que afirmo que os selvagens, os africanos, os índios da América do Sul ou os australianos são criminosos?

Piza (1998, p. 168) explica que os criminólogos positivistas frequentemente faziam comparações como: selvagem/criminoso ou moral civilizada/amoralidade. Essas comparações eram aceitas sem questionamento como um conhecimento comum.

Com efeito, o chamado racismo científico, que legitimava as diferenças raciais, também era aplicado à questão criminal, onde a manutenção da ordem era (e ainda é) exaltada.

Ocorre que o discurso criminológico positivista pregava que essa ordem não é necessária apenas dentro das sociedades centrais, mas também na relação dessas sociedades com as colonizadas, que passaram a fazer parte do sistema capitalista global. Torna-se, assim, possível falar em recepção dessa corrente de pensamento por parte dos países periféricos.

É preferido frisar que a recepção do positivismo criminológico e das teorias raciais, no Brasil não dizia respeito apenas a uma suposta necessidade de manutenção da ordem. Para Lilia Schwarcz (1993, p. 247), essas recepções, para além da questão da substituição da mão-de-obra, deveriam atender à necessidade de estabelecimento

de critérios diferenciados de cidadania, o que frequentemente foi objeto de argumento em legislações (como em leis sobre imigração e até contravenções penais).

Não obstante, a compreensão do processo de recepção, segundo menciona Evandro Piza, também parte da análise da perspectiva de Thomas Kuhn sobre a estrutura das revoluções científicas:

[...] um paradigma científico, em sentido sociológico, indica toda a constelação de crenças, valores, técnicas, etc, partilhadas pelos membros de uma comunidade determinada; ou seja, um paradigma governa, em primeiro lugar, não um objeto de estudos, mas um grupo de praticantes de ciência, devendo qualquer estudo sobre a existência de um paradigma começar pela localização do grupo ou responsáveis pela sua sustentação (Duarte, 1998, p. 181).

De acordo Evandro Piza (1998, p. 181), a criação de um ambiente institucional no Brasil para a importação de ideias científicas estrangeiras começou a partir da década de 1930, período em que houve um impulso significativo com a instituição universitária e a criação de centros de investigação e até mesmo a ocorrência de debates independentes.

No entanto, conforme explicado por Schwarcz (1993, p. 20-70), mesmo antes desse período, na década de 1870, houve esforços para demarcar um espaço social para o discurso científico sob os auspícios do positivismo. Embora esse período não tenha representado uma massificação das instituições de saber, ele provavelmente marcou o início do processo de institucionalização do conhecimento científico no Brasil.

A existência dessa institucionalização permite, assim, entender quais foram os processos de criminalização que construíram um discurso falseador da realidade que culminava na inserção do elemento “raça” como uma variável no âmbito do sistema criminal brasileiro¹⁶.

Com efeito, durante o processo de consolidação do Estado Nacional no Brasil, normas e políticas específicas foram implementadas para garantir a criminalização das populações negras. Contudo, essa criminalização não era apenas uma

¹⁶ É válido mencionar a lição de Zaffaroni (1993, p. 15) ao dizer que o sistema penal não é um sistema único e coeso, pois é composto por diferentes agências que frequentemente competem entre si. Além disso, existe um “sistema penal paralelo” composto por agências de menor hierarquia que, apesar de formalmente terem menos poder, acabam exercendo muita arbitrariedade e discricionariedade devido à falta de hierarquização. Dentro desse sistema penal formal, alguns de seus integrantes mantêm um controle social punitivo “para-institucional” ou “subterrâneo”, ou seja, através de práticas não institucionais (e muitas vezes ilícitas) que são bastante comuns e regulares em termos estatísticos.

consequência, mas sim um objetivo claro das políticas programáticas da época. As ações policiais e as instituições envolvidas no controle social aplicavam medidas específicas direcionadas a essas populações, com o intuito de estabelecer um novo modelo de controle social que perpetuava a marginalização e a repressão das comunidades negras, o que demonstra como o racismo institucionalizado estava entrelaçado com as estruturas legais e sociais desde o início da formação do Estado Brasileiro (Duarte, 1998, p. 192).

3.3.1 Influências para o surgimento do pensamento criminológico no Brasil

Evandro Piza (1998, p. 251), conforme destacado, menciona que a introdução da criminologia positivista no Brasil faz parte de uma transformação mais ampla da sociedade e do controle social no país. Essa mudança pode ser resumida como a transição do período de escravidão para um capitalismo dependente. Esse processo se baseava nas condições materiais da época, mas também na habilidade das elites brasileiras em criar meios para lidar com essas condições. Portanto, a modernização foi um projeto conduzido por essa elite, e é possível afirmar que a criminologia brasileira emergente e a transformação das estruturas repressivas foram parte desse projeto de modernização.

Buscava-se, segundo o mencionado autor, legitimar práticas racistas pré-existentes, conforme mencionado em:

Ou seja, de fato também aqui a Criminologia positivista poderia nascer de uma relação de poder concreta que vinculava o olhar do especialista à exposição da degradação provocada pelo próprio sistema penal e partia para a construção de um discurso racionalizador dessa prática seletiva. Entretanto, a bem da verdade, a exposição do “corpo negro” aos agentes do controle social, como ocorreu nos países centrais em relação aos encarcerados, não representou novidade significativa na produção de discursos racistas, pois ser negro já era ser objeto, “coisa”, para a mentalidade escravista (Duarte, 1998, p. 254).

É também preciso mencionar que não havia necessariamente uma corrente unânime de pensamento criminológico no cenário Brasileiro, tendo em vista a existência de divergências no que diz respeito ao discurso criminológico. Evandro Piza (1998) considera, nesse sentido, três obras de grande relevância para compreender as estruturas repressivas da época: 1) *Menores e Loucos*, de Tobias Barreto (1884);

2) As Raças Humanas e a Responsabilidade Penal no Brasil, de Raimundo Nina Rodrigues (1894); e 3) Criminologia e Direito, de Clóvis Beviláqua (1896).

Tobias Barreto, ao escrever Menores e Loucos, analisou o Código Criminal do Império (1830), tecendo críticas às razões determinantes do processo de inclusão ou exclusão de determinados grupos de indivíduos na categoria de criminosos (especialmente os menores e os considerados loucos). Trata-se de uma obra que discutia a relação entre as ciências sociais e o direito, conforme pode ser visto em:

É um facto que a sociedade se desenvolve; porém as leis desse desenvolvimento não estão descobertas, o que importa dizer que a sciencia social existe ainda apenas como uma aspiração, e, em taes condições, não tem, não pôde ter princípios seus, princípios próprios, com os quaes possam conformarse os direitos, quaesquer direitos do homem (Barreto, 1926, p. 16).

Barreto (1926, p. 16-20) faz uma análise de como esses grupos eram tratados pelo sistema judicial, de modo a propor mudanças a partir de seus ideais liberais. Como uma obra positivista, Barreto faz uma crítica à questão da alienação mental e aos poderes dos magistrados em determinar o destino dos inimputáveis responsáveis pela prática de delitos, defendendo a atuação de médicos que fossem integrados à estrutura judiciária (Duarte, 1988, p. 269).

É válido mencionar que, no que pese a influência italiana, ao tratar do tema loucura, Barreto faz crítica à obra de Lombroso, acusando-o de realizar generalizações e de carecer de rigor científico ao realizar determinadas associações. Segundo Barreto (1926):

Por minha parte já se vê, que nisto vai um grande merecimento. Mas também, por outro lado, importa reconhecer que o auctor alargou de mais as suas vistas e é excessivo nas suas apreciações. Tudo tem seus limites. O conhecimento exacto do criminoso não se compõe somente de dados psychologicos, fornecidos pela observação interna, directa ou indirecta; mas é igualmente certo que não se compõe só de dados craniométricos, dynamométricos, ophtalmoscópicos e todos os mais epithetos sesquipedaes, de que sóe usar a tecnologia medical.

Além de acusar o autor de italiano de não ter evidências concretas de todas suas alegações na obra “O Homem Delinquente”, conforme argumenta o autor (1926, p. 16-17), reforça que o direito penal deve ser baseado em princípios científicos e observação empírica, em vez de autoridade e especulação filosófica.

É preciso mencionar que embora Tobias Barreto não tenha desenvolvido um discurso focado em questões raciais, conforme leciona Evandro Piza (1998, p. 284-285), ele acabou assumindo uma perspectiva alinhada aos interesses dos senhores de escravos. Sua obra não era especulativa sobre a realidade brasileira como as de Clóvis Beviláqua e Nina Rodrigues. Apesar disso, em alguns trechos, Barreto reconhecia as diferenças entre a realidade local e a europeia, mencionando "distâncias geográficas e intelectuais" e influências ambientais, climáticas e sociais.

Pode-se alegar que Barreto inovou no tratamento do delinquente de uma forma patológica, considerando a participação das mulheres na criminalidade, além de preocupações com menores e pessoas com transtornos mentais. Contudo, conforme salientado, ele não abordava o controle social através de argumentos raciais (de modo que era possível perceber até mesmo a defesa de um livre-arbítrio relativo), embora sua compreensão estivesse de acordo com a visão das sociedades escravocratas da época (Duarte, 1998, p. 285).

Por sua vez, a obra seminal do médico e antropólogo brasileiro Raimundo Nina Rodrigues, publicada em 1894, examina a relação entre as características raciais e a imputabilidade penal. É certo que a obra reflete as teorias racistas predominantes de sua época, uma vez que foi fortemente influenciada pelas ideias do determinismo biológico e da antropologia criminal de Cesare Lombroso. Contudo, é importante observar que a obra também é marcada pelo contexto socioeconômico e científico do final do século XIX no Brasil, período em que o debate sobre raça e criminalidade estava em alta. Segundo Nina Rodrigues, as condições de raça eram fundamentais para que se pudesse individualizar a responsabilidade penal dos agentes delitivos, conforme menciona em:

Pode-se exigir que todas estas raças distintas respondam por seus atos perante a lei com igual plenitude de responsabilidade penal? Acaso, no célebre postulado da escola clássica e mesmo abstraindo do livre arbítrio incondicional dos metafísicos, se pode admitir que os selvagens americanos e os negros africanos, bem como os seus mestiços, já tenham adquirido o desenvolvimento físico e a soma de faculdades psíquicas, suficientes para reconhecer, num caso dado, o valor legal do seu ato (discernimento) e para se decidir livremente a cometê-lo ou não (livre arbítrio)? – Por ventura pode-se conceder que a consciência do direito e do dever que tem essas raças inferiores, seja a mesma que possui a raça branca civilizada? – ou que, pela simples convivência e submissão, possam aquelas adquirir, de um momento para o outro, essa consciência, a ponto de se adotar para elas conceito de responsabilidade penal idêntico ao dos italianos, a quem fomos copiar o nosso código? (Rodrigues, 2011, p. 43)

Segundo Evandro Piza (1998, p. 287), Raimundo Nina Rodrigues pretendia demonstrar como o estudo da responsabilidade penal das chamadas "raças inferiores" poderia justificar a reorganização do controle social de forma a manter a supremacia de uma elite branca no Brasil. Dos autores nacionais mencionados, ele foi quem mais se focou em entender os efeitos do debate entre as teorias clássicas e positivistas no Brasil, particularmente no contexto racial, fazendo uso das discussões dos já citados especialistas estrangeiros como referência. É possível afirmar que Nina Rodrigues tentou criar um modelo que complementasse a ciência europeia.

A influência italiana é muito clara na obra do autor brasileiro. A noção de piedade e probidade, conforme desenhada por Garófalo, também se fazia presente em "As Raças Humanas e a Responsabilidade Penal no Brasil", conforme demonstrado no seguinte trecho:

O segundo ponto de vista, que é o do professor Garofalo, pressupõe a existência de um critério fundamental da criminalidade, pelo menos do delito natural, na violação do senso moral médio, representado pelos dois sentimentos básicos da probidade e da piedade, existentes em todos os povos chegados a certa fase de desenvolvimento (Rodrigues, 2011, p. 8).

Percebe-se na obra de Nina Rodrigues, tal como na obra de Garofalo, a ideia de que o desenvolvimento de um certo senso moral era algo que se alinhava ao desenvolvimento das chamadas raças humanas, de forma que o fator diferencial entre elas estava na forma como esses sentimentos (piedade e probidade) se faziam presentes (havendo, portanto, uma correspondência clara entre o desenvolvimento orgânico das populações e seu desenvolvimento moral).

Nina Rodrigues, após estabelecer seus pressupostos no início de sua obra, dedicou os segundo e terceiro capítulos para criticar a aceitação do livre-arbítrio relativo pelos criminalistas brasileiros, especialmente Tobias Barreto. Ele argumentava que tal posição era insustentável, pois acreditava que a ideia de liberdade era ilusória:

Ante o receio da imputabilidade geral que havia de ser, no seu conceito, a consequência da condenação do livre arbítrio, ele constituiu-se adversário intransigente daqueles que tendem a identificar o crime com a loucura, os patólogos do crime na sua frase, e lançou-se em formal contradição, admitindo um livre arbítrio relativo, parcial, a ponto de fazer carga ao legislador brasileiro de não ter afirmado, expressamente, no código do império, o momento da liberdade como condição da imputação (Rodrigues, 2011, p. 14).

Da mesma forma similar a Lombroso, para Rodrigues, as escolhas dos indivíduos eram sempre determinadas por suas necessidades, que, por sua vez, eram fruto de sua constituição biológica. Portanto, não havia espaço para a ideia de vontade livre e consciente.

Em termos mais simples, Nina Rodrigues usava as ideias da ciência europeia para defender um determinismo biológico rígido, onde a ordem natural era totalmente hierarquizada. Seu modelo racista apontava não apenas para a desigualdade entre indivíduos, mas entre grupos inteiros.

Conforme identifica Evandro Piza (1998, p. 291), o autor maranhense demonstrava, ao analisar a sociedade brasileira, que a falta de homogeneidade era evidente devido à herança biológica distinta, o que certamente resultava em uma posição de subordinação para as populações não brancas no processo de modernização do país. Contudo, seu modelo não era apenas uma cópia das teorias estrangeiras, pois colocava no centro do debate a responsabilidade penal das raças consideradas inferiores, um problema que, para os teóricos europeus, seria apenas marginal.

O terceiro jurista evidenciado no presente tópico é Clóvis Beviláqua, autor de “Criminologia e Direito”, publicada em 1896, uma obra marcada por ser uma coletânea de artigos relacionados às ciências criminais.

Muito embora a obra esteja segmentada em diversos escritos, é possível ter uma noção do pensamento do autor a partir de alguns temas que foram abordados. É possível, por exemplo, mencionar a tentativa de conciliação entre o pensamento clássico e positivista a partir de seus dois primeiros artigos “Criminologia e Direito” e “Sobre uma Nova Teoria da Responsabilidade”.

Sobre esse segundo artigo, para Evandro Piza (1998, p. 327), não há aprofundamento na discussão sobre o fundamento da responsabilidade penal, mas uma escrita teórica na qual o autor propõe um modelo integrado de ciência penal com a conciliação das escolas penais e com o surgimento de um princípio novo de defesa social.

É interessante mencionar que o autor brasileiro também fazia críticas aos clássicos italianos, muito embora tenha sido influenciado por eles. Menciona Beviláqua (2019, p. 26):

Abramos a Criminologia de Garofalo e ahi leremos: “responsabilidade moral e proporção penal, eis justamente dois princípios combatidos pela nova escola naturalística”. Em outra ocasião disse o mesmo autor que “o princípio da responsabilidade não representava mais do que um escolho lançado pelo legislador diante da pena, para impedi-la de atingir o delinquente”. E ainda mais: “é talvez lícito concluir que há uma contradição manifesta entre o fim da tutela ou defesa social e a condição da responsabilidade moral”.

Em linhas gerais, conforme explicado por Piza (1998, p. 330), Clóvis Beviláqua criticava Garofalo, focando na necessidade de uma reação social ao crime que não levasse em consideração se o ato criminoso era resultado de uma patologia ou de uma manifestação da vontade. Para Garofalo, qualquer “ação perigosa” deveria ser combatida, mas Beviláqua, entretanto, argumentava que era essencial diferenciar entre indivíduos imputáveis e imputáveis, já que a reação deveria ser ajustada de acordo com o caso.

Não obstante, Beviláqua defendia um sistema de controle social especializado e uma finalidade clara para distinguir crimes culposos de dolosos, além de lidar com a questão das tentativas de crime. As críticas não param por aí, tendo em vista que a própria noção de “delito natural” também era tratada (e rechaçada) por Beviláqua em seus artigos (Beviláqua, 2019, p. 30-34; Duarte, 1998, p. 330-331).

O autor brasileiro, no terceiro artigo de sua coletânea, “Confrontos Étnicos e Históricos”, também buscava analisar como diferentes grupos étnicos influenciavam a criminalidade no Ceará, resumindo as ideias de Nina Rodrigues. No quarto artigo, “O suicídio na Capital Federal”, ele utilizava um trabalho anterior de Viveiros de Castro para examinar as taxas de suicídio em várias regiões do Brasil, não apenas no Rio de Janeiro e no Ceará. No conjunto dos artigos, Beviláqua frequentemente comparava estatísticas brasileiras com europeias para validar teorias estrangeiras, mas também reconhecia as diferenças entre os contextos nacionais e internacionais. Além disso, ele apresentava diversas hipóteses, que nem sempre eram confirmadas e, em alguns casos, se contradiziam (Duarte, 1998, p. 332).

Nesse contexto, Beviláqua apresentava diferentes causas para a criminalidade, que não estaria apenas limitada a fatores biológicos. Sua visão das causas da criminalidade incluía, portanto, tanto fatores primários quanto secundários. Como dito, ele não descartava os aspectos biológicos como causas principais da criminalidade, mas considerava fatores sociais como causas secundárias, que incluíam a falta de rigor nas medidas repressivas, o alcoolismo, a dificuldade de adaptação ao meio, a

pobreza, o contato com criminosos, a educação negligenciada, e as crises econômicas e políticas (Duarte, 1998, p. 333).

É preciso mencionar, conforme demonstra Evandro Piza (1998, p. 355), que Beviláqua, ao contrário de Nina Rodrigues, não se detinha em descrições detalhadas sobre aspectos cerebrais e nervosos dos criminosos. Ele simplificava as ideias de Rodrigues, tornando-as mais acessíveis e menos acadêmicas, removendo os detalhes patológicos excessivos e a ressalva sobre a possível instabilidade dos mestiços. A essência da argumentação de Beviláqua era direta: quanto mais próximo do "tipo branco", menor a criminalidade; quanto mais próximo do "tipo negro", maior a criminalidade.

3.4 A CRIMINOLOGIA CRÍTICA E O DESLOCAMENTO DA QUESTÃO CRIMINAL PARA OS PROCESSOS DE CRIMINALIZAÇÃO

O que o trabalho atual demonstra, no presente capítulo, é que o chamado "retorno à Lombroso" presente no título da tese diz respeito ao retorno a um pensamento determinista. De forma similar, as novas tecnologias, ainda que permeadas por um racismo histórico, institucional e estrutural, são apresentadas como inovações de um progresso científico, supostamente sem vieses.

É preciso ainda mencionar, contudo, que, a partir da criminologia crítica, tornou-se possível evidenciar a existência de uma clara seletividade penal (não apenas no Brasil), um fenômeno essencial para uma melhor compreensão do objeto da presente pesquisa.

Com efeito, o presente tópico tem o intuito de demonstrar quando ocorreu a inserção do pensamento criminológico crítico no cenário brasileiro, o que fez com que a questão criminal se deslocasse do estudo do delinquente para os verdadeiros processos de criminalização. Entretanto, é importante ressaltar que a presente tese não tem o escopo de apresentar toda jornada do pensamento criminológico crítico (seja no mundo ou no Brasil), nem de aprofundar autores ou teorias específicas.

Conforme já ressaltado, o intuito do presente tópico é demonstrar que, com a superação do pensamento criminológico positivista e o foco na macrocriminalidade e nos processos de criminalização, foi possível desvelar a existência de um sistema criminal altamente seletivo e violento.

Dessa forma, ao tratar do surgimento do pensamento criminológico crítico no Brasil, torna-se também possível frisar a inexistência de um determinismo biológico para a prática da criminalidade, da mesma maneira que é possível afirmar a inexistência de um determinismo tecnológico e, assim, apontar com maior precisão as formas com que as novas ferramentas de policiamento preditivo têm violado direitos fundamentais das populações marginalizadas.

3.4.1 A crítica ao direito penal no Brasil e um novo pensamento criminológico

A criminologia crítica começou a ganhar força no Brasil no final da década de 1960, com o surgimento de diversos movimentos dentre os quais se destacam autores como Roberto Lyra Filho e Juarez Cirino dos Santos. Segundo Vera Malaguti (2012, p. 117), suas obras foram fundamentais para o desenvolvimento dessa corrente do pensamento criminológico no país.

Segundo a mesma autora (2012, p. 117), a criminologia crítica latino-americana se posicionou como uma barreira utópica contra a violência dos ciclos militares presentes em todo continente na segunda metade do século XX. Lyra Filho e Cirino dos Santos forneceram as interpretações mais profundas do marxismo aplicadas às questões criminais no Brasil. É importante, portanto, mencionar que esse enfoque discutiu, entre outras coisas, a relação entre poder, violência estatal e justiça social.

Na década de 1970, durante o rigor do regime militar, havia, segundo menciona Salo de Carvalho (2021, p. 5), uma ala de juristas criminalistas no Brasil que se posicionavam firmemente contra a arbitrariedade e a brutalidade perpetrada pelo Estado.

Vale ressaltar que essa resistência não era apenas política, mas também acadêmica, envolvendo figuras como os já mencionados Roberto Lyra Filho e Juarez Cirino dos Santos, mas também outros intelectuais, como Nilo Batista, Zahidé Machado Neto, Gizlene Neder e Gisálio Cerqueira. Os trabalhos desses autores se tornaram clássicos das ciências criminais brasileiras (Carvalho, 2021, p. 5).

Quanto a essa fase do pensamento criminológico crítico no território nacional, Carvalho (2021, p. 6) reforça a importância de Lyra Filho e Cirino dos Santos a partir de seus escritos (respectivamente: “Criminologia Dialética”, de 1972; e “Criminologia Radical”, de 1981), demonstrando que tiveram relevância no cenário nacional da

mesma forma que “Punição e Estrutura Social”, publicada em 1939 por Rushe e Kirchheimer, teve em âmbito mundial.

Para Carvalho (2021, p. 6), já no contexto da criminologia crítica brasileira, a década de 1970 marcou uma transição de uma abordagem ensaística para uma proposição mais sistematizadora e científica. Essa distinção entre as duas perspectivas se justifica pelo fato de que Roberto Lyra Filho não organizou seu pensamento criminológico de maneira rigorosa como fez Juarez Cirino dos Santos.

“Criminologia Radical”, de Cirino dos Santos, pode ser compreendida como uma tentativa de definir os fundamentos científicos e os objetivos programáticos da criminologia a partir de uma teoria do crime e do controle social comprometida com o projeto político das classes trabalhadoras nas sociedades capitalistas. Conforme mencionado pelo próprio autor (Santos, 2018, p. 36):

O compromisso primário da Criminologia Radical é com a abolição das desigualdades sociais em riqueza e poder [...] Essa posição política evita a degeneração da Criminologia Radical em mera “moralização”, ou no correccionalismo repressivo da “reabilitação pessoal”, que identifica crime com patologia e, nas posições mais liberais, propõe reformas de superfície, ou mais serviços sociais, modificando alguma coisa para deixar tudo como está – ou seja, preservando o sistema de dominação e de exploração do homem pelo homem.

É possível concluir que a criminologia radical de Santos se diferencia das criminologias tradicionais a partir dos seguintes elementos: 1) pela natureza do objeto de estudo; 2) pelo método dialético de estudo de seu objeto; 3) pela base social de seus compromissos ideológicos; e 4) por seus objetivos políticos estratégicos e táticos, que são expressos, no decorrer da obra, em programas alternativos de política criminal.

Por outro lado, o estilo de Lyra Filho é caracterizado por uma densidade reflexiva e complexidade, que se manifestava de forma dispersa em diversos textos, manifestos, manuscritos, apostilas, conferências, seminários e aulas. Grande parte desse material foi transcrito livremente por seus alunos e colegas. A obra “Criminologia Dialética”, segundo Carvalho (2021, p. 6) é um exemplo dessa consolidação de trabalhos, sendo resultado de repensamentos e rearranjos apresentados em cursos e artigos desde o final dos anos 1960.

É preciso salientar, contudo, que uma liberdade quanto aos procedimentos metodológicos do autor não pode ser confundida por uma ausência do método.

Conforme demonstra Carvalho (2021, p. 6), em todos seus escritos, o autor respeita rigorosamente o método dialético.

Assim, pode-se dizer que as obras de Lyra Filho e Cirino dos Santos marcaram um ponto irreversível na criminologia brasileira. Por meio das contribuições desses autores, é possível afirmar que qualquer tentativa de explicar o crime apenas por fatores individuais e por métodos estatísticos, sem considerar as violências estruturais e institucionais geradas pelo próprio sistema de justiça, incluindo a distribuição desigual da criminalidade na sociedade (criminalizações e imunidades seletivas), será vista, no mínimo, como uma "ingenuidade científica" (Carvalho, 2021, p. 7).

3.4.2 A questão racial no pensamento criminológico crítico brasileiro

No que pese os referidos autores (Lyra Filho e Cirino dos Santos) tenham marcado o surgimento de uma criminologia crítica no cenário brasileiro, é preciso mencionar que essa corrente foi marcada por outras “ondas” nas décadas seguintes.

Segundo Prando (2018, p. 74), uma “segunda onda” da criminologia crítica ocorreu entre as décadas de 1980 e 2000. Esse momento foi marcado por um diálogo com autores de outros países da América Latina, como Raúl Zaffaroni, Lola Anyiar de Castro e Rosa Del Olmo. Prando também menciona a importância das obras de Alessandro Baratta que, apesar de ser italiano, realizou diversas conexões com as discussões latino-americanas.

Frisa-se que as obras desse período rondavam em torno de temas já retratados no primeiro momento da criminologia crítica brasileira, mas também abordavam questões como a periferia e o poder colonial. Debateu-se também, com mais amplitude, questões de gênero.

A “terceira onda” criminológica crítica data do início dos anos 2000 até os dias atuais. Para Prando (2018, p. 74), questões raciais, que tinham um singelo espaço na primeira fase e que começaram a ser debatidas na segunda, ganharam mais espaço. Essas questões passaram a ser discutidas no âmbito da seletividade em conjunto com os marcadores de classe social.

A obra “Criminologia e Racismo”, de Evandro Piza, defendida como dissertação, foi publicada pela primeira vez em 2002 e tem sua marca como um dos trabalhos que aprofunda a análise estrutural sobre a questão racial no âmbito de processos de criminalização. Não obstante, Ana Flauzina, com “Corpo Negro caído

no chão: o sistema penal e o projeto genocida do Estado Brasileiro”, faz uma crítica ao fato de que a noção de raça é utilizada comumente como uma variável no processo de explicação da seletividade e controle penal, embora seja, para a autora, um elemento que faz parte de toda estrutura do controle punitivo (Flauzina, 2008).

Mais uma vez, é preciso salientar que o presente capítulo não se propõe a fazer uma análise de todo processo de ascensão do pensamento criminológico crítico no cenário nacional, mas as referências às obras e autores mencionados é relevante para a compreensão de que a questão racial é, hoje, uma questão central no estudo do sistema criminal brasileiro.

No intuito de ilustrar melhor a questão, é importante mencionar a pesquisa de Barreto e Preussler (2023, p. 190) que demonstram, a partir de dados do sistema de informações estatísticas do sistema penitenciário brasileiro, que apenas no ano de 2019, o número de pessoas negras assassinadas foi até 2,6 vezes maior do que a de outros grupos raciais, representando 75,7% das vítimas de homicídios. Tais números não sofreram reduções significativas nos anos seguintes.

Esse tipo de informação, conforme destacado pelos autores (2023, p. 190), demonstram o que a criminologia crítica brasileira vem apontando nas décadas recentes: trata-se do resultado de todo um longo período de opressão racial destinada ao fortalecimento das estruturas sociais de poder.

Conforme demonstrado no decorrer da tese, sustenta-se que a adoção acrítica de ferramentas de policiamento preditivo pode aprofundar e naturalizar ainda mais a violência contra populações marginalizadas no país, o que também pode ser visto como resultado de todo um processo de expansão do direito penal.

Esse processo de expansão do direito penal (ou hipertrofia penal), para Krohling e Boldt (2010, p. 243), é marcado pelo fenômeno do direito penal simbólico¹⁷ e pelo fortalecimento do punitivismo.

Tal expansionismo do direito penal pode ter impactos desproporcionais sobre a população marginalizada (o que já vem acontecendo), especialmente a população

¹⁷ O direito penal simbólico refere-se a um “[...] argumento legitimador de reformas legislativas e administrativas, voltadas ao esvaziamento das garantias processuais do suspeito [...] e ao recrudescimento dos poderes investigatórios e punitivos do Estado” (Krohling; Boldt, 2010, p. 244). Pode-se dizer que se trata utilização do sistema penal não necessariamente para punir comportamentos de forma eficaz, mas para transmitir uma mensagem à sociedade. Em vez de abordar as causas reais dos problemas sociais, o direito penal simbólico busca criar uma sensação de segurança e controle através da criação de leis e penalidades que são mais demonstrativas do que práticas.

negra. Esse fenômeno tende a intensificar a seletividade do sistema penal, exacerbando a criminalização de comportamentos frequentemente associados a esses grupos.

Como já ressaltado, defende-se que tal expansão não é apenas marcada pela ampliação das normas punitivas e o uso crescente do encarceramento como resposta às questões sociais, mas também pelo uso crescente da tecnologia por meio de um discurso cientificista que, em verdade, é carregado de vieses.

A consequência é, portanto, um maior número de prisões e sentenças severas para indivíduos negros (que já enfrentam discriminação sistêmica e desigualdades socioeconômicas, perpetuando e naturalizando o ciclo de pobreza e exclusão social), reforçando estigmas e estereótipos raciais e dificultando ainda mais a inclusão e a igualdade de oportunidades, conforme demonstrado no capítulo seguinte.

4 RACISMO ALGORÍTMICO E SELETIVIDADE PENAL: UMA ANÁLISE A PARTIR DE PROJETOS CONCRETOS DE POLICIAMENTO PREDITIVO

No capítulo introdutório desta tese, foram apresentados alguns conceitos fundamentais para embasar a análise do problema proposto. Este capítulo, por sua vez, direciona o foco para o problema central: o estudo do racismo algorítmico e do colonialismo digital no contexto das ferramentas de policiamento preditivo.

Embora o termo "racismo algorítmico" seja frequentemente utilizado como sinônimo de "racismo digital", é importante fazer uma distinção entre os dois. O racismo digital abrange um fenômeno mais amplo de discriminação no ambiente virtual, enquanto o racismo algorítmico constitui uma manifestação específica desse fenômeno, relacionada ao uso de algoritmos em determinados sistemas (como demonstrado a seguir).

Com o objetivo de aprofundar a compreensão do problema de pesquisa, este capítulo busca esclarecer essa diferenciação, além de analisar a interseção existente entre racismo algorítmico e seletividade penal (o que ocorre na criminalização primária, secundária e terciária¹⁸).

O capítulo é responsável por trazer exemplos concretos que ilustram como o uso dessas novas tecnologias pode resultar na violação de direitos, especialmente de grupos historicamente marginalizados. Contudo, para discutir racismo algorítmico de forma adequada, é necessário compreender também dois conceitos-chave: 1) racismo estrutural-algorítmico e 2) colonialismo digital.

Demonstra-se, ao longo do texto, com base em autores como Simone Browne, que o uso de tecnologias para violar os direitos de minorias, em especial da população negra, não é um fenômeno novo, mas sim uma prática histórica.

¹⁸ A criminalização primária pode ser evidenciada pela tipificação penal de determinada conduta, ou seja, sua inserção no ordenamento jurídico. A criminalização secundária, por sua vez, é verificada a partir da ação da força policial e do Ministério Público (no cenário estadunidense, a seletividade latente na criminalização secundária foi muito debatida após o assassinato de George Floyd em Minneapolis). Fala-se também em criminalização terciária com o ingresso do condenado no sistema prisional (Grosner, 2008).

4.1 RACISMO ALGORÍTMICO E UMA POLÍTICA CRIMINAL CIBERLOMBROSIANA

É importante ressaltar que, ao se falar em racismo digital, não se fala apenas em racismo algorítmico e, da mesma forma, é preciso esclarecer que o racismo digital também existe muito além de discursos racistas perpetrados na rede mundial de computadores.

Na perspectiva de Silva (2022, p. 18), há uma espécie de senso comum responsável por qualificar o racismo digital apenas como as ofensas que são realizadas por pessoas físicas no ambiente virtual. Essa visão, contudo, serve apenas para enfraquecer a luta contra o racismo que é estrutural, sub-reptício e que afeta a seara digital da mesma forma que atinge as esferas econômicas, do conhecimento e até da política institucional.

Importa dizer que a própria existência do fenômeno do racismo no ambiente virtual faz com que caia por terra a ideia de que a neutralidade da rede tornaria o ciberespaço um ambiente indiferente a pautas identitárias (tal como defendido na mencionada “Declaração de Independência do Ciberespaço”). Essa perspectiva equivocada, conforme demonstrado por Silva (2022, p. 18), vem de um momento no qual: 1) os ambientes digitais eram escassos; 2) havia pouca pesquisa e pesquisadores no campo da tecnologia advindas de populações minorizadas; e 3) dominava o pensamento de pretensão de neutralidade tecnológica derivado de um tecnoliberalismo que acabou por se consolidar.

O que se percebe, contudo, é o fato de que os processos de construção das tecnologias digitais são racializados. Assim, em um primeiro momento, é necessário diferenciar racismo digital de racismo algorítmico, apontando como o primeiro é um fenômeno mais amplo que engloba o segundo.

Racismo digital engloba práticas racistas que se tornam possíveis devido ao uso de artefatos tecnológicos. Menciona-se aqui, por exemplo, o racismo discursivo (e online) entendido como “[...] um sistema de práticas contra minorias racializadas que privilegiam e mantêm poder político, cultural e econômico em prol de brancos no espaço digital” (Silva, 2022, p. 26).

Dentro desse contexto, para além das ofensas perpetradas digitalmente, é preciso compreender que há modos sutis pelos quais atos racistas se vinculam às tecnologias digitais em diferentes recursos, como recomendações de conteúdo; moderação; reconhecimento facial e processamento de imagens (Silva, 2022, p. 27).

Esses modos, auxiliados por ferramentas algorítmicas e aprendizado de máquina, formam o que pode ser chamado de racismo estrutural-algorítmico.

O racismo se difere do mero preconceito ou da discriminação. Apesar de “racismo” comumente estar associado a esses outros termos, ele extrapola o conceito criado sobre uma pessoa ou grupo a partir de questões sociais e históricas. Para Foucault (2002, p. 304), o racismo é um poder de morte dentro de um sistema que tem como escopo central a “vida e o fazer viver”. Trata-se de um poder que insere a vida dentro do cálculo político de gestão da própria população.

Segundo o autor francês, o termo ainda constitui um mecanismo fundamental inerente à existência dos Estados-nação da modernidade, na medida em que estes fazem uso de políticas de “conformação do corpo-espécie” de sua população nacional (Foucault, 2002, p. 304).

Foi a partir do séc. XV que a Europa passou a problematizar a questão da diversidade humana, de modo que ao final do séc. XVIII a ideia de “raça” ganhou status científico. Vários pensadores, como Agassiz, Morton, Goubineau e Chamberlain, usaram o conceito de raça para justificar políticas racistas, colonialistas e eugênicas. Esses teóricos compartilhavam a ideia de que existiam hierarquias raciais, com raças superiores e inferiores, além de acreditarem que a miscigenação representava uma ameaça à civilização branca ocidental (Silva; Araújo, 2020, p. 4).

Em referência ao título da presente pesquisa, na obra “O Homem Delinquente” (publicado originalmente em 1876), Cesare Lombroso desenvolveu a teoria do “criminoso nato”, sustentando que alguns indivíduos possuíam características biológicas inatas que os predispõem ao crime. A base de sua argumentação está em uma análise (hoje considerada) pseudocientífica, na qual ele correlacionou traços físicos, como formato do crânio e características faciais, com tendências criminosas (Lombroso, 2001, p. 256)¹⁹. Lombroso utilizou medições de crânios e observações de indivíduos considerados criminosos para afirmar que certos grupos raciais e étnicos tinham maior propensão à criminalidade. Esse pensamento, embora contestado,

¹⁹ Em um determinado momento, na mencionada obra, Lombroso compara “500 pessoas honestas com 500 criminosos” da região de Piemonte. Quanto aos criminosos, analisa, por exemplo, que pessoas de cabelos negros eram responsáveis por 40% dos crimes de assassinato, enquanto pessoas de cabelos castanhos eram responsáveis por 42% e pessoas loiras por 12,5%. Quanto ao crime de “vagabundagem”, pessoas de cabelos negros eram responsáveis por 42,1% dos delitos, enquanto 55,9% deles eram atribuídos a pessoas de cabelos castanhos e apenas 1,3% a pessoas de cabelos loiros. De forma similar, Lombroso prossegue seu estudo com outros crimes e outras características físicas (como cor dos olhos e cor da pele), de modo a demonstrar que pessoas não brancas eram responsáveis por mais crimes (Lombroso, 2001, pp. 256-268).

alimentou a ideia de que o crime era um fenômeno biológico, não social e que vinculava diretamente “raça” e comportamento criminoso.

O racismo científico de Lombroso (e daqueles que o sucederam) é, portanto, evidente no reforço de estereótipos, colocando indivíduos de certas etnias, especialmente negros e indígenas, como predispostos ao crime devido a um atavismo e uma suposta inferioridade biológica. De modo a exemplificar essa afirmação, destaca-se a seguinte passagem da obra do criminólogo italiano:

Um violador siciliano, um ladrão, três assassinos apresentavam obliquidade do orbital, arredondamento do crânio, saliência e distância dos zigomas maxilar quadrado e compacto e uma cor amarelada da pele, apresentando assim uma exata reprodução do tipo mongol. Alguns, como Cartouche e O., pela fronte fugidia, a pequenez do crânio e a saliência da face, aproximam-se do tipo simiesco. Nos criminosos, tal reunião de anomalias, sobretudo atávicas (algumas são patológicas: por exemplo, a assimetria) nos dizem que eles têm o tipo criminal (Lombroso, 2001, p. 266).

É importante destacar que suas ideias influenciaram não só o campo da criminologia, mas também políticas sociais e criminais, especialmente no âmbito de controle populacional e repressão de minorias raciais. A obra de Lombroso (e inúmeras outras que afirmavam a criminologia positivista) perpetuou uma visão eurocêntrica e discriminatória que contribuiu para justificar o racismo e a segregação, oferecendo uma suposta “base científica” para práticas eugenistas e de controle social ao longo do século XX até mesmo em países de população miscigenada como o Brasil.

As ferramentas de policiamento preditivo contemporâneas (estudadas no presente capítulo) podem ser vistas como uma espécie de retorno às crenças de Cesare Lombroso porque, assim como a teoria do “criminoso nato”, elas baseiam suas previsões em dados que refletem preconceitos raciais enraizados e muitas vezes baseados em tais compreensões pseudocientíficas. Como demonstrado, Lombroso acreditava que características físicas determinavam a propensão ao crime, categorizando certos grupos raciais como predispostos à criminalidade. De maneira análoga, os algoritmos de policiamento preditivo, treinados com dados tendenciosos, acabam por reforçar estereótipos raciais e a focalizar a vigilância em grupos marginalizados, fomentando uma política criminal que pode ser chamada de “ciberlombrosiana”.

Com efeito, assim como as ideias de Lombroso foram usadas para justificar o controle de certas populações a partir de uma roupagem científica, essas ferramentas perpetuam a discriminação, reforçando a seletividade penal e o racismo estrutural disfarçados sob o véu de objetividade tecnológica. Ambos os sistemas, ainda que separados por mais de um século, exemplificam o uso de pseudociência ou tecnologia para legitimar práticas discriminatórias e desvelam que Lombroso, em seu tempo, não foi capaz de encontrar nas prisões a figura do criminoso, mas do criminalizado.

Assim, falar de racismo pressupõe falar de um fenômeno complexo que extrapola as condutas individuais, sendo possível até mesmo relacionar o aspecto estrutural do fenômeno com a natureza agencial de sistemas algorítmicos para que seja possível discutir o que pode ser chamado de racismo estrutural-algorítmico (Silva; Araújo, 2020, p. 8).

Para Silva e Araújo (2022, p. 8), os novos modelos de interação tecnológica não podem ser separados do racismo estrutural²⁰, tendo em vista que ao serem construídos em um contexto social já marcado pela estrutura do racismo, as ferramentas acabam refletindo e ampliando as dinâmicas discriminatórias.

Assim, o racismo estrutural-algorítmico é um dispositivo de constituição de dados e de arranjos estruturais desses artefatos tecnológicos. No caso de plataformas digitais, por exemplo, é possível verificarmos a contribuição de usuários na reprodução de preconceitos²¹ e estereótipos raciais e de gênero, bem como decisões e direcionamentos intencionais realizados pela gestão das empresas responsáveis por esses sistemas²².

Os referidos autores (2022, p. 9) exemplificam a afirmativa acima a partir dos casos de filtragem algorítmica praticados pelo Facebook. Em agosto de 2014, na ocorrência de manifestações contra a violência policial contra a população negra na cidade estadunidense de Ferguson, as postagens a respeito do evento foram invisibilizadas na plataforma. Além das plataformas digitais, a influência de sistemas

²⁰ Segundo Cathy O'Neil, autora de "Algoritmos de destruição em Massa", o próprio racismo é um modelo de previsão construído tendo como base informações incompletas, defeituosas e generalizadas. O racismo é construído a partir de uma coleta equivocada de dados e correlações rasas, que são reforçadas por injustiças estruturais contaminadas por vieses de confirmação (O'Neil, 2020, p. 37).

²¹ Lembremos do *Chatbot Tay*, desenvolvido pela Microsoft.

²² Em um exemplo recente, redes digitais como *X*, *Facebook* e *Instagram* têm suprimido conteúdo postado por usuários e grupos de defesa dos direitos humanos em relação à solidariedade com o povo palestino (Espanol, 2023).

algorítmicos de classificação com viés racializado se estende ao sistema penal de diversos países. Entre eles, os Estados Unidos. Isso pode ser evidenciado na utilização do software COMPAS²³ para prever reincidência criminal, o que resultou em prejuízos para réus negros e vantagens para réus brancos.

Conforme demonstrado no decorrer dos tópicos seguintes, o racismo estrutural-algorítmico não é um problema técnico, mas a transferência de um problema social ao campo das tecnologias de comunicação e informação. Ele pode reforçar e amplificar as desigualdades existentes na sociedade, perpetuando, por meio de políticas criminais, a exclusão, marginalização e possivelmente extermínio de grupos historicamente discriminados²⁴.

4.2 O RACISMO, O COLONIALISMO E O DIGITAL

A complexidade do tema fica ainda mais acentuada a partir da inserção de outro elemento, o colonialismo digital. Faustino e Lippold (2023, p. 52) sugerem que para compreender profundamente o colonialismo digital é preciso entender o contexto histórico mais amplo do capitalismo. É possível que, no entendimento dessa relação entre colonialismo e capitalismo, pode-se confrontar os traços históricos identificados com as tendências atuais de participação do desenvolvimento tecnológico informacional nas formas contemporâneas de exploração e dominação.

A discussão quanto o colonialismo certamente não é recente. Muito embora Frantz Fanon não tenha vivido para presenciar a criação de muitas das ferramentas tecnológicas utilizadas contemporaneamente, é possível fazer uso de seus estudos sobre o tema para explicar alguns fenômenos recentes, em especial, sua análise

²³ Conforme destacado, o COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*) é um sistema utilizado por determinados tribunais estadunidenses para auxiliar na tomada de decisões ligadas à execução da pena, como a concessão de liberdade condicional. O *software* foi alvo de controversas quando uma pesquisa independente realizada pela PROPUBLICA (Larson; Mattu; Kirchner; Angwin, 2016) demonstrou que havia uma tendência a atribuir pontuações de risco a reincidência mais altas a réus negros do que a réus brancos, ainda quando todos os demais fatores (exceto a cor da pele) eram semelhantes.

²⁴ É válido ressaltar, conforme analisado nos exemplos citados por Silva (2022), que o racismo algorítmico pode ser identificado tanto na alimentação do banco de dados, como no aprendizado da inteligência artificial (dependendo de como os sistemas foram projetados e treinados). Por exemplo, se uma saboneteira automatizada foi treinada predominantemente com imagens de mãos brancas, pode não reconhecer mãos de outras tonalidades. Mais amplamente, se o conjunto de dados usado para treinar a inteligência artificial não for representativo, ela perpetuará esses vieses. Portanto, o racismo algorítmico reflete tanto o design dos algoritmos quanto os dados utilizados para treinar a IA.

quanto às colônias da primeira metade do século XX e o uso de tecnologias pelos colonizados.

Não obstante, podemos afirmar que a compreensão das revoluções tecnológicas não é completa sem um exame cuidadoso do papel histórico do colonialismo e do racismo (Faustino; Lippold, 2023, p. 52).

Segundo Fanon (1979), o colonialismo e o racismo não podem ser vistos apenas como sistemas de dominação política e social, mas também como um fenômeno que afeta as estruturas econômicas. Assim, as revoluções tecnológicas não surgem em um vácuo, mas são influenciadas e moldadas por todos esses fatores históricos.

Em concordância, Faustino e Lippold entendem que o colonialismo é uma forma particular de exploração econômica, mas sua reprodução seria impossível sem a utilização de ferramentas de dominação nas quais o racismo é um elemento essencial. Dessa forma, sustentam também que as inovações tecnológicas emergem e se consolidam dentro de uma estrutura social já existente, que eles chamam de *hardware* social, que é essencialmente a base material sobre a qual essas inovações são construídas e operadas (Faustino; Lippold, 2023, p. 52). Assim, o capitalismo não pode ser entendido separadamente do colonialismo, pois o colonialismo desempenhou um papel fundamental na formação e no desenvolvimento desse modo social de produção.

Os processos históricos de colonização e dominação não são, portanto, aspectos de um passado remoto, mas continuam a moldar e ditar as estruturas sociais e econômicas contemporâneas, atuando por meio de novos aparatos tecnológicos. Logo, há uma relação essencial entre racismo, ferramentas de policiamento preditivo e o que se conhece como colonialismo digital.

De modo a exemplificar, as ferramentas algorítmicas estudadas na presente tese possuem certas limitações que, para Sales (2023, p. 15), podem resultar em discriminações em diferentes níveis, agravando disparidades devido à composição desigual dos conjuntos de dados que armazenam as informações utilizadas. Em casos de ferramentas de reconhecimento facial, por exemplo, geralmente há uma predominância significativa de imagens de pessoas com tons de pele mais claros em comparação com indivíduos de pele mais escura.

Sales (2023, p. 16), em suas pesquisas, sustenta que até pelo próprio preconceito inconsciente inserido nos sistemas pelos seus responsáveis, o problema

do viés algorítmico se manifesta primeiro sempre em comunidades historicamente marginalizadas devido a determinadas características físicas. Isso evidencia limitações que necessitam de abordagens no campo da tecnologia para promover práticas consideradas éticas e não discriminatórias entre aqueles que operam as redes neurais responsáveis por interpretar imagens.

Não se trata, contudo, apenas da necessidade de identificação de questões puramente técnicas nesse processo, mas também se torna necessário realizar considerações mais profundas sobre transparência e a maneira como esses processos podem vir a ser implementados.

Quanto ao viés atribuído de forma sub-reptícia (ou às vezes até intencionalmente) pelos responsáveis pela manutenção do banco de dados, Silva (2022, p. 108) afirma que o supremacismo branco nunca abandonou as instituições, sejam elas prisões, cidades, empresas e, certamente, as forças policiais. Para o autor (2022, p. 109), a vigilância e a categorização social baseada em hierarquias raciais formam o cerne da distribuição desigual de humanidade, facilitando a sustentação do capitalismo em escala global.

Essas transgressões contra direitos estabelecidos podem ser vistas como parte do colonialismo contemporâneo e, mais especificamente, do colonialismo digital, um fenômeno que descreve as maneiras pelas quais nações mais avançadas exercem sua influência e controle sobre países em desenvolvimento através da tecnologia.

Importa destacar que, para Faustino e Lippold (2023, p. 68), o colonialismo digital não é uma fase posterior a um neocolonialismo tardio, mas sim sua expressão contemporânea tecnológica informacional. Dessa forma, é por esse colonialismo que surgem as condições para o surgimento da dominação por meios tecnológicos.

Não obstante, Sérgio Amadeu da Silveira (2021, p. 35) observa que, embora o colonialismo histórico e tradicional tenha formalmente cessado, a colonialidade permanece intacta e pode ser empiricamente observada e mapeada.

Segundo a visão dos autores citados, é possível afirmar que o colonialismo digital opera de diversas maneiras, entre elas: 1) pela disparidade no acesso à tecnologia, onde países mais desenvolvidos desfrutam de maior acesso e recursos para inovação tecnológica; 2) pela dependência tecnológica dos países em desenvolvimento, que muitas vezes dependem de serviços oferecidos por empresas (ou Estados) estrangeiras; 3) pela exploração de dados e da privacidade, mediante a

coleta de dados pessoais e sensíveis dos usuários, frequentemente sem consentimento adequado.

Dessa forma, um Estado, ao fazer uso de ferramentas de policiamento preditivo baseadas em tecnologias estrangeiras (seja por não ter condições próprias de pesquisa ou pela adoção do discurso de superioridade tecnológica do outro), acaba por incorporar os preconceitos, o racismo e a própria seletividade inerente a essas ferramentas. Contudo, isso não significa que os Estados, ao não fazerem usos dessas tecnologias, são naturalmente livres do preconceito institucional, que passaria a existir apenas com a adoção de determinadas tecnologias (como ferramentas de policiamento preditivo).

O racismo é uma parte relevante dessa nova face do colonialismo (como sempre foi no colonialismo tradicional). De acordo com Silva (2022, p. 110), são inúmeras as formas pelas quais as novas tecnologias digitais têm impactado os direitos da população negra, o que inclui a amplificação de estereótipos negativos, a discriminação algorítmica em processos de seleção de emprego e práticas invasivas de vigilância digital que afetam desproporcionalmente certas comunidades, como o monitoramento e uso discriminado de tecnologias de reconhecimento facial.

A vigilância sobre os corpos negros não é um fenômeno recente, mas uma estrutura histórica redesenhada e perpetuada pelo colonialismo contemporâneo. Faustino e Lippold (2023, p. 52), ao analisarem os textos de Fanon, afirmam que o capitalismo não existe sem colonialismo, assim como o colonialismo não existe sem racismo.

O que se pode concluir a partir da análise da evolução do colonialismo é que as tecnologias, como ferramentas de dominação e segregação, não são um evento novo, mas algo histórico que tem se redesenhado a partir do surgimento de ferramentas mais aprimoradas.

Como foi evidenciado, a falta de imparcialidade no desenvolvimento de ferramentas e métodos (e até no próprio conhecimento científico) já tem sido objeto de pesquisa por muitos anos. Simone Browne (2015, p. 5) ilustra de diversas formas e por meio de vários relatos históricos como as tecnologias foram empregadas contra a população negra nos Estados Unidos. Não é surpreendente, portanto, que o controle sobre certas tecnologias seja usado como meio de coerção e dominação.

No Brasil, a relação entre colonialismo digital e racismo algorítmico pode ser evidenciada em inúmeras situações. Um exemplo de 2019 envolveu a busca pela

expressão “tranças bonitas” e “tranças feias” no Google. Enquanto a primeira pesquisa mostrava majoritariamente mulheres brancas de cabelo liso, a segunda apresentava mulheres negras com cabelo cacheados ou crespos. Para Silva (2019), tais situações não são automaticamente intencionais, tendo em vista que muitas vezes o problema surge por omissão e pela falta de atenção aos vieses existentes nos dados ou nos processos. Essa ausência de ações corretivas, contudo, pode ser tão prejudicial quanto ações intencionais.

Outro caso, no âmbito da segurança pública, foi apontado em um estudo feito pela Rede de Observatórios de Segurança e publicado em reportagem do *Intercept* Brasil (Nunes, 2019), cujo levantamento demonstrou que 90,5% dos presos por monitoramento facial no país são negros.

Esses exemplos ilustram como o que foi mencionado no trabalho sobre como o racismo algorítmico e o colonialismo digital estão interligados e impactam os direitos humanos de populações marginalizadas no Brasil. Não é exagero afirmar que essas ferramentas tecnológicas importadas ou inspiradas em modelos de países desenvolvidos carregam consigo os mesmos preconceitos e vieses presentes em seus contextos de origem.

4.2.1 A Tecnologia e vigilância histórica sob a perspectiva de Fanon e Browne

Uma vez que o trabalho abordou a relação entre racismo algorítmico e colonialismo digital, busca-se, com o presente tópico, aprofundar a ideia de que racismo e colonialismo sempre caminharam juntos.

Os autores selecionados para o enriquecimento da discussão foram escolhidos diante da relevância de seus trabalhos. Frantz Fanon, muito embora não tenha presenciado os fenômenos contemporâneos já mencionados, é um grande expoente da luta antirracista e suas obras são de grande importância até os dias de hoje. Foi a partir delas que Browne foi capaz de discutir questões como racismo histórico e formas de resistência.

Nesse sentido, Browne traça seus estudos a partir de conceitos estabelecidos por Fanon, como, por exemplo, na ideia de “epidermização” (*epidermalization*), que pode ser vista em obras como “Pele Negra, Máscaras Brancas”, de 1952.

A expressão é uma referência ao processo de interiorização, que Fanon (2008, p. 28) chama de “epidermização” de uma inferioridade. Trata-se de um processo pelo

qual o racismo e a opressão são responsáveis por reduzir os indivíduos negros à cor de sua pele (e aparência física). O processo, portanto, indica que a própria identidade ou o valor de alguém é determinado primeiramente pela cor de sua derme.

Compreende-se a partir de Fanon (2008, p. 28) que a “epidermização” é um tipo de alienação na qual o indivíduo negro é forçado a ver-se pelos olhos de uma sociedade racista que o defina pelo seu fenótipo. Trata-se de um fenômeno que gera uma série de impactos psicológicos, como a mencionada internalização de uma inferioridade²⁵.

Assim, sustenta Browne (2015, p. 26):

Through Frantz Fanon’s concept of epidermalization — that being the imposition of race on the body — I trace and provide a genealogy of modern, digital epidermalization by focusing on branding and the role of prototypical whiteness in the development of contemporary biometric information technology²⁶.

A autora estadunidense, em sua obra, demonstra como aparatos tecnológicos foram historicamente responsáveis por reforçar hierarquias raciais e contribuíram para a constante vigilância de corpos negros nas américas.

Vale ressaltar que essa imposição de hierarquias encontra respaldo no que Silveira, Souza e Cassino (2021, p. 35-36) entendem ser um desdobramento do já citado colonialismo. Esse fenômeno ainda se mantém em tempos de colonialismo digital. E para além da questão racial, os países colonizados podem também se sentir compelidos a adotar comportamentos, linguagens e valores impostos por seus colonizadores em um processo fútil de busca por aceitação e visibilidade.

Em sua obra *“Dark Matters: On the surveillance of blackness”*, a autora norte-americana recorre não apenas a Fanon, mas também menciona o “panóptico²⁷” de

²⁵ Fanon (2008, p. 16) explica que a expressão “máscara branca” está ligada ao modo como os valores culturais da população branca são levados à população negra como superiores, forçando-as a enquadrarem-se em uma sociedade que não lhes pertence.

²⁶ “Através do conceito de epidermização de Frantz Fanon — que é a imposição da raça no corpo — eu traço e forneço uma genealogia da epidermização moderna e digital, focando no branding e no papel da branquitude prototípica no desenvolvimento da tecnologia contemporânea de informação biométrica” (Tradução nossa).

²⁷ Browne (2015, p. 33), ao estudar Bentham, explica o panóptico como um modelo arquitetônico criado para permitir que um único observador seja capaz de vigiar os prisioneiros sem que estes saibam que estão sendo observados. A estrutura é formada por uma torre central e por células periféricas. O panóptico, como instrumento de reforma social e moral, foi analisado também pelo filósofo Michel Foucault.

Jeremy Bentham como um modelo criado para supervisionar a força de trabalho dos detentos (Browne, 2015, p. 33), assim como também analisou Michel Foucault.

Foucault, por sua vez, utiliza o panóptico de Bentham para explorar a forma como o poder é exercido na sociedade moderna. Trata-se, para o autor francês, de um modelo geral que revela dinâmicas de poder em inúmeras instituições e aspectos da vida cotidiana, não se limitando às prisões, mas estendendo-se a escolas, hospitais, fábricas e outros centros (Foucault, 2007, p. 170).

A partir de tais estudos, Browne (2015, p. 42) mostrou como a arquitetura, o registro, a documentação e a inspeção nos fortes e portos usados para o tráfico de escravos (incluindo a "Porta do Não Retorno"²⁸ e os navios negreiros durante a travessia do Atlântico, da África aos mercados de escravos e plantações no Novo Mundo) eram métodos usados para controlar e definir os escravos, submetendo-os a violências constantes. Dessa forma, muito embora Foucault possa ter explicado como o espetáculo deu lugar ao controle e vigilância dos corpos dóceis, pode-se afirmar que a regulação violenta da negritude historicamente combinou elementos de espetáculo, disciplina e vigilância.

Um dos pontos estratégicos da obra de Browne é explicar que, historicamente, a negritude tem sido fundamental na elaboração de ferramentas e infraestruturas de vigilância. Para a autora (Browne, 2015, p. 43), a omissão das narrativas atuais de controle tende a levar as discussões sobre o assunto de forma desvinculada da questão racial.

Dessa forma, podemos observar o navio negreiro como um exemplo de vigilância que ultrapassa o modelo panóptico. Enquanto o panóptico de Bentham foca na vigilância disciplinar em instituições fechadas, como prisões e hospitais, Browne ressalta que o navio negreiro funcionava como uma forma móvel de vigilância em alto mar, onde os negros escravizados eram mantidos sob estrito confinamento.

A tecnologia, para a autora, sempre tem buscado afetar os corpos negros. O navio, conforme sustenta sua obra, não é o único exemplo, de modo que é possível também citar a lei das lanternas, que eram carregadas de forma obrigatória pelos escravizados durante a noite. Tratava-se de clara forma de controle visual e terror.

²⁸ A "Porta do Não Retorno" (*"Door of No Return"*) se refere a determinados locais em fortalezas na costa da África Ocidental, como na Ilha de Corée, em Senegal. Em tais localidades, escravos africanos eram mantidos antes de serem enviados para as Américas. Tratava-se do ponto final na África para milhões de capturados, que jamais retornariam (Browne, 2015, p. 42).

Segundo a autora, até mesmo instrumentos “abolicionistas” tinham clara intenção em manter um controle, como foi o caso do “*Book of Negroes*”²⁹ (Browne, 2015, p. 66).

Por fim, a autora menciona outros exemplos de controle, como a utilização de iluminação artificial no confinamento solitário, além do uso de flashes de câmeras e programas de televisão que documentam a violência contra pessoas negras. Essas tecnologias são consideradas ferramentas de terror visual (Browne, 2015, p. 21) e são responsáveis por tornar os corpos negros mais visíveis, evidenciando o papel da alta visibilidade na opressão e na continuidade do racismo.

Essas novas ferramentas (nas quais se insere a identificação facial) podem ser consideradas novas formas de “*branding*” (marcação), existentes desde o período de escravidão (Browne, 2015, p. 22).

4.3 O PODER DE PUNIR E A SELETIVIDADE PENAL

O poder de punir sempre exerceu um papel crucial nas diferentes formas de colonização. Segundo Fanon (1974, p. 54-55):

[...] o mundo colonizado é um mundo cortado em dois. A linha de corte, a fronteira, é indicada pelas casernas e pelos postos policiais. Nas colônias, o interlocutor legítimo e institucional do colonizado, o porta-voz do colono e do regime de opressão, é o policial ou o soldado.

O que se pretende demonstrar com o presente tópico é que a seletividade em relação aos processos de criminalização é uma constante (não apenas no cenário nacional) nos três âmbitos de criminalização e que o uso de ferramentas algorítmicas representa nada mais que uma continuidade desse processo sob um discurso de neutralidade tecnológica que não se sustenta.

Os processos de criminalização, como pontuado no início do capítulo, envolvem suas vertentes primária, secundária e terciária. Segundo Alessandro Baratta (1999, p. 161), a criminalização primária é marcada pelo mecanismo de

²⁹ O “*Book of Negroes*” foi um registro meticuloso criado pelos britânicos em 1783, documentando aproximadamente 3.000 afro-americanos que haviam sido escravizados e que obtiveram sua liberdade ao se aliarem aos britânicos durante a Guerra de Independência dos Estados Unidos. O documento continha detalhes pessoais como nome, idade, características físicas, local de origem e provas de sua liberdade, sendo usado para garantir que essas pessoas pudessem ser reassentadas, principalmente na Nova Escócia, no Canadá, como indivíduos livres. Para Browne, este registro também funcionava como uma ferramenta de catalogação racializada, permitindo às autoridades exercerem controle sobre os ex-escravizados.

produção de normas; a criminalização secundária, por sua vez, compreende a atuação dos órgãos de investigação e culmina com o juízo; por fim, fala-se em criminalização terciária com a execução da pena ou das medidas de segurança. A seletividade contra certos grupos marginalizados pode ser evidenciada em todos esses níveis.

A criminalização primária é marcada pela elaboração das normas. Existe a partir do momento em que o Estado define qual bem jurídico merece a tutela penal e um maior rigor na responsabilização daqueles que lesionam ou ameaçam tal bem.

Ocorre que a desigualdade no direito penal já existe a partir desse momento, conforme demonstrado pelas proposições estabelecidas por Alessandro Baratta. O autor italiano alega que o direito penal não cumpre suas funções declaradas, quais sejam: 1) proteger igualmente todos os cidadãos contra ofensas aos bens considerados juridicamente essenciais (consoante demandado pelo princípio do interesse social); e 2) ser igualmente aplicado a todos comportamentos ilícitos e violadores, com igual consequência àqueles que praticarem delitos similares e nas mesmas circunstâncias (Baratta, 1999, p. 161-162).

Dessa forma, Baratta (1999, p. 162), ao analisar a realidade da elaboração e aplicação da lei penal (não apenas no cenário italiano), chega a três conclusões: 1) não é verdade que o direito penal defende todos os indivíduos e protege apenas bens jurídicos essenciais, além de punir as ofensas aos bens realmente essenciais com intensidade diferenciada e de forma fragmentária; 2) não é verdade que a lei penal é aplicada igualmente a todos, tendo em vista que o status de criminoso é distribuído de forma desigual entre os indivíduos que compõe a sociedade, afetando com mais rigor os marginalizados; e 3) A distribuição do status de criminoso ocorre de maneira independente da danosidade social das ações individuais ou da gravidade das infrações cometidas.

No cenário jurídico brasileiro, a desigualdade na criminalização primária fica evidente no maior rigor punitivo nas condutas que normalmente são praticadas por indivíduos de classes econômicas menos favorecidas, de modo que uma ilustração clara persiste (no que pese a existência de críticas) no art. 28, §2º, da Lei de Drogas (Lei 11.346/06) quando há menção direta a circunstâncias “pessoais” e “sociais” no processo de determinação de um indivíduo como usuário ou traficante:

Art. 28. Quem adquirir, guardar, tiver em depósito, transportar ou trazer consigo, para consumo pessoal, drogas sem autorização ou em desacordo com determinação legal ou regulamentar será submetido às seguintes penas:

- I - advertência sobre os efeitos das drogas;
- II - prestação de serviços à comunidade;
- III - medida educativa de comparecimento a programa ou curso educativo.

[...]

§ 2º Para determinar se a droga destinava-se a consumo pessoal, o juiz atenderá à natureza e à quantidade da substância apreendida, ao local e às condições em que se desenvolveu a ação, às circunstâncias sociais e pessoais, bem como à conduta e aos antecedentes do agente (Brasil, 2006).

Para Chaves Júnior e Mendes (2008, p. 10), as malhas dos delitos são ainda mais sutis nos chamados crimes de “colarinho branco” praticados pela população mais abastada. Nesses casos, a legislação elenca vastas hipóteses de redução de pena ou a garantia de benefícios, como ocorre no arrependimento posterior, previsto no art. 16, do Código Penal:

Art. 16 – Nos crimes cometidos sem violência ou grave ameaça à pessoa, reparado o dano ou restituída a coisa, até o recebimento da denúncia ou da queixa, por ato voluntário do agente, a pena será reduzida de um a dois terços (Brasil, 1940).

Se é evidente a existência de um direcionamento do poder punitivo ainda na fase de elaboração das normas e de sua “aplicação abstrata”, o mesmo ocorre no que diz respeito à criminalização secundária (e com clareza na violência exercida, seja real ou simbólica³⁰).

Segundo Rezende (2011, p. 40), a criminalização secundária é marcada em grande parte pelo poder exercido pelas forças policiais, que desempenham a tarefa de selecionar a ocorrência do desvio e o desviante.

Importa ressaltar que apesar de ter grande poder no processo de “propulsão da máquina punitiva estatal”, as agências policiais são as que possuem o menor poder discursivo ou mesmo capital intelectual, além de terem baixas remunerações e de

³⁰ A diferença entre violência real e simbólica pode ser entendida a partir de suas formas de manifestação e impacto. A violência real envolve o uso direto da força física ou agressão para impor a vontade de um agente sobre outro, resultando em danos tangíveis, como lesões ou morte. Já a violência simbólica, segundo a análise de Pierre Bourdieu, refere-se a uma forma de dominação mais sutil, em que o dominado aceita a sua submissão como algo natural ou inevitável. Isso ocorre porque o dominado utiliza os mesmos códigos e instrumentos de conhecimento que o dominador para interpretar a sua posição no mundo, reforçando a desigualdade sem a necessidade de coerção física (Bourdieu, 2001, p. 208).

terem seus membros originados da mesma população na qual ela recolhe os desviantes (Rezende, 2011, p. 40).

Trata-se de um processo histórico e já criticado por Fanon (como citado no início do presente tópico), que nos relembra desse fenômeno como uma herança da colonialidade:

Nas regiões coloniais, em contrapartida, o policial e o soldado, por sua presença imediata, suas intervenções diretas e frequentes, mantêm o contato com o colonizado e lhe aconselham, com coronhadas e napalm, que fique quieto. Como vemos, o intermediário do poder utiliza uma linguagem de pura violência (Fanon, 1974, p. 55).

A clara seletividade nos processos de elaboração das normas e na atuação das agências policiais (e no próprio processo penal) fez com que fosse possível sustentar a eventual adoção de um modelo do “Direito Penal do Inimigo”, tal como proposto por Günther Jakobs³¹.

Carvalho (2006, p. 219-220), contudo, sustenta que o Brasil sequer caminha para a adoção de um modelo processual rígido pautado na existência de figuras “inimigas” e outro garantista para os cidadãos (tal como o “Direito Penal do Inimigo” de Jakobs). Segundo o autor, o que ocorre é o aprofundamento de um modelo caótico e autoritário de exceção sem que tenha ocorrido sequer a etapa de construção de um modelo de garantias. Ocorreu, assim, um abandono definitivo da busca pela efetivação de garantias básicas e uma naturalização da seletividade penal.

Ainda quanto à criminalização secundária, a naturalização da violência e ostensividade contra determinadas populações (em evidência a população negra) tem sido graduada a partir do uso de aparatos algorítmicos, conforme demonstrado no decorrer do presente capítulo.

No que diz respeito à criminalização terciária, é necessário mencionar a já seletiva ideia de periculosidade criminal e novos instrumentos que geram a continuidade dessa estigmatização a partir de algoritmos.

³¹ O modelo do Direito Penal do Inimigo propõe uma aplicação de tratamento mais rigoroso nos processos de criminalização contra determinados infratores que são considerados riscos para a sociedade. Tal modelo possui características como a antecipação da punibilidade, a criação de legislações combativas e a supressão de garantias processuais. O tratamento diferenciado dos seres humanos é, conforme Zaffaroni (2011, p. 11), algo próprio do Estado absoluto e, portanto, incompatível com a teoria política do Estado de direito. A crítica também é feita por Carvalho (2006, p. 219), ao mencionar que apesar de o Estado de direito poder tentar se preservar a partir da distinção entre os cidadãos (na dicotomia cidadão – inimigo), Jakobs abre grande espaço para a adoção de um modelo de exceção permanente.

Sobre o termo “periculosidade criminal”, trata-se de conceito jurídico aproximado às ciências do comportamento, sendo uma referência originalmente adotada no processo de aplicação de institutos como medidas de segurança; medidas cautelares; avaliação de prevenção de delitos e nas várias camadas da execução da pena (como na progressão de regime).

Sánchez-Arjona nos recorda que esse conceito nasceu nas bases do positivismo criminológico de Lombroso, Ferri e Garofalo (ao final do século XIX) como forma de analisar a predisposição do sujeito a cometer delitos futuros. O conceito (e os instrumentos jurídicos-penais ligados a ele) foram se aperfeiçoando, o que cedeu lugar ao que ficou chamado de *violence risk assessment* (avaliação do risco de reincidência ou de violência), que focava não em determinar a presença ou ausência da qualidade subjetiva de perigoso, mas em avaliar uma série de fatores (pessoais, ambientais, temporais, etc.) que resultaria na comissão de novos crimes (Sánchez-Arjona, 2024, p. 107).

Como conceito advindo da criminologia lombrosiana com mais de uma centena de anos, é de se esperar que elementos como racismo e seletividade já estejam presentes na ideia de periculosidade criminal, mas é necessário mencionar que esses elementos não estão ausentes nas normas formas atuariais de avaliação de risco. É preciso esclarecer que novos aparatos tecnológicos geram novas formas de investigação. Entretanto, ao mesmo tempo em que podem gerar maior eficácia em determinados assuntos, podem também provocar sérios danos a garantias fundamentais individuais. Conforme menciona Sánchez-Arjona (2024, p. 110), as ferramentas atuariais não são tão recentes e já existem em diversos campos:

Certamente, o uso desses recursos tecnológicos se estende a uma já infinidade de atividades que enquadram a vida das pessoas, como são as decisões que estão relacionadas com a concessão de um empréstimo, a contratação de um seguro privado de vida, o reconhecimento de um crédito, a automatização e melhora da produtividade do tecido empresarial, no âmbito sanitário... chegando inclusive em países como a China a estabelecer um sistema que classifica o civismo de seus cidadãos através de um sistema de crédito social no qual ganham ou perdem pontos em função das normas estabelecidas e de seu comportamento na sociedade, e tudo isso com base nas ações que ficam registradas em rede.

No que pese as violações que possam ocorrer nas outras esferas do direito, no campo criminal os danos são bem conhecidos. O já mencionado COMPAS (instrumento de análise de risco para perfilamento e conceção de sanções

alternativas), ao ser analisado em 2016 por Jeff Larson, Surya Mattu, Lauren Kirchner e Julia Angwin, gerou uma pesquisa com os seguintes resultados:

- Black defendants were often predicted to be at a higher risk of recidivism than they actually were. Our analysis found that black defendants who did not recidivate over a two-year period were nearly twice as likely to be misclassified as higher risk compared to their white counterparts (45 percent vs. 23 percent).
- White defendants were often predicted to be less risky than they were. Our analysis found that white defendants who re-offended within the next two years were mistakenly labeled low risk almost twice as often as black re-offenders (48 percent vs. 28 percent).
- The analysis also showed that even when controlling for prior crimes, future recidivism, age, and gender, black defendants were 45 percent more likely to be assigned higher risk scores than white defendants.
- Black defendants were also twice as likely as white defendants to be misclassified as being a higher risk of violent recidivism. And white violent recidivists were 63 percent more likely to have been misclassified as a low risk of violent recidivism, compared with black violent recidivists.
- The violent recidivism analysis also showed that even when controlling for prior crimes, future recidivism, age, and gender, black defendants were 77 percent more likely to be assigned higher risk scores than white defendants (Larson; Mattu; Kirchner; Angwin, 2016).

Em linhas gerais, é possível traduzir e explicar tais resultados da seguinte maneira: 1) réus negros frequentemente eram avaliados com um risco de reincidência mais alto do que realmente tinham. A análise dos pesquisadores descobriu que réus negros que não reincidiram ao longo de um período de dois anos tinham quase o dobro de chances de serem incorretamente classificados como de maior risco em comparação com seus equivalentes brancos (45% contra 23%); 2) réus brancos frequentemente eram avaliados com menos risco do que realmente apresentavam. A análise descobriu que réus brancos que reincidiram dentro dos dois anos seguintes eram erroneamente rotulados como de baixo risco quase duas vezes mais frequentemente do que reincidentes negros (48% contra 28%); 3) a análise também mostrou que, mesmo controlando por crimes anteriores, reincidência futura, idade e gênero, réus negros tinham 45% mais chances de receberem pontuações de risco mais altas do que réus brancos; 4) réus negros também tinham duas vezes mais chances do que réus brancos de serem incorretamente classificados como de maior risco de reincidência violenta. E reincidentes violentos brancos tinham 63% mais chances de serem incorretamente classificados como de baixo risco de reincidência violenta em comparação com reincidentes violentos negros.

Não obstante, a análise de reincidência violenta também mostrou que, mesmo controlando por crimes anteriores, reincidência futura, idade e gênero, réus negros tinham 77% mais chances de receberem pontuações de risco mais altas do que réus brancos.

Demonstra-se, dessa maneira, a perpetuação da seletividade penal e racismo em todos os âmbitos dos processos de criminalização (da primária à terciária). O capítulo, contudo, visa apontar com maior aprofundamento a seletividade nas ferramentas presentes no processo de criminalização secundária a partir da atuação policial com instrumentos de policiamento preditivo.

4.4 AS NOVAS FERRAMENTAS DE SELETIVIDADE PENAL

Algoritmos e aprendizado de máquina são uma realidade no Poder Judiciário de inúmeros países. Conforme já destacado, no Brasil, o Supremo Tribunal Federal anunciou a criação, em 2018 de um projeto de Inteligência Artificial (VICTOR), no intuito de facilitar a busca de temas de repercussão geral, o agrupamento de processos por tema, entre outras matérias (Maia Filho; Junquilha, 2018, p. 221).

Da mesma maneira, é de se esperar que também haja o uso de ferramentas similares em órgãos auxiliares, como nas forças policiais (ostensivas ou investigativas).

Em países como os Estados Unidos, as chamadas ferramentas de policiamento preditivo já fazem parte de quase toda estrutura do processo de criminalização secundária, o que tem gerado um extenso debate em relação à eficácia desses sistemas e a possibilidade de violação a direitos fundamentais.

A título de exemplo, a Geolitica, anteriormente conhecida como PredPol (até uma reformulação em 2021), produziu um software que processa dados de relatórios de incidentes criminais e gera previsões diárias sobre onde e quando os crimes são mais propensos a ocorrer. Contudo, conforme o resultado de alguns levantamentos, as predições de crimes realizadas pela ferramenta não são sempre alinhadas (em verdade, raramente são) com os crimes que são efetivamente reportados.

Segundo pesquisa feita pelo grupo The Markup³², foram analisadas 23.631 predições geradas pela Geolitica entre 25 de fevereiro a 18 de dezembro (para o

³² Pode ser considerada uma redação sem fins lucrativos que investiga como instituições de poder fazem uso de ferramentas tecnológicas para mudar aspectos da sociedade (The Markup, 2018).

Departamento de Polícia de Plainfield). Cada previsão da empresa indicava que um tipo de crime era provável de ocorrer em um local não patrulhado pelo departamento policial. No final, a taxa de sucesso foi inferior a meio por cento. Menos de 100 das previsões coincidiram com um crime na categoria prevista (Sankin; Mattu, 2023).

Uma análise mais aprofundada examinou as previsões para roubos ou agressões graves que provavelmente ocorriam em Plainfield. Percebeu-se uma taxa de sucesso igualmente baixa: 0,6 por cento. O padrão foi ainda pior quando analisadas as previsões de arrombamentos, que tiveram uma taxa de sucesso de 0,1 por cento (Sankin; Mattu, 2023).

O problema, contudo, não está apenas na presença de erros. Há pesquisas desde o ano de 2021 demonstrando que o software da Geolitica tendia a direcionar desproporcionalmente bairros de baixa renda, negros e latinos em 38 cidades do país. A citada investigação foi baseada em dados baixados em janeiro de 2021 de um servidor de armazenamento em nuvem desprotegido, vinculado a uma página no site do Departamento de Polícia de Los Angeles. Esse servidor armazenava previsões que a Geolitica havia fornecido a departamentos de polícia em dezenas de cidades do país ao longo de quase três anos. O acesso a esse servidor foi posteriormente restrito (Sankin; Mattu, 2023).

Não obstante, é preciso ressaltar que a citada investigação analisou apenas dois dos 38 departamentos de polícia da localidade, tendo em vista que foram os únicos que forneceram aos pesquisadores dados sobre quando os policiais patrulharam as áreas previstas. A Geolitica afirmou, à época, que enviar policiais para um local previsto dissuadiria crimes apenas com a presença dos agentes. É impossível, assim, determinar com precisão a eficácia do programa sem saber a quais previsões os policiais responderam e a quais não responderam (Sankin; Mattu, 2023).

Com efeito, os registros mostraram que nenhuma das prisões que ocorreram nos locais previstos durante o período de análise poderia ter sido concebivelmente devido a policiais respondendo a uma previsão da Geolitica, o que foi confirmado pelo departamento responsável. Da mesma forma, outros 37 departamentos investigados foram indagados se eles poderiam apontar alguma prisão que tenha sido resultado direto de uma previsão da Geolitica, mas nenhum deles foi capaz de responder (Sankin; Mattu, 2023).

Diante da demonstração de ineficácia do sistema, a reportagem aponta que a Geolitica encerrou as operações estudadas no final do ano de 2023.

O que pode ser dito é que há um grande problema com o sistema da Geolítica, conforme foi usado em Plainfield, tendo em vista que havia um número enorme de previsões em comparação com um número relativamente pequeno de crimes. Plainfield, que tem uma população de cerca de 54.000 habitantes, teve uma média de cerca de 80 previsões por dia para apenas alguns tipos de crimes, enquanto o número máximo de crimes de qualquer tipo relatado em um único dia foi 22 durante o período de análise. Esse volume de previsões está abaixo da mediana de 107 previsões diárias para as outras 37 localidades analisadas na pesquisa original (Sankin; Mattu, 2023).

Quanto à latente seletividade, é preciso mencionar que o sistema analisado pode ainda ter corroborado para o aumento de comportamentos delinquentes por parte da população marginalizada, corroborando com a noção de profecia autorrealizável (*self fulfilling prophecy*³³). Essa conclusão é possível, tendo em vista a existência de um estudo feito em 2019 e publicado na *Proceeding of the National Academy of Sciences* (PNAS) que desvelou que adolescentes negros e latinos abordados pela polícia posteriormente experimentaram níveis elevados de angústia emocional, levando a um aumento do comportamento delinquente no futuro.

O estudo, cujo título pode ser traduzido como: “Os efeitos criminogênicos e psicológicos das abordagens policiais em adolescentes negros e latinos”³⁴ investigou os impactos das paradas policiais em adolescentes. Os resultados, baseados em dados coletados de meninos predominantemente negros e latinos no nono e décimo ano escolar, revelaram que os adolescentes parados pela polícia relataram maior envolvimento em comportamentos delinquentes 6, 12 e 18 meses após a abordagem, independentemente de terem histórico de delinquência anterior. Além disso, a

³³ Cunhado por Robert K. Merton, a profecia autorrealizável é um conceito sociológico que busca descrever o fenômeno em que uma crença falsa influencia o comportamento das pessoas, moldando a realidade conforme essa crença. No campo da criminologia, Rotulação e a profecia autorrealizável explicam comportamentos criminosos e antissociais em termos de fatores sociais. A profecia mostra como uma crença ou expectativa sobre uma pessoa/grupo, correta ou não, afeta a forma como essa pessoa/grupo vai se comportar em momentos futuros. O conceito de rotulação é, portanto, necessariamente relacionado. Rotular é usar um termo genérico e amplo para descrever membros de um grupo minoritário. Os rótulos utilizam termos inferiores e implicam conotações negativas. O termo está relacionado ao conceito de estereotipagem, que é uma visão generalizada e simplificada de um indivíduo ou situação baseada em evidências limitadas. Rotular alguém como criminoso e então tratar essa pessoa como tal pode levar ao comportamento criminoso na pessoa que é submetida a essa expectativa. Isso ocorre porque eles passam a se ver em termos do rótulo criminoso. A pessoa cumpre o rótulo de criminoso e se torna o que o rótulo diz. A profecia, dessa forma, autorrealiza-se (Bearman; Hedström, 2009).

³⁴ “*The criminogenic and psychological effects of police stops on adolescent black and Latino boys*” no título original.

angústia psicológica foi identificada como um fator mediador nessa relação (Del Toro; Lloyd; Buchanan, Goff, 2019).

4.4.1 A polêmica quanto à utilização da ferramenta PredPol

O já referido PredPol (simplificação do termo *Predictive Policing*) foi cofundado pelos antropólogo Jeff Brantingham e matemático George Mohler. Não se trata de uma tecnologia recente, tendo em vista que foi concebida e utilizada entre os anos de 2011 e 2012 em cidades como Los Angeles (Lucena, 2020, p. 5; Yang, 2019, p. 4).

A ferramenta fazia uso de informações de crimes passados para orientar ações policiais futuras (como também fez seu sucessor, mencionado no tópico anterior). Esses dados eram fornecidos pela polícia e normalmente eram constituídos pelo tipo de crime, local e horário, que eram utilizados para treinar seu algoritmo de aprendizado de máquina e prever onde e quando certos tipos de delitos teriam maior probabilidade de ocorrer no futuro.

Em relação a características mais técnicas, Yang (2019, p. 4) menciona que, com base nas previsões geradas pelo computador, o PredPol marcava uma interface web por meio do Google Maps com pequenos quadrados vermelhos de 150m × 150m que indicavam as áreas de maior risco de crimes futuros. Em seguida, notificava os oficiais de patrulha, instruindo-os a patrulhar essas áreas por aproximadamente 10% de seu turno. Além de fazer previsões para fins de controle de crimes, o PredPol também auxiliava em tarefas administrativas na gestão policial.

Apesar das Críticas, Yang (2019, p. 5), em artigo intitulado *Predictive Policing*, sustenta que o PredPol coletava apenas dados de tipos de crime, localização e momento, motivo pelo qual possuiria reduzido viés racial.

Surge, daí, uma indagação fundamental: se a ferramenta faz uso apenas de dados objetivos, como delitos praticados e horários, como podem ter surgido as discussões a respeito do enviesamento de suas informações?

É preciso mencionar que a preocupação com o enviesamento racial nos dados do policiamento preditivo não é infundada porque, mesmo que a ferramenta utilize apenas dados objetivos como local, hora e crimes anteriores, esses dados podem (e muitas vezes o fazem) refletir desigualdades sistêmicas e práticas de policiamento desproporcionais que já existem (haja vista a já comentada tradicional seletividade penal). Dessa forma, se uma área específica foi historicamente mais policiada do que

outras por motivações raciais, isso pode levar a um número maior de relatos de crimes nessa área, não necessariamente porque mais crimes ocorrem lá, mas simplesmente porque há mais policiamento.

Portanto, mesmo que a intenção seja usar dados objetivos, o resultado pode ser um sistema que inadvertidamente perpetua preconceitos existentes. É por isso que muitos especialistas, conforme demonstrado no trabalho, enfatizam a importância de analisar e ajustar os algoritmos de policiamento preditivo para garantir que eles não reforcem esses padrões.

É válido destacar que o intenso debate ocorrido nos últimos anos a respeito de tal ferramenta fez com que sua utilização fosse descontinuada em diversos departamentos, enquanto outros ainda fazem uso de ferramentas similares, algumas reformuladas a partir da ideia original, como a Geolítica (Sankin; Mattu, 2023).

Algumas características e problemas relacionados a esse tipo de ferramentas já foram bem apontados por alguns autores, como Lucena (2019) e comentados no primeiro capítulo, mas merecem algum aprofundamento.

Havia muito alvoroço em relação à ferramenta no momento de sua criação. Desenvolvido a partir de um sistema de prevenção de terremotos, o PredPol foi considerado ainda no início da década de 2010 como uma das 50 melhores invenções anuais pela revista *Time* (Lucena, 2020, p. 5).

Em verdade, alguns pesquisadores enxergam benefícios no uso e desenvolvimento de ferramentas preditivas, como Fei Yang (2019) e até mesmo Cathy O'Neil. Entretanto, Lucena (2020, p. 6), ao ressaltar a pesquisa desta última, alerta que, no caso do PredPol, há clara presença de um caráter discriminatório (o que foi evidenciado a partir da análise de grupos como o *Human Rights Data Analysis Group*, formado por hackers experientes).

De certa forma, a análise desses grupos demonstra que dificilmente o sistema estará livre de certos vieses se não houver uma alteração drástica no seu modelo estruturante, que no caso do PredPol, são as políticas criminais estadunidenses de Tolerância Zero (Lucena, 2020, p. 6), historicamente responsáveis por potencializar a criminalização secundária na população pobre, negra e latina.

Falar do modelo é sempre relevante porque ele reflete determinados objetivos e até ideologias, muito embora possam passar uma certa impressão de imparcialidade (o que não há, como já discutido no primeiro capítulo).

Cathy O’Neil deixa claro que os modelos nada mais são que opiniões e valores expressos por meio de equações matemáticas (O’Neil, 2020, p. 35).

Com o intuito de esclarecer ainda mais, é preciso que se tenha em mente que não apenas quem é o responsável pelo desenho do modelo, mas também qual a finalidade que essa pessoa (ou grupo) deseja alcançar.

A título de exemplo, é possível imaginar a criação de um modelo de “vida saudável a ser seguida”. É certo que se uma pessoa criar tal modelo, ainda que ela seja especialista em nutrição ou educação física, as dicas que ela passará não são completamente baseadas em manuais ou literatura especializada. É bem provável que ela indicará algum esporte que goste, alguma comida que ache boa ou mais saudável que outra. Ainda que seus gostos estejam dentro de certos parâmetros técnicos, eles ainda são gostos e isso estará necessariamente atrelado a um eventual sistema algorítmico a ser criado.

Como já destacado, segundo a autora, o próprio racismo pode ser considerado um modelo criado a partir de informações rasas e equivocadas que indicam falsamente que todas as pessoas de um determinado grupo se comportam de uma determinada maneira (O’Neil, 2020, p. 37).

A autora, portanto, deixa claro que o PredPol dificilmente pode ser dissociado da histórica questão racial no sistema de justiça estadunidense e na atuação concreta da atividade policial³⁵.

O mesmo ocorre em sistemas de previsão de reincidência, como o LSI-R (*Level of Service Inventory – Revised*). Não é possível remover o viés humano, muito embora isso seja tentado por meio da camuflagem da tecnologia.

Nesse sistema de reincidência, busca-se uma suposta imparcialidade a partir de um questionário que deve ser respondido pelo detento, com perguntas como o número de condenações anteriores, as influências e abusos de substâncias como álcool e outras drogas.

Conforme O’Neil menciona, o conjunto de pressupostos humanos está sempre lá, embutido na ferramenta:

³⁵ Para além da criminalização secundária no âmbito policial, a autora também aponta dados da Universidade de Maryland, demonstrando que os promotores de justiça do estado do Texas eram três vezes mais inclinados a realizar pedidos de pena de morte para condenados afro-americanos e até quatro vezes para condenados hispânicos, se comparados aos condenados brancos acusados dos mesmos delitos (O’Neil, 2020, p. 38).

Mas conforme as questões continuam indo mais a fundo na vida da pessoa, é fácil imaginar como detentos de origem privilegiada iriam responder de um jeito e aqueles vindos das árduas ruas do centro da cidade de outro. Pergunte a um criminoso que cresceu num confortável bairro nobre sobre “a primeira vez em que você se envolveu com a polícia”, e ele pode não ter um único incidente a relatar além daquele que o fez ser preso. Homens negros jovens, em contrapartida, provavelmente já foram parados pela polícia dezenas de vezes mesmo sem ter feito nada de errado (O’Neil, 2020, p. 41).

No caso do policiamento preditivo, a questão se torna igualmente complexa com o uso do algoritmo SSL (*Strategic Subject Lists*) nas ferramentas como o PredPol. Muito embora haja alegação por parte da polícia de não utilização de tal lista (que ranqueia indivíduos em pontuações de 0 a 500 a depender de sua proximidade com a criminalidade), Lucena (2020, p. 6) menciona que há documentos oficiais que se tornaram público demonstrando exatamente o contrário. Segundo menciona, ao menos na cidade de Chicago, mais de 50% das pessoas negras das idades que variam de 20 a 29 anos estão inseridas na lista nos maiores índices, enquanto as pessoas brancas de mesma idade representam apenas 2%.

Menciona-se, por fim, a pesquisa intitulada “*Racism Repeats Itself: AI Racial Bias in Predictive Policing Algorithms*”, realizada pelas pesquisadoras Alexia Gallon e Emily Liu (ambas de Irvine Valley College) e publicada em 2023.

Gallon (2023, p. 1) desenvolveu seu próprio algoritmo de policiamento preditivo usando aprendizado supervisionado por máquina. Ela treinou esse algoritmo com três tipos de dados dos arquivos do Departamento de Polícia de Los Angeles (LAPD): tipo de crime, horário e local. Em seguida, ela comparou a precisão das previsões do seu algoritmo com as do PredPol.

Ela observou tendências no comportamento dos algoritmos de policiamento preditivo (PPAs) e argumenta que esses sistemas perpetuam disparidades raciais devido a três fatores principais: 1) *implicit racial biases*; 2) *tech-washing phenomenon*; e 3) *disproportionate minority contact*³⁶.

³⁶ Os fatores alegados por Gallon (2023) podem ser respectivamente traduzidos e explicados como: 1) vieses raciais implícitos - haja vista que os dados históricos usados para treinar os algoritmos podem conter vieses raciais, que são então replicados nas previsões do algoritmo (como já exaustivamente explicado no trabalho); 2) fenômeno de “lavagem tecnológica” - que nada mais é além da ideia equivocada de que a tecnologia é neutra e objetiva, podendo mascarar os vieses embutidos nos algoritmos, o que leva à falsa percepção de que as previsões são imparciais; e 3) contato desproporcional com minorias - já que as previsões dos PPAs (*Predictive Policing Algorithms*) podem levar a uma aplicação da lei desproporcionalmente direcionada a minorias, exacerbando as desigualdades pré-existentes.

Em conclusão, Gallon (2023, p. 17) em seu estudo, encontrou uma disparidade evidente entre os crimes documentados nos bairros de Los Angeles e as comunidades historicamente marginalizadas (negros e latinos). Isso levou à conclusão de que os dados de treinamento usados nos algoritmos de policiamento preditivo (do PredPol) são incompletos desde sua concepção.

Ao comparar o algoritmo PredPol com o seu próprio, ela demonstrou que os algoritmos tendem a favorecer previsões para americanos brancos e a desfavorecer previsões para pessoas de origem afro-americana, resultando em menos previsões de crimes para os primeiros e uma quantidade significativamente maior para os últimos, perpetuando assim as disparidades raciais.

4.4.2 Nova ferramenta e antigos problemas: o Palantir

O problema mencionado não foi verificado apenas no PredPol. Em texto publicado no MIT Technology Review, Will Douglas Heaven, pesquisador e PhD em Ciência da Computação, defende uma radical abolição de todas as ferramentas de policiamento preditivo (Heaven, 2020).

Essa defesa de um encerramento em tais atividades decorre desses históricos e recorrentes abusos de poder policial que ocorrem não apenas nas ruas, mas também em manifestações e protestos em escolas e universidades.

O pesquisador também aponta que há dois sistemas de predição delituosa³⁷. Há os algoritmos baseados em localização que se baseiam em ligações entre lugares, eventos, taxas históricas de crimes e prisões para prever onde e quando delitos são mais prováveis de acontecer (a ferramenta pode, por exemplo, indicar a provável ocorrência de crimes em situações pontuais, como eventos esportivos). Muito embora esses dados sejam mais objetivos (como demonstrado por Yang, nesse mesmo capítulo), sabemos que isso não significa que tais dados não estejam viciados.

O segundo tipo de instrumentos se baseia em dados sobre pessoas, como idade, gênero, estado civil, histórico de abuso de substâncias e registro criminal, para prever quem tem uma alta chance de se envolver em atividades criminosas futuras.

³⁷ É preciso afirmar que predição delituosa é diferente de policiamento preditivo. Enquanto este último se refere exclusivamente às ferramentas utilizadas pelas agências policiais em suas atividades típicas, a primeira expressão engloba também as ferramentas utilizadas durante a execução criminal, como aquelas utilizadas para determinar a probabilidade de reincidência.

Essas ferramentas baseadas em pessoas podem ser usadas tanto pela polícia, para intervir antes que um crime ocorra, quanto pelos tribunais, para determinar durante audiências pré-julgamento ou sentenças se alguém que foi preso é provável de reincidir (o mencionado COMPAS ilustra esse sistema muito bem).

Em relação ao que já foi afirmado, Heaven (2020) ainda acrescenta que o crescente uso de algoritmos por departamentos policiais significa que há um verdadeiro corte no orçamento da segurança público dos Estados Unidos e não um aumento. Isso ocorre porque a presença física do policial está sendo substituída por sistemas de informação.

A partir do que foi estabelecido pelos pesquisadores mencionados, é possível perceber que o problema não está na estrutura do instrumento, mas no vício da própria informação que alimenta esses sistemas. Desse modo, pode-se concluir que eventuais vícios presentes no PredPol podem estar presentes em outros instrumentos similares. Isso é o que outras pesquisas têm demonstrado.

As acusações de racismo algorítmico também ocorreram em departamentos policiais que fizeram uso de instrumentos similares, como o Palantir³⁸, utilizado na década passada pela polícia de Nova Orleans.

A utilização do sistema já causou certa preocupação em seu início, tendo em vista que foram disponibilizadas informações limitadas sobre o sistema e sua parceria com a cidade de Nova Orleans, pois o contrato e suas subseqüentes extensões foram firmados com pouca publicidade (Richardson; Schultz; Crawford, 2019, p. 36).

Mais uma vez, as evidências sugerem uma grande chance de que o sistema Palantir tenha se baseado em alguma forma de “*dirty data*” (dados sujos³⁹) do departamento policial de Nova Orleans (Richardson; Schultz; Crawford, 2019, p. 36). Pode-se dizer que a análise do sistema refletiu disparidades raciais semelhantes e outros preconceitos das práticas e políticas das forças policiais locais. Os pesquisadores demonstraram que documentos do governo da cidade destacaram que o sistema identificou vítimas e perpetradores de crimes violentos ou de gangues como

³⁸ Curioso é o fato de que o nome remete a um artefato da literatura de John Ronald Reuel Tolkien. Trata-se de um instrumento de previsão criado pelos Noldor e mencionado em “O Senhor dos Anéis: As Duas Torres”. Nesta obra, Gandalf afirma que os *palantíri* foram tocados e maculados pelo Senhor do Escuro, Sauron, o que causou a subseqüente queda do mago Saruman, o Branco, que fez uso do artefato para fins malignos (Tolkien, 2002, p. 205).

³⁹ No caso mencionado, os autores sugerem (2019, p. 36) que o sistema Palantir pode ter se baseado em dados considerados problemáticos. Esses dados podem incluir registros errôneos, duplicados, incompletos ou obtidos de forma ilegítima ou discriminatória.

"predominantemente jovens, afro-americanos, homens, pouco escolarizados e subempregados".

O Contrato com a empresa responsável pelo Palantir foi encerrado na cidade no ano de 2018 após grande repercussão negativa do sistema e dos acordos firmados a respeito de sua utilização.

Em síntese, o que se identificou com o Palantir não foi diferente do que o identificado com o PredPol. Trata-se de mais um exemplo: quando tais sistemas utilizam os dados de prisões previamente ocorridas em determinados locais, codificam padrões de comportamento policial racista. Como resultado, é mais provável que prevejam um alto potencial de crime em bairros de minorias ou entre pessoas pertencentes a esses grupos. Mesmo quando os dados de prisões e crimes coincidem, há uma miríade de razões socioeconômicas pelas quais certas populações e certos bairros têm taxas históricas de crimes mais altas do que outros. Alimentar esses dados em ferramentas preditivas permite que situações de injustiças possam ocorrer com mais frequência (sob o sempre presente discurso da neutralidade algorítmica).

Por fim, vale dizer que o uso de dados prisionais para treinar ferramentas preditivas não permitem que haja um retrato preciso da atividade criminal que é efetivamente praticada.

Como já mencionado, Lombroso, ao buscar o homem delinquente na criminologia positivista do séc. XIX, encontrou não o criminoso, mas o criminalizado. Da mesma forma, as ferramentas de policiamento algorítmico não encontram necessariamente os perpetradores de ilícitos, mas aqueles que são normalmente alvos da atividade policial.

4.4.3 O Brasil e a tecnologia na segurança: o projeto Smart Sampa

A discussão a respeito do uso de dados sensíveis no campo da segurança pública e a possibilidade de ocorrência de perfilamento racial no já complexo fenômeno da criminalização secundária também ocorre no Brasil.

Tal discussão também está acompanhada de outro fenômeno: o das cidades inteligentes (smart cities), que podem ser compreendidas como grandes regiões urbanas que fazem uso de avançadas ferramentas de coleta de dados com o intuito

de melhoria da qualidade de vida de seus habitantes, levando em consideração, ao mesmo tempo, a sustentabilidade e a eficiência da prestação de serviços públicos.

Podem ser consideradas “cidades inteligentes” aquelas que fazem uso integrado de TIC (Tecnologia de Informação e Comunicação) a partir da noção de Internet das Coisas (*Internet of Things*, ou IoT) em diferentes searas como infraestrutura, energia, transporte e, certamente, segurança pública.

Conforme a descrição de Weiss, Bernardes e Consoni (2017, p. 4):

As cidades inteligentes fornecem serviços interoperáveis, que facilitam a conectividade para transformar processos de governo internamente entre as agências e departamentos e, externamente, para cidadãos e empresas. Para criar cidades inteligentes, as tecnologias devem estar prontamente integradas, conectando diferentes sistemas em diferentes organizações. Nas cidades inteligentes, a tecnologia da informação desempenha o papel de tornar os dados da vida urbana tangíveis, por meio da criação e execução de projetos voltados para a sua captura e tratamento em tempo real.

Há uma visão positiva por parte de tais projetos, tendo em vista a possibilidade de maior conectividade e contato entre gestores e munícipes, além de busca de ferramentas para uma melhoria de qualidade de vida, com facilidades de diminuição de riscos em trânsito, combate à criminalidade e até mesmo incentivo de redução de emissões de carbono.

Entretanto, não é possível esquecer que a adoção acrítica desse tipo de ferramenta também está ligada ao que foi conceituado no início do presente capítulo como colonialismo digital.

O fenômeno das cidades inteligentes está, portanto, indissociável de um controle exercido por grandes empresas de tecnologia estrangeiras sobre os dados e sobre a própria infraestrutura digital de países em desenvolvimento. Conforme Faustino e Lippold (2023, p. 52) mencionam, não há como se falar em colonialismo sem falar em capitalismo.

Os efeitos dessa nova forma de controle sobre as colônias ocorrem de diferentes maneiras, o que já foi abordado anteriormente no trabalho, mas é preciso frisar as seguintes ocorrências: 1) a influência de grandes empresas de tecnologia pode ser capaz de afetar as políticas locais; 2) a implementação de tecnologias digitais avançadas pode deixar ainda mais escancarada a desigualdade existente nos países e cidades emergentes (o que acarreta benefícios para as áreas urbanas mais abastadas, em detrimento das comunidades periféricas); 3) o eventual controle dos

dados pessoais compartilhados pela população por essas empresas, tendo em vista que legislações locais de proteção de dados podem não ser observadas (ainda mais pelas *big techs*); e 4) a dependência de tecnologias estrangeiras, bem como a naturalização do uso de aparatos de grandes empresas tecnológicas como “única opção viável”.

Apesar da relevância dos quatro pontos mencionados, o último guarda uma estreita relação com o objeto central da presente tese, haja vista que a dependência de tecnologias estrangeiras e uma possível percepção de que essas tecnologias são a única opção viável (ou disponível) podem contribuir para a naturalização do racismo (e das violências realizadas pelo Estado), especialmente por meio de ferramentas de policiamento preditivo.

Como se sabe, tais aparatos utilizam algoritmos que analisam dados históricos para prever onde crimes futuros podem ocorrer. Como esses dados contêm viés racial (vide PredPol e Palantir), os algoritmos tendem a perpetuar e amplificar esse viés, resultando na já apresentada e constante vigilância e policiamento de comunidades negras e marginalizadas. Também se sabe que muitas dessas tecnologias são caixas-pretas, de modo que seus processos internos não são transparentes, o que dificulta a identificação e correção de vieses raciais embutidos nos algoritmos.

Quando cidades na América Latina (notadamente São Paulo) importam essas tecnologias de policiamento preditivo desenvolvidas em outros contextos, elas também importam os problemas e vieses desses modelos.

O uso de policiamento preditivo, portanto, tende a reforçar estereótipos raciais, pois as áreas e indivíduos identificados como “de risco” são frequentemente aqueles que já são alvo de discriminação racial. Isso pode levar a um ciclo vicioso onde a presença policial aumentada em certas comunidades gera mais dados de suposta criminalidade, que, por sua vez, reforçam a necessidade de cada vez mais policiamento.

Quanto aos projetos de cidades inteligentes no Brasil, um expoente é o Smart Sampa, na cidade de São Paulo. O projeto original inclui a instalação de 40 mil câmeras de videomonitoramento, o que inclui câmeras panorâmicas, câmeras com analítico de leitura de placa e câmeras com visão térmica (São Paulo, 2023).

O sistema a ser adotado tem como objetivo permitir o monitoramento em tempo real, além de realizar integração entre serviços públicos, como a Guarda Civil

Metropolitana, além de Polícias Civil e Militar, facilitando a identificação de indivíduos com prisões decretadas e veículos furtados ou roubados (São Paulo, 2023).

O projeto se baseia nos preceitos da Lei nº 14.129/2021, conhecida como a Lei de Governo Digital. Esta legislação tem como objetivo modernizar a administração pública, promovendo o acesso rápido a informações em tempo real, buscando facilitar a tomada de decisões e incentivando a integração e interoperabilidade entre os diversos órgãos governamentais. Além disso, a lei visa simplificar o acesso dos cidadãos aos serviços públicos digitais, estabelecendo um sistema onde um único login permita o uso de diversos serviços de acordo com o perfil do usuário (São Paulo, 2023).

No contexto do programa incentivador de vigilância urbana, que busca consolidar em uma plataforma o maior número possível de câmeras públicas e privadas na cidade de São Paulo, essa legislação é considerada pelos seus defensores como fundamental. A plataforma, que era mantida por uma doação cujo prazo terminou, é, portanto, uma prioridade para o monitoramento de câmeras pela Guarda Civil Metropolitana, e sua reativação foi considerada essencial para garantir a continuidade desse serviço (São Paulo, 2023).

No que pese a discussão a respeito dos novos modelos de colonialismo digital, a integração entre ferramentas digitais como meio de agilizar a prestação de serviços públicos essenciais é algo a ser celebrado, principalmente quando se fala em segurança pública em grandes centros.

Ocorre que é discutível, contudo, a efetividade desse novo sistema de segurança. Pergunta-se se ele é capaz de incorrer nos mesmos problemas que as ferramentas já mencionadas (ainda que as informações dos bancos de dados não sejam alimentadas pelos departamentos policiais estadunidense, não há que se esquecer o racismo e a seletividade penal tipicamente brasileira, também encontrada na atuação das políticas e poder judiciário nacionais).

Tais indagações chegaram a ser discutidas judicialmente, haja vista que, por determinação do Tribunal de Justiça de São Paulo (em pedido fruto de mandato coletivo do Partido Socialismo e Liberdade), o próprio edital do projeto (que estabelecia a compra de milhares de câmeras) foi suspenso (Arreguy; Ribeiro, 2023).

Não foi essa, contudo, a primeira vez que a tecnologia foi questionada, haja vista que o Tribunal de Contas do Município também havia determinado tal suspensão a partir de pedido de diversas organizações que indicavam a possibilidade de

perfilamento racial a partir das câmeras e do sistema adotado (Arreguy; Ribeiro, 2023).

Conforme destaca a decisão:

Inúmeros pesquisadores de diversas áreas (direito, psicologia, sociologia, tecnologia de informação entre outras) e instituições renomadas (do Brasil e do exterior) apontam para riscos concretos de reprodução do racismo estrutural no uso do sistema de monitoramento por reconhecimento facial porque essa tecnologia reproduz padrões de discriminação incorporados na cultura e na dinâmica institucional das sociedades sem permitir qualquer revisão desses graves comportamentos (Arreguy; Ribeiro, 2023).

A sentença ainda destaca sobre a ausência de informações cristalinas a respeito de como será feito o processamento de dados da população. Muito embora a LGPD (Lei Geral de Proteção de Dados), em seu art. 4º, III, “b”, estipule que não há incidência da legislação em atividades investigativas e ligadas a repressão penal, o projeto Smart Sampa também utiliza dados da população também para outros fins.

Vale ressaltar que a suspensão do edital não perdurou muito, tendo em vista que a decisão que impedia a contratação de empresas para a implementação das câmeras foi revertida. Conforme destacado na nova decisão, as evidências de racismo algorítmico não eram concretas para o sistema que sequer tinha sido efetivamente adotado.

Consoante disposto por representantes da prefeitura de São Paulo, o projeto determina que seja criado um conselho de gestão e transparência com a finalidade de garantir o cumprimento das diretrizes da LGPD (G1, 2023).

O Smart Sampa, portanto, deu seguimento a sua implementação, mas isso não impediu que determinados órgãos e associações ainda enxergassem sua aplicação com certa cautela.

O Ministério Público estadual, por sua vez, demandou da capital a realização de medidas concretas para garantir a segurança das informações coletadas e evitar a discriminação no programa (Cardoso, 2024).

A preocupação da promotoria excede o campo do racismo algorítmico nas abordagens policiais, tendo em vista que a documentação encaminhada à prefeitura também tenha apontado riscos de violação à intimidade por meio de tratamento de dados sensíveis, bem como possibilidade de inibição de liberdades como associação e expressão pela possível “vigilância pública exacerbada” (Cardoso, 2024).

A solução apontada pelo Ministério Público foi no sentido de recomendar a criação de ouvidorias com funcionários especializados e capacitados na apuração de relatos discriminatórios, além de também recomendar o estabelecimento de critérios mais rigorosos para o compartilhamento de informações coletadas a partir do programa entre as diversas instituições envolvidas.

4.4.4 Outro breve exemplo brasileiro: a Guarda Civil de Vitória

A prefeitura de Vitória é outro exemplo de capital que busca, no uso de supercomputadores, reduzir os índices de criminalidade. Conforme mencionado por Bento (2024, p. 89), Vitória foi uma das primeiras cidades a implementar um sistema de Infraestrutura Hiperconvergente⁴⁰ (HCI). Conforme apontado pelo mencionado pesquisador, o computador do município (com 900 processadores, 3,7 *terabytes* de memória RAM e 2,5 *petabytes* para armazenamento) possui a capacidade de analisar e processar mais de mil imagens derivadas de diferentes câmeras de videomonitoramento e forma simultânea, o que possibilita o reconhecimento facial de pessoas envolvidas em atividades criminosas, com sua célere identificação e verificação de existência de mandados de prisão.

O objetivo em adotar essas novas tecnologias não está apenas no reconhecimento facial de eventuais suspeitos, mas na busca por uma eficiência em situações como como colisões, furtos ou roubos de veículos, além de pretender detectar comportamentos incomuns em áreas específicas e em horários determinados, como nas estações da Bike Vitória durante o período da madrugada (A Gazeta, 2024).

O autor mencionado, em sua pesquisa, explica que o sistema de vigilância do município será capaz de suportar a instalação de 800 novas câmeras internas em escolas e unidades de saúde. Inicialmente, o que se pretende é a implementação de dez pontos de reconhecimento facial em locais públicos, como praças, escolas e unidades de saúde. Após a fase preliminar de avaliação, a vigilância deve ser

⁴⁰ Infraestrutura Hiperconvergente (HCI) é uma tecnologia que combina computação, armazenamento e rede em um único sistema integrado, gerenciado por um *software* centralizado. No lugar de gerenciar cada componente separadamente, a HCI permite administrar tudo de maneira conjunta, o que simplifica a operação. Essa integração é possível graças ao processo de virtualização, que permite alocar recursos conforme necessário, facilitando a escalabilidade e reduzindo custos (Bigelow, 2022).

expandida para 50 pontos itinerantes, cobrindo mais áreas do município (Bento, 2024, p. 89).

Mais uma vez, no que pese o uso de novos aparatos tecnológicos ser algo imprescindível e necessário no campo da segurança pública, é preciso que suas implicações sejam devidamente discutidas.

Além do que já foi exposto a respeito de preconceitos inerentes às forças de segurança que podem estar inseridos sub-repticiamente nos bancos de dados de tais instrumentos, Bento ainda faz uma abordagem a respeito da ferramenta a partir da perspectiva de Arendt, ao mencionar que o cidadão, ao sujeitar-se às relações de poder legitimamente constituídas, acaba por desconsiderar eventuais atentados aos direitos fundamentais (notadamente à dignidade da pessoa humana), o que muitas vezes ocorre por práticas sexistas ou racistas (Bento, 2024, p. 94).

Em conclusão, a citada pesquisa não difere muito das outras já realizadas e citadas no presente trabalho. As autoridades tendem a negar ou minimizar a possibilidade de que algoritmos possam ser racistas (mantendo a crença em um determinismo tecnológico). O mencionado pesquisador, contudo, baseia-se em teóricos como Michel Foucault e Hannah Arendt para levantar preocupações sobre como a automatização pode perpetuar padrões normativos, o que certamente inclui vieses raciais e sexistas.

Conforme destaca (Bento, 2024, p. 94), as autoridades focam nas propriedades matemáticas dos algoritmos, como se isso garantisse neutralidade, mas essa ênfase matemática ignora as complexidades éticas, especialmente os possíveis vieses de gênero e raça que podem estar embutidos nos algoritmos e demonstrado nas demais ferramentas estudadas.

4.5 O POLICIAMENTO PREDITIVO COMO A APROXIMAÇÃO DE UMA DISTOPIA

Em um artigo chamado “*This Is Not Minority Report: Predictive Policing and Population Racism*”, Joshua Scannel, professor assistente na School of Media Studies (Nova York), faz menção a uma expressão: “*garbage in, garbage out*” (Scannel, 2019, p. 107), que pode ser traduzida, dentro do contexto apropriado, como “lixo entra, lixo sai”.

O que o professor pretende afirmar com essa expressão diz respeito a qualquer ferramenta com sistemas automatizados. Significa dizer que um determinado sistema

apenas será tão bom quanto às informações nas quais ele se baseia. Se o banco de dados de um software apresentar problemas, certamente haverá reflexos disso em sua tarefa final.

No contexto do presente trabalho, é possível concluir, portanto, que se os dados nos quais as ferramentas de policiamento preditivo se baseiam estão viciados devido a problemas raciais já históricos nas agências de polícia estadunidenses, certamente esses vícios poderão ser encontrados na execução da tarefa fim de tais ferramentas.

Nos últimos anos, houve intensos debates e publicações (algumas já citadas) sobre as tecnologias de policiamento preditivo. Torna-se até possível argumentar que um sistema desse tipo, baseado em práticas policiais passadas, inevitavelmente reproduzirá as estruturas de violência estatal relacionadas a raça, gênero e classe. Os desenvolvedores, devem, portanto, ampliar sua rede de vigilância para abranger uma variedade cada vez maior de práticas cotidianas e efeitos, na tentativa de minimizar qualquer viés.

Para Scannel (2019, p. 108), contudo, alguns pontos fundamentais merecem destaque. Em especial: o policiamento não tem uma “história racista”. O policiamento cria raça e é inseparável dela. Algoritmos não podem eliminar a raça do policiamento americano porque a raça é uma tecnologia policial originária, assim como o policiamento é uma tecnologia fundamental de racialização.

Em linhas gerais, isso nos leva a crer que o policiamento, ou seja, a própria atuação policial, é moldada pelas percepções raciais e pelas estruturas de poder que perpetuam as desigualdades. A atuação policial no policiamento, portanto, não apenas reflete preconceitos raciais pré-existentes, mas ativamente contribui para a construção e manutenção de categorias raciais.

Dessa forma, não se pode retirar a ideia de raça do policiamento estadunidense porque raça é uma parte fundamental da própria tecnologia policial.

O entendimento do autor vai ao encontro do que já foi estabelecido a respeito da existência de uma latente seletividade racial não apenas nas ferramentas de policiamento preditivo, mas em todo processo de criminalização: primária (processo legislativo); secundária (ação policial e persecução); e terciária (encarceramento).

Certos grupos, portanto, são mais propensos a serem alvo de policiamento e criminalização com base em sua raça, classe ou outros fatores. A seletividade é uma das maneiras pelas quais o policiamento “produz” o crime e o criminoso, perpetuando

desigualdades e reforçando estruturas que, conforme já salientado por Browne e Almeida, são intrinsecamente racistas.

É interessante destacar a forma como o autor mencionado até mesmo faz uma análise de tais ferramentas a partir de uma comparação com a obra *Minority Report*, escrita por Philip K. Dick, em 1956.

Minority Report é um conto de ficção científica em que a narrativa se passa em uma sociedade futurista onde crimes de homicídio são prevenidos antes de acontecerem, graças aos “*precogs*” (indivíduos com habilidades de previsão do futuro). Esses “*precogs*” conseguem ver os crimes antes que ocorram (Dick, 2009).

A obra é comumente apresentada como um enredo futurístico e distópico que poderia servir como um aviso (ou alerta) a respeito dos perigos da concentração de uma tecnologia poderosa de previsão nas mãos do Estado. Scannel, contudo, menciona que essa leitura da obra é rasa e não destaca seu propósito fundamental: o intrincado envolvimento do capitalismo americano com a eugenia, o encarceramento e a escravidão (Scannel, 2019, p. 108).

Na obra original, o estado nacional sequestra crianças nascidas com um tipo específico de hidrocefalia e as transforma em computadores vivos. Elas são confinadas pelo resto da vida nos porões de repartições governamentais e instituições financeiras, onde máquinas extraem previsões de curto prazo.

Os mutantes “*precogs*” na história de Dick são chamados de “*idiotas tagarelas*” (“*babbling idiots*”) e “*macacos*” (“*Monkeys*”). Suas prisões são chamadas de “*asas de macaco*” (“*monkey wings*”). Presume-se que eles não sintam dor, emoções ou questionem sua escravidão. São considerados apenas ferramentas (Dick, 2009). Eles definham rapidamente e morrem jovens. A sociedade construída com base em seu trabalho não pode sobreviver sem eles. Da narrativa do protagonista Anderton, eles desaparecem rapidamente sem seres notados.

Conforme Scannel (2019, p. 109), é possível traçar uma trajetória específica de pensamento liberal-reparador nessa narrativa. No futuro de Dick, há um pesadelo implacável no qual a acumulação de riqueza, a encarceramento em massa e o poder estatal se articulam mutuamente. O trabalho ciborgue dos escravizados, justificado pelo capacitismo eugênico, é dado como certo e não é questionado.

Assim, conclui que se a ficção científica aponta para uma realidade de encarceramento em massa e desapropriação organizada, ela também realiza uma “falsificação de memória e significado” (Scannel, 2019, p. 110) que reescreve a história

da democracia americana como algo separado da história da opressão racial e exploração, que é tanto a condição real quanto imaginada do mundo.

Por fim, é possível alcançar às seguintes conclusões a partir do Texto de Scannel. Em primeiro lugar, o policiamento preditivo oferece um verdadeiro paradoxo: quanto mais confiamos em previsões baseadas em dados históricos, mais perpetuamos as desigualdades existentes. Se esses dados históricos refletem práticas discriminatórias, o policiamento preditivo pode reforçá-las.

Em segundo lugar, o policiamento preditivo pode distorcer a própria compreensão da realidade. Ao confiarmos cegamente em previsões algorítmicas, podemos ignorar contextos sociais, econômicos e históricos mais amplos, o que pode levar a uma visão limitada e até mesmo simplificada da criminalidade e da segurança pública.

5 A POSSÍVEL CRIAÇÃO DE TECNOLOGIAS NÃO VIOLADORAS DE DIREITOS: COMO EVITAR UM ETERNO RETORNO A CESARE LOMBROSO

Se as ferramentas de policiamento preditivo analisadas (e aquelas que, apesar de não terem sido mencionadas no trabalho, funcionam de forma similar) representam uma verdadeira possibilidade de violação a direitos, o presente capítulo busca analisar o possível desenvolvimento de sistemas algorítmicos no campo da segurança pública que não reproduzam preconceitos raciais e outras violações. A referência a Cesare Lombroso, serve, mais uma vez, como uma metáfora para o perigo de um "retorno" às práticas discriminatórias por meio das tecnologias contemporâneas.

Também se insere como objetivo deste capítulo investigar o que tem sido feito globalmente, com especial ênfase na América Latina, para regulamentar o uso de ferramentas de policiamento preditivo e outras tecnologias baseadas em inteligência artificial.

Questões sobre como essas ferramentas podem ser desenvolvidas e aplicadas de maneira que respeitem os direitos humanos, sem perpetuar estigmas históricos, são o ponto central nesta discussão.

Consequentemente, busca-se verificar se sistemas de segurança pública realmente podem ser projetados sem viés racial. Tecnologias que pretendem prever comportamentos futuros baseadas em dados passados estão, conforme demonstrado exaustivamente no presente trabalho, ligadas a contextos históricos de opressão, especialmente em sociedades onde o racismo estrutural moldou as políticas de segurança. Este capítulo, portanto, examina soluções propostas por pesquisadores e desenvolvedores de tecnologia para minimizar esses vieses, buscando avaliar até que ponto tais esforços têm sido eficazes e se há esperança de criação de tecnologias verdadeiramente não violadoras.

Destacou-se que países como o Brasil têm buscado importar de forma acrítica ferramentas similares ao PredPol ou Palantir, o que se trata de um verdadeiro desdobramento do colonialismo digital, fenômeno que se refere ao poder exercido por grandes empresas de tecnologia (ligadas aos países desenvolvidos) sobre a infraestrutura, economia, política e até em relação à cultura e (definitivamente) ao poder punitivo de países considerados menos desenvolvidos.

Destaca-se, mais uma vez, que ao contrário do colonialismo histórico, no qual potências coloniais controlavam territórios e recursos, o colonialismo digital inclui, em

suas características, questões como o racismo algorítmico, a soberania digital (que se refere à dependência de tecnologias estrangeiras que limita a autonomia digital de um país) e a extração massiva de dados de dados pessoais por essas mesmas empresas (Faustino; Lippold, 2023, p. 52-53).

As mencionadas violações, no caso das ferramentas de policiamento preditivo, podem potencializar a já existente seletividade penal a partir da criminalização secundária. Há, portanto, um paradoxo evidente relacionado ao policiamento algorítmico. Como já dito, se por um lado, confiar em previsões baseadas em dados históricos pode ser útil para antecipar eventos e tomar decisões, por outro, essa confiança pode naturalizar desigualdades, uma vez que esses dados provavelmente refletem práticas discriminatórias (como a abordagem policial direcionada a grupos historicamente marginalizados).

Torna-se, assim, possível sustentar que sistemas com a finalidade de auxiliar as forças policiais na mitigação de ações delituosas são inerentemente racistas (Scannel, 2019, p. 108), tendo em vista que são baseadas em práticas policiais passadas e por isso perpetuam as estruturas de violência estatal relacionadas a raça, gênero e classe, independente de quem seja responsável pela alimentação de seus bancos de dados. Conforme destaca Scannel (2019, p. 108), os algoritmos não conseguem simplesmente eliminar a raça do policiamento americano, pois a própria ideia de raça é uma tecnologia policial originária.

Contudo, é relevante que se demonstre que a preocupação com os limites éticos do desenvolvimento da inteligência artificial tem tomado alcance global, com regulamentações (ou tentativas de algo próximo a isso) tanto em países considerados desenvolvidos como aqueles ainda em desenvolvimento.

5.1 A REGULAMENTAÇÃO DA IA PELO MUNDO E SEUS PRINCIPAIS ASPECTOS

Apesar de não ser recente o desenvolvimento das chamadas inteligências artificiais (como já foi destacado no primeiro capítulo), nos últimos anos, esse campo experimentou um crescimento exponencial. É possível conjecturar que esse crescimento foi impulsionado (e se tornou possível) por fatores como a disponibilidade de grandes volumes de dados (*Big Data*), o aumento da capacidade de processamento dos computadores e o desenvolvimento de algoritmos cada vez mais sofisticados.

A título de exemplo, a OpenAI, quando ainda desenvolvia o GPT-3 (*Generative Pre-Trained Transformer 3*) conseguiu bons resultados com a escrita de textos, resposta de perguntas e até criação de histórias verossímeis. Criada em 2015, no ano de 2016, a empresa começou a trabalhar com o sistema Gym em 2016, que ensinava sua IA a tomar decisões com base em estímulos e recompensas (Coutinho, 2023).

Com o intuito de ilustrar esse aumento, a presente tese fez uso de pesquisa desenvolvida pela McKinsey (empresa de consultoria estratégica global), no ano de 2023. Conforme demonstra, tal ano foi responsável pelo aumento vertiginoso no uso de ferramentas de IA, em especial a IA generativa (CHUI, et. al., 2023).

Em linhas gerais, a pesquisa, que foi realizada em abril, revelou que, embora a disponibilidade pública da IA generativa ainda esteja em estágio inicial, muitas pessoas já estão experimentando suas ferramentas em diferentes campos de profissionais e em diferentes áreas do saber. Os entrevistados alegaram grandes expectativas de que essas novas capacidades sejam cruciais em suas áreas de atuação em um futuro próximo.

Conforme menciona Chui (et. al., 2023), a GenAI despertou interesse de profissionais de diferentes regiões, setores e níveis hierárquicos. O uso da GenAI também já extrapolou o ambiente laboral e tem sido comumente utilizada para fins meramente pessoais. Conforme identificou o estudo, cerca de 79% dos entrevistados tiveram algum contato com a IA generativa, sendo que 22% a utilizam regularmente no ambiente de trabalho. Notavelmente, aqueles que trabalham no setor de tecnologia, especialmente na América do Norte, são os que mais intensamente adotaram a GenAI em suas tarefas cotidianas.

É importante ressaltar que o aumento do uso de ferramentas de inteligência artificial também acarreta no aumento dos riscos intrínsecos a elas, tais como: a automação exacerbada em determinadas áreas; o uso malicioso de tais ferramentas em campanhas de desinformação ou em ciberataques; e, certamente, aqueles riscos ligados a direitos fundamentais, como a privacidade e segurança de dados e o risco de viés algorítmico amplamente debatido neste trabalho.

A pesquisa citada aponta que, para além do entusiasmo de profissionais e empresas, poucas são as companhias totalmente preparadas para o uso generalizado de instrumentos de IA (como as ferramentas de GenAI). Conforme sustenta Chui (et al., 2023) Apenas 21% dos entrevistados que relataram a adoção de inteligência artificial afirmam que suas organizações têm diretrizes específicas para orientar o uso

dessas ferramentas pelas suas equipes. Não obstante, quando questionados sobre os riscos envolvendo o uso de inteligência artificial, apenas 32% das empresas possuíam efetivamente diretrizes para mitigá-los.

Apesar de o colonialismo digital impactar de maneira mais drástica os países colonizados, os riscos associados à IA afetam toda a sociedade global, independentemente do grau de desenvolvimento dos Estados nacionais. Sabemos que países menos desenvolvidos frequentemente sofrem com a exploração de seus dados (e até com a dificuldade de uma efetiva inclusão digital), mas até as nações mais ricas enfrentam desafios sérios quanto à regulação desses instrumentos.

A centralização do poder nas grandes empresas tecnológicas, em sua maioria situadas em países desenvolvidos (como no Vale do Silício), cria uma assimetria no controle das ferramentas de IA, o que intensifica preocupações com as questões já apontadas. Assim, países desenvolvidos também estão tomando consciência dos potenciais riscos.

Nesse contexto é importante destacar as já existentes diretrizes internacionais de regulamentação da IA que vem sendo feitas ao redor do globo, que buscam estabelecer padrões mínimos para a produção, uso e compartilhamento de tecnologias, em especial no que diz respeito ao policiamento preditivo. Uma dessas diretrizes, que será analisada no próximo tópico, surge como uma resposta da União Europeia a essa necessidade urgente.

5.1.1 A União Europeia e sua regulamentação sobre IA

A falta de transparência nos algoritmos, o uso inadequado de dados e a capacidade de a IA amplificar desigualdades e vieses estruturais são temas que vêm ganhando destaque em debates públicos, políticos e jurídicos, e foram constantemente citados no presente trabalho.

Não é incomum, portanto, que governos e órgãos reguladores desses países busquem formular políticas que garantam o desenvolvimento de IAs seguro e, na medida do possível, ético, com atenção especial à sua aplicação em setores como justiça, saúde, economia e segurança pública.

A União Europeia, dentro de sua estratégia digital, tem buscado regular a inteligência artificial para criar um ambiente mais favorável ao desenvolvimento e ao uso responsável desse tipo de tecnologia. Como já destacado, a IA possui um grande

potencial para gerar benefícios significativos, como avanços na saúde, melhorias em sistemas de transporte, aumento da eficiência produtiva e uso mais sustentável de energia. Entretanto, os riscos também têm colocado em xeque a própria soberania e controle dos Estados europeus. Em resposta a isso, em abril de 2021, a Comissão Europeia apresentou uma proposta de regulamentação pioneira que estabelece um quadro normativo para a IA. Esse regulamento prevê a classificação dos sistemas de IA com base no grau de risco que representam para os usuários, garantindo que os setores em que são aplicados operem de maneira efetiva (União Europeia, 2023). Conforme estabelece em suas diretrizes, o Parlamento apresenta as seguintes propostas legislativas:

A prioridade do Parlamento Europeu consiste em garantir que os sistemas de IA utilizados na UE sejam seguros, transparentes, rastreáveis, não discriminatórios e respeitadores do ambiente. Os sistemas de IA devem ser supervisionados por pessoas, em vez de serem automatizados, para evitar resultados prejudiciais.

O Parlamento quer também estabelecer uma definição uniforme e neutra em termos tecnológicos para a IA, de modo a ser aplicada em futuros sistemas de IA (União Europeia, 2023).

A legislação busca classificar os riscos das ferramentas que podem ser desenvolvidas. Apesar de muitos instrumentos terem riscos mínimos ou aceitáveis, outros podem ter riscos elevados e até mesmo inaceitáveis.

A União Europeia classifica como “inteligência artificial de risco inaceitável” aqueles sistemas que são considerados uma ameaça real à segurança ou aos demais direitos das pessoas (como dignidade, liberdade e até mesmo vida) e, portanto, estão sujeitos à proibição na maioria dos casos.

Entre eles estão tecnologias que manipulam o comportamento de grupos vulneráveis, como brinquedos ativados por voz que podem incitar crianças a agir de forma perigosa, e sistemas de pontuação social, que avaliam indivíduos com base em fatores como comportamento, status socioeconômico ou outras características pessoais (União Europeia, 2023). Também estão incluídos sistemas de identificação biométrica e de categorização de pessoas, especialmente aqueles que utilizam reconhecimento facial em tempo real.

É importante destacar que apesar deste último tipo de sistema estar relacionado às ferramentas estudadas na presente tese, não significa que o policiamento preditivo está completamente banido nos países europeus, tendo em

vista que algumas exceções são permitidas para fins de “aplicação da lei”, em casos considerados graves e mediante autorização judicial (como nos sistemas de pós-identificação biométrica à distância).

Já os “sistemas de IA de risco elevado” são aqueles que podem impactar negativamente um ou mais direitos fundamentais, mas de forma menos drástica. Eles são divididos em duas categorias: 1) sistemas usados em produtos cobertos pela legislação de segurança da UE, como brinquedos, automóveis e dispositivos médicos; 2) sistemas aplicados em áreas como gestão de infraestruturas, educação, emprego, serviços públicos e privados essenciais, controle migratório e interpretação jurídica (União Europeia, 2023).

Diferente da categoria de risco anterior, esses sistemas (risco elevado) devem ser registrados e monitorados continuamente, desde sua introdução no mercado até o fim de seu ciclo de vida. Não obstante, é necessário que os cidadãos tenham o direito de contestar seu uso perante as autoridades.

Ainda quanto à regulação europeia, é válido mencionar que inteligências artificiais generativas (de modo que o ChatGPT é o exemplo mais evidente) não serão catalogadas junto às ferramentas de alto risco, o que parece ser uma inconsistência, haja vista que as consequências de tais sistemas generativos podem causar danos comparáveis aos das ferramentas mencionadas no risco elevado.

Ainda assim, nesses sistemas deverá haver o cumprimento de uma série de requisitos relacionados à transparência, como a necessidade de divulgação de que um determinado conteúdo foi gerado por IA, bem como conceber o modelo adotado com o intuito de impedir a geração de conteúdo ilícito (União Europeia, 2023).

O Parlamento aprovou a legislação referente à inteligência artificial em março de 2024, com a aprovação subsequente pelo Conselho ocorrendo em maio desse mesmo ano. Conforme destacado, a norma deve entrar em vigor plenamente 24 meses após sua publicação, embora algumas de suas disposições tenham prazos diversos.

Muito embora seja interessante destacar a preocupação europeia, o foco do trabalho consiste no tratamento da IA que é (ou será) dado pelos países mais expostos ao colonialismo digital.

5.1.2 A regulamentação da IA no Brasil

No Brasil, a regulamentação da Inteligência Artificial (IA) tem gerado inúmeros debates devido à sua relevância crescente e aos impactos dessa tecnologia em diversos setores (como no caso das eleições em que se discutem riscos não apenas ligados à criação de notícias falsas pelas ferramentas generativas, mas também criação de imagens e deep fakes⁴¹). Muito embora haja uma preocupação em não obstruir a inovação tecnológica, é preciso que sejam integralmente garantidos os direitos fundamentais e assegurada a transparência e a segurança no uso de tais ferramentas.

No âmbito legislativo, no Senado, a Comissão Temporária sobre Inteligência Artificial no Brasil (CTIA) analisou um relatório do projeto de lei que visa regulamentar o uso da IA no país (PL 2.338/24). O presidente da comissão, senador Carlos Viana, anunciou que o texto seria votado após a concessão de vista aos demais membros da comissão (Agência Senado, 2024).

O relator do projeto, senador Eduardo Gomes, apresentou a versão mais recente do PL 2.338/24, que havia sido proposto originalmente pelo senador Rodrigo Pacheco, que atualmente preside o Senado. O tema tem chamado a atenção dos parlamentares, tendo em vista que após a apresentação do relatório inicial, o texto recebeu aproximadamente 30 emendas, das quais 15 foram aceitas total ou parcialmente (Agência Senado, 2024).

O projeto em questão tem se mostrado o maior avanço legislativo brasileiro sobre o tema. É chamado coloquialmente de “PL da Inteligência Artificial” e possui o escopo de criação de um marco legal para o uso dessas ferramentas no Brasil.

Dentre as preocupações do projeto, estão as técnicas para indução de comportamento prejudiciais à saúde e segurança individual, além do cuidado com sistemas que possam explorar vulnerabilidade de determinados grupos (como crianças). Busca-se também dar ênfase à transparência dos algoritmos, principalmente daqueles utilizados em plataformas de comunicação social (Senado Federal, 2023).

⁴¹ O *deep fake* tem se mostrado uma preocupação real da Justiça Eleitoral, uma vez que possui a capacidade de tornar as eleições ainda mais complexas, tendo em vista que, com a tecnologia, torna-se possível realizar montagens de vídeo e até mesmo trocar o rosto de indivíduos e gerar falas que nunca foram ditas, mas simulando o tom real da pessoa afetada (Mendes, 2024).

Segundo a Agência Senado (2024), o relator também destacou que determinados campos da sociedade civil ainda buscam novas alterações na regulamentação, motivo pelo qual destacou que a normatização (em seu texto final) deve ser genérica, já que caberá às futuras legislações tratarem da temática de forma específica em cada um de seus aspectos (como no caso da discussão de direitos autorais).

O contato com a sociedade civil é (ou deveria ser) um aspecto fundamental do projeto. Nos dias 1º, 2 e 3 de julho foram realizadas audiências com a presença de especialistas para debater o tema. O relator também citou as audiências promovidas pela comissão temporária de juristas, que elaboraram sugestões de projeto para o presidente do Senado (Agência Senado, 2024).

Um ponto do projeto que merece destaque é a “classificação de risco” das ferramentas a serem regulamentadas, algo que se assemelha à normatização europeia, já citada no tópico anterior.

O projeto estabelece diferentes regras regulatórias com base no nível de risco que um sistema de IA representa para os indivíduos. De forma similar ao regramento europeu, o sistema de IA pode ser classificado como de “risco excessivo”, sujeito a uma regulação mais rígida; de “alto risco”; ou não se enquadrar em nenhuma das classificações mencionadas (Senado Federal, 2023). A classificação de risco, conforme consta no projeto, é determinada após uma avaliação inicial realizada pelos desenvolvedores, fornecedores ou especialistas no campo.

Quanto às ferramentas que podem estar inseridas em cada uma dessas categorias, o projeto estabelece que tecnologias como armas autônomas estão inseridas em “risco excessivo” e, portanto, terão sua implementação proibida. Não obstante, em uma das audiências já realizadas, a pesquisadora da organização Coalizão Direitos na Rede, Paula Guedes, defendeu que sistemas de identificação biométrica (tais como os estudados na presente tese) fossem proibidos no âmbito da segurança pública ou Justiça (Agência Senado, 2024).

No que pese a discussão sobre o tema nas audiências públicas citadas, o uso de ferramentas de policiamento preditivo no Brasil permanece com um grupo amplo de apoiadores em cargos políticos.

Importa dizer que, diferente das ferramentas de risco excessivo, as ferramentas de “alto risco” podem ser implementadas, conforme destacado no Projeto (Senado Federal, 2023), desde que haja medidas de controle, operações e testes de

confiabilidade. Defende-se, por exemplo, a realização de avaliação de impacto algorítmico, com medidas preventivas e que sejam capazes de reverter (ou ao menos mitigar) os efeitos nocivos de tais ferramentas.

Dois outros pontos que merecem destaque no projeto apresentado dizem respeito: 1) à Autoridade competente para a fiscalização da legislação (caso haja conversão do projeto em lei efetiva) e 2) à autorregulação das empresas.

Quanto ao primeiro ponto, parece correto afirmar que determinados órgãos farão parte do chamado Sistema Nacional de Regulação e Governança de Inteligência Artificial (SIA). As normas específicas sobre o tema serão expedidas pela Autoridade Nacional de Proteção de Dados, já responsável pela fiscalização sobre tratamentos de dados pessoais no Brasil (Senado Federal, 2023).

No que diz respeito ao segundo, o projeto prevê uma “autorregulação” de modo complementar à fiscalização dos órgãos estatais. Há previsão de incentivos na modalidade de associação voluntária das corporações por meio de certificadores de boas práticas (Agência Senado, 2024).

Por fim, para além dos princípios como transparência e crescimento inclusivo, há uma série de garantias individuais (tal como feito no âmbito da Lei Geral de Proteção de Dados) aos indivíduos afetados por sistemas de inteligência artificial. É possível mencionar os seguintes: 1) direito à informação em relação à interação com sistemas de IA; 2) direito à privacidade e proteção de dados (como já previsto em outras legislações); 3) direito à determinação e participação humana em decisões de determinadas ferramentas de IA; 4) direito a não discriminação e correção de vieses discriminatórios; 3) direito ao uso a uma linguagem simples e clara em ferramentas destinadas a grupos específicos (como crianças e adolescentes).

Sobre o tema, Bioni, Rielli e Zanatta (2024), pesquisadores do DataPrivacy Brasil, resumem os avanços na regulamentação da IA do país, a partir da versão mais recente do projeto mencionado, da seguinte maneira:

SISTEMA DE FISCALIZAÇÃO HÍBRIDO: a nova proposta é uma síntese entre os modelos regulatórios do PL 21/20 e PL 2338/23, que pode ser ilustrada pelo chamado Sistema de Regulação e Governança de Inteligência Artificial. Uma estrutura híbrida que busca valorizar o trabalho de regulação e de fomento já feito pelas agências reguladoras setoriais existentes, mas que devem cooperar entre si e com uma autoridade que seja coordenadora deste sistema. Com isso, a princípio, a proposta brasileira conjuga as abordagens estadunidenses e britânica à europeia, esta última que já era uma clara inspiração do PL 2338/23;

HARMONIA ENTRE REGULAÇÃO & POLÍTICAS PÚBLICAS: há uma nova gramática que busca posicionar a regulação como sendo um componente de fomento para um progresso tecnológico que seja realmente inclusivo. Há a criação de (02) dois capítulos novos, quais sejam, de fomento e sobre poder público que traçam diretrizes programáticas e obrigações ao Estado que vão desde mais investimento em pesquisa até padrões interoperáveis para que não haja dependência tecnológica e, por conseguinte, mais soberania.

VALORIZAÇÃO DO HUMANO: toda a atuação da Data é pautada por uma dupla de vetores: justiça social e redução de assimetrias de poder. O novo texto se mostra preocupado com o que há de material por trás das linhas de código de sistemas de IA. Além de ter um capítulo dedicado à proteção do trabalho, também acena para os impactos ambientais destes por consumirem muitos recursos naturais, o que afeta principalmente o chamado Sul Global. E, principalmente, busca conciliar a proteção de criatividade-inventividade, que são objeto da proteção de direitos autorais, a um uso justo sobre os dados que são essenciais para a modelagem de tais sistemas.

PRESERVAÇÃO DE UMA ESTRUTURA DE DIREITOS E ESCRUTÍNIO PÚBLICO: de forma geral, foi preservada a estrutura de direitos, antes contida no PL 2338/23, em especial o que chamamos, quando do julgamento do caso IBGE no Supremo Tribunal Federal, de devido processo informacional. Em especial, os direitos de não discriminação, de transparência, de explicação e de revisão quando um sistema de IA pode gerar impactos materiais significativos sobre uma pessoa ou grupo. A esse respeito, nota-se, ainda, uma preocupação com o fenômeno da desinformação que tem desestabilizado a nossa e diversas outras democracias ao redor do mundo. Por fim, há a manutenção de mecanismos como avaliação de impacto e uma base de dados públicas sobre sistemas de IA para uma efetiva accountability e escrutínio público sobre quais são os riscos aceitáveis do lançamento desta tecnologia em nosso meio ambiente.

O que merece destaque negativo, contudo, é o fato de que certas ferramentas de alto ou excessivo risco, mesmo que com certas limitações, ainda encontram respaldo.

Conforme destacam Bioni, Rielli e Zanatta (2024), apesar da proibição da elaboração de armas letais, o texto está acompanhado de uma exceção vinculada à legislação de defesa nacional. Não obstante, apesar das audiências públicas a respeito do tema, há legalização de IAs relacionadas ao reconhecimento facial no âmbito da segurança pública e policiamento preditivo, mesmo que com certas condições pré-estabelecidas.

Dessa maneira, embora o Projeto de Lei sobre Inteligência Artificial no Brasil possa vir a ser considerado um passo importante (e até necessário na regulação de certas ferramentas), ainda há pouca concretização em termos de regulamentação efetiva no país, bem como também está ausente uma maior preocupação com fenômenos como o racismo algorítmico nas ferramentas de policiamento.

Para além de projetos de legislação, é válido destacar o Plano Brasileiro de Inteligência Artificial 2024-2028, proposto pelo Governo Federal durante a 5ª Conferência Nacional de Ciência.

O PBIA tem o objetivo de fomentar o desenvolvimento de ferramentas com o intuito de melhoria na qualidade de vida da população brasileira (Brasil, 2024). Entretanto, muito embora tenha, dentre seus objetivos, a busca pela transformação da vida de brasileiros por meio de inovações sustentáveis e inclusivas baseadas em IA, o plano não faz menção direta à proteção de direitos humanos e possui um foco maior em tornar o país uma referência global no desenvolvimento desse tipo de tecnologia.

O foco na soberania nacional pode ser identificado no próprio plano quando são mencionadas as iniciativas de aplicação de desenvolvimento de IA por empresas, haja vista que as áreas de aplicação se limitam a: 1) educação; 2) desenvolvimento social; 3) meio ambiente, clima e sustentabilidade; 4) indústria, comércio e serviços (Brasil, 2024).

Apesar de alguns desses projetos estarem relacionados (de forma direta ou indireta) a direitos humanos como educação, desenvolvimento social e meio ambiente, eles representam menos da metade dos exemplos citados (sete projetos, se somados todos relacionados a esses temas). A maior parte, entretanto, diz respeito ao campo da indústria e comércio (nove projetos no total). Não apenas isso, mas várias dessas iniciativas contam com a participação de grandes corporações de tecnologia, como a OpenAI; NeuralMind e Lenovo.

O presente trabalho não busca criticar o PBIA, que se mostra como uma iniciativa relevante para o desenvolvimento nacional, mas é importante destacar que, em seu âmbito, a proteção contra violações a direitos humanos e a busca por sua efetivação é algo que aparece de modo secundário.

5.1.3 A regulamentação da IA na América Latina

A regulamentação da inteligência artificial, como um fenômeno mundial, também tem se tornado uma prioridade em diversos países da América Latina. Nos últimos anos (em especial em 2024), a preocupação com segurança e ética no âmbito dessa tecnologia levou vários governos à elaboração (ou ao menos a estudarem o desenvolvimento) de legislações específicas para abordar os desafios que a IA

apresenta em diferentes campos. Conforme mencionado no tópico anterior, o Brasil caminha com o Projeto de Lei nº 2.338/2023, que tem como escopo central criar um marco legal abrangente para o uso da IA (muito embora o projeto não esteja livre de críticas, tal como já destacado).

Na Argentina, iniciativas legislativas estão sendo formuladas para alinhar-se às diretrizes da União Europeia. O objetivo, conforme já destacou o atual presidente do país, é criar um ambiente regulatório que permita a inovação tecnológica. Entretanto, parece haver uma tentativa maior de aproximação com as empresas do Vale do Silício, local onde o presidente se reuniu com executivos de empresas como Apple e Google (Viana, 2024).

O Chile também tem buscado avançar com propostas legislativas que classificam os riscos da IA em diferentes categorias, desde riscos inaceitáveis até riscos limitados⁴². O governo do atual presidente, Gabriel Boric, propôs a criação de uma agência dedicada à proteção de dados pessoais, que terá a responsabilidade de monitorar e regular a indústria de IA (Viana, 2024).

Outros países da região, como Colômbia, Costa Rica, México e Uruguai, estão seguindo caminhos semelhantes (em termos de alinhamento com a regulação feita pela União Europeia). Segundo informa Viana (2024) o desenvolvimento dessas tecnologias nesses países busca estabelecer a segurança na sua criação acima das meras oportunidades econômicas.

O Índice Latinoamericano de Inteligência Artificial (ILIA), desenvolvido pela CEPAL e CENIA (Comissão Econômica para América Latina e Caribe e Centro Nacional de Inteligencia Artificial de Chile), avalia a preparação dos países da região em termos de infraestrutura tecnológica, desenvolvimento de talentos e governança. O relatório de 2024, após a avaliação de 19 países da região, destaca que, embora alguns países estejam avançando significativamente, outros ainda enfrentam desafios importantes. A disparidade na preparação tecnológica e na capacidade de governança é mencionada como um dos principais obstáculos para a implementação eficaz da IA na região (Nações Unidas, 2024).

⁴² Segundo o projeto chileno, que também busca se alinhar às diretrizes europeias, podem ser classificadas como IAs de riscos inaceitáveis aqueles que são incompatíveis com o respeito aos direitos fundamentais mais básicos, como a criação de *deepfakes* com conteúdo sexual ou mesmo a identificação biométrica em tempo real (Viana, 2024). Ainda assim, há exceção para a utilização dessas ferramentas para fins de segurança pública, o que não pode ser considerado um avanço na perspectiva de criação de tecnologias não violadoras de direitos.

Segundo o Índice Latino-Americano de Inteligência Artificial, os países foram examinados (quanto ao desenvolvimento de normas sobre IA em áreas como pesquisa, desenvolvimento e governança. Vale destacar que esses aspectos abrangem tanto a infraestrutura tecnológica quanto a formação de profissionais e os marcos regulatórios necessários para integrar a IA nas economias locais (Nações Unidas, 2024).

Conforme o relatório, embora os países mais bem colocados tenham alcançado avanços notáveis, outros ainda encontram dificuldades significativas, como a complexidade em reter profissionais especializados em IA e a criação de estruturas de governança eficazes para promover a inovação.

O ranking é liderado pelo Chile, que obteve 73,07 pontos em uma escala de 100, seguido pelo Brasil (69,30) e pelo Uruguai (64,98). Esses países, juntamente com outros que também se destacam no desenvolvimento de IA, implementaram estratégias nacionais para a incorporação dessas tecnologias em diferentes setores, conforme destacado (Nações Unidas, 2024).

Por outro lado, muitos países da região ainda estão atrasados em comparação aos mencionados. Não obstante, também é percebido que a União Europeia tem sido considerada um verdadeiro espelho no qual as normas de diferentes países latino-americanos têm se baseado.

O que se pode concluir a partir da análise do ILIA é que, muito embora o racismo algorítmico surja como uma preocupação em diversos países latino-americanos, não houve uma proibição completa da utilização de ferramentas de IA no âmbito da segurança pública, no máximo uma busca na limitação de seu uso.

5.2 A MITIGAÇÃO DO RACISMO ALGORÍTMICO E UMA POSSÍVEL RELAÇÃO COM DIREITOS HUMANOS INSTITUINTES

A questão principal do presente capítulo envolve verificar se é possível pensar o desenvolvimento de ferramentas de IA (em especial aquelas ligadas ao campo da segurança pública) sem vieses discriminatórios.

A utilização de tais ferramentas já é algo presente, no que pese a existência de inúmeros estudos que demonstrem erros e uma clara presença de um racismo algorítmico. O uso desses instrumentos fez com que órgãos policiais estrangeiros reduzissem seus custos e simplificou tarefas que antes eram mais custosas e

complexas. Essa realidade chamou a atenção até mesmo de países do Sul Global, como o Brasil, que tem buscado importar tais ferramentas ou desenvolver modelos similares.

O uso dessa tecnologia, como se apontou, tem trazido uma série de problemas relacionados a já existente seletividade penal. Lucena (2020, p. 8) destaca que, quando correntes como a teoria das janelas quebradas⁴³ são aplicadas no policiamento (o que frequentemente ocorre com um processo de etiquetamento, tal como explicado na perspectiva do labelling approach⁴⁴), há um foco em reprimir crimes menores, como o tráfico de pequenas quantidades de drogas, a prostituição e a vandalização de bens públicos.

O autor sustenta que de acordo com os defensores de políticas como as de “Lei e Ordem”, o enfrentamento rigoroso desses delitos menores resultaria na redução de crimes mais graves.

Entretanto, pesquisas indicam que a eficácia dessa teoria é, no máximo, limitada em termos de redução da criminalidade. O que ocorre é que ela tende a intensificar a criminalização de grupos sociais específicos, sendo uma abordagem simplista que tenta oferecer uma solução direta para problemas extremamente complexos (Lucena, 2020, p. 8).

Dessa maneira, é extremamente difícil programar esses sistemas para evitar discriminação, já que eles podem acabar associando características de certos grupos sociais com comportamentos criminosos.

⁴³ A teoria das janelas quebradas, ou *broken windows theory*, surgiu com a publicação de um artigo de James Wilson e George Kelling em 1982 na revista *The Atlantic Monthly*, e defende que sinais de desordem em espaços urbanos, como janelas quebradas ou pichações, estão associados a um aumento na criminalidade, pois sugerem uma falta de controle social. Inspirada na Escola de Chicago, que analisava as subculturas periféricas e a sociologia do crime, essa teoria também rompeu com explicações biológicas do comportamento criminoso e adotou uma perspectiva sociológica, focada na interação e etiquetagem social para entender as causas do crime. Em suas interpretações e aplicações, essa teoria foi associada a diferentes abordagens como “Lei e Ordem,” “tolerância zero,” “realismo de direita” e “neorretribucionismo”, evidenciando sempre uma postura punitivista. Amplamente aplicada em políticas de policiamento como a de “tolerância zero,” a teoria tem sido criticada pela possível estigmatização de áreas e grupos sociais específicos, reforçando preconceitos e reforçando a seletividade penal em regiões já vulneráveis (Schecaira, 2004, p. 331).

⁴⁴ O *labeling approach*, ou teoria do etiquetamento social, é uma perspectiva criminológica que sugere que o comportamento desviante é definido não pelo ato em si, mas pela forma como ele é rotulado pela sociedade. Diferente das teorias anteriores que tentavam identificar as causas do crime, o *labeling approach*, desenvolvido na década de 1960, direciona o foco para os processos de criminalização, ou seja, como determinados comportamentos e indivíduos são socialmente construídos como desviantes. Segundo essa teoria, o comportamento desviante é aquele que as pessoas rotulam como tal, e esse rótulo pode perpetuar a marginalização, levando os indivíduos a adotarem um “papel desviante” em resposta ao estigma social (Becker, 2008, p. 22).

Essa situação se agrava quando os algoritmos não apenas analisam crimes passados ou áreas marginalizadas, mas também utilizam dados de redes sociais ou outras informações geradas pelos indivíduos, ampliando o risco de discriminação. O autor citado reconhece o potencial do Big Data para melhorar a segurança pública, mas enfatiza que seu uso precisa ser fundamentado em princípios filosóficos sólidos, evitando a hipercriminalização de indivíduos que já vivem à margem da sociedade (Lucena, 2020, p. 8).

5.2.1 É possível erradicar vieses ocultos (ou não) em inteligências artificiais?

A todo momento, algoritmos estão tomando decisões que podem mudar vidas completamente, o que particularmente perigoso em sistemas em que é possível negar liberdade condicional, antecipar a prática de delitos ou até mesmo conceder empréstimos.

É preciso, no presente tópico, reforçar a pergunta principal da presente tese: é possível evitar um eterno retorno a Cesare Lombroso a partir da adoção da inteligência artificial como instrumento de auxílio à persecução penal?

Conforme Douglas Heaven, as ferramentas de policiamento preditivo são e continuarão sendo racistas, não importa quais dados sejam utilizados. O pesquisador em Ciências da Computação da Imperial College London menciona que não é nenhum segredo que as ferramentas de policiamento preditivo são racialmente tendenciosas. O autor cita que diversos estudos mostraram que loops de retroalimentação racistas podem surgir se os algoritmos forem treinados com dados policiais, como prisões. No entanto, também sustenta que o treinamento dessas ferramentas preditivas de uma maneira destinada a reduzir o viés tem pouco efeito (Heaven, 2021).

Desenvolvedores de ferramentas de policiamento preditivo ainda ativas afirmaram que começaram a usar relatos de vítimas para obter uma imagem mais precisa das taxas de criminalidade em diferentes bairros. Em teoria, os relatos de vítimas deveriam ser menos tendenciosos, pois não são afetados por preconceitos policiais ou os citados loops de retroalimentação (Heaven, 2021).

No entanto, o autor faz uma breve referência ao trabalho das pesquisadoras Nil-Jana Akpinar e Alexandra Chouldechova, da Universidade Carnegie Mellon, e Maria De-Arteaga, da Universidade do Texas em Austin, que mostram que a visão fornecida pelos relatos de vítimas também é distorcida. A equipe construiu seu próprio

algoritmo preditivo usando o mesmo modelo encontrado em várias ferramentas populares, incluindo o PredPol, o sistema que mais foi utilizado nos EUA. Elas treinaram o modelo com dados de relatos de vítimas de Bogotá, na Colômbia, uma das poucas cidades para as quais há dados independentes de relatórios de crimes disponíveis em nível distrital (Akpinar; De-Arteaga; Choudechova, 2021, p. 1).

Conforme a pesquisa citada, quando as previsões da ferramenta criada foram comparadas com os dados reais de crimes da região estudada, descobriu-se que ela também cometeu erros. Por exemplo, em um distrito onde poucos crimes foram relatados, a ferramenta previu cerca de 20% dos pontos críticos reais, locais com alta taxa de criminalidade. Por outro lado, em um distrito com muitos relatos, a ferramenta previu 20% a mais de pontos críticos do que realmente existiam (Akpinar; De-Arteaga; Choudechova, 2021, p. 1).

Segundo a análise de Heaven, a razão pela qual os algoritmos erram tanto reside exatamente no fato de que o racismo algoritmo decorre de um fenômeno já presente e estruturado na sociedade. Alterar os relatórios dos policiais para os relatórios das vítimas é algo que surtirá pouco efeito. Dessa maneira, o problema com os relatórios de vítimas é que pessoas negras são mais propensas a serem denunciadas por um crime do que pessoas brancas. Não obstante, pessoas brancas mais ricas têm mais probabilidade de denunciar uma pessoa negra mais pobre do que o contrário. Não apenas isso, mas pessoas negras também são mais propensas a denunciar outras pessoas negras. Assim como nos dados de prisões, isso leva bairros negros e pobres a serem classificados como pontos críticos de criminalidade com mais frequência do que verdadeiramente deveriam (Heaven, 2021).

Outros fatores indiretamente relacionados também distorcem o quadro. Para Heaven (2021), a confiança nas forças policiais também pode ser um fator determinante. O que ele pretende dizer é que em comunidades onde a polícia tem um histórico de corrupção ou preconceito racial, as pessoas podem ser menos propensas a relatar crimes. Isso resulta em uma subestimação da criminalidade real naquela área, levando a uma alocação inadequada de recursos policiais. Assim, a ferramenta preditiva pode deixar de identificar áreas que realmente precisam de mais atenção.

O autor menciona, por fim, que a pressão pública para dismantelar ferramentas racistas e as políticas que as acompanham tem sido vista como a única solução. Localidades que têm sido pioneiras na adoção de ferramentas de policiamento preditivo, como Santa Cruz, anunciaram que não farão mais uso de tais sistemas.

Outros locais, como Chicago, contudo, muito embora tenham suspenso o uso do policiamento preditivo, reinvestiram em um banco de dados para o policiamento de gangues, o que, na perspectiva do autor e de advogados na área, apresenta muitos dos mesmos problemas (Heaven, 2021).

É preciso demonstrar, contudo, que há pesquisadores que ainda enxergam um campo para utilização de ferramentas de policiamento preditivo, mas, certamente, não da forma como elas têm sido tradicionalmente utilizadas.

Conforme destacado, Lucena sustenta que é essencial que os sistemas de policiamento preditivo sejam projetados com uma forte barreira contra o viés discriminatório. Defende-se uma inversão ao que é proposto pela teoria das janelas quebradas e as políticas de *Law and Order*. Ou seja, que o *Big Data* seja preferencialmente aplicado no combate e prevenção de macrocrimes (Lucena, 2020, p. 8).

Dessa forma, pode ser desenvolvido um sistema cuja base seja a análise de sonegações fiscais de grande escala, conexões entre pequenos traficantes e figuras influentes, como políticos e empresários. Além disso, a utilização do *Big Data* deve também, nesse contexto, focar em crimes que envolvem violência direta, como homicídios e crimes sexuais. O autor também sugere que essas tecnologias podem trazer grandes avanços na prevenção e punição de crimes de trânsito, especialmente em casos de mortes ou comportamentos reincidentes que ameaçam a segurança pública.

Essa linha de pensamento pode ser considerada um importante passo para se repensar o uso de tais ferramentas, o que nos leva a um dos pontos mais relevantes do trabalho e a um possível “início de resposta” à pergunta que marca o problema da tese.

Como dito, Lucena apresenta uma diferente perspectiva que propõe um redirecionamento das ferramentas de policiamento preditivo, rompendo com a prática tradicional que tem focado na microcriminalidade e perpetuado discriminações raciais. O posicionamento de Douglas Heaven (além dos estudos de Akpinar, De-Arteaga e Choudechova) é fundamental, pois corrobora com essa posição ao demonstrar que, enquanto essas ferramentas forem aplicadas em atividades historicamente marcadas pelo racismo, como o policiamento e a ostensividade contra delitos frequentemente vinculados a populações marginalizadas, elas continuarão reproduzindo padrões racistas.

O algoritmo, ao ser alimentado por dados que refletem preconceitos enraizados, inevitavelmente reforça essas distorções, independentemente se os bancos de dados são alimentados pelas forças policiais ou pelas informações das vítimas. Nesse contexto, o foco deveria se deslocar esse aparato para a macrocriminalidade, buscando alcançar crimes que afetam a sociedade em maior escala, como evasão fiscal, tráfico de influência e grandes redes de corrupção. Esse deslocamento de foco não apenas buscaria evitar a perpetuação de preconceitos relacionados à microcriminalidade, mas, ao mesmo tempo, poderia gerar maior eficiência na persecução de delitos complexos que realmente dependem de aparatos de inteligência em sua apuração.

Além disso, essa mudança de paradigma vai além da simples redistribuição de recursos ou de áreas de atuação das ferramentas preditivas. Nesse aspecto, é fundamental reestruturar o modo como os dados são coletados e processados. Um dos principais pontos levantados por Heaven é que, se a alimentação dos algoritmos continuar sendo realizada por dados tendenciosos, o resultado será invariavelmente preconceituoso. Portanto, é preciso destacar a necessidade de se ter uma diversidade significativa nos grupos responsáveis por essa alimentação, de modo a garantir que diferentes perspectivas sejam representadas. Isso significa, na prática, incluir pessoas de diferentes origens sociais, econômicas e raciais na concepção e na gestão dos dados.

Em momento algum a presente tese apresenta essa visão como uma solução definitiva ao enviesamento das ferramentas de policiamento preditivo, mas a apresenta como um início de uma discussão na qual seja possível repensar um cenário em que as ferramentas de predição podem ser eficazes sem reforçar desigualdades e preconceitos. Para além do âmbito da segurança pública, o estudo sobre o desenvolvimento de IA deve também ser proporcionado a grupos tradicionalmente afetados por violações de direitos, de modo que possam fazer uso da tecnologia para questões relacionadas a seus próprios interesses e a efetivação de suas garantias.

Importa relembrar os estudos de Simone Browne em sua obra: *“Dark Matters: On the surveillance of blackness”*. É possível afirmar que as novas ferramentas de policiamento preditivo configuram uma continuação da vigilância sobre os corpos negros da mesma forma que outras ferramentas no passado, como as leis das lanternas e o *“Book of Negroes”*, mencionadas pela autora (Browne, 2015, p. 43).

Um ponto relevante da obra de Browne diz respeito ao termo “*dark sousveillance*”, que trata da possibilidade de enfrentamento da vigilância por meio de estratégias que podem ser de anti-vigilância ou contra-vigilância⁴⁵.

A menção à obra de Browne é relevante neste momento, pois se as ferramentas de policiamento existente invariavelmente são capazes de violar direitos de populações marginalizadas, o ideal seria a utilização de ferramentas de inteligência artificial em um processo contrário: o de identificação e fiscalização de violações de direitos fundamentais.

Projetos com esse intuito já existem, como no caso de ferramentas de IA que buscam identificar discursos de ódio em plataformas digitais⁴⁶ e outros projetos que buscam fazer uso da tecnologia para a afirmação de direitos humanos e fundamentais (o que será abordado nos tópicos seguintes).

Não apenas Browne, mas o presente tópico também nos remete a Franz Fanon, que tratou do papel da tecnologia o contexto de processos de descolonização, negando a ideia de existência de tecnologias neutras, conforme pode ser visto no texto de Faustino e Lippold (2023, p. 185):

Do uso colonial das ciências médicas para a estigmatização da cultura autóctone até o emprego das técnicas agrícolas (plantation) voltadas à produção e à exportação de commodities, da participação da psiquiatria no desenvolvimento de técnicas de tortura até a introdução do rádio como forma de dismantelar a cultura muçulmana, Fanon problematizou com bastante cuidado a função colonial da tecnologia no contexto da dominação europeia, recusando qualquer pretensão a uma ciência ou tecnologia neutras.

O autor martinicano criticou de maneira clara a ideia de que a tecnologia é livre de intencionalidades, de modo que ela pode ser encarada, em verdade, como uma extensão do poder colonial utilizada no intuito de subjugar e controlar populações colonizadas.

Dessa maneira, Fanon destacou que o processo de libertação do colonialismo não se trata apenas de uma luta contra o domínio físico ou militar, mas também uma

⁴⁵ A autora não utiliza esses termos como sinônimos, tendo em vista que anti-vigilância remete a técnicas e medidas adotadas para evitar ou para resistir a um processo de vigilância (como as técnicas para evitar reconhecimento facial). Contra-vigilância, por outro lado, refere-se a resposta direta à vigilância existente, com o intuito de exposição de práticas violadoras de direitos (Browne, 2015, p. 21).

⁴⁶ Artigos como “*Decoding Digital Hostility: Examining and Addressing Hate Speech on Social Media Platforms*” (Siddiqui; Yuhani; Memon, 2024) e “*Detection of Hate Speech by Employing Support Vector Machine with Word2Vec Model*” (Sevani; Soenandi; Wijaya, 2021) buscam explicar como o aprendizado de máquina pode ser utilizado para a elaboração de ferramentas eficientes para identificação de discurso de ódio em plataformas digitais.

luta pelo domínio do saber e da tecnologia. De modo similar, o autor rejeita a ideia de que a tecnologia possa ser imposta como uma solução externa, sem que os indivíduos envolvidos participem ativamente de sua concepção e construção (Fanon, 1979, p. 164).

Em “Os condenados da terra”, Fanon afirma que uma ponte, por exemplo, não deve ser construída se isso não contribuir para o enriquecimento da consciência e autonomia dos cidadãos que estão diretamente ligados ao projeto. A ponte, no caso, serve como uma metáfora para qualquer tipo de tecnologia ou infraestrutura. Ele defende que as soluções tecnológicas não podem surgir como um *deus ex machina*, mas precisam emergir da coletividade, do esforço conjunto e do engajamento das pessoas (Fanon, 1979, p. 164).

Essa crítica nos interessa, pois ela se opõe ao determinismo tecnológico, que sugere que a tecnologia é um fator externo que molda a sociedade de forma inevitável, sem considerar o papel das pessoas na criação e uso dessa tecnologia. Fanon, ao contrário, advoga por uma visão onde as ferramentas e tecnologias não devem ser apenas introduzidas, mas devem ser moldadas.

Há, portanto, uma comunicação inegável entre Fanon e Browne. Se a IA tem sido utilizada para monitorar e prender populações marginalizadas, como no caso de comunidades negras e periféricas sendo alvo de ferramentas preditivas que reforçam estereótipos racistas, então, conforme o raciocínio de Fanon, essas mesmas populações devem ser protagonistas na redefinição do uso dessas tecnologias.

A tecnologia, dessa maneira, no lugar de continuar a violar direitos, pode ser reconfigurada para garantir sua efetivação. A IA, por exemplo, poderia ser empregada para identificar padrões de discriminação sistêmica, mapear injustiças sociais e promover mecanismos de denúncia e fiscalização de abusos.

Essa perspectiva guarda uma clara vinculação com a ideia de direitos humanos instituintes, conforme demonstrado a seguir.

5.2.2 O desenvolvimento de IA e sua conexão com direitos humanos instituintes

Conforme sustentado pelo autor espanhol David Sanchez Rúbio (2022, p. 17), é impossível negar a importância dada aos direitos humanos da forma como são convencionalmente concebidos. Há uma relevância simbólica, que se presta a legitimar a atuação dos estados civilizados. Não obstante, no campo discursivo, são

tidos como necessários e que não podem ser desvinculados do âmbito democrático, em especial, direitos como liberdade, igualdade e dignidade da pessoa humana.

Sabe-se, contudo, que tais direitos precisam de efetivação e o cumprimento de seus ditames, na prática, enfrenta obstáculos relacionados às mais diversas questões, como limitação de recursos, interesses econômicos e até mesmo a colonialidade contemporânea. Conforme destacado por Sánchez Rubio (2022, p. 18):

[...] mas estamos sobre uma base trágica e receosa de exclusões concretas marcadas pela nacionalidade, pelo racismo, pelo sentido de pertencimento, pela condição de classe, pela defesa do direito à propriedade privada e absoluta, pelo machismo ou pelo conceito de cidadania.

A efetivação de direitos humanos e fundamentais enfrenta, assim, desafios provenientes de diferentes direções. Percebe-se, de maneira clara, uma condição cultural contraditória, tendo em vista que, embora haja discursos e valores que promovam a igualdade e a justiça, também persistem práticas enraizadas que fomentam violações por meio de ações ou inações.

Inúmeros são os exemplos, como a tolerância ao sofrimento de imigrantes e mulheres vítimas de violência, bem como a impunidade para crimes cometidos por indivíduos de colarinho branco. Adicionalmente, grupos étnicos e ativistas que lutam contra violações de direitos frequentemente enfrentam criminalização, mesmo quando buscam algo tão fundamental quanto uma saúde pública universal, trabalho digno ou simplesmente a não discriminação por parte do Estado.

Além das questões já discutidas, Sánchez Rubio (2022, p. 19) também destaca que quando é conveniente ou do interesse de certos atores, a legalidade muitas vezes é usada para justificar injustiças. Isso é particularmente evidente na Europa e nos Estados Unidos em relação à imigração e ao uso da força policial. Quando a legalidade se torna um obstáculo para a aplicação dos direitos humanos, a força dos fatos (mesmo que injusta) é frequentemente aceita como um consolo ou como uma forma de legitimar sacrifícios de vidas humanas ou a negação de direitos que deveriam ser universalmente reconhecidos.

A maneira como frequentemente compreendemos os direitos humanos, especialmente quando os limitamos à dimensão institucional e burocrática (comumente aprendida nas graduações do curso de Direito), acaba por restringir nossa capacidade de agir como agentes ativos na construção da realidade.

Essa visão restrita é, para Sánchez Rubio (2022, p. 21), desempoderadora, pois retira a habilidade de significar e ressignificar o mundo. Assim, entende o autor espanhol que quando os direitos humanos são reduzidos a meras formalidades teóricas e normativas, perde-se a oportunidade de reconhecimento como sujeitos capazes de moldar nossa própria cultura, política, etnia e até mesmo sua própria tecnologia.

O conceito de direitos humanos instituintes, proposto por David Sánchez Rubio (2022, p. 37), é, portanto, fundamental para uma compreensão maior sobre o trabalho, tendo em vista que o autor espanhol nos lembra que os direitos não são apenas normas estáticas, mas também processos em constante construção a partir de tramas sociais. Ao aplicarmos essa ideia no desenvolvimento de tecnologias, é possível questionar como essas ferramentas podem ser projetadas para promover e proteger os direitos humanos, em vez de prejudicá-los.

Em sua origem, os direitos humanos estão intrinsecamente ligados a lutas instituintes, a processos de reivindicação que ocorrem antes mesmo da criação de normas ou leis. No entanto, ao delegarmos o cumprimento desses direitos às instituições, muitas vezes seu caráter histórico de defesa e de contínua construção é esquecido. É nesse contexto que David Sánchez Rubio propõe sua abordagem (2022, p. 37).

No livro “Direitos humanos instituintes”, o autor busca devolver aos direitos humanos essa característica instituinte. Ele argumenta que devemos vinculá-los a uma práxis contínua que atenda às necessidades e às realidades humanas sem exceções, e com o menor grau possível de abstração. O autor é responsável por questionar quando exatamente a participação social foi sequestrada dos processos constituintes e legislativos. Ele defende que a sociedade civil deve recuperar sua capacidade de implementar sistemas de garantias, não para desafiar as normas, mas para complementar a dimensão institucional dos direitos humanos. Essa abordagem ética e comprometida com a dignidade humana busca insuflar a proteção e a construção contínua dos direitos.

Compreende-se, no presente trabalho, que a noção instituinte dos direitos humanos se conecta diretamente à ideia de empoderamento dos grupos marginalizados (e colonizados) por meio da tecnologia (algo sustentado por Fanon e atualmente mencionado por Browne).

Assim como o autor argumenta que os direitos humanos devem ser mais do que meras abstrações legais (Sanchez Rúbio, 2022, p. 38-39), a ideia de que esses grupos devem ter voz no desenvolvimento de ferramentas tecnológicas também transcende a passividade. Quando comunidades marginalizadas têm a capacidade de criticar a violação de direitos por ferramentas tecnológicas e fomentar a criação de soluções tecnológicas, elas podem vislumbrar alternativas que atendam às suas necessidades específicas, sem depender exclusivamente das grandes corporações e sem que tenham seus direitos violados.

5.2.3 Falsas tecnologias garantidoras: inclusões abstratas e exclusões concretas

Conforme tem sido explorado, David Sanchez Rubio propõe uma visão de direitos humanos que se opõe às imposições formais e abstratas dos Estados, centrada na criação e emergência desses direitos a partir das tramas sociais e das demandas das próprias comunidades.

Nesse contexto, a imposição de direitos de maneira vertical, vinda do Estado, revela-se abstratamente inclusiva quando, em verdade, mostra-se excludente, pois parte de uma lógica de poder que historicamente ignora e marginaliza as populações mais vulneráveis. Essa abordagem pode ser claramente ilustrada pelo uso de ferramentas de policiamento preditivo pelos Estados, que, ao invés de proteger direitos constitucionalmente garantidos e efetivamente contribuir para a segurança pública, acabam perpetuando a exclusão e naturalizando desigualdades.

Assim, quando observamos a utilização do policiamento preditivo, é evidente como ele reproduz o padrão ocidental de controle e dominação descrito por Sanchez Rubio. As ferramentas tecnológicas atuais, criada e aplicada sob a égide do capitalismo e da racionalidade científica ocidental, carregam consigo os mesmos elementos de dominação, exploração e conflito. Isso acontece porque essas ferramentas são desenvolvidas e empregadas por sistemas de poder que seguem uma lógica de colonialidade (Sanchez Rúbio, 2022, p. 109). Elas não levam em consideração as realidades e especificidades das populações marginalizadas, mas, ao contrário, reforçam estereótipos e preconceitos que alimentam as estruturas opressivas de dominação racial, econômica e social.

A partir dessa lógica, é possível afirmar que o policiamento preditivo é utilizado como um mecanismo que classifica e categoriza pessoas com base em dados historicamente enviesados. Ele se alinha àquilo que Sanchez Rubio descreve como a divisão do ser e do saber no interior do capitalismo ocidental. A racionalidade científica, ao tratar os indivíduos como meros objetos de controle e vigilância, sem considerar suas realidades sociais e culturais, contribui para a perpetuação de uma ordem excludente (Sanchez Rúbio, 2022, p. 110). As ferramentas tecnológicas, ao invés de serem neutras ou emancipatórias, tornam-se instrumentos que reafirmam as desigualdades preexistentes, particularmente no que se refere à racialização e criminalização de grupos historicamente marginalizados.

Sanchez Rubio destaca que os sistemas de dominação modernos operam de maneira mais intensa nas periferias, onde as opressões são mais severas. Isso se reflete na forma como o policiamento preditivo tem sido comumente aplicado, afetando desproporcionalmente as populações de periferias urbanas, negras e pobres, que são, na prática, os “corpos sacrificiais” de um sistema que os vê como alvo preferencial. O uso dessas tecnologias, em vez de garantir direitos, reforça a vulnerabilidade dessas comunidades.

O ocidente, como argumenta Sanchez Rubio, constrói uma “paisagem fictícia de harmonia” (Sanchez Rúbio, 2022, p. 115) ao sustentar discursos de democracia, segurança e direitos humanos, que, na realidade, estão fundamentados em exclusões concretas. As ferramentas de policiamento preditivo, desenvolvidas em nome da segurança pública, exemplificam essa contradição: ao prometer maior proteção e justiça, elas, de fato, intensificam as práticas de controle e criminalização secundária, especialmente sobre aqueles que já são historicamente marginalizados.

Assim, as figuras dos direitos humanos, conforme impostas de cima para baixo, tornam-se insuficientes para transformar verdadeiramente as estruturas opressivas. Como coloca Sanchez Rubio, esses direitos institucionais, praticados de forma simplificada e formalizada, acabam por alimentar o “Minotauro faminto” que precisa constantemente de novas vítimas das camadas mais vulneráveis.

Esse discurso vazio, conforme relembra o autor espanhol, também está presente nas próprias constituições, sejam elas liberais, sociais-democratas ou socialistas, mas que são estabelecidas de cima para baixo e praticadas institucionalmente de maneira insuficiente e reduzida, de forma que são incapazes de gerar qualquer transformação ou de enfrentar estruturalmente a divisão desigual

racial, sexual, socioeconômica e epistêmica (Sanchez Rúbio, 2022, p. 115). O autor prossegue ao sustentar que:

O resultado desse imaginário construído de acordo com uma lógica de poder que se considera superior e que opera como instrumento de expansão, domínio e controle, inclusive sob o véu de uma aparente universalidade e respeito pela dignidade, é o desenvolvimento de produções humanas que anulam a dimensão plural humana e fazem dele um ser prescindível e sacrificável em todos os sentidos, pertença ou não pertença à cultura ocidental, seja cidadão nacional ou imigrante irregular, por mais que a democracia e os direitos humanos apareçam em seu discursos. Por essas e outras razões, sobre esse padrão, o Ocidente construiu uma sociabilidade assimétrica, de exclusões particulares e desiguais, mas camuflada por inclusões teórico-formais e nominais (Sanchez Rúbio, 2022, p. 116).

As ferramentas de policiamento preditivo (e outras utilizadas no âmbito do Poder Judiciário) se encaixam perfeitamente nessa lógica ilustrada pelo autor, pois são apresentadas como instrumentos de segurança, mas, na realidade, operam sob um véu de neutralidade tecnológica que esconde seu caráter discriminatório. Assim como o poder descrito no parágrafo, essas ferramentas são desenvolvidas e aplicadas de maneira a anular a pluralidade humana e tratar certas populações como prescindíveis e sacrificáveis.

5.2.4 Direitos humanos instituintes para além de garantias impostas “de cima para baixo”

O assunto abordado também guarda uma relação com a noção de democracia. Sanchez Rúbio sustenta que vivemos em “democracias minimalistas e sem democratas”, uma vez que parece ter se consolidado um imaginário de uma legitimidade popular que é incapaz de governar a partir de si mesma, de modo que o faz por meio de profissionais políticos que observam de maneira cega os ditames do mercado (Sanchez Rúbio, 2022, p. 120).

A consequência dessa possível realidade é a de que os cidadãos se tornam servos obedientes, de modo que o exercício da cidadania se limita à votação nos representantes (que poucas vezes realmente representam) e não há efetiva participação da população nas esferas políticas, econômicas e até sociais.

O entendimento do autor é interessante a partir do momento em que verificamos os resultados expostos neste trabalho sobre a regulamentação da IA. Percebe-se um discurso unânime de garantia de direitos, mas os projetos pouco

fazem para se afastar do alcance das *big techs*, também agentes do colonialismo digital (a regulamentação argentina, por exemplo, busca uma maior proximidade com as corporações do Vale do Silício).

Não apenas isso, mas projetos como o brasileiro até chegam a apontar como determinadas ferramentas podem ser perigosas e hostis à população (ou a certas parcelas), mas não impedem que elas sejam construídas e utilizadas (ainda que diante do atendimento a certos requisitos).

Os alertas feitos por diversos pesquisadores durante as audiências públicas parecem ter escasso ou nenhum efeito prático. Isso reforça a visão de Sanchez Rúbio (2022, p. 121), que argumenta que a população é, em grande parte, excluída das decisões que afetam diretamente sua vida. A política e a participação social, em vez de serem formas de governança inclusiva, são muitas vezes tratadas como meras ferramentas de administração ou como técnicas de engenharia social, implementadas por tecnocratas e burocratas sem um verdadeiro envolvimento popular.

Conclui-se, a partir da leitura da obra do autor, que não há uma efetiva cidadania. Os cidadãos, pelo contrário, são meros objetos do consumo, verdadeiras mercadorias. É neste ponto que o autor destaca a importância de se compreender os direitos humanos como algo não impositivo:

Em relação aos direitos humanos, as dimensões normativas, procedimentais, formais, instituídas e pós-violatórias devem se reenquadrar e se reposicionar num marco mais complexo, relacional, sócio-material e sócio-histórico. Isso porque há outros elementos dos direitos humanos iguais ou mais importantes, que giram em torno das práticas e ações, particulares e contextualizadas espacial-temporalmente, que articulam e constroem espaços não de exclusões concretas contra os que lutam, mas de inclusões particulares e cotidianas que dão o sustento e a base fática, real e material para os discursos de direitos humanos baseados em inclusões abstratas.

O que o autor propõe é que a efetivação dos direitos humanos não deve depender apenas das estruturas formais do Estado, mas também das práticas e iniciativas cotidianas dos próprios cidadãos. É fundamental que os indivíduos, em seus contextos particulares, se engajem ativamente na construção e defesa desses direitos, garantindo sua efetividade de forma concreta, e não apenas na eterna espera por uma imposição estatal.

O que a parte atual do trabalho pretende demonstrar é que essa noção de direitos humanos instituintes pode ser perfeitamente ilustrada por desenvolvedores de ferramentas de IA que, no que pese a tentativa de regulação estatal, atuam

proativamente para garantir que suas criações respeitem e promovam os direitos humanos. Esses desenvolvedores têm buscado criar projetos que incorporem princípios básicos de direitos em suas tecnologias de forma independente. Dessa forma, não se espera apenas a ação do Estado, mas a existência de um papel ativo na criação de ferramentas que possa fortalecer os direitos humanos a partir de soluções tecnológicas que beneficiem a sociedade e que, preferencialmente, auxiliem grupos marginalizados.

Nesse aspecto, conforme destaca Sanchez Rúbio (2022, p. 131):

Por isso, é necessário retomar e repensar os direitos humanos mudando o ponto de vista para um paradigma pré-violatório, para que existam não somente quando são violados e, como consequência, quando são acionados os circuitos judiciais e garantias estatais, mas também para que haja infinitos mundos prévios de construção, reconhecimento e efetividade, onde possam convergir as políticas públicas e as relações sociais individuais, comunitárias, privadas e coletivas de consideração e respeito [...]

Assim, é essencial repensar os direitos humanos a partir de um paradigma que anteceda as violações, de modo que eles não existam apenas quando são transgredidos e acionados os mecanismos judiciais e garantias estatais. Nesse contexto, conforme discutido anteriormente, as tecnologias não são neutras, nem é possível falar em um determinismo tecnológico. Torna-se, portanto, possível a criação de novas ferramentas, direcionadas por novas intencionalidades, que possam atuar de forma preventiva. Assim, um dos objetivos centrais deste trabalho é justamente apresentar projetos que busquem desenvolver tecnologias que não apenas respeitem os direitos humanos, mas os fortaleçam e os efetivem desde a sua concepção.

5.3 NOVAS FERRAMENTAS NA BUSCA PELA EFETIVAÇÃO DE DIREITOS

A relação entre o desenvolvimento de ferramentas de inteligência artificial que efetivem e não violem direitos humanos e a noção de direitos humanos instituintes pode ser demonstrada a partir de movimentos que buscam reposicionar a tecnologia como um agente de emancipação e inclusão, e não de exclusão e dominação.

O que o trabalho pretende apontar, no presente momento, é que ferramentas de IA, quando desenvolvidas com uma intencionalidade inclusiva, têm o potencial de se alinhar com a ideia mencionada de direitos humanos que enfatizam práticas que

antecedem a violação dos direitos, ao invés de apenas respondê-las (muitas vezes de forma ineficiente) após o dano ter ocorrido.

Com base nesse pretexto, o tópico atual busca apresentar, de forma breve, alguns projetos que se alinham com a ideia de promoção de direitos. Nesse sentido, os projetos apresentados a seguir não possuem necessariamente o intuito de desenvolvimento de uma ferramenta concreta que vise solucionar questões ligadas a direitos humanos, mas o escopo de realizar uma democratização do conhecimento a respeito de IA para populações de países não desenvolvidos e/ou grupos marginalizados.

5.3.1 Projetos do Sul Global

Conforme demonstrado no decorrer do trabalho, o uso indiscriminado e não (devidamente) regulamentado de determinadas tecnologias (como as ferramentas de policiamento preditivo) pode levar à já estudada discriminação algorítmica, onde os padrões e dados usados para "prever" comportamentos criminais tendem a se basear em históricos de dados enviesados, resultando em ações policiais desproporcionais em comunidades já historicamente estigmatizadas.

No entanto, a sustentação principal do trabalho reside no fato de que as IAs não precisam necessariamente seguir esse caminho de violação de direitos. Elas podem, ao contrário, ser projetadas com o objetivo de fortalecer os direitos humanos, promover inclusão e reduzir as desigualdades existentes. Um exemplo que reforça essa abordagem é o projeto "DecolonizAI", que visa desafiar as estruturas de poder coloniais (em especial o colonialismo digital) que ainda moldam a indústria tecnológica e as relações de poder em todo globo.

Segundo as informações da própria plataforma, seus objetivos são (DecolonizAI, 2024):

- 1) agregar pesquisadores com interesse e produção acadêmica que se relaciona com a diversidade cultural, étnica, identitária em estudos sobre a sociedade e tecnologia;
- 2) reunir referências que colaborem para compreensão e reflexão sobre o tema;
- 3) apresentar exemplos visuais que colaborem para compreensão do tema por um público mais amplo, através de memes, fotos, vídeos, gráficos, etc.;
- 4) tornar-se um ambiente interativo para provas de conceito de projetos relacionados ao tema como IHC inclusiva, IHC feminista, APIs de filtragem de conteúdo semântico a partir de premissas éticas, assim como planos futuros de construir uma base de dados

decentralizada com o controle dos usuários proprietários dos dados postados na plataforma.

Importa destacar que o termo IHC (Interação Humano-Computador) se refere ao estudo e design de interfaces e interações entre usuários e sistemas computacionais. A ideia central dos desenvolvedores é garantir que a tecnologia seja intuitiva e acessível para todas as pessoas, levando em consideração fatores como usabilidade, acessibilidade e a experiência do usuário. No contexto de "IHC inclusiva" e "IHC feminista", a proposta é desenvolver interfaces que levem em conta as necessidades de grupos tradicionalmente excluídos, como mulheres e minorias étnicas (DecolonizAI, 2024).

Por sua vez, API (Interface de Programação de Aplicações) é um conjunto de ferramentas e protocolos que permite que diferentes softwares se comuniquem entre si. Uma API facilita o desenvolvimento de aplicativos ao fornecer componentes pré-configurados que podem ser integrados em sistemas maiores. No caso mencionado, "APIs de filtragem de conteúdo semântico a partir de premissas éticas" se refere a interfaces que permitiriam filtrar e organizar dados com base em padrões éticos, ajudando a eliminar conteúdos que poderiam violar direitos básicos (DecolonizAI, 2024).

O projeto mencionado é fruto de uma iniciativa acadêmica desenvolvida pelo Instituto de Estudos Avançados da Universidade de São Paulo. É inspirado por uma visão crítica sobre as formas tradicionais de desenvolvimento tecnológico e busca aproximar o campo acadêmico-científico da sociedade civil, promovendo, como resultado, a conscientização sobre o uso de ferramentas digitais. Um de seus principais objetivos, conforme consta na página do projeto (DecolonizAI, 2024) é repensar o uso de dados e algoritmos sob uma perspectiva que valorize a diversidade cultural e que seja sensível aos desafios sociais e ambientais, especialmente no contexto brasileiro.

É importante destacar que a iniciativa também se posiciona contra o uso de tecnologias que reforçam desigualdades e negações de problemas como o racismo. Ao invés disso, o projeto busca promover ações afirmativas que utilizem a tecnologia para resolver esses problemas, e não para perpetuá-los (DecolonizAI, 2024).

De forma semelhante, a Estratégia Latino-Americana de Inteligência Artificial é um projeto conjunto que envolve vários pesquisadores de países da América Latina para promover o desenvolvimento e a implementação de tecnologias de inteligência

artificial (IA). O objetivo principal deste projeto é utilizar a IA como uma ferramenta para superar desafios sociais, econômicos e ambientais, enquanto se garante a proteção dos direitos humanos (Ela-IA, 2024). De forma mais específica, o projeto elenca, dentre seus objetivos, o seguinte:

Artigo 4º O objetivo social da ELA-IA é promover o uso ético de tecnologias digitais, por meio do desenvolvimento tecnológico, da pesquisa e do ensino, que visem a proteção de dados pessoais de indivíduos, comunidades e coletivos populacionais, enquanto cria condições para a apropriação, livre e consciente, de dados digitais pelos indivíduos possuintes de tais dados, conforme o direito fundamental à sua proteção prevista no art. 5º, LXXIX, da Constituição da República Federativa do Brasil, especialmente nos meios digitais (Ela-IA, 2024).

Como um projeto inter-regional, envolve a colaboração entre diferentes países da América Latina, o que inclui Chile, Brasil e Uruguai, que são considerados pioneiros na implementação de IA na região (consoante demonstrado anteriormente no presente trabalho).

A estratégia dos pesquisadores é garantir que a IA seja desenvolvida e utilizada de forma a respeitar e promover os direitos humanos. Isso inclui a participação ativa na implementação de políticas e regulamentações que protejam os cidadãos contra abusos e discriminação que possam ser causadas por essas ferramentas.

De forma similar ao DecolonizAI, o projeto latino-americano de IA visa o desenvolvimento de profissionais especializados no campo, o que envolve a criação de programas educacionais e de treinamento para capacitar esses profissionais a trabalharem com essas tecnologias a partir de uma perspectiva em oposição ao colonialismo.

É possível afirmar que a Estratégia Latino-Americana de Inteligência Artificial está diretamente relacionada à efetivação de direitos a partir da noção de direitos humanos instituintes, tendo em vista que busca utilizar a IA como uma ferramenta para melhorar a qualidade de vida das pessoas na região latino-americana a partir da atuação de pesquisadores e profissionais que também buscam promover e democratizar esse conhecimento.

Além disso, a estratégia prega por uma governança completa e transparente para a IA, que inclui o apoio à implementação de políticas e regulamentações que protejam os direitos fundamentais dos cidadãos, motivo pelo qual são responsáveis por uma nota técnica em relação ao mencionado PL 2.338/24 (Ela-IA, 2024)..

A nota traz preocupações quanto ao projeto de lei. Algumas já tratadas no presente trabalho. Os pesquisadores do projeto destacam, por exemplo, a importância que deve ser dada à intervenção e revisão humana nos sistemas de IA, o que deve ocorrer de forma contínua. Não obstante, também militam pela implementação de uma rigorosa avaliação preliminar de riscos antes da utilização de qualquer ferramenta pelo Estado. Como se espera de pesquisadores, a nota também alerta para os sistemas de IA de risco excessivo, demonstrando as vulnerabilidades de certos grupos que podem ser exploradas a partir do uso de determinados instrumentos (Ela-IA, 2024).

Por fim, como exemplo final, o projeto "InteliGente", da Universidade Tecnológica Federal do Paraná (2024), pode ser trazido como mais uma ideia de democratização do conhecimento a respeito da inteligência artificial, pois demonstra como o tema pode ser ensinado para impactar positivamente a sociedade ou mesmo comunidades específicas.

O projeto foi criado em 2001 (UTFPR, 2024) com o objetivo de lidar com uma desigualdade global evidente: o desenvolvimento e a aplicação da IA que estão majoritariamente concentrados em países tecnologicamente mais avançados, deixando o Sul Global, e regiões carentes em particular, à margem dos benefícios proporcionados por essas inovações (além de tecnologicamente dependentes).

A partir dessa premissa, o "InteliGente" tem realizado capacitações de mais de 300 estudantes no campo da inteligência artificial aplicada a áreas diversas, como saúde, educação e igualdade de gênero. Segundo a instituição promotora do projeto, os alunos são, no decorrer da capacitação, direcionados a aplicar esse conhecimento para atender necessidades reais de grupos em necessidade (UTFPR, 2024).

Conforme diretrizes da Universidade Tecnológica Federal do Paraná, o projeto fomenta o pensamento crítico, o trabalho em equipe, e a responsabilidade cidadã. As competências desenvolvidas durante o programa englobam áreas como argumentação, comunicação e planejamento, mas um diferencial importante é a ênfase na empatia e no uso consciente da IA (UTFPR, 2024).

5.3.2 Um novo rumo para a inteligência artificial

Apesar de ter a tese abordado alguns projetos que relacionam a inteligência artificial (ou seu ensino) com a promoção dos direitos humanos (demonstrando como o conceito de direitos humanos instituintes se relaciona com as políticas de

desenvolvimento e adoção de ferramentas de IA), o objetivo deste trabalho não foi esgotar o tema, nem trazer uma (ou mais) ferramentas criadas para “resolver o problema” da violação de direitos humanos, tampouco se buscou realizar uma análise completa de todos os projetos existentes ou tratar de maneira exaustiva aqueles que aqui foram mencionados.

O foco foi, principalmente, destacar como há iniciativas que têm buscado não apenas desenvolver tecnologias, mas também em ensinar a própria população a desenvolver ferramentas voltadas para a solução de problemas enfrentados por esses grupos. Esses projetos são exemplos que ilustram a relação entre os chamados “direitos humanos instituintes”, aqueles que se constroem e consolidam por meio de práticas cotidianas e transformações sociais, e a criação de ferramentas de IA que podem ajudar a corrigir desigualdades e impedir violações.

Este trabalho, portanto, conseguiu responder de forma breve à questão da possibilidade de desenvolver tecnologias que não sejam violadoras de direitos. Mostrou-se que o conhecimento técnico e ético pode ser orientado para uma aplicação correta em diferentes áreas, incluindo a segurança pública.

Nesse contexto, é possível que novas soluções sejam pensadas e criadas localmente, sem a necessidade de importar tecnologias já conhecidas por seus problemas, como o caso do PredPol. A construção de novas ferramentas precisa, contudo, ser feita com um direcionamento certo: a promoção de direitos, para que o avanço tecnológico contribua de fato para a redução de desigualdades, e não para a sua perpetuação.

6 CONCLUSÃO

De modo a relembrar os principais pontos da presente tese, buscou-se, na pesquisa, explicar o racismo existente quando determinados sistemas de inteligência artificial reproduzem ou ampliam preconceitos raciais. Como pode ser notado, esses sistemas, muitas vezes utilizados em áreas diversas como segurança, finanças e saúde, são treinados em grandes volumes de dados. Dessa forma, se esses dados processados contêm vieses (até mesmo históricos), os algoritmos tendem a reproduzir essas mesmas desigualdades.

As ferramentas primordialmente estudadas foram as de policiamento preditivo, que são sistemas que analisam dados pregressos de crimes e comportamentos delituosos para tentar prever onde e quando crimes podem ocorrer no futuro. Essas ferramentas frequentemente são utilizadas sob um discurso de neutralidade tecnológica e foram adotadas não apenas nos países economicamente centrais, mas também por inúmeros países em desenvolvimento, como o Brasil (como um reflexo direto do chamado colonialismo digital).

Não foram poucos os críticos que afirmaram que essas ferramentas amplificam os preconceitos existentes no sistema de justiça (já marcado historicamente por uma seletividade penal), de forma que há um certo temor em relação a um fomento à violência de Estado e hipervigilância contra determinada camada da população.

Vale destacar que o racismo algorítmico no campo da persecução penal não se encontra presente apenas no âmbito da criminalização secundária nas ferramentas de policiamento preditivo, mas também na criminalização terciária, após o ingresso do condenado no sistema penitenciário. Mostrou-se, no decorrer do segundo capítulo, que ferramentas como o COMPAS, usada para determinar a reincidência de detentos, demonstrou uma tendência de superestimar o risco de reincidência para pessoas negras e subestimar para pessoas brancas.

Não se trata, conforme sustentado pela base teórica do trabalho (composta por autores como Franz Fanon e Simone Browne), de uma política recente ou de um problema possível apenas com o avanço das tecnologias digitais, tendo em vista que, como mencionado por Browne, a vigilância sobre as pessoas não brancas é algo histórico e decorre de um colonialismo que, apesar de ter recebido uma nova roupagem, nunca deixou de existir. O que as pesquisas elencadas no trabalho demonstraram é que os algoritmos de sistemas como o PredPol e Palantir parecem

tentar retomar a noção de “criminoso nato” ou “homem delinquente” das premissas lombrosianas do século XIX.

Como questão central, buscou-se compreender se é possível que tais ferramentas, no âmbito da persecução penal, podem ser desenvolvidas e utilizadas de forma a não violar direitos constitucionalmente consagrados.

O debate sobre o uso de ferramentas de inteligência artificial no policiamento é marcado, conforme demonstrado no trabalho, por opiniões nem sempre convergentes. Por exemplo, Will Douglas Heaven defende o desmantelamento dessas tecnologias, pois argumenta que essas ferramentas sempre terão tendência à discriminação, pois os algoritmos são alimentados por dados já historicamente racistas de um sistema de justiça que sempre foi permeado por desigualdades. Dessa forma, inevitavelmente esses sistemas reproduzem as discriminações estruturais presentes na realidade material, reforçando a seletividade penal e até mesmo fomentando ainda mais a vigilância e violência.

Por outro lado, o presente trabalho concorda com Lucena ao defender que as ferramentas de IA na segurança pública não devem ser totalmente descartadas, mas redirecionadas para o combate à macrocriminalidade (entram em cena os crimes contra o sistema financeiro ou contra a Administração Pública). Dessa forma, é possível que o problema do viés possa ser mitigado se a IA for aplicada em áreas onde o viés racial não é um fator determinante. Concorde-se com Heaven e Scannel no sentido de que a utilização do policiamento algorítmico em atividades em que há tradicionalmente a discriminação poderá, inevitavelmente, gerar mais discriminação.

Dessa forma, em vez de focar na microcriminalidade urbana, cenário no qual as práticas discriminatórias se intensificam, a IA poderia ser útil em investigações de delitos de maior escala, onde o foco é menos suscetível a preconceitos individuais e estruturais. Não obstante, os autores também mencionam a possibilidade de seu uso em delitos de trânsito.

Concordando com a visão de Lessig sobre a “regulação do ciberespaço” (mencionada no primeiro capítulo), é possível argumentar que a intencionalidade de uma ferramenta de policiamento preditivo pode ser alterada a partir de modificações no código que a governa. Mencionou-se que o entendimento de Lessig a respeito do fato de que a criação tecnológica transcende as legislações e normas formais, sendo fundamentalmente determinada pelo próprio código e pelos protocolos subjacentes aos sistemas digitais. Esses códigos, que são desenvolvidos pelos “arquitetos” do

sistema, definem, por exemplo, os limites e possibilidades das interações e comportamentos permitidos em uma plataforma, funcionando, de certa forma, como uma "lei" interna ao sistema e, certamente, também estão envolvidos de uma intencionalidade por parte do programador.

Assim, no caso das ferramentas de IA no âmbito da segurança pública, seria possível reconfigurar o código para redirecionar seu foco para delitos da macrocriminalidade e reduzir sua tendência a discriminações estruturais, modificando intencionalmente o escopo e os parâmetros do sistema.

De certa forma, esse entendimento também se comunica com o que é mencionado por Habermas (também citado no primeiro capítulo do trabalho). Sua visão sugere que a tecnologia, enquanto instrumento, não é intrinsecamente atrelada a um objetivo específico. Sua função prática pode ser alterada conforme uma orientação política ou social dada a ela. No entanto, isso não implica uma neutralidade plena das tecnologias, pois a concepção de qualquer artefato ou código já carrega intenções de seus criadores.

No que diz respeito à finalidade de efetivação de direitos em tecnologias digitais, é possível afirmar que as ferramentas de policiamento preditivo (citadas no trabalho), adotadas sob o discurso da segurança, acabam se encaixando na crítica de Sanchez Rúbio ao funcionamento insuficiente dos direitos humanos quando impostos "de cima para baixo". Essas tecnologias, que na maioria das vezes são apresentadas como neutras, ocultam (ou escancaram, dependendo do ponto de vista) seu potencial de reforçar a opressão de grupos já vulneráveis, demonstrando como o poder punitivo funciona como uma espécie de "Minotauro faminto" que depende de vítimas específicas para sua sustentação. Com isso, certas populações são tratadas como substituíveis e estão frequentemente sujeitas a um monitoramento invasivo e seletivo, que opera sob uma suposta objetividade técnica que, na verdade, não existe.

A relação entre o tema principal do trabalho (as IAs no campo do policiamento) e a noção de direitos humanos instituintes está evidente, primeiramente, na demonstração de que o Estado, muitas vezes, permanece inerte ou até mesmo conivente diante de violações de direitos que envolvem interesses de controle social. Em segundo lugar, e talvez mais importante, há a necessidade de uma movimentação popular que reconheça esses padrões e busque efetivar direitos de forma prática. Isso se conecta com uma visão de direitos humanos que transcende a abordagem formalista e burocrática. Dessa forma, não é exagero afirmar que a IA poderia servir

como ferramenta para reforçar os direitos humanos e combater discriminações, ao invés de alimentar estruturas opressivas.

Por fim, é sempre importante ressaltar que o trabalho em questão não se presta a tentar resolver de forma definitiva o problema entre o desenvolvimento de ferramentas algorítmicas e a eventual violação a direitos. Demonstra-se, a partir da pesquisa realizada, que o futuro dos instrumentos ligados à inteligência artificial não está necessariamente ligado endurecimento da criminalização secundária (ou terciária).

É necessário alertar a respeito da possibilidade de um aumento às violações de garantias fundamentais, o que inclui a violência estatal contra grupos específicos e até mesmo a perpetuação de uma naturalização da desigualdade, mas, ao mesmo tempo, também é possível verificar novas possibilidades para além de atividades de investigação de delitos que são normalmente praticados pela população periférica.

Com efeito, o trabalho conclui com o apontamento de duas vertentes para o futuro da IA não violadora de garantias: 1) seu uso no âmbito da persecução penal em relação à macrocriminalidade; e 2) seu uso geral, a partir das próprias populações marginalizadas, na tarefa de fiscalização a violação de direitos e efetivação de demais garantias desses grupos.

REFERÊNCIAS

A GAZETA. **Guarda começa a utilizar inteligência artificial no videomonitoramento de Vitória.** 2024. Disponível em:

<https://www.agazeta.com.br/es/policia/guarda-comeca-a-utilizar-inteligencia-artificial-no-videomonitoramento-de-vitoria-0724>. Acesso em: 02 ago. 2024.

AGÊNCIA SENADO. **Relator apresenta relatório atualizado sobre regulamentação da IA,** 2024. Disponível em:

<https://www12.senado.leg.br/noticias/materias/2024/07/04/relator-apresenta-relatorio-atualizado-sobre-regulamentacao-da-ia>. Acesso em: 14 out. 2024.

AGUIRRE, Katherine; BADRAN, Emile; MUGGAH, Robert. **Future crime:** assessing twenty first century crime prediction. Igarapé Institute, Rio de Janeiro, Strategic Note 33, jul. 2019. Disponível em: https://igarape.org.br/wp-content/uploads/2019/07/2019-07-12-NE_33_Future_Crime.pdf. Acesso em: 4 jan. 2025.

AKPINAR, Nil-Jana; DE-ARTEAGA, Maria; CHOUDECHOVA, Alexandra. The effect of differential victim crime reporting on predictive policing systems. In: **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)**, Virtual Event, Canada, 3-10 Mar. 2021. ACM, New York, NY, USA, 17 pages. Disponível em: <https://dl.acm.org/doi/10.1145/3442188.3445877>. Acesso em: 17 out. 2024.

ALVARENGA, Rúbia Zanotelli de. Direitos humanos e direitos fundamentais: conceito, objetivo e diferença. In: **Revista eletrônica [do] Tribunal Regional do Trabalho da 9ª Região**, Curitiba, v. 8, n. 78, p. 22-31, maio 2019. Disponível em: <https://juslaboris.tst.jus.br/handle/20.500.12178/159631>. Acesso em: 20 mar. 2023.

ANITUA, Gabriel Ignacio. **Histórias dos pensamentos criminológicos.** Trad. Sérgio Lamarão. Rio de Janeiro: Revan, 2008.

ANDERSEN, Lindsey. **Human rights in the age of artificial intelligence,** 2018. Disponível em: <https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf>. Acesso em: 02 mar. 2023.

ARREGUY, Juliana; RIBEIRO, Bruno. **Justiça suspende compra de câmeras com reconhecimento facial em SP.** 2023. Disponível em: <https://www.metropoles.com/sao-paulo/justica-suspende-compra-de-cameras-com-reconhecimento-facial-em-sp>. Acesso em: 01 ago. 2024.

BARATTA, Alessandro. **Criminologia crítica e crítica do Direito Penal:** Introdução à Sociologia do Direito Penal. Rio de Janeiro: Revan, 1999.

BARRETO, Brena Lohane Monteiro; PREUSSLER, Gustavo de Souza. Sistema penal e violência racial: Uma discussão teórica decolonial. In: **R. Dir. Gar. Fund.**, Vitória, v. 24, n. 3 p. 189-219, set./dez. 2023.

BARRETO, Tobias. **Menores e Loucos e fundamento do direito de punir**. Edição do Estado de Sergipe, 1926. Disponível em: <https://www.stf.jus.br/bibliotecadigital/DominioPublico/146962/pdf/146962.pdf>. Acesso em: 18, jan. 2024.

BATISTA, Vera Malaguti. A Escola Crítica e a Criminologia de Juarez Cirino dos Santos. In: ZILIO, Jacson; BOZZA, Fábio (org.). **Estudos Críticos sobre o Sistema Penal**: homenagem ao Professor Juarez Cirino dos Santos. Curitiba: LedZe, 2012.

BEARMAN, Peter; HEDSTRÖM, Peter. **The Oxford Handbook of Analytical Sociology**. Oxford: Oxford University Press, 2009. Disponível em: <https://users.ox.ac.uk/~sfos0060/prophecies.shtml>. Acesso em: 12 jul. 2024.

BECCARIA, Cesare Bonesana, Marchesi di. **Dos delitos e das penas**. Trad. J. Cretella Jr.; Agnes Cretella. 2. ed. São Paulo: Revista dos Tribunais, 1999.
BECKER, Howard. **Outsiders**: estudos de sociologia do desvio. Rio de Janeiro: Jorge Zahar Editor, 2008.

BENTO, Eudson Ferreira. **Qual a expectativa do uso de reconhecimento facial pela Guarda Civil Municipal de Vitória/ES?** um estudo de caso sobre os riscos do racismo e do sexismo algorítmico. Dissertação (Mestrado em Segurança Pública) – Universidade Vila Velha, Vila Velha, 2024.

BEVILÁQUA, Clóvis. **Criminologia e Direito**. Fortaleza: TJCE, 2019.

BIGELOW, Stephen. **Infraestrutura hiperconvergente ou HCI**. 2022. Disponível em: <https://www.computerweekly.com/br/definicoe/Infraestrutura-hiperconvergente-ou-HCI>. Acesso em: 02 ago. 2024.

BIONI, Bruno; Rielli, Mariana; ZANATTA, Rafael. **Regulação de IA no Brasil: onde estávamos, onde estamos e onde podemos estar**, 2024. Disponível em: <https://www.dataprivacybr.org/regulacao-de-ia-no-brasil-onde-estavamos-onde-estamos-e-onde-podemos-estar/>. Acesso em: 15 out. 2024.

BOURDIEU, Pierre. **Meditações Pascalianas**. Rio de Janeiro: Bertrand Brasil, 2001.

BRASIL. **Decreto-Lei nº. 2.848, de 7 de dezembro de 1940**, 1940. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 26 jun. 2024.

BRASIL. **Lei nº. 11.343, de 23 de agosto de 2006**, 2006. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Lei/L11343.htm. Acesso em: 26 jun. 2024.

BRASIL. **Lei nº. 13.709, de 14 de agosto de 2018**, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 27 mar. 2023.

BRASIL, Neil Patel. O que é algoritmo e como ele é utilizado na internet? In: **Olhar digital**. Disponível em: <https://olhardigital.com.br/2022/07/05/internet-e-redes-sociais/o-que-e-algoritmo/>. Acesso em: 27 fev. 2023.

BRASIL. **Plano Brasileiro de Inteligência Artificial (PBIA) 2024-2028**, 2024. Disponível em: <https://www.gov.br/lncc/pt-br/assuntos/noticias/ultimas-noticias-1/plano-brasileiro-de-inteligencia-artificial-pbia-2024-2028>. Acesso em: 21 out. 2024.

BROWNE, Simone. **Dark Matters**: On the surveillance of blackness. Durham: Duke University Press, 2015.

CARDOSO, Willian. **Câmeras em SP**: promotor sugere medidas contra abuso e discriminação. 2024. Disponível em: <https://www.metropoles.com/sao-paulo/cameras-em-sp-promotor-sugere-medidas-contrabuso-e-discriminacao>. Acesso em: 01 ago. 2024.

CARNIEL, Guadalupe. O que é ICQ? Conheça mais sobre o mensageiro. In: **Canaltech**, 2021. Disponível em: <https://canaltech.com.br/apps/icq-o-que-e/>. Acesso em: 14 fev. 2023.

CARVALHO, Salo de. Perspectivas metodológicas na criminologia crítica brasileira: Diretrizes fundacionais e mapeamento de fontes de referência. In: **Revista Brasileira de Sociologia do Direito**, v. 8, n. 2, maio/ago. 2021.

CARVALHO, Thiago Fabres de. O "direito penal do inimigo" e o "direito penal do homo sacer da baixada": exclusão e vitimação no campo penal brasileiro. In: **Revista da Procuradoria Geral do Estado do Espírito Santo**. Vitória, v. 5, n. 5, p. 209 - 258, 1º/2º sem. 2006.

CASTELLS, Manuel. **A sociedade em rede**. Trad. Roneide Venâncio Majer. 6. ed. São Paulo: Paz e Terra, 2002.

CASTILHO, Marcos Alexandre; SILVA, Fabiano; WEINGAERTNER, Daniel. **Algoritmos e estruturas de dados**. Curitiba: Universidade Federal do Paraná, 2020.

CHAVES JUNIOR, Airto; MENDES, Marisa Schmitt Siqueira. A Criminalização Primária e a Norma Penal Brasileira: considerações acerca da sua seletividade. In: **Revista Eletrônica Direito e Política**, [S. l.], v. 3, n. 3, p. 1–21, 2008. P. 1-21. Disponível em: <https://periodicos.univali.br/index.php/rdp/article/view/7289>. Acesso em: 26 jun. 2024.

CHUI, Michael; HALL, Bryce; SINGLA, Alex; SUKHAREVSKY, Alexander; YEE, Lareina. **O estado da inteligência artificial em 2023: o ano do crescimento explosivo da IA Generativa**, 2023. Disponível em: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year/pt-BR>. Acesso em: 12 set. 2024.

CORRÊA, Fabiano Simões. **Um estudo qualitativo sobre as representações utilizadas por professores e alunos para significar o uso da Internet.**

Dissertação – Universidade de São Paulo, 2013. Disponível em: https://www.teses.usp.br/teses/disponiveis/59/59137/tde-08102013-162610/publico/Fabiano_Correa_Mestrado.pdf. Acesso em: 06 fev. 2023.

COUTINHO, Flávio Motta. **A origem do ChatGPT:** conheça a história da OpenAI. 2023. Disponível em: <https://www.tecmundo.com.br/internet/260413-origem-chatgpt-conheca-historia-openai.htm>. Acesso em: 06 set. 2024.

DECLARAÇÃO DOS DIREITOS DO HOMEM E DO CIDADÃO, 1789. Universidade de São Paulo: Biblioteca Virtual de Direitos Humanos, 2015.

DECOLONIZAI. **Sobre o projeto**, 2024. Disponível em: <https://decolonizai.com/sobre-o-projeto/>. Acesso em: 21 out. 2024.

DEL TORO, Juan; LLOYD, Tracey; BUCHANAN, Kim; GOFF, Philip Atiba. The criminogenic and psychological effects of police stops on adolescent black and Latino boys. In: **PNAS**. Abr, 2019. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.1808976116>. Acesso em: 12 jul. 2024.

DICK, Philip K. **Minority Report**. 1. ed. London: Gollancz, 2009.

DUARTE, Evandro Charles Piza. **Criminologia e racismo:** Introdução ao processo de recepção das teorias criminológicas no Brasil. 1998. 415 f. Tese (Doutorado em Direito). Centro de Ciências Jurídicas, Universidade Federal de Santa Catarina, Florianópolis, 1998.

ELA-IA. **Estratégia Latino-Americana de Inteligência Artificial**, 2024. Disponível em: <https://ela-ia.org/quem-somos>. Acesso em: 21 out. 2024.

ESPAÑOL, Marc. Censorship of pro-Palestinian voices on social media soars amid war in Gaza. In: **El País**, 2023. Disponível em: <https://english.elpais.com/international/2023-12-27/censorship-of-pro-palestinian-voices-on-social-media-soars-amid-war-in-gaza.html>. Acesso em: 23 abr. 2024.

ESZTERI, Dániel. Liability for Operation and Damages Caused by Artificial Intelligence – with a Short Outlook to Online Games, 2015. In: **HeinOnline**. Disponível em: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/studia153&div=7&id=&page=>. Acesso em: 02 mar. 2023.

FANON, Frantz. **Os condenados da terra**. 2. ed. Tradução de José Laurênio de Mello. Rio de Janeiro: Civilização Brasileira, 1979.

FANON, Frantz. **Pele negra, máscaras brancas**. Tradução de Renato da Silveira. Salvador: EDUFBA, 2008.

FAUSTINO, Deivison; LIPPOLD, Walter. **Colonialismo digital**: por uma crítica hacker-fanoniana. 1. ed. São Paulo: Boitempo, 2023.

FERRI, Enrico. **La Scuola Criminale Positiva**, 1885. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=24122. Acesso em: 15 jan. 2025.

FLAUZINA, Ana Luiza Pinheiro. **Corpo Negro caído no chão**: o sistema penal e o projeto genocida do Estado brasileiro. Rio de Janeiro: Contraponto, 2008.

FOUCAULT, Michel. **Em defesa da sociedade**: Curso dado no Collège de France (1975-1976). São Paulo: Martins Fontes. 2002.

FOUCAULT, Michel. **Vigiar e Punir**: história da violência nas prisões. 33. ed. Petrópolis: Editora Vozes, 2007.

FORNASIER, Mateus de Oliveira; BECK, Cesar. Cambridge Analytica: Escândalo, Legado e Possíveis Futuros para a Democracia. In: **Revista do Departamento de Ciências Jurídicas e Sociais da Unijuí**, nº. 53, jan.-jun. 2020.

FREITAS, Marcio Coelho de. Entre tecnodeterminismo e interesse público: limites e possibilidades de regulação da internet. In: **Revista de Direito, Estado e Telecomunicações**, Brasília, v. 10, nº. 1, p. 125-146, maio 2018.

G1. **Justiça suspende liminar e Prefeitura de SP retoma pregão para contratar sistema de reconhecimento facial. 2023**. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2023/05/23/justica-suspende-liminar-e-prefeitura-de-sp-retoma-pregao-para-contratar-sistema-de-reconhecimento-facial.ghml>. Acesso em: 01 ago. 2024.

GALLON, Alexia; LIU, Emily. **Racism Repeats Itself: AI Racial Bias in Predictive Policing Algorithms**. The Proceedings of the Bay Honors Research Symposium, 2023. Disponível em: https://thinkyou.bayhonors.org/wp-content/uploads/BHS_Papers/2023_Symposium/AlexiaGallon_RacismRepeatsItself.pdf. Acesso em: 26 jul. 2024.

GAROFALO, Rafaelle. **Criminologia**. Trad. Danielle Maria Gonzaga. Campinas: Péritas Editora, 1997.

GIBSON, William. **Neuromancer**. 5. ed. Tradução de Fábio Fernandes. São Paulo: Aleph, 2016.

GROSNER, Marina Quezado. **A Seletividade do Sistema Penal na Jurisprudência do Superior Tribunal de Justiça**: O Tratamento da Criminalização Secundária por Decisões em Habeas Corpus. São Paulo: IBCCRIM, 2008.

GUSTIN, Miracy Barbosa de Sousa; DIAS, Tereza Fonseca; NICÁCIO, Camila Silva. **(Re)pensando a pesquisa jurídica: teoria e prática**. 5. ed. São Paulo: Almedina, 2020.

HABERMAS, Jürgen. A ciência e a técnica como ideologia. In: BENJAMIN, Walter, HORKHEIMER, Max, ADORNO, Theodor, HABERMAS, Jürgen. **Textos escolhidos**. São Paulo: Brasil Cultural, 1980.

HEAVEN, Will Douglas. Predictive policing algorithms are racist. They need to be dismantled. In: **MIT Technology Review**, jul. 2020. Disponível em: <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/>. Acesso em: 27 jul. 2024.

HEAVEN, Will Douglas. Predictive policing is still racist—whatever data it uses. In: **MIT Technolgy Review**, 2021. Disponível em: <https://www.technologyreview.com/2021/02/05/1017560/predictive-policing-racist-algorithmic-bias-data-crime-predpol/>. Acesso em: 17 out. 2024.

HERRERA FLORES, Joaquín. **Teoria Crítica dos Direitos Humanos: Os Direitos Humanos como produtos culturais**. Rio de Janeiro: Lumen Juris, 2009.

KROHLING, Aloísio; BOLDT, Raphael. Entre cidadãos e inimigos. O discurso midiático e a expansão do direito penal como instrumentos de consolidação da subcidadania. In: **Revista Direitos Humanos Fundamentais**, [S. l.], v. 10, n. 1, 2010.

KUBRICK, Stanley. **2001: Uma Odisseia no Espaço**. Estados Unidos: Metro-Goldwyn-Mayer (MGM), 1968.

LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren; ANGWIN, Julia. How we analyzed the COMPAS recidivism algorithm. In: **PROPÚBLICA**, 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em: 11 abr. 2024.

LESSIG, Lawrence. **The Code Version 2.0**. Cambridge: Basic Books, 2006.

LÉVY, Pierre. **As tecnologias da inteligência: O futuro do pensamento na era da informática**. Trad. COSTA, C. I. Editora 34: São Paulo, 1993.

LOMBROSO, Cesare. **O homem delinquente**. Trad. Maristela Bleggi Tomasini; Oscar Antonio Corbo Garcia. Porto Alegre: Ricardo Lenz, 2001.

LUCENA, Pedro Arthur Capelari de. Policiamento preditivo, discriminação algorítmica e racismo: potencialidades e reflexos no Brasil. In: **VI Simpósio Internacional**

LAVITS, 2019. Disponível em: <https://lavits.org/wp-content/uploads/2019/12/Lucena-2019-LAVITSS.pdf>. Acesso em: 11 mar. 2023.

LUCENA, Pedro Arthur Capelari de. Viés e racismo no policiamento preditivo: Casos estadunidenses e os reflexos de conexão com o Brasil. In: **Revista de Direito e as Novas Tecnologias**. V. 7, abr.-jun. 2020.

MAIA FILHO, Mamede Said; JUNQUILHO, Tainá Aguiar. Projeto Victor: perspectivas de aplicação da inteligência artificial ao direito. **Revista de Direitos e Garantias Fundamentais**, [S. l.], v. 19, n. 3, p. 218–237, 2018. Disponível em: <https://sisbib.emnuvens.com.br/direitosegarantias/article/view/1587>. Acesso em: 18 jan. 2024.

MAGRANI, Eduardo. **A internet das coisas**. Rio de Janeiro: FGV Editora, 2018.

MENDES, Lucas. **TSE proíbe “deep fakes” nas eleições e amplia responsabilidade de big techs**, 2024. Disponível em: <https://www.cnnbrasil.com.br/politica/tse-proibe-deep-fakes-nas-eleicoes-e-amplia-responsabilidade-de-big-techs/>. Acesso em: 14 out. 2024.

MEZZAROBA, Orides; MONTEIRO, Cláudia Servilha. **Manual de Metodologia da Pesquisa no Direito**. 5. ed. São Paulo: Saraiva.

MCLHUAN, Herbert; FIORE, Quentin. **O meio são as mensagens**: inventário de efeitos. Rio de Janeiro: Record. 1969.

MIRC. **About mIRC**, 1995-2023. Disponível em: <https://www.mirc.com/about.html>. Acesso em: 14 fev. 2023.

MOREIRA, Nelson Camatta; MOREIRA JÚNIOR, Ronaldo Felix. Os replicantes de nosso tempo – a violência estatal e a negação da igualdade e dignidade humana a partir da perspectiva da teoria crítica e da distopia na ficção científica. In: **Revista Brasileira de Direito**, Passo Fundo-RS, v. 13, n. 3, p. 277-294, set./dez. 2017.

MOREIRA JÚNIOR, Ronaldo Félix; LOPES, Jaime Reis. O poder dos fluxos de informação: análise sociológica do exercício político pela rede mundial de computadores. In: **Revista da Semana Discente de Sociologia Política do Instituto Universitário de Pesquisas do Rio de Janeiro**, v. 2, p. 33-35, 2019.

MOREIRA JÚNIOR, Ronaldo Félix; NEVES, Otávio Augusto Fernandes. O papel da rede mundial de computadores na ocupação da UEMG em 2017. In: **Revista do Centro Acadêmico Afonso Pena**, v. 24, p. 115-131, 2019.

NAÇÕES UNIDAS. **Declaração Universal dos Direitos Humanos**, 1948. Disponível em: <<https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos>>. Acesso em: 12 mar. 2023.

NAÇÕES UNIDAS. ILIA 2024: **Evaluando la Preparación y el Progreso de la IA en América Latina**, 2024. Disponível em: <https://www.cepal.org/es/notas/ilia-2024-evaluando-la-preparacion-progreso-la-ia-america-latina>. Acesso em: 15 out. 2024.

NAKAGAWA, Liliane. **Pausar desenvolvimento de IA não é suficiente. ‘Desliguem tudo’, alerta um dos fundadores da área**. 2023. Disponível em: <https://blog.kabum.com.br/tecnologia/inteligencia-artificial/ia-desliguem-tudo-alerta-um-fundadores-area/>. Acesso em: 01 out. 2024.

O’NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Trad. Rafael Abraham. 1. ed. Santo André: Rua do Sabão, 2020.

PINHEIRO, Patricia Peck. **Proteção de dados pessoais**: comentários à Lei n. 13.709/18 (LGPD). 2. ed. São Paulo: Saraiva, 2020.

PRANDO, Camila Cardoso de Mello. A Criminologia Crítica no Brasil e os estudos críticos sobre branquidade. In: **Revista Direito e Práxis**, v. 9, n. 1, p. 70–84, jan. 2018.

RAMOS, Debora Oliveira; COSTA, Ernane José Xavier. Encruzilhada tecnológica: Um diálogo entre duas pessoas negras acerca da inteligência artificial e racismo. In: **Jornal da Usp**, 2022. Disponível em: <https://jornal.usp.br/?p=556833>. Acesso em: 08 mar. 2023.

REZENDE, Beatriz Vargas Ramos Gonçalves de. **A ilusão do proibicionismo**: estudo sobre a criminalização secundária do tráfico de drogas no Distrito Federal. 2011. Tese (Doutorado em Direito) - Universidade de Brasília, Faculdade de Direito. Disponível em: http://icts.unb.br/jspui/bitstream/10482/9856/1/2011_BeatrizVargasRamosGon%c3%a7alvesdeRezende.pdf. Acesso em: 26 jun. 2024.

RICHARDSON, Rashida; SCHULTZ, Jason; CRAWFORD, Kate. Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. In: **New York University Law Review**, v. 94, n. 15, maio 2019. Disponível em: https://www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson_et al-FIN.pdf. Acesso em: 27 jul. 2024.

RODRIGUES, Marcelo. Microsoft explica episódio com chatbot racista e diz que Tay deve voltar, 2016. In: **TecMundo**. Disponível em: <https://www.tecmundo.com.br/inteligencia-artificial/102835-microsoft-explica-episodio-chatbot-racista-diz-tay-deve-voltar.htm>. Acesso em: 23 fev. 2023.

RODRIGUES, Raimundo Nina. **As raças humanas e a responsabilidade penal no Brasil** [online]. Rio de Janeiro: Centro Edelstein de Pesquisa Social, 2011.

SALES, Caio César Galdino. **Impactos sociais do reconhecimento facial: Privacidade e vigilância**. Dissertação de mestrado, Iscte - Instituto Universitário de Lisboa, 2021. Disponível em: <http://hdl.handle.net/10071/24838>, acesso em 30 mar. 2024.

SÁNCHEZ RUBIO, David. Ciencia-ficción y derechos humanos. Una aproximación desde la complejidad, las tramas sociales y los condicionales contrafácticos. In: **Revista PRAXIS**, [S.l.], n. 64-65, p. 51-72, dec. 2010. Disponible en: <<http://www.revistas.una.ac.cr/index.php/praxis/article/view/4066>>. Fecha de acceso: 20 mar. 2023.

SÁNCHEZ RUBIO, David. **Direitos humanos instituintes**. Rio de Janeiro: Lumen Juris, 2022.

SÁNCHEZ-ARJONA, Mercedes Llorene. Inteligência artificial, avaliação de risco e direito ao devido processo. In: **Conexões globais: desafios para o direito na "era digital"**. BOLZAN, José Luis; CAMPUZANO, Alfonso de Julios; MAFRA, Lúgia Kunzendorff (Org.). 1. ed. São Paulo: Tirant lo Blanch, 2024.

SANKIN, Aaron; MATTU, Surya. Predictive Policing Software Terrible At Predicting Crimes. Wired, 2023. Disponível em: <https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes>. Acesso em: 12 jul. 2024.

SANTOS, Juarez Cirino dos. **A criminologia radical**. 4. ed. Florianópolis: Tirant Lo Blanch, 2018.

SANTOS, Marcel Koenigkam [et. al.]. Inteligência artificial, aprendizado de máquina, diagnóstico auxiliado por computador e radiômica: avanços da imagem rumo à medicina de precisão. In: **Radiol Bras**, n. 52, nov-dez, 2019. Disponível em: <https://www.scielo.br/j/rb/a/9yX6w83KDDT33m6G9ddCqBn/?lang=pt>. Acesso em: 17 out. 2022.

SÃO PAULO. **Prefeito assina contrato para o início do Smart Sampa, maior programa de videomonitoramento da cidade com até 40 mil câmeras**, 2023. Disponível em: <https://www.capital.sp.gov.br/w/noticia/prefeito-assina-contrato-para-o-inicio-do-smart-sampa-maior-programa-de-videomonitoramento-da-cidade-com-ate-40-mil-cameras-2>. Acesso em: 31 jul. 2024.

SEVANI, Nina; SOENANDI, Iwan; WIJAYA, Jeremy. Detection of Hate Speech by Employing Support Vector Machine with Word2Vec Model. In: **7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)**,

Malang, Indonesia, 2021. Disponível em:
<https://ieeexplore.ieee.org/document/9616721>. Acesso em: 18 out. 2024.

SENADO FEDERAL. **Projeto de Lei nº 2.338, de 2023**, 2023. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>. Acesso em: 14 out. 2024.

SCANNEL, Joshua. This is not Minority Report. In: BENJAMIN, Ruha (Editor). **Captivating Technology: Race, Carceral Technoscience and liberatory imagination in everyday life**. Durham: Duke University Press, 2019.

SCHECAIRA, Sérgio Salomão. **Criminologia**. São Paulo: Revista dos Tribunais, 2004.

SICHMAN, Jaime Simão. Inteligência Artificial e sociedade: avanços e riscos. In: **Rev. Estudos Avançados**, v. 35, n. 101, p. 37–50, jan. 2021.

SIDDIQUI, Jawaid; YUHANIZ, Siti Sophiayati; ALI, Zulfiqar. Decoding Digital Hostility: Examining and Addressing Hate Speech on Social Media Platforms. In: **1st Karachi Section Humanitarian Technology Conference (KHI-HTC)**, Tandojam, Pakistan, 2024. Disponível em: <https://ieeexplore.ieee.org/abstract/document/10481686>. Acesso em: 18 out. 2024.

SILVA, Mozart Linhares da; ARAÚJO, Willian Fernandes. Biopolítica, racismo estrutural-algorítmico e subjetividade. In: **Educação Unisinos**, v. 24, 2020. Disponível em:
<http://revistas.unisinos.br/index.php/educacao/article/view/edu.2020.241.40>. Acesso em: 26 jul. 2023.

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

SILVA, Tarcízio. **“Tranças feias” e “tranças bonitas”**: como Google intensifica racismo, 2019. Disponível em: <https://tarciziosilva.com.br/blog/trancas-feias-e-trancas-bonitas-como-google-intensifica-racismo/>. Acesso em: 18 out. 2024.

SILVEIRA; Sérgio Amadeu da; SOUZA, Joyce; CASSINO, João Francisco. **Colonialismo de dados: como opera a trincheira algorítmica na guerra neoliberal**. São Paulo: Autonomia Literária, 2021.

SULOCKI, Victoria de. Novas tecnologias, velhas discriminações: ou da falta de reflexão sobre o sistema de algoritmos na justiça criminal. In: **Inteligência artificial e direito: ética, regulação e responsabilidade**. FRAZÃO, Ana; MULHOLLAND, Caitlin [Coord.]. 2. ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

SCHWARCZ, Lilia Moritz. **O espetáculo das raças**: cientistas, instituições e questão racial no Brasil, 1870-1930. São Paulo: Companhia das Letras, 1993.

TELLES, Primonata Silva Brilhante. Inteligência artificial e polícia preditiva: limites e possibilidades. In: **Boletim Científico ESMPU**, Brasília, a. 20, n. 57, jul./dez. 2021.

THE MARKUP. **2018 Announcement**, 2018. Disponível em: <https://themarkup.org/press-release-20180923>. Acesso em: 24 out. 2024.

TOLKIEN, John Ronald Reuel. **O senhor dos anéis**: segunda parte: as duas torres. Trad. Lenita Maria Rímoli Esteves; Almiro Pisetta; BORGES, Luís Carlos (Coord.). 2.ed. São Paulo: Martins Fontes, 2002.

TOMAZELLA, Marlon. Como a ciência e a técnica se tornam ideologia: Uma reflexão a partir de Jürgen Habermas. In: **Educação Pública**, 2011. Disponível em: <https://educacaopublica.cecierj.edu.br/artigos/11/34/como-a-ciencia-e-a-tecnica-se-tornam-ideologia-uma-reflexao-a-partir-de-juergen-habermas>. Acesso em: 07 abr. 2023.

UNIÃO EUROPEIA. **Lei da EU sobre IA**: primeira regulamentação de inteligência artificial, 2023. Disponível em: <https://www.europarl.europa.eu/topics/pt/article/20230601STO93804/lei-da-ue-sobre-ia-primeira-regulamentacao-de-inteligencia-artificial>. Acesso em: 12 set. 2024.

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ. **InteliGente**: Construindo Caminhos de Igualdade com Inteligência Artificial, 2024. Disponível em: <https://www.utfpr.edu.br/noticias/cornelio-procopio/inteligente-construindo-caminhos-de-igualdade-com-inteligencia-artificial?form=MG0AV3>. Acesso em: 21 out. 2024.

VENTURINI, Jamila; CAPPI, Juliano. Encontro 1: Declaração de Independência do Ciberespaço. In: **Ciclo de Leituras Comentadas**, 2018. Disponível em: <https://www.nic.br/publicacao/uma-declaracao-de-independencia-do-ciberespaco/>. Acesso em: 23 jan. 2023.

VERASZTO, Estéfano Vizconde [et. al.]. Tecnologia: Buscando uma definição para o conceito. In: **Prisma.com**, n. 7, 2008. Disponível em: https://www.researchgate.net/publication/266374098_Tecnologia_Buscando_uma_definicao_para_o_conceito_Technology_Looking_for_a_definition_for_the_concept. Acesso em: 25 jan. 2023.

VIANA, Natalia. **Como está a regulação de inteligência artificial na América Latina?** 2024. Disponível em: <https://apublica.org/2024/08/como-esta-a-regulacao-de-inteligencia-artificial-na-america-latina/>. Acesso em: 15 out. 2024.

VIEIRA, Ronaldo da Mota. World Wide Web: Terra encantada onde tudo se encontra? In: **Educação, Gestão e Sociedade**: Revista da Faculdade Eça de Queirós. Ano 4, n. 16, nov. 2014. Disponível em:

http://uniesp.edu.br/sites/_biblioteca/revistas/20170509161129.pdf. Acesso em: 06, fev. 2023.

WEISS; Marcos Cesar; BERNARDES, Roberto Carlos; CONSONI, Flávia Luciane. Cidades inteligentes: casos e perspectivas para as cidades brasileiras. In: **Revista Tecnológica da Fatec Americana**, vol. 05, n. 01, 2017. Disponível em: https://ric.cps.sp.gov.br/bitstream/123456789/516/1/cidades_inteligentes.pdf. Acesso em: 30 jul. 2024.

WINNER, Langdon. Artefatos têm política? Trad. Ferreira, Debora Pazetto; ABRAHÃO, Luiz Henrique de Lacerda. In: **Analytica**, Rio de Janeiro, v. 21, n. 2, 2017, p. 195-218. Disponível em: <https://revistas.ufrj.br/index.php/analytica/article/viewFile/22470/12527>. Acesso em: 30 jan. 2023.

YANG, Fei. Predictive Policing. In: **Oxford Research Encyclopedia, Criminology and Criminal Justice**. Oxford University Press, 2019. Disponível em: https://www.researchgate.net/publication/342216493_Predictive_Policing_Oxford_Research_Encyclopedia_Criminology_and_Criminal_Justice_Oxford_University_Press. Acesso em: 24 jul. 2024.

ZAFFARONI, Eugenio Raúl. **Criminología**: aproximación desde um margen. Bogotá, Colômbia: Temis, 1993.

ZAFFARONI, Eugenio Raúl. **O inimigo no direito penal**. 3. ed. Trad. Sérgio Lamarão. Rio de Janeiro: Revan, 2011.

ZENHA, Luciana. Redes sociais online: o que são as redes sociais e como se organizam? In: **Caderno de Educação**, n. 49, v. 1, 2018, p. 19-42. Disponível em: <https://revista.uemg.br/index.php/cadernodeeducacao>. Acesso em: 14 jan. 2023.

ZULMAR, Fachin. Desafios da regulação do ciberespaço e a proteção dos direitos da personalidade. In: **Revista Jurídica (FURB)**, [S. l.], v. 25, n. 56, jan./abr. 2021. Disponível em: <https://bu.furb.br/ojs/index.php/juridica/article/view/10081>. Acesso em: 23 jan. 2023.