



**FACULDADE BAIANA DE DIREITO**  
**CURSO DE GRADUAÇÃO EM DIREITO**

**CAROLINE MARIA VELANES VILELA**

**DA TRANSPARÊNCIA DOS ALGORITMOS UTILIZADOS EM  
SISTEMAS AUTOMATIZADOS DE APOIO À DECISÃO: UMA  
ANÁLISE SOB A PERSPECTIVA DA SALVAGUARDA DOS  
DIREITOS FUNDAMENTAIS DE GRUPOS  
HISTORICAMENTE SUBALTERNIZADOS**

**Salvador**  
**2025**

**CAROLINE MARIA VELANES VILELA**

**DA TRANSPARÊNCIA DOS ALGORITMOS UTILIZADOS EM  
SISTEMAS AUTOMATIZADOS DE APOIO À DECISÃO: UMA  
ANÁLISE SOB A PERSPECTIVA DA SALVAGUARDA DOS  
DIREITOS FUNDAMENTAIS DE GRUPOS  
HISTORICAMENTE SUBALTERNIZADOS**

Monografia apresentada ao curso de graduação em  
Direito, Faculdade Baiana de Direito, como requisito  
parcial para obtenção do grau de bacharel em  
Direito.

Orientador: Prof. Gabriel Dias Marques da Cruz

Salvador  
2025

## **TERMO DE APROVAÇÃO**

**CAROLINE MARIA VELANES VILELA**

### **DA TRANSPARÊNCIA DOS ALGORITMOS UTILIZADOS EM SISTEMAS AUTOMATIZADOS DE APOIO À DECISÃO: UMA ANÁLISE SOB A PERSPECTIVA DA SALVAGUARDA DOS DIREITOS FUNDAMENTAIS DE GRUPOS HISTORICAMENTE SUBALTERNIZADOS**

Monografia aprovada como requisito parcial para obtenção do grau de bacharel em Direito,  
Faculdade Baiana de Direito, pela seguinte banca examinadora:

Nome: \_\_\_\_\_

Titulação e instituição: \_\_\_\_\_

Nome: \_\_\_\_\_

Titulação e instituição: \_\_\_\_\_

Nome: \_\_\_\_\_

Titulação e instituição: \_\_\_\_\_

Salvador, \_\_\_\_ / \_\_\_\_ / 2025.

Aos meus avós, Helena Velanes Faria e Antenor Fontes de Faria, cuja doçura, amor e compaixão moldaram meu coração e meus valores — existo como sou porque vocês existiram, e seguirão existindo em tudo o que faço.

## AGRADECIMENTOS

Agradeço, antes de tudo, ao mistério divino — a Deus — por me sustentar nos silêncios, nos cansaços e nas retomadas.

Ao meu orientador, Gabriel Marques, por semear cuidado e inspiração ao longo da minha trajetória acadêmica.

À minha mãe, Maria Aparecida, e ao meu pai, Renê, que, com amor e esforço, me ofereceram a dádiva de uma educação digna em um país onde ela ainda é privilégio.

À minha irmã, Isabele, por me inspirar a me lançar no mundo e a ser, sem reservas, quem eu sou.

Aos meus avós, Helena e Antenor, cuja incansável dedicação ao trabalho tornou possível que eu estivesse aqui hoje.

Às minhas tias, Maria Conceição e Maria das Graças, pelo apoio generoso nos momentos difíceis.

Aos meus primos, João Victor e Maria de Fátima Velanes, por estarem sempre ao meu lado, com carinho e apoio incondicional.

A João Lucas Lima, pelo amor generoso, pela companhia constante e pela paciência em traduzir um mundo de códigos para alguém das humanidades.

A todos os meus amigos e amigas — em especial Alice Kimie, Ana Dulce Leal, Angela Cordeiro, Beatriz Macário, Beatriz Paim, Carolina Luckesi, Clara Passos, Danilo Magalhães, Glória Moreno, Juliana Melo, Júlia Caribé, Júlia Sombra, Liliana Py, Lucas Bonafini, Mariana Parcero, Matheus Souto e Vinícius Quadros — por escutarem meus lamentos e celebrarem comigo cada conquista.

Aos professores Marlo Vieira Souza e Gustavo Prazeres, pelas relevantes contribuições intelectuais ao meu percurso de pesquisa.

Aos meus gatos, Lili, Tom e Nicolau, que dormiam ao meu lado enquanto eu escrevia, lembrando-me de que o afeto também é método.

*“Quem tem consciência para ter coragem  
Quem tem a força de saber que existe  
E no centro da própria engrenagem  
Inventa a contra mola que resiste”.*

Secos & Molhados

## RESUMO

A presente monografia analisa criticamente os impactos jurídicos e sociais da opacidade algorítmica em Sistemas Automatizados de Apoio à Decisão, com foco nos efeitos produzidos sobre os direitos fundamentais de grupos historicamente marginalizados. A pesquisa parte da constatação de que a crescente utilização desses sistemas em processos decisórios públicos e privados vem configurando as estruturas de poder contemporâneas ao deslocar a centralidade da deliberação humana para mecanismos automatizados e, muitas vezes, opacos. O estudo destaca os riscos de reprodução de discursos discriminatórios por meio desses mecanismos tecnológicos, uma vez que seu desenvolvimento é envolto de decisões às quais se associam à artefatos políticos, com potencial de reproduzir e ampliar desigualdades estruturais. Para tanto, a investigação adota uma metodologia qualitativa e abordagem hipotético-dedutiva, fundamentada em revisão bibliográfica interdisciplinar e análise crítica de documentos jurídicos, relatórios técnicos e evidências empíricas. O objetivo é demonstrar que a transparência algorítmica não deve ser tratada apenas como requisito técnico, mas como um imperativo jurídico, indispensável à proteção dos direitos constitucionais à igualdade e à não discriminação. Para isso, utiliza-se a interseccionalidade como ferramenta analítica central, evidenciando como raça, gênero, classe e territorialidade se combinam na produção de injustiças tecnicamente mediadas. Os resultados obtidos revelam a existência de um déficit estrutural de auditabilidade e inteligibilidade desses sistemas, o que compromete a possibilidade de escrutínio democrático e a responsabilização institucional por decisões automatizadas com efeitos discriminatórios. A ausência de marcos regulatórios específicos e a fragilidade dos dispositivos legais existentes, como a Lei Geral de Proteção de Dados Pessoais, revelam a necessidade de medidas mais robustas, que incorporem o princípio da explicabilidade, a revisão humana qualificada e o controle social sobre sistemas decisórios automatizados. Conclui-se pela essencialidade da construção de uma agenda jurídica orientada pela transparência, pela justiça distributiva e pela dignidade humana para salvaguardar os direitos fundamentais frente à automatização opaca.

**Palavras-chave:** Transparência Algorítmica. Direitos Fundamentais. Discriminação Algorítmica. Inteligência Artificial. Sistemas de Apoio à Decisão. Interseccionalidade.

## ABSTRACT

This monograph critically analyzes the legal and social impacts of algorithmic opacity in Automated Decision Support Systems, with a focus on the effects produced on the fundamental rights of historically marginalized groups. The research begins with the observation that the growing use of such systems in both public and private decision-making processes is shaping contemporary power structures by shifting the centrality of human deliberation toward automated and often opaque mechanisms. The study highlights the risks of reproducing discriminatory discourses through these technological tools, as their development is enmeshed in decisions tied to political artifacts, with the potential to reproduce and amplify structural inequalities. To achieve this objective, the investigation adopts a qualitative methodology and a hypothetical-deductive approach, grounded in an interdisciplinary literature review and critical analysis of legal documents, technical reports, and empirical evidence. The objective is to demonstrate that algorithmic transparency should not be treated merely as a technical requirement, but as a legal imperative, essential to safeguarding constitutional rights to equality and non-discrimination. Intersectionality is used as a central analytical tool to highlight how race, gender, class, and territoriality intersect in the production of technologically mediated injustices. The findings reveal a structural deficit in the auditability and intelligibility of these systems, which undermines the possibility of democratic scrutiny and institutional accountability for automated decisions with discriminatory effects. The absence of specific regulatory frameworks and the fragility of existing legal mechanisms, such as Brazilian General Data Protection Law, underscore the need for more robust measures that incorporate the principle of explainability, qualified human review, and social oversight over automated decision-making systems. The conclusion emphasizes the essential need to construct a legal agenda guided by transparency, distributive justice, and human dignity, in order to safeguard fundamental rights in the face of opaque automation.

**Keywords:** Algorithmic Transparency. Fundamental Rights. Algorithmic Discrimination. Artificial Intelligence. Decision Support Systems. Intersectionality

## LISTA DE ABREVIATURAS E SIGLAS

§	Parágrafo
AIA	Avaliação de Impacto Algorítmico
ANPD	Autoridade Nacional de Proteção de Dados
CADH	Convenção Americana sobre Direitos Humanos
CESeC	Centro de Estudos de Segurança e Cidadania
COMPAS	Correctional Offender Management Profiling for Alternative Sanctions
COR	Centro de Operações Rio
FBSP	Fórum Brasileiro de Segurança Pública
GDPR	General Data Protection Regulation
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
IPEA	Instituto de Pesquisa Econômica Aplicada
IP.rec	Instituto de Pesquisa em Direito e Tecnologia do Recife
LGBTQIA+	Lésbicas, Gays, Bissexuais, Transexuais, Queer, Intersexo, Assexuais, e outras
LGPD	Lei Geral de Proteção de Dados Pessoais
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
OIT	Organização Internacional do Trabalho
ONU	Organização das Nações Unidas
PL	Projeto de Lei
SRFs	Sistemas de Reconhecimento Facial
SSD	Sistemas de Suporte à Decisão
TICs	Tecnologias da Informação e Comunicação
XAI	Explainable AI

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	<b>12</b>
<b>2 A COMPLEXIDADE DOS MECANISMOS DECISÓRIOS INTELIGENTES DIANTE DA (IMPERATIVA) TRANSPARÊNCIA ALGORÍTMICA</b>	<b>16</b>
2.1 ALGORITMOS E TRANSPARÊNCIA ALGORÍTMICA NA ERA DA INTELIGÊNCIA ARTIFICIAL	17
2.2 UTILIZAÇÃO E IMPACTOS DOS SISTEMAS AUTOMATIZADOS DE APOIO À DECISÃO	27
2.3 PERIGOS DERIVADOS DA OPACIDADE DISCRIMINATÓRIA DOS ALGORITMOS NA MODERNIDADE	33
<b>3 SOBRE A VULNERABILIDADE DE GRUPOS HISTORICAMENTE MARGINALIZADOS NA UTILIZAÇÃO DE ALGORITMOS</b>	<b>39</b>
3.1 DEFINIÇÃO E PECULIARIDADES RELATIVAS AOS SEGMENTOS MINORITÁRIOS	40
3.2 A INTERSECCIONALIDADE APLICADA A GRUPOS VULNERABILIZADOS COMO ELEMENTO DE AGRAVAMENTO OPRESSIVO	45
3.3 MANIFESTAÇÕES DA DISCRIMINAÇÃO ALGORÍTMICA NA CONJUNTURA DAS MINORIAS SOCIAIS	49
<b>4 A IMPERIOSIDADE DE ASSEGURAR OS DIREITOS FUNDAMENTAIS DAS MINORIAS NA ERA ALGORÍTMICA</b>	<b>58</b>
4.1 A CORRELAÇÃO ENTRE TRANSPARÊNCIA ALGORÍTMICA E A TUTELA DOS DIREITOS FUNDAMENTAIS DAS MINORIAS	58
4.2 O PAPEL DO CONTROLE PREVENTIVO-REPRESSIVO NA DEFESA DOS GRUPOS MINORITÁRIOS	66
<b>5 CONCLUSÃO</b>	<b>74</b>
<b>REFERÊNCIAS</b>	<b>77</b>

## 1 INTRODUÇÃO

A crescente incorporação de Sistemas Automatizados de Apoio à Decisão em diferentes esferas organizacionais, tanto no setor público quanto no privado, tem promovido transformações profundas nas arquiteturas decisórias contemporâneas. Fundamentados em análises massivas de dados e técnicas de aprendizado de máquina, esses sistemas operam inferências automatizadas que progressivamente deslocam a centralidade da deliberação humana, atribuindo à técnica computacional um papel decisivo na mediação de escolhas com impactos concretos sobre a vida social e os direitos fundamentais dos indivíduos aos quais estão submetidos esses sistemas.

Não obstante a crescente prevalência de discursos que conferem legitimidade aos Sistemas Automatizados de Apoio à Decisão ancorados em atributos como eficiência procedimental, acurácia computacional e uma presumida neutralidade algorítmica, torna-se imprescindível submetê-los a uma crítica epistemológica e sociopolítica mais rigorosa. Tais sistemas, comumente apresentados como instrumentos tecnicamente objetivos, não operam de maneira dissociada das estruturas sociais, políticas e históricas em que são concebidos. Pelo contrário, são produzidos e implementados em contextos profundamente marcados por relações assimétricas de poder, o que os torna suscetíveis à incorporação de racionalidades normativas e interesses específicos dos grupos socialmente hegemônicos.

Nesse contexto, desponta como questão central o fenômeno da opacidade algorítmica, compreendido como a ausência de transparência, rastreabilidade e inteligibilidade dos processos decisórios automatizados, a qual compromete gravemente a possibilidade de controle democrático sobre os algoritmos, dificultando a identificação dos critérios normativos, estatísticos e técnicos que orientam as inferências computacionais.

Em sociedades atravessadas por desigualdades estruturais, essa ausência de inteligibilidade transforma-se em vetor de injustiças silenciosas, particularmente em indivíduos atravessados por vulnerabilidades históricas e estruturais, como aqueles enfrentados por populações racializadas, povos originários, mulheres, pessoas LGBTQIA+, sujeitos em situação de precariedade socioeconômica, dentre outros grupos, de modo que a impedir a possibilidade da apropriação crítica das decisões por parte dos sujeitos impactados, esvaziando seu poder de agência e, por conseguinte, sua capacidade de influenciar decisões que os afetam diretamente. Nessa dinâmica, o poder automatizado desprovido de controle

social e revisão substantiva converte-se em instrumento de dominação, promovendo a continuidade de hierarquias sociais sob o disfarce da racionalidade algorítmica.

Diante dessas problematizações, o presente trabalho tem como escopo central a análise crítica das implicações jurídico-sociais da opacidade algorítmica sobre os direitos fundamentais de sujeitos estruturalmente vulnerabilizados. Parte-se do pressuposto de que a automação decisória, quando desvinculada de mecanismos efetivos de controle democrático, revisão crítica e participação social, representa uma ameaça concreta à efetividade dos direitos fundamentais estabelecidos pela Constituição Federal de 1988 e ao próprio ideal de justiça distributiva.

Assim, o objetivo geral desta pesquisa consiste em avaliar em que medida a ausência de inteligibilidade e auditabilidade dos Sistemas Automatizados de Apoio à Decisão compromete a garantia dos direitos fundamentais de grupos historicamente marginalizados. Para tanto, estabelecem-se os seguintes objetivos específicos: (i) reconstruir os fundamentos teóricos, filosóficos e jurídicos que sustentam o princípio da transparência algorítmica no âmbito do Estado Constitucional de Direito; (ii) identificar os obstáculos técnicos, institucionais e metodológicos que impedem a efetivação da auditabilidade e da responsabilização nas decisões automatizadas; (iii) examinar criticamente os efeitos discriminatórios concretos produzidos por tais decisões sobre sujeitos coletivos vulneráveis; e (iv) propor diretrizes normativas e procedimentais voltadas à promoção de uma governança algorítmica pautada na conformidade com a Constituição Federal de 1988 pelos princípios da equidade material, da justiça distributiva e da dignidade humana.

A metodologia adotada é de natureza qualitativa, com abordagem exploratória e crítica, baseada em revisão bibliográfica e análise documental, construída pelo diálogo com referenciais teóricos provenientes do Direito Constitucional, da Filosofia Política, da Teoria Crítica da Tecnologia e dos Estudos Sociais da Ciência e da Tecnologia por meio de anais de congresso, periódicos, monografias, dissertações de mestrado, teses de doutorado e livros. Além disso, são mobilizados documentos normativos nacionais e internacionais, relatórios de instituições de pesquisa e decisões judiciais que ilustram a materialidade dos problemas analisados a partir de uma perspectiva interdisciplinar, articulando conhecimentos jurídicos, técnicos e sociológicos, a fim de compreender a complexidade dos sistemas algorítmicos e seus impactos normativos.

O presente trabalho estrutura-se em cinco seções: introdução, três capítulos de desenvolvimento e conclusão.

A segunda seção realiza uma análise crítica do processo de consolidação dos sistemas algorítmicos no contexto da hiperdigitalização contemporânea, com ênfase em seu papel estruturante na reconfiguração dos regimes decisórios nas esferas social, econômica e institucional e a importância da transparência algorítmica para o desenvolvimento ético das tecnologias inteligentes. Nessa perspectiva, apresentam-se os Sistemas Automatizados de Apoio à Decisão, destacando-se como um avanço tecnológico devido a suas capacidades adaptativas de aprendizado autônomo e inferência estatística contínua, as quais ampliam significativamente a eficiência procedimental.

Contudo, tais ganhos vêm acompanhados de dilemas ético-normativos de elevada complexidade, especialmente quando esses sistemas são implementados em contextos marcados por desigualdades históricas e vulnerabilidades interseccionais, de modo que a transferência de decisões destituídos de intencionalidade ética e insensíveis à contextualidade social impõe sérios desafios à legitimidade das decisões automatizadas.

O terceiro capítulo aprofunda a análise das formas contemporâneas de vulnerabilização a que estão submetidos grupos historicamente subalternizados, especialmente diante da crescente adoção de sistemas algorítmicos em processos decisórios públicos e privados. Essa análise é aprofundada a partir do referencial da interseccionalidade, compreendida como instrumento teórico-metodológico de natureza estrutural e relacional, capaz de captar a articulação entre sistemas de dominação como racismo, patriarcado, capacitismo e exploração econômica, de modo que a seção evidencia que tais estruturas operam conformando regimes de exclusão persistentes que exigem abordagens críticas para além de leituras identitárias reducionistas. Nesse escopo, discute-se o fenômeno da discriminação algorítmica, entendido como uma reconfiguração tecnocientífica das desigualdades sociais que, ao ocultar lógicas operacionais, passam a desempenhar, por meio de sistemas algorítmicos, um papel ativo na reprodução de preconceitos estruturais.

O quarto capítulo volta-se à análise dos desafios jurídicos e políticos que decorrem da integração intensiva de sistemas algorítmicos aos processos decisórios contemporâneos, com especial atenção às implicações para a efetividade dos direitos fundamentais de sujeitos historicamente subalternizados. Parte-se do diagnóstico de que a emergência de uma

racionalidade algorítmica, caracterizada por opacidade estrutural, autonomia técnica e propensão à reprodução de desigualdades, impõe uma releitura crítica do projeto constitucional de 1988, este fundado nos princípios da dignidade da pessoa humana e da igualdade substancial.

Diante desse cenário, defende-se a formulação de um arcabouço jurídico específico para a governança algorítmica, estruturado em um modelo de controle preventivo e repressivo que articule os princípios constitucionais a mecanismos eficazes de responsabilização institucional e a diretrizes regulatórias sensíveis às especificidades socioculturais brasileiras. Conclui-se, por fim, que é imprescindível a construção de uma agenda normativa sólida, juridicamente vinculante e socialmente emancipadora, alicerçada na justiça social, na equidade epistêmica e na plena efetivação dos direitos fundamentais em face dos desafios postos pela era algorítmica.

## **2 A COMPLEXIDADE DOS MECANISMOS AUTÔMATOS DECISÓRIOS DIANTE DA (IMPERATIVA) TRANSPARÊNCIA ALGORÍTMICA**

A consolidação das tecnologias digitais avançadas, em particular os sistemas computacionais inteligentes estruturados sobre arquiteturas algorítmicas preditivas e dotados de elevada complexidade lógico-matemática, tem promovido uma reconfiguração substantiva nas dinâmicas estruturais que definem a relação entre técnica e sociedade. Esses artefatos, outrora concebidos como ferramentas auxiliares, passaram a ocupar posições estratégicas nos processos decisórios que atravessam distintas esferas da vida social, desempenhando funções normativas e operacionais que influenciam de maneira decisiva a formulação de condutas, a distribuição de recursos e o acesso a direitos fundamentais (Castells, 2013; Schwab, 2016).

Essa inflexão tecnológica sinaliza uma transição paradigmática na ontologia da técnica, que deixa de ser concebida como entidade neutra e exterior à ordem social para assumir a condição de instância co-constitutiva das práticas simbólicas, institucionais e políticas (Barocas; Selbst, 2016). Essa mudança torna-se particularmente evidente nos contextos de aplicação desses sistemas em domínios sensíveis, tais como os setores da saúde, justiça, segurança pública e finanças, nos quais sua operatividade adquire relevância sociopolítica e impacto jurídico-normativo consideráveis.

Entretanto, a ascendente centralidade dos sistemas algorítmicos na condução de aspectos fundamentais da vida social contemporânea evidencia um dilema ético persistente, a saber, a opacidade inerente às suas lógicas internas de funcionamento. Desenvolvidos no seio de ecossistemas sociotécnicos historicamente situados, permeados por assimetrias estruturais de poder e orientações normativas específicas, esses sistemas deixam de operar como instâncias neutras e tendem a cristalizar intencionalidades políticas, filtragens e mediações atravessadas por hierarquias sociais, estigmas consolidados e vieses (O’Neil, 2020).

Essa opacidade algorítmica compromete não apenas a inteligibilidade pública dos processos decisórios automatizados, aspecto crucial para a legitimidade democrática, mas também inviabiliza em grande medida a atribuição de responsabilidade por seus efeitos concretos. Diante desse cenário, tornam-se imperativas formas de transparência e explicabilidade algorítmica que ultrapassem a mera acessibilidade técnica a códigos-fonte ou infraestruturas computacionais e se constituam como instrumentos epistêmico-críticos capazes de viabilizar uma análise aprofundada das implicações jurídicas, normativas e

políticas decorrentes da atuação desses sistemas no interior das instituições democráticas (Roscher *et al.*, 2020; Fernandes, 2023).

## 2.1 ALGORITMOS E TRANSPARÊNCIA ALGORÍTMICA NA ERA DA INTELIGÊNCIA ARTIFICIAL

A intensificação do desenvolvimento tecnológico nas últimas décadas, aliada à disseminação em larga escala das tecnologias da informação e comunicação (TICs), consolidaram um novo paradigma técnico-econômico caracterizado pela convergência de inovações nos campos da microeletrônica, da computação e das telecomunicações (Castells, 2013). Esse fenômeno, descrito por Schwab (2016) como a Quarta Revolução Industrial, inaugura uma era marcada pela hiperconectividade, pela centralidade dos dados e pela integração entre sistemas físicos e digitais, com a consequente reconfiguração das dinâmicas econômicas, socioculturais e organizacionais.

Delineia-se, assim, um ecossistema informacional interconectado, cuja sustentação técnica se ancora na infraestrutura da *World Wide Web* e cuja lógica operacional é marcada por fluxos informacionais contínuos, dinâmicos e descentralizados (Castells, 2013). Nesse panorama, os sujeitos tornam-se progressivamente integrados a circuitos de retroalimentação informacional, nos quais a informação assume uma função ambivalente, sendo, ao mesmo tempo, produto das interações sociais mediadas por dispositivos digitais e vetor estruturante que modula essas mesmas interações, de modo a orientar condutas, antecipar preferências e viabilizar práticas de monitoramento (Bioni, 2021).

Segundo Floridi (2015), a crescente interdependência entre tecnologia e práticas cotidianas inaugura uma nova condição existencial denominada *onlife*, a qual representa a dissolução das fronteiras entre os domínios físico e digital, instaurando um ambiente híbrido no qual os indivíduos operam em constante interação com sistemas digitais, dados e práticas sociais informacionalmente mediadas. Como resultado, a subjetividade passa a ser modulada em função de ecossistemas nos quais as interações, escolhas e emoções são constantemente atravessadas por processos de quantificação e codificação, de modo que os regimes de informação e os dispositivos digitais não apenas mediam a experiência, mas também

estruturam a própria constituição do sujeito e o funcionamento das infraestruturas sociais (Castells, 2013).

Inserido nessa lógica, o vínculo entre usuários e ambientes digitais passa a ser mediado por exigências crescentes de disponibilização de dados pessoais, frequentemente estabelecidas como condição para o acesso a bens e serviços. Nesse sentido, Paulichi e Cardin (2025) argumentam que essa exigência configura uma relação assimétrica, na qual os indivíduos, ao priorizarem conectividade e funcionalidade, consentem com a entrega de informações sensíveis, muitas vezes sem plena consciência do alcance, profundidade e persistência dos mecanismos de monitoramento aos quais estão submetidos.

Conforme Hoffmann-Riem (2020), os dados podem ser compreendidos como composições estruturadas de sinais ou símbolos, cuja principal finalidade reside na veiculação de informações, permitindo desde interações entre indivíduos até configurações mais complexas, como aquelas mediadas por sistemas tecnológicos ou mesmo realizadas de forma autônoma entre essas estruturas. Nessa perspectiva, Bioni (2021) assinala que os dados constituem a forma mais básica da informação e, por si sós, carecem de significado e relevância epistemológica, sendo considerados recursos iniciais que apenas adquirem sentido e funcionalidade após passarem por etapas de organização, análise e contextualização, as quais tornam possível a produção de conteúdo informativo.

A sua aplicação contemporânea, contudo, transcende as funções comunicacionais tradicionais ao assumir um caráter estratégico na criação e consolidação de novos modelos de negócios. Isso ocorre porque, ao serem compostos por registros detalhados de interações e padrões comportamentais dos usuários, os dados oferecem uma base sólida de informações, capazes de subsidiar a formulação de estratégias mercadológicas voltadas tanto para o aumento da eficácia organizacional quanto para a maximização dos lucros (Akamine Jr.; Abrão; Kashiura Jr., 2023).

Esse processo é intensificado pela consolidação do fenômeno do *Big Data*, compreendido como o conjunto de tecnologias e metodologias voltadas à coleta, ao tratamento e à análise de grandes volumes de informações, gerados de forma contínua e provenientes de diversas fontes. A complexidade e a rapidez desses fluxos informacionais ultrapassam as capacidades analíticas humanas tradicionais, exigindo, assim, o emprego de soluções baseadas em tecnologias computacionais sofisticadas e técnicas de inteligência

artificial aptas a extrair valor e significado desses vastos repositórios de informação (Paulichi; Cardin, 2025).

Nesse panorama, configura-se uma nova lógica econômica orientada pela centralidade dos dados digitais, os quais deixam de ser interpretados como produtos colaterais das interações tecnológicas para serem reconhecidos como recursos intangíveis dotados de elevado potencial estratégico. Assim, atribui-se à informação um papel fundamental nos mecanismos de criação de valor, e, conseqüentemente, nas estruturas institucionais, nas dinâmicas mercadológicas e nas orientações que norteiam o desenvolvimento social e econômico (Bioni, 2021; Frazão; Oliva; Tepedino, 2019).

É nesse contexto que se insere o conceito de Capitalismo de Vigilância, formulado por Zuboff (2021), para caracterizar uma nova forma de acumulação baseada na extração sistemática de dados individuais como insumos não onerosos. Essa dinâmica compreende a conversão de experiências subjetivas em registros quantificáveis, os quais são analisados com o objetivo de antecipar, direcionar e condicionar comportamentos futuros, com fins comerciais. A partir dessa lógica, consolida-se um novo modelo de exploração informacional, em que a observação contínua das condutas humanas se torna um vetor privilegiado de geração de valor no mercado.

Dentro dessa lógica, Zuboff introduz o conceito de Instrumentalismo, que, embora esteja profundamente vinculado ao Capitalismo de Vigilância, remete a uma dimensão específica da assimetria de poder: trata-se do uso de dispositivos e ambientes inteligentes para monitorar, influenciar e direcionar ações individuais em benefício de terceiros. Enquanto o Capitalismo de Vigilância define a infraestrutura econômica que impulsiona a coleta e comercialização de dados, o Instrumentalismo descreve a ideologia sociotécnica que legitima e naturaliza essa forma de dominação, esvaziando a subjetividade humana em nome da eficiência e do controle. Assim, embora interdependentes, os dois conceitos operam em níveis distintos: o primeiro, como um regime econômico; o segundo, como uma lógica de poder e governança informacional (Zuboff, 2021).

Para Zuboff (2021), o problema central não reside na tecnologia em si, mas na racionalidade econômica que orienta o seu uso. Através da coleta massiva, do processamento automatizado e da análise detalhada de dados pessoais, busca-se identificar padrões de comportamento capazes de induzir os indivíduos à conformidade com interesses

predominantemente comerciais. Nesse contexto, o indivíduo passa a ser gradualmente redefinido como um agente passivo, desprovido de controle efetivo sobre os dados que o representam, consolidando-se, assim, uma lógica de vigilância e dominação informacional (Bioni, 2021).

Com o avanço dessa lógica extrativa de dados, autores como Varoufakis (2021) argumentam que o modelo capitalista tradicional estaria sendo substituído por uma nova configuração socioeconômica, o tecno-feudalismo. Nesse arranjo, corporações digitais como Amazon, Google e Meta deixam de operar como empresas capitalistas convencionais e passam a se comportar como feudos digitais privados, onde os usuários, enquanto produtores-consumidores, geram gratuitamente os insumos informacionais que alimentam a acumulação de capital dessas plataformas. Nessa dinâmica, a apropriação de valor deixa de ocorrer prioritariamente por meio das trocas mercantis e passa a ocorrer através do controle de infraestruturas digitais e da gestão da atenção.

Como destacam Akamine Jr., Abrão e Kashiura Jr. (2023), a cessão de dados pessoais ocorre, em grande parte, de forma automatizada, sutil e contínua, o que dificulta a percepção, por parte dos usuários, da real extensão da coleta e do controle exercido sobre suas ações. Um caso emblemático que ilustra essa problemática é o escândalo envolvendo o Facebook e a empresa britânica de consultoria política Cambridge Analytica. Em 2018, revelou-se que dados comportamentais e informações sensíveis de ao menos 50 milhões de usuários haviam sido coletados sem consentimento explícito, com o objetivo de construir perfis psicográficos usados para a segmentação e disseminação de mensagens eleitorais personalizadas (The Guardian, 2018).

Esse episódio ilustra como as ferramentas digitais, embora frequentemente associadas à inovação e ao progresso tecnológico, também podem ser mobilizadas como instrumentos de intensificação de práticas históricas de vigilância e controle social, agora potencializadas pelos avanços tecnológicos (Eubanks, 2018). Assim, ainda que a digitalização tenha promovido ganhos na personalização de serviços e na inovação, ela também acentuou desigualdades estruturais de poder entre grandes corporações e usuários, contribuindo para a consolidação de um sistema de vigilância constante (Paulichi; Cardin, 2023).

No Brasil, a conformação de um arcabouço normativo voltado à regulação dos fenômenos jurídicos decorrentes da transformação digital deu-se de maneira gradual, em um

contexto originalmente caracterizado por uma legislação fragmentada, desarticulada e carente de fundamentos técnicos compatíveis com as novas dinâmicas sociotecnológicas. Até a metade da década de 2010, o ordenamento jurídico vigente apresentava limitações estruturais significativas para lidar com a complexidade crescente resultante da digitalização das esferas social, econômica e institucional, em especial quanto à intensificação do uso de tecnologias voltadas ao tratamento automatizado de dados pessoais (Bioni, 2019).

Nesse cenário de déficit normativo, a promulgação do Marco Civil da Internet, pela Lei nº 12.965/2014, constituiu um marco inaugural ao estabelecer diretrizes principiológicas que orientam a atuação estatal e privada no ambiente digital. A lei consagrou direitos como a proteção da privacidade, a inviolabilidade dos dados pessoais e a transparência no ambiente digital, inaugurando, assim, uma estrutura normativa mínima para a regulação do ciberespaço no país. Contudo, diante da crescente sofisticação das tecnologias contemporâneas, o Marco Civil revelou-se insuficiente para garantir a efetiva proteção dos indivíduos frente às novas formas de tratamento de dados baseadas em inteligência artificial e aprendizagem de máquina.

Com o objetivo de suprir tais lacunas e conferir maior densidade normativa ao campo da proteção de dados, sobreveio a promulgação da Lei Geral de Proteção de Dados Pessoais (LGPD), pela Lei nº 13.709/2018, concebida como um instrumento normativo de caráter abrangente e sistematizador, orientado pelos fundamentos do constitucionalismo democrático. A LGPD disciplina de modo minucioso todas as etapas do ciclo de vida dos dados pessoais, da coleta à eliminação, e ancora-se em princípios estruturantes, tais como a dignidade da pessoa humana, a autodeterminação informativa e o livre desenvolvimento da personalidade.

Nesse panorama regulatório, os algoritmos ocupam uma posição central, em razão, principalmente, de sua capacidade em processar grandes volumes de dados de forma sistemática, o que permite não apenas ampliar as possibilidades de inferência e análise em domínios complexos, mas também fortalecer mecanismos de poder e tomada de decisão que frequentemente operam fora do alcance do escrutínio público (Hoffmann-Riem, 2020).

O termo "algoritmo" tem sua origem no matemático persa do século IX, Abdullah Muhammad Ibn Musa Al-Khwarizmi, cujo nome, ao ser latinizado como *Algoritmi*, foi inicialmente associado ao sistema de numeração decimal da época. Contudo, com o advento dos primeiros computadores na década de 1950, a definição do conceito foi revisitada e

ampliada, resultando na interpretação contemporânea que prevalece atualmente (Kowalkiewicz, 2019).

Atualmente, algoritmos são definidos como arranjos sistemáticos de operações matemáticas, estruturados de forma lógica e implementados por meio de linguagens computacionais com vistas à realização de funções específicas com elevado grau de precisão, cujo desempenho está diretamente vinculado à adequação, completude e confiabilidade das informações utilizadas como insumo (Mittelstadt *et al.*, 2016). Para que produzam resultados válidos, é imprescindível que sua concepção siga critérios técnicos rigorosos em todas as fases, abrangendo desde a arquitetura dos sistemas de entrada de dados, responsáveis pela captação fidedigna de informações, até o delineamento dos mecanismos de saída, que devem representar com fidelidade os resultados esperados (Valentini, 2021).

Entre as diversas aplicações algorítmicas contemporâneas, destaca-se a prática do perfilamento (*profiling*), que, fundamentado nas técnicas de mineração de dados (*data mining*), visa identificar correlações significativas que possibilitem a projeção de atributos ou condutas futuras em sujeitos que compartilham estruturas similares àquelas mapeadas previamente (Hildebrandt; Koops, 2010). A sua operacionalização ocorre majoritariamente por meio de modelos de aprendizado de máquina (*machine learning*), que conferem aos sistemas a capacidade de identificar autonomamente regularidades complexas e, a partir delas, produzir decisões computacionalmente orientadas por probabilidades inferidas (Barocas; Selbst, 2016; Zuiderveen Borgesius, 2022).

A construção desses modelos segue uma arquitetura metodológica de natureza bifásica, composta por etapas interdependentes e articuladas. A primeira etapa consiste na coleta e no pré-processamento de dados empíricos, como registros de navegação, transações financeiras, dados de geolocalização e interações digitais, submetidos a procedimentos de higienização e categorização. O objetivo é constituir base de dados coerente, íntegra e compatível com os requisitos técnicos dos sistemas de processamento algorítmico. Em seguida, esse conjunto de dados organizados é utilizado como base para o treinamento dos modelos computacionais, que identificam, por meio de técnicas inferenciais, padrões latentes e interdependências estruturais não perceptíveis à observação empírica direta, permitindo que os algoritmos generalizem estruturas estatísticas e estimem, com alto grau de probabilidade, atributos ou comportamentos de instâncias ainda não observadas (Barocas; Selbst, 2016).

É nesse cenário que se insere o campo do aprendizado de máquina, um subdomínio da inteligência artificial dedicado ao desenvolvimento de sistemas autônomos capazes de aperfeiçoar suas respostas com base na análise extensiva de grandes volumes de dados. A construção desses modelos, portanto, exige a aplicação rigorosa de critérios estatísticos e computacionais, com vistas à geração de respostas precisas, de modo que os resultados produzidos alimentam os próprios sistemas em ciclos contínuos de otimização, promovendo a elevação progressiva de sua capacidade adaptativa (Cozman; Kaufman, 2022).

Para compreender o surgimento e a evolução desse campo, é necessário retornar às origens da própria inteligência artificial. O termo "inteligência artificial" foi cunhado em 1955 por John McCarthy, então professor do Dartmouth College, ao organizar a primeira conferência dedicada ao tema. Ainda que Alan Turing já houvesse explorado conceitos similares ao investigar a possibilidade de as máquinas processarem informações e tomarem decisões de maneira análoga ao pensamento humano, foi McCarthy quem sistematizou a área, atribuindo-lhe uma nomenclatura própria. A partir desse marco, a IA evoluiu consideravelmente, com foco na simulação de processos cognitivos humanos, viabilizando o surgimento de sistemas capazes de assimilar dados, adaptar-se a diferentes contextos e operar de forma autônoma (Wedy; Campos, 2023).

Nesse cenário, a IA pode ser compreendida como o resultado da convergência entre técnicas computacionais avançadas e infraestruturas informacionais especializadas. Trata-se de uma área voltada à simulação da cognição humana, com ênfase na capacidade de análise, decisão e adaptação, em que um dos seus principais diferenciais, conforme apontam Protázio *et al.* (2023), reside em sua capacidade preditiva em processar grandes quantidades de dados e identificar padrões complexos, viabilizando a formulação de estratégias operacionais altamente precisas.

Entre os pilares fundamentais da IA, destaca-se o aprendizado de máquina, cujo conceito foi introduzido por Arthur Samuel, em 1959. Esse campo de estudo permite que sistemas computacionais adquiram conhecimento, identifiquem padrões e tomem decisões sem a necessidade de programação explícita. Frequentemente descrito como uma forma de autoaprendizagem, o aprendizado de máquina estrutura-se em três categorias principais: aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço, cada uma com abordagens distintas para a resolução de problemas e otimização de desempenho (Turner, 2019).

Na abordagem supervisionada, os modelos são treinados com dados previamente rotulados, permitindo que o sistema estabeleça associações entre variáveis de entrada e suas respectivas saídas, o que favorece a construção de modelos preditivos com elevado grau de acurácia. Por outro lado, a aprendizagem não supervisionada exige ao sistema a identificação de agrupamentos, padrões e estruturas intrínsecas, sem a necessidade de diretrizes explícitas, ou seja, sem dados rotulados. Complementando essas abordagens, o aprendizado por reforço fundamenta-se na interação contínua entre o agente e seu ambiente, com base em um sistema de recompensas e penalizações, permitindo o aprimoramento progressivo das decisões tomadas, à medida que o sistema aprende com os resultados de suas próprias ações (Turner, 2019).

Ainda que essas técnicas permitam estimar certos resultados com um grau razoável de precisão, a realização de tarefas mais complexas continua sujeita a significativas incertezas. Isso ocorre devido à imprevisibilidade inerente a sistemas de aprendizado contínuo e à ausência de padrões predefinidos, especialmente em contextos em que a transparência dos critérios decisórios é limitada e a rastreabilidade das escolhas realizadas se torna inviável (Rossetti; Angeluci, 2021).

Impõe-se reconhecer, ainda, que os algoritmos não configuram entidades autônomas, tampouco resultam de processos espontâneos no seio dos sistemas computacionais. Constituem, ao contrário, artefatos técnico-científicos elaborados a partir da agência de sujeitos sociais, cujas decisões, metodológicas e operacionais, exercem influência política sobre sua arquitetura funcional e sua dinâmica de execução que, ainda que legitimadas por critérios de natureza técnica, são capazes de refletir e internalizar desigualdades, estigmas sociais e assimetrias estruturais nas práticas algorítmicas (Oliveira; Arantes, 2024).

Embora frequentemente concebidos sob a égide de uma suposta neutralidade axiológica e de uma racionalidade instrumental orientada à eficiência, os algoritmos são, na realidade, produtos de complexos processos de mediação humana, uma vez que sua formulação está ancorada em matrizes teóricas, normativas e técnicas que orientam suas possibilidades de aplicação e delimitam seus horizontes interpretativos. Esses condicionantes estruturais, assim, revelam a presença de fundamentos ideológicos subjacentes ao desenho e à operacionalização dos modelos computacionais, evidenciando que tais sistemas não estão isentos de valores, nem tampouco são epistemicamente neutros (O’Neil, 2020).

No âmbito do desenvolvimento de sistemas computacionais, destaca-se o viés algorítmico, fenômeno estrutural cuja gênese transcende falhas técnicas pontuais, enraizando-se em decisões metodológicas, organizacionais e epistemológicas tomadas ao longo de todo o ciclo de vida dos algoritmos. Desde a curadoria dos dados até a modelagem das arquiteturas lógico-formais e sua implementação, essas escolhas, embora apresentadas como neutras, refletem valores sociais, capazes de reproduzirem desigualdades estruturais sob a aparência de objetividade técnica (Russell; Norvig, 2021).

Além disso, nas fases de modelagem e parametrização, por exemplo, a exclusão deliberada ou contingente de determinadas variáveis, em função de julgamentos sobre sua pertinência, de limitações dos métodos empregados ou de restrições computacionais, justificada em termos de ganho de eficiência ou simplificação do modelo, pode culminar na incorporação inadvertida de vieses sistêmicos e na produção de externalidades negativas, comprometendo não apenas a equidade material dos processos decisórios mediados por algoritmos, mas também a coerência lógico-formal dos resultados produzidos (O’Neil, 2020).

Segundo a análise de Mittelstadt *et al.* (2016), os dilemas éticos relacionados aos sistemas algorítmicos podem ser classificados em três dimensões centrais, cuja compreensão requer um enfoque crítico e interdisciplinar.

A primeira refere-se às consequências epistêmicas, especialmente no que se refere à geração de inferências inadequadas, parciais ou distorcidas. Um exemplo ilustrativo é o estudo conduzido por Rich Caruana, da Microsoft, no qual um algoritmo utilizado para prever complicações em casos de pneumonia produziu, de forma errônea, a estimativa de menor risco de mortalidade entre indivíduos asmáticos. Esse equívoco resultou de vieses presentes nos dados de treinamento, que indicavam maior taxa de sobrevivência para esse grupo em razão do acesso mais ágil a cuidados médicos, o que foi interpretado de forma equivocada como uma relação causal (Caruana *et al.*, 2015).

Esse episódio evidencia a importância de analisar sistemas de inteligência artificial a partir da noção de incerteza ontológica, conforme argumenta Arriagada-Bruneau (2024), uma vez que essas tecnologias operam em ambientes sociotécnicos permeados por valores, normatividades implícitas e relações de poder. Assim, suas falhas não podem ser somente compreendidas apenas como disfunções técnicas, mas como expressões de processos históricos e sociais que influenciam a estruturação dos dados e suas interpretações.

A segunda categoria diz respeito aos efeitos sociais decorrentes da utilização desses sistemas em diferentes contextos institucionais e operacionais. Ao serem desenvolvidos com base em conjuntos de dados historicamente marcados por desigualdades, os algoritmos tendem a manter e, em alguns casos, a acentuar padrões de exclusão, sobretudo quando aplicados sem uma leitura crítica dos contextos históricos que moldam tais informações.

Por fim, a terceira vertente diz respeito à dificuldade de rastreabilidade e à complexidade envolvida na identificação de responsabilidades. A sofisticação técnica e o funcionamento parcialmente autônomo desses sistemas dificultam a determinação de agentes claramente responsáveis pelas decisões automatizadas, tornando a atribuição de responsabilidade por eventuais prejuízos um desafio substancial (Mittelstadt *et al.*, 2016).

Nesse cenário, evidencia-se a relevância do princípio da transparência algorítmica, compreendido como a capacidade de acessar e analisar criticamente os fundamentos teóricos, os procedimentos técnicos e os condicionantes contextuais que sustentam o funcionamento de sistemas computacionais. A clareza quanto à estrutura interna dos algoritmos, aos critérios que regem sua lógica decisória e às fontes de dados envolvidas constitui um requisito fundamental para a compreensão dos processos automatizados, ao mesmo tempo em que favorece a identificação de limitações metodológicas, a mitigação de desequilíbrios informacionais, o aprimoramento contínuo dos modelos e a legitimação de seu uso em diferentes contextos sociais (Tintarev; Masthoff, 2007).

Com o intuito de tornar o conceito de transparência mais aplicável, Roscher *et al.* (2020) propõem sua estruturação em três dimensões complementares. A primeira, simulabilidade, refere-se à capacidade de compreender o modelo como um todo e identificar as inter-relações entre suas partes. A segunda, decomponibilidade, diz respeito à possibilidade de examinar separadamente os componentes e parâmetros que compõem o sistema. Por fim, a terceira dimensão, compreendida como transparência em sentido estrito, envolve a explicitação dos processos de treinamento, validação e tomada de decisão que sustentam o funcionamento do algoritmo.

Dessa forma, compreende-se que a tecnologia não é uma força autônoma e inevitável, mas um fenômeno intrinsecamente ligado às escolhas e ações humanas, exigindo uma abordagem que vá além da aceitação irrestrita ou da rejeição absoluta (Schwab, 2016). Ainda que a transparência não represente uma solução completa para os dilemas éticos associados a

esses sistemas, ela desempenha um papel central na análise de sua legitimidade e confiabilidade, elementos fundamentais para assegurar justiça e eficácia em sua implementação (Floridi *et al.*, 2018).

## 2.2 UTILIZAÇÃO E IMPACTOS DOS SISTEMAS AUTOMATIZADOS DE APOIO À DECISÃO

Entre as décadas de 1970 e 1980, foram desenvolvidos os primeiros programas de computador com o objetivo de auxiliar os seres humanos na resolução de tarefas e na tomada de decisões (Power, 2024). Esses sistemas, designados como Sistemas de Suporte à Decisão, firmaram-se como ferramentas capazes de realizar análises detalhadas e solucionar problemas ao se ajustarem com precisão às particularidades de cada cenário decisório (Joshi, 2018; Santos, 2023).

Contudo, suas limitações tornaram-se evidentes frente a questões de elevada complexidade, as quais requeriam não apenas um conhecimento técnico avançado, mas também uma maior capacidade para gerenciar incertezas, flutuações e ambiguidades intrínsecas a diversas situações de tomada de decisão, os quais frequentemente excedem o alcance preditivo dos modelos tradicionais (Santos, 2023). Nesse cenário, os Sistemas Automatizados de Apoio à Decisão emergem como um aprimoramento dos modelos tradicionais, ao incorporarem atributos que caracterizam uma "inteligência comportamental" (Turban; Aronson, 1998), como a habilidade de aprender com experiências anteriores, interpretar ambiguidades e contradições e aplicar raciocínios lógicos para solucionar questões complexas (Phillips-Wren, 2013).

De acordo com Desmoulin-Canselier (2020), esses sistemas diferenciam-se pelo grau de autonomia. Enquanto os Sistemas Automatizados de Apoio à Decisão operam de forma autônoma, adaptando-se dinamicamente ao ambiente, identificando padrões, prevendo eventos e automatizando escolhas a partir de grandes volumes de dados, os Sistemas de Suporte à Decisão atuam como ferramentas baseadas em regras fixas, cabendo ao especialista humano a responsabilidade pela decisão final.

Nos Sistemas Automatizados de Apoio à Decisão, a etapa inicial compreende a coleta e o tratamento preliminar de dados, estes extraídos de diversas fontes, como banco de dados e

dispositivos de monitoramento. Esses dados, contudo, frequentemente apresentam distorções, como lacunas, inconsistências sintáticas e erros de entrada, os quais exigem procedimentos especializados voltados à garantia da sua qualidade, o que abrange desde a imputação de valores ausentes até a uniformização de formatos e conversões técnicas compatíveis com os requisitos da etapa analítica (Desmoulin-Canselier, 2020).

Posteriormente, durante o desenvolvimento dos modelos, são aplicadas metodologias avançadas, incluindo aprendizado supervisionado e não supervisionado, redes neurais profundas e sistemas especialistas. Esse processo exige a definição criteriosa dos hiperparâmetros, elementos responsáveis por ajustar a dinâmica interna dos métodos utilizados, assim como a aplicação de validação cruzada, técnica que assegura a análise do desempenho dos modelos em diferentes divisões dos dados, favorecendo a estabilidade e a capacidade de generalização das soluções propostas (Kaggwa *et al.*, 2024).

A capacidade adaptativa dos Sistemas Automatizados de Apoio à Decisão constitui um dos seus diferenciais mais relevantes, uma vez que permite a atualização contínua de seus componentes analíticos a partir do monitoramento do desempenho em tempo real. Esse aprimoramento ocorre por meio da detecção automática de anomalias, da análise de novos dados incorporados ao sistema e da retroalimentação oriunda das decisões anteriores, resultando na elevação progressiva da acurácia e da confiabilidade dos resultados, além de consolidar o papel estratégico desses sistemas em cenários complexos e voláteis que demandam respostas ágeis e previsibilidade operacional (Antoniadi *et al.*, 2021).

A aplicação desse sistema inteligente tem se expandido para diversos setores organizacionais e institucionais, cada qual explorando as potencialidades específicas dessas tecnologias conforme as suas demandas informacionais. No domínio da saúde, os sistemas têm sido empregados na construção de diagnósticos clínicos, na definição de protocolos terapêuticos e na gestão integrada dos serviços assistenciais por meio da análise de grandes volumes de dados heterogêneos, como imagens médicas, registros eletrônicos e literatura científica, com vistas ao reconhecimento de padrões e ao aprimoramento da tomada de decisão clínica (Rocha *et al.*, 2022).

No ambiente industrial, observa-se a incorporação dos Sistemas Automatizados de Apoio à Decisão na automação de processos produtivos, no controle operacional de equipamentos e na adoção de estratégias de manutenção baseadas em predição, contribuindo

para a elevação da eficiência e para a redução de custos, ao viabilizar a antecipação de falhas e a otimização dos mecanismos de controle da qualidade. No setor financeiro, por sua vez, esses sistemas vêm sendo utilizados na análise dinâmica de mercados, na formulação de estratégias de investimento e na identificação de práticas fraudulentas, como transações atípicas e tentativas de falsificação de identidade (Oliveira, 2023).

Apesar das vantagens práticas desses sistemas em contextos técnicos e operacionais, surgem preocupações quanto ao seu uso em domínios que envolvem complexidade social e ética. Em “Automatizando as Desigualdades”, Virginia Eubanks (2018) examina os impactos das tecnologias digitais nos processos decisórios em áreas sensíveis, como finanças pessoais, mercado de trabalho, políticas públicas, saúde e assistência social. Segundo a autora, decisões anteriormente mediadas por profissionais capazes de considerar elementos contextuais e subjetivos passaram a ser transferidas a sistemas algorítmicos, cuja lógica de funcionamento apresenta limitações para lidar com as nuances das dinâmicas sociais e estruturais que permeiam tais decisões.

Nesse sentido, importa destacar que o processo decisório mediado por sistemas algorítmicos difere substancialmente do raciocínio humano, distinguindo-se, de modo particular, por sua rigidez operativa e pela inaptidão para incorporar variações interpretativas contextuais. Enquanto sujeitos humanos mobilizam capacidades cognitivas complexas, aptas a articular variáveis situacionais, critérios ético-normativos e julgamentos morais, os dispositivos algorítmicos operam com base em estruturas lógico-formais estritamente definidas, sustentadas por arquiteturas computacionais que privilegiam a precisão técnica, a previsibilidade de resultados e a otimização da eficiência operacional.

Na obra “A Condição Humana”, Hannah Arendt (1999) propõe uma análise crítica acerca da crescente matematização dos fenômenos sociais. A autora observa que a inserção da matemática nas questões humanas estabelece uma tensão, dado que seus conceitos não se deixam traduzir integralmente em linguagem discursiva, de modo que todo conhecimento ou experiência humana só adquire sentido na medida em que pode ser discutido. Em um cenário em que os modelos matemáticos alcançam níveis cada vez mais avançados, especialmente com o apoio de tecnologias computacionais, há o risco de que essas representações da realidade, embora altamente precisas, se tornem excessivamente complexas e, conseqüentemente, incompreensíveis.

Observa-se que modelos algorítmicos, ao serem empregados em distintos contextos de aplicação, podem ser programados para atender a finalidades específicas, conforme os critérios definidos por seus desenvolvedores ou pelas instituições que os utilizam. No âmbito da condução autônoma, por exemplo, as escolhas implementadas nos sistemas de decisão podem oscilar entre posturas mais utilitaristas, orientadas pela maximização de eficiência, e estratégias mais restritivas, voltadas à redução de riscos. De modo análogo, no campo das operações financeiras, especialmente na concessão de crédito, é possível que agentes autorizem financiamentos em áreas com maior índice de inadimplência, adotando tal medida como instrumento de estímulo ao crescimento econômico regional, ainda que envolva certo grau de exposição a perdas (O'Neil, 2020).

Contudo, conforme adverte O'Neil (2020), a pretensão de neutralidade atribuída aos modelos matemáticos revela-se ilusória. Embora projetados para descrever e intervir em fenômenos sociais, esses sistemas invariavelmente refletem os valores morais e os julgamentos subjetivos de seus projetistas. Isso se deve, em grande medida, aos critérios adotados para definir o desempenho dos algoritmos, os quais influenciam diretamente as prioridades e os resultados promovidos em sua aplicação prática.

Essa problemática adquire particular relevância em contextos de alta sensibilidade, sobretudo quando envolve a inferência de características pessoais, a exemplo do gênero, a orientação sexual, condições psicológicas, conduta delituosa ou desempenho profissional (Frazão, 2019). Nesses cenários, é recorrente que algoritmos incorporem vieses decorrentes tanto de sua estrutura técnica quanto das bases de dados que os alimentam, o que pode acarretar a reprodução de desigualdades historicamente consolidadas, sobretudo quando essas distorções se originam de dados já enviesados (Silva, 2020; Chesterman, 2020).

O incidente ocorrido em 2015 com a plataforma Google Fotos, em que indivíduos negros foram incorretamente identificados como "gorilas", ilustra os potenciais impactos negativos decorrentes do uso de conjuntos de dados com viés implícito (Salas, 2018). Embora posteriormente ajustado, o caso evidencia a capacidade dos sistemas algorítmicos de incorporar distorções presentes nos processos de desenvolvimento, ressaltando a necessidade de práticas rigorosas, pautadas por critérios éticos, críticos e transparentes ao longo do ciclo de vida dessas tecnologias (Rossetti; Angeluci, 2021).

Não obstante os riscos associados, a credibilidade frequentemente atribuída aos sistemas algorítmicos sustenta-se, de modo geral, em uma perspectiva acrítica que os posiciona como instrumentos de elevada eficácia técnica, capazes de oferecer decisões mais precisas, neutras e livres de distorções. Essa suposição, identificada por Broussard (2018) como tecnochovinismo, expressa uma confiança exacerbada na imparcialidade epistemológica das tecnologias computacionais, desconsiderando os múltiplos fatores históricos, institucionais e culturais que permeiam tanto sua elaboração quanto sua implementação.

Essa compreensão dialoga com a crítica desenvolvida por Morozov (2013), por meio do conceito de tecnosolucionismo, que problematiza a tendência de converter questões sociopolíticas complexas em desafios essencialmente técnicos, cuja resolução se daria exclusivamente por meio de soluções tecnológicas. Nessa linha de análise, Bossewitch *et al.* (2022) apontam para um deslocamento analítico preocupante, no qual ganham destaque abordagens pragmáticas e tecnocráticas, em detrimento de leituras críticas voltadas à compreensão dos mecanismos estruturais que sustentam as desigualdades sociais, como as relações de poder, os condicionantes históricos e os processos econômicos que configuram a sociedade contemporânea.

Assim, a incorporação progressiva de tecnologias automatizadas nos mais variados domínios da vida social e econômica tem engendrado uma dissociação estrutural entre os critérios técnico-computacionais que fundamentam o funcionamento desses sistemas e os processos humanos de avaliação crítica, discernimento moral e sensibilidade contextual. Essa cisão epistemológica e operacional favorece a consolidação de uma racionalidade instrumental e impessoal, característica das lógicas tecnocráticas contemporâneas, na qual decisões derivadas de processos algorítmicos passam a ser investidas de legitimidade automática. Essa legitimação, muitas vezes acrítica, tende a esvaziar os espaços de deliberação ética e restringir a agência reflexiva dos sujeitos afetados por essas decisões (Chesterman, 2020).

Esse deslocamento do locus decisório se evidencia com especial nitidez no setor financeiro, em que a automação dos processos decisórios alcançou elevado grau de difusão. Conforme dados do Instituto de Pesquisa Econômica Aplicada (IPEA), em 2016 mais de 40% das transações realizadas na bolsa de valores brasileira foram conduzidas por algoritmos, sinalizando uma crescente transferência da tomada de decisão para dispositivos destituídos de capacidade interpretativa ou empatia.

No entanto, as fronteiras da confiança na inteligência artificial revelam-se notoriamente frágeis quando sistemas algorítmicos são confrontados com cenários que extrapolam os limites dos parâmetros operacionais para os quais foram originalmente projetados. Em contextos nos quais a previsibilidade é substituída por ambiguidade, e a estabilidade por variabilidade semântica e situacional, tornam-se evidentes limitações que não se restringem à engenharia computacional, mas que alcançam dimensões heurísticas, axiológicas e ontológicas (Pierro, 2018).

A Lei Geral de Proteção de Dados Pessoais (LGPD), em seu artigo 20, consagra aos titulares de dados o direito à revisão de decisões tomadas exclusivamente por meios automatizados, desde que tais decisões produzam efeitos jurídicos ou afetem de forma relevante seus direitos ou interesses legítimos. Todavia, a efetivação concreta dessa prerrogativa jurídica é obstada por uma série de fatores estruturais, entre os quais se destacam a ambiguidade terminológica, a ausência de critérios normativos densos e objetivos, e a fragilidade regulatória no que tange à exigência de intervenção humana qualificada nos fluxos decisórios mediados por tecnologias algorítmicas.

A primeira dessas insuficiências reside na ambiguidade terminológica que envolve os conceitos estruturantes do referido dispositivo legal, notadamente os termos "decisão automatizada" e "impacto significativo", cuja carência de parâmetros interpretativos densos e sistematizados compromete a delimitação precisa do campo de incidência da norma. Essa indefinição conceitual é agravada pela inexistência de garantias institucionais adequadas para assegurar uma intervenção humana qualificada, dotada de conhecimento técnico, autonomia funcional e capacidade crítica para avaliar os resultados produzidos por sistemas automatizados. Em vez de oferecer dispositivos regulatórios capazes de viabilizar a efetividade desse controle, a LGPD adota uma abordagem essencialmente procedimental, limitando-se a reconhecer a existência do direito à revisão sem estabelecer, de forma concreta, os meios normativos necessários para sua realização prática (Silveira; Oliveira, 2022).

Ao adotar uma abordagem procedimental e minimalista do direito à revisão, a LGPD se distancia do modelo normativo mais robusto previsto pelo Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia. O artigo 22 do GDPR garante, além da possibilidade de revisão, o direito à compreensão inteligível dos fundamentos da decisão automatizada, bem como à contestação substantiva por meio de intervenção humana efetiva

(Voigt; Von Dem Bussche, 2017). A legislação brasileira, em contraposição, permanece carente de mecanismos normativos que assegurem tais salvaguardas de forma plena.

### 2.3 PERIGOS DERIVADOS DA OPACIDADE DISCRIMINATÓRIA DOS ALGORITMOS NA MODERNIDADE

No livro *“The Black Box Society”*, Frank Pasquale (2015) mobiliza a alegoria da caverna de Platão como matriz interpretativa para problematizar criticamente os modos de atuação dos sistemas algorítmicos na sociedade contemporânea. Assim como os prisioneiros da caverna experienciavam apenas projeções sombreadas e ilusórias da realidade, os sujeitos imersos na era digital se veem confinados a representações parciais, filtradas por arquiteturas algorítmicas cujo funcionamento interno é, em grande medida, inacessível à compreensão pública.

Consolida-se, na contemporaneidade, uma ruptura estrutural em relação ao paradigma técnico-industrial tradicional, no qual a materialidade dos artefatos tecnológicos e a relativa transparência de seus mecanismos operacionais permitiam certo grau de inteligibilidade funcional e apropriação interpretativa dos processos por parte dos usuários. Com a consolidação de sistemas algorítmicos marcados por elevada complexidade computacional e autonomia decisória progressiva, o saber técnico é transferido para esferas hiperespecializadas que aprofundam a distância epistemológica entre os desenvolvedores e os sujeitos por eles impactados, o que intensifica as assimetrias cognitivas e institucionais (Cabrera Altieri, 2024).

Nesse sentido, Layos (2021) propõe uma tipologia que distingue dois regimes de operação algorítmica: os sistemas transparentes, que possibilitam o escrutínio público e a auditoria independente de seus fundamentos matemáticos; e os sistemas opacos, ou “caixas-pretas”, cuja complexidade computacional tornam seus processos decisórios indecifráveis para observadores externos, comprometendo a rastreabilidade dos encadeamentos lógicos e, por conseguinte, a detecção de vieses, falhas sistêmicas e inconsistência dos seus modelos (Bibal *et al.*, 2020; Fernández, 2023).

Nesse panorama, Burrell (2016) conceitua a opacidade algorítmica como um fenômeno multidimensional, resultante da interação entre o sigilo institucional, a complexidade técnico-científica e os limites cognitivos humanos.

A primeira dessas dimensões, de caráter institucional, refere-se à imposição de barreiras ao acesso às informações técnicas e operacionais que regem o funcionamento dos algoritmos, o que inclui suas lógicas decisórias, parâmetros de modelagem e bases de dados. Essa restrição é frequentemente legitimada por argumentos relativos à proteção da propriedade intelectual, à segurança funcional dos sistemas ou à preservação de vantagens competitivas (Burrell, 2016).

Contudo, essa racionalidade pode ocultar práticas de fraude algorítmica intencional ou expressões discriminatórias. Nesse sentido, o escândalo envolvendo a empresa Volkswagen constitui um exemplo paradigmático dessa problemática. A companhia empregou um software especificamente desenvolvido para detectar condições de inspeção ambiental e, a partir disso, manipular temporariamente os dados de emissão de poluentes, simulando conformidade regulatória (Chesterman, 2020). De maneira análoga, uma análise crítica de um sistema de predição de risco em saúde, amplamente adotado nos Estados Unidos, evidenciou que o algoritmo em questão subestimava sistematicamente a gravidade clínica de pacientes negros. Essa distorção derivava do uso de gastos médicos anteriores como variável *proxy* de severidade, uma métrica metodologicamente falha, sobretudo diante das assimetrias históricas no acesso à saúde por populações racializadas naquele país (Obermeyer *et al.*, 2019).

Esses episódios evidenciam que o sigilo institucional, mesmo quando respaldado por fundamentos jurídicos e econômicos, não pode ser concebido como uma prerrogativa absoluta e inquestionável, de modo que sua legitimidade normativa deve ser continuamente reavaliada à luz dos princípios estruturantes do Estado Democrático de Direito, do interesse público e da proteção dos direitos fundamentais (Hoffmann-Riem, 2020). Como adverte Quinn (2022), torna-se imprescindível estabelecer um equilíbrio normativo entre a tutela da propriedade intelectual e a salvaguarda de interesses empresariais, de um lado, e, de outro, os imperativos ético-políticos de transparência, responsabilidade e controle democrático, sobretudo em contextos nos quais decisões algorítmicas produzem efeitos diretos, substanciais e, por vezes, irreversíveis sobre a dignidade e a autonomia dos indivíduos.

Ademais, a opacidade manifesta-se, em sua segunda vertente, como uma barreira resultante da limitação ao acesso e à apropriação dos conhecimentos técnico-científicos que fundamentam os sistemas automatizados contemporâneos. Esses sistemas pressupõem um nível de especialização que ultrapassa as capacidades do público leigo, promovendo a concentração do saber e do poder decisório em núcleos restritos de especialistas,

majoritariamente situados nos domínios da engenharia, da ciência da computação e das ciências exatas em geral, dificultando a compreensão crítica e a participação informada de outros setores da sociedade no debate sobre os impactos, limitações e riscos dessas tecnologias (Burrell, 2016).

Por fim, a terceira e mais crítica manifestação da opacidade dos sistemas algorítmicos manifesta-se sob a forma da opacidade cognitiva, categoria que designa a dificuldade constitutiva de compreender os mecanismos internos de funcionamento e as lógicas decisórias subjacentes aos sistemas de inteligência artificial de última geração, especialmente aqueles baseados em arquiteturas de redes neurais profundas. Em contraste com modelos estatísticos mais interpretáveis, como classificadores lineares ou árvores de decisão, essas estruturas operam mediante múltiplas camadas de abstração não linear, que processam informações de forma não transparente e cujas dinâmicas internas resistem à auditabilidade, mesmo por parte de especialistas experientes (Burrell, 2016).

A sofisticação técnico-computacional que configura os sistemas algorítmicos contemporâneos, ainda que não implique, em termos estritos, a manifestação explícita de uma intenção discriminatória por parte de seus idealizadores, estabelece obstáculos significativos à promoção da transparência, da auditabilidade e do escrutínio crítico de seus processos decisórios. Essa opacidade estrutural dificulta de maneira substancial a identificação e a mitigação de vieses epistêmicos e normativos que se inserem, de forma muitas vezes imperceptível, ao longo das distintas etapas de seu desenvolvimento. Em consequência, delinea-se um cenário no qual condições sistêmicas favorecem a cristalização de práticas de discriminação algorítmica, frequentemente reproduzidas por meio de decisões tecnocientíficas tomadas desde os momentos inaugurais do ciclo de vida desses sistemas, abarcando a coleta, curadoria e pré-processamento dos dados, bem como a escolha de arquiteturas computacionais, métricas de desempenho e critérios de otimização (Burrell, 2016).

Nesse arcabouço, a curadoria de dados emerge como um vetor de centralidade crítica no processo de constituição das lógicas inferenciais dos algoritmos. Longe de se tratar de uma operação meramente técnica ou instrumental, a curadoria assume um papel fundante, ao requerer a formulação de um aparato metodológico sensível às determinações históricas e sociopolíticas que condicionam os contextos de desenvolvimento e aplicação dos sistemas automatizados (Jackson, 2021). A ausência de um posicionamento reflexivo e crítico nesse estágio compromete não apenas a legitimidade sociotécnica das soluções algorítmicas, mas

também amplia a probabilidade de que desigualdades estruturais sejam reproduzidas e reforçadas e reconfiguradas sob novas formas.

Com efeito, quando os modelos são treinados sobre bases de dados atravessadas por pressupostos homogêneos ou por assimetrias históricas, há uma tendência à reprodução automatizada de práticas discriminatórias, frequentemente camufladas por uma retórica de objetividade técnico-científica, operando como mecanismo de legitimação das inferências geradas, de modo a obscurecer as estruturas excludentes que lhes são subjacentes. A expressão *bias in, bias out* condensa de maneira sintética essa dinâmica tecnopolítica, na medida em que dados moldados por estruturas de desigualdade inevitavelmente engendram inferências enviesadas, independentemente do grau de complexidade dos modelos empregados (McFarland; McFarland, 2015).

A etapa de modelagem algorítmica, nesse contexto, representa um dos núcleos críticos mais sensíveis do processo de desenvolvimento, pois nela se definem os objetivos operacionais dos sistemas, os atributos considerados relevantes e os esquemas classificatórios adotados. Embora frequentemente apresentados como atos puramente técnicos, tais procedimentos são atravessados por racionalidades políticas e morais que, em contextos orientados por uma lógica instrumental e performativa da inovação, tendem a ser naturalizadas, atuando como mecanismo de invisibilização das dimensões ético-políticas das escolhas envolvidas ao legitimar a reprodução silenciosa de estruturas excludentes sob o véu da eficiência e da neutralidade tecnológica (Yarger *et al.*, 2019).

Um exemplo paradigmático dessa lógica encontra-se no sistema automatizado de triagem de currículos desenvolvido pela empresa Amazon. Treinado a partir de um conjunto de dados enviesado, composto majoritariamente por perfis masculinos, o algoritmo passou a penalizar sistematicamente candidaturas femininas, reproduzindo padrões de exclusão de gênero historicamente consolidados nas dinâmicas corporativas de recrutamento. Esse caso evidencia, de maneira contundente, como a racionalidade algorítmica pode operar como mecanismo sofisticado de legitimação e intensificação de desigualdades estruturais preexistentes (Miasato; Silva, 2019).

Em face da crescente centralidade desses sistemas em processos com impactos sociais profundos, torna-se imperativa a consolidação de mecanismos voltados à transparência e à responsabilização algorítmica. Nesse escopo, destaca-se o emergente campo da Inteligência

Artificial Explicável (*Explainable Artificial Intelligence*), que visa ao desenvolvimento de soluções metodológicas e técnicas capazes de conferir inteligibilidade, auditabilidade e justificabilidade às decisões produzidas por sistemas automatizados (Wachter; Mittelstadt; Russell, 2021).

Importa salientar, contudo, que a explicabilidade não deve ser concebida em termos meramente técnicos ou reducionistas, tais como o simples acesso ao código-fonte ou à estrutura computacional do algoritmo. Trata-se, antes, de uma prerrogativa que envolve o acesso à obtenção de informações substantivas, contextualizadas e compreensíveis acerca dos dados de entrada, dos critérios inferenciais mobilizados e dos parâmetros decisórios empregados pelos sistemas algorítmicos, de modo a exigir o estabelecimento de práticas sistemáticas que envolvam a documentação rigorosa das fases de modelagem, a visualização estruturada dos fluxos decisórios, a reconstrução lógica dos fundamentos das decisões e a avaliação retrospectiva dos resultados produzidos (Kroll *et al.*, 2017; Mittelstadt *et al.*, 2016).

Todavia, a efetivação desses instrumentos enfrenta obstáculos relevantes, decorrentes, sobretudo, das tensões inerentes aos interesses do setor privado, particularmente no que diz respeito à proteção do segredo industrial, à defesa da propriedade intelectual e à preservação de vantagens competitivas em um cenário econômico orientado pela inovação tecnológica (Frazão, 2021). No âmbito do ordenamento jurídico brasileiro, tais inquietações encontram respaldo normativo na Constituição da República Federativa do Brasil de 1988, cuja matriz principiológica, consubstanciada no artigo 170, institui a livre iniciativa, a livre concorrência e a valorização do trabalho humano como fundamentos da ordem econômica.

À luz dessa arquitetura constitucional, o segredo comercial se apresenta como instituto jurídico de natureza instrumental, cuja função precípua consiste na proteção de ativos informacionais sensíveis, tais como processos técnicos e organizacionais, estratégias comerciais e dados operacionais, cuja revelação não autorizada compromete não apenas a eficiência operacional de determinadas empresas, mas também a estabilidade e a integridade do regime concorrencial em sua dimensão sistêmica (Peixoto, 2020).

Todavia, a proteção conferida a tais bens jurídicos, embora constitucionalmente assegurada, não ostenta caráter absoluto ou imune à incidência de outros direitos fundamentais. Em contextos específicos, a imposição de deveres de revelação técnica, como a publicização de códigos-fonte, parâmetros estatísticos ou lógicas decisórias de sistemas

automatizados, pode revelar-se juridicamente admissível, desde que presente a demonstração da estrita necessidade da medida para a efetiva tutela de valores igualmente protegidos pela ordem constitucional. Essa hipótese se concretiza, por exemplo, quando a transparência técnica se mostra indispensável à promoção da igualdade material, à prevenção de discriminações algorítmicas ou à responsabilização jurídica de agentes que se valem de processos decisórios automatizados de natureza opaca.

Nessa perspectiva, a categoria da explicabilidade emerge como elemento conceitual central à arquitetura regulatória dos sistemas algorítmicos, devendo ser compreendida como uma propriedade analítica complexa, dotada de densidade normativa e conformada por uma estrutura multidimensional. Sua configuração abrange, de modo articulado e indissociável: a dimensão jurídico-normativa, responsável pela proteção de direitos fundamentais como a privacidade, a autodeterminação informacional e a não discriminação; a dimensão técnico-operacional, que se refere à rastreabilidade, auditabilidade e reprodutibilidade dos processos decisórios automatizados; e a dimensão cognitivo-comunicativa, voltada à promoção da inteligibilidade institucional e social das decisões algorítmicas, como condição essencial para sua legitimidade em regimes democráticos (Kaminski, 2018).

### **3 SOBRE A VULNERABILIDADE DE GRUPOS HISTORICAMENTE MARGINALIZADOS NA UTILIZAÇÃO DE ALGORITMOS**

A intensificação do emprego de sistemas algorítmicos nas dinâmicas sociotécnicas contemporâneas configura uma inflexão paradigmática nas modalidades de exercício do poder, na medida em que tais tecnologias, ancoradas na racionalidade estatística e nos imperativos da eficiência computacional, operam como dispositivos de mediação normativa

que tendem a automatizar e a reproduzir, de maneira sistemática, desigualdades históricas enraizadas nas estruturas sociais (Silva, 2020, Marmelstein, 2021).

É importante destacar que a lógica algorítmica não incide de forma isonômica sobre o tecido social. Seus efeitos tendem a ser desproporcionalmente deletérios para grupos historicamente marginalizados, como pessoas negras, mulheres, povos originários, indivíduos LGBTQIA+, pessoas com deficiência, migrantes e populações periféricas, entre outros. A vulnerabilidade estrutural que atravessa esses sujeitos compromete tanto o acesso efetivo a direitos fundamentais quanto sua visibilidade epistêmica e sua agência política, inclusive nos próprios sistemas digitais que os afetam (Piovesan, 2011).

A adoção de tecnologias algorítmicas em áreas sensíveis, como segurança pública, justiça penal e gestão do trabalho, tem se dado, de modo geral, de forma acrítica, sendo esses sistemas alimentados por bases de dados marcadas por vieses históricos e operando sem levar em conta as especificidades socioculturais dos sujeitos impactados. Essa dinâmica resulta na naturalização das desigualdades, que passam a ser legitimadas como parâmetros técnicos sob uma lógica técnico-instrumental que compromete a dimensão ética dos direitos e da dignidade humana (Eubanks, 2018).

Para uma análise crítica dos impactos produzidos por sistemas algorítmicos, torna-se imprescindível a mobilização do referencial teórico da interseccionalidade, formulado por Crenshaw (2002), uma vez que essa matriz analítica permite apreender a articulação simultânea e interdependente de múltiplos vetores de opressão, incluindo raça, gênero, classe, deficiência e territorialidade, na constituição de dinâmicas complexas de dominação estrutural.

Superar o regime de governança algorítmica excludente demanda mais do que reformas regulatórias pontuais, pois implica uma reconfiguração epistemológica e política profunda, orientada à desestabilização dos fundamentos eurocentrados que sustentam o saber técnico-científico e à reinscrição dos direitos humanos como eixo central das agendas de inovação tecnológica. Nesse sentido, torna-se imperativo descolonizar os dispositivos de produção de saber e poder algorítmico (Quijano, 2008; Collins, 2016) e reafirmar a dignidade humana como princípio estruturante de projetos tecnopolíticos comprometidos com a justiça social, a equidade e uma concepção substantiva de democracia (Mendes, Mattiuzzo, 2020; Arendt, 1999).

### 3.1 DEFINIÇÃO E PECULIARIDADES RELATIVAS AOS SEGMENTOS MINORITÁRIOS

O conceito de “minoria” deve ser apreendido como um instrumento analítico de natureza relacional, cuja formulação está indissociavelmente vinculada a contextos históricos atravessados por assimetrias estruturais e processos de exclusão sistematicamente institucionalizados. Em oposição a abordagens de cunho estritamente quantitativo, essa concepção evidencia a constituição histórica de posições de subalternidade e marginalização ocupadas por determinados grupos sociais, os quais, ao longo do tempo, foram reiteradamente alocados em situações de desvantagem no âmbito das dinâmicas sociais e políticas (Chaves, 1971).

Essa lógica excludente se manifesta por meio de articulações interseccionais que atravessam grupos historicamente vulnerabilizados, mobilizando múltiplas camadas de opressão. Entre esses grupos, incluem-se os povos originários submetidos à desterritorialização forçada; as populações negras atravessadas pelos efeitos persistentes do colonialismo e da escravidão; os migrantes expostos a políticas de vigilância e controle racializado; os sujeitos periféricos das zonas urbanas; bem como mulheres, pessoas LGBTQIA+, indivíduos com deficiência, idosos, pessoas em situação de rua e egressos do sistema penal, entre outros grupos (ONU, 2010).

A despeito da multiplicidade empírica que singulariza as experiências vividas pelos grupos minoritários, é possível identificar, entre tais coletividades, uma convergência estrutural caracterizada por mecanismos reiterados de deslegitimação social, marginalização político-institucional e restrição sistemática no acesso a direitos fundamentais. Evidencia-se a persistência de estruturas históricas de dominação que, ao configurarem de modo assimétrico os campos político, econômico e sociocultural, operam como vetores de reprodução de desigualdades e consolidação de hierarquias excludentes (Hunter, 2000; Mazza, 2005).

À luz desse panorama, o Direito Internacional dos Direitos Humanos se estabelece como um arcabouço normativo vocacionado à proteção de sujeitos historicamente situados em contextos de vulnerabilidade estrutural, impondo aos Estados obrigações positivas, não apenas no sentido de se absterem de práticas discriminatórias, mas também no dever de adotar

medidas concretas e eficazes para assegurar, promover e proteger os direitos culturais, religiosos, linguísticos e identitários das minorias (Piovesan, 2011).

O alicerce ético-jurídico desse regime encontra respaldo no princípio da dignidade da pessoa humana, concebido sob uma perspectiva que se contrapõe à tradição liberal-individualista hegemônica. Ao reconhecer a dimensão coletiva como constitutiva da subjetividade e da identidade humanas, essa abordagem introduz uma inflexão teórica que desloca o centro de gravidade do sujeito de direitos, passando a conferir à diversidade o estatuto de valor normativo fundante da ordem internacional contemporânea de proteção dos direitos humanos (Piovesan, 2011).

Essa concepção adquire expressão normativa concreta no artigo 27 do Pacto Internacional sobre Direitos Civis e Políticos (1966), o qual, com caráter vinculante, assegura aos membros de minorias étnicas, religiosas ou linguísticas o direito de preservar, manifestar e desenvolver suas tradições culturais, práticas religiosas e línguas originárias. Ao reconhecer a identidade coletiva como manifestação indissociável da dignidade humana, esse dispositivo inscreve tais direitos no âmago dos valores fundamentais do Direito Internacional dos Direitos Humanos (Piovesan, 2011).

Todavia, a concretude desse regime jurídico permanece fortemente condicionada à disposição política dos Estados-parte, bem como à existência de estruturas institucionais nacionais dotadas de competência normativa, capacidade técnica e eficácia operativa. A inexistência de mecanismos internacionais providos de autoridade coercitiva, somada à persistência de uma racionalidade soberanista na configuração da governança jurídica global, limita de forma substancial a aplicabilidade plena dessas normas (Piovesan, 2011).

Diante dessas limitações estruturais, a Declaração das Nações Unidas sobre os Direitos das Pessoas Pertencentes a Minorias Nacionais ou Étnicas, Religiosas e Linguísticas, adotada em 1992, apresenta-se como instrumento normativo de natureza diretiva voltado ao fortalecimento da proteção internacional dessas populações. Alicerçada nos princípios da valorização da identidade cultural, da igualdade e da participação política efetiva, a Declaração visa estabelecer uma articulação mais eficaz entre os marcos normativos internacionais e as realidades sociopolíticas locais, com o objetivo de conferir maior densidade normativa e justiciabilidade aos direitos das minorias (Martínez; Cascón; Liesa, 2000).

No âmbito do ordenamento jurídico brasileiro, essa perspectiva encontra ressonância na Constituição da República Federativa do Brasil de 1988, cuja promulgação, inserida no contexto de redemocratização do Estado, institui uma racionalidade jurídico-política inovadora, sustentada por premissas democráticas. Essa nova racionalidade estrutura-se a partir de fundamentos centrais, notadamente a dignidade da pessoa humana, a justiça social e o reconhecimento da diversidade sociocultural, que conformam o núcleo normativo da ordem constitucional, consagrados nos artigos 3º e 5º, os quais impõem ao Estado o dever de reduzir as desigualdades sociais e institucionais, vedar toda forma de discriminação e consagrar a igualdade material como critério orientador da interpretação jurídica e da formulação de políticas públicas (Brasil, 1988).

Todavia, não obstante o avanço normativo, este ainda tem se revelado insuficiente para subverter as estruturas de exclusão social que atravessam a formação histórica do Estado nacional, moldada por matrizes coloniais, escravocratas e patriarcais, cujas estruturas simbólicas e materiais ainda operam na atualidade, perpetuando dinâmicas de exclusão social, seletividade institucional e hierarquização racializada (Groff; Pagel, 2009; Tachy, 2023).

Nesse sentido, a reprodução sistemática de dinâmicas excludentes no interior das estruturas estatais contemporâneas não pode ser reduzida à noção de falhas pontuais ou disfunções administrativas episódicas. Trata-se, antes, da expressão de uma racionalidade política hegemônica que estrutura os dispositivos de governança a partir da institucionalização da precariedade como tecnologia de poder aplicada às populações historicamente subalternizadas, operando pela naturalização das desigualdades estruturais e pela conversão da violência sistêmica (Carvalho, 2024).

É nesse horizonte teórico-crítico que a formulação da necropolítica, concebida por Achille Mbembe, emerge como uma crítica à biopolítica foucaultiana. Ao deslocar o foco analítico da administração da vida para o exercício soberano do poder sobre a morte, Mbembe evidencia que a soberania moderna se define pela capacidade de decidir, de maneira seletiva, sobre a continuidade ou interrupção da existência de determinados grupos sociais. Assim, o poder estatal não se limita à negligência, mas também à permissão de dispositivos de exclusão e extermínio, ao relegar corpos racializados, dissidentes de gênero e sexualidade ou pertencentes a segmentos empobrecidos à condição de abandono institucional e à violência sistemática (Mbembe, 2011).

Assim, a necropolítica não apenas complementa o projeto moderno-colonial, como o atualiza, instituindo regimes de distribuição assimétrica da vida e da morte baseados em dispositivos de exclusão e hierarquização da dignidade humana, articulando racismo, sexismo, cisheteronormatividade e desigualdade econômica como eixos estruturantes da produção de sujeitos descartáveis, corpos vulnerabilizados cuja existência é tornada politicamente irrelevante e juridicamente dispensável (Collins, 2016).

É nesse marco teórico que se insere o álbum *Sobrevivendo no Inferno* (1997), do grupo Racionais MC's, cuja densidade crítica transcende os limites da produção cultural e se projeta como dispositivo epistemológico e contra-hegemônico. A obra articula saberes oriundos das periferias racializadas, elaborando uma crítica enraizada na experiência vivida e socialmente situada da população negra urbana. A escolha do título, ao evocar a ideia de sobrevivência em um “inferno” social, expressa de forma simbólica e incisiva a condição de precariedade imposta aos corpos negros periféricos, cujas existências são atravessadas por dinâmicas de extermínio material e apagamento subjetivo.

A poética insurgente presente no álbum constrói uma verdadeira cartografia afetiva e política das margens urbanas, desvelando o modo como a cidade opera como tecnologia racializada de regulação social. Essa configuração é conceitualmente elaborada por Alves e Vargas (2010), que designa a lógica histórica do urbanismo brasileiro como um projeto de supressão sistemática da presença negra nos circuitos de visibilidade, mobilidade e poder simbólico. Sob os imperativos da racionalidade neoliberal, esse ordenamento espacial se intensifica, reconfigurando-se como um dispositivo de contenção territorial e política de corpos considerados excedentes ou indesejáveis (Alves, 2011).

A crítica engendrada pelo álbum estrutura-se em três dimensões interdependentes. A primeira, tematizada na faixa “Diário de um Detento”, denuncia o sistema carcerário como herança direta do regime escravocrata, operando como tecnologia contemporânea de aprisionamento dos corpos negros e de continuidade da lógica de contenção racial. A segunda dimensão, expressa em “Fim de Semana no Parque”, problematiza a exclusão simbólica e afetiva dos sujeitos negros dos espaços de lazer e prestígio social, revelando a existência de um apartheid urbano difuso, no qual o direito à cidade é condicionado por códigos raciais de pertencimento e normatividade. A terceira dimensão, articulada na faixa “Mágico de Oz”, aborda a banalização da violência letal nas periferias, onde o desejo por dignidade, segurança

e estabilidade material se expressa como um clamor ético-político diante de um Estado que se omite ou age ativamente como executor da necropolítica (James, 2010; Oliveira, 2022).

Esse arcabouço teórico encontra respaldo em dados empíricos que evidenciam a persistência e intensificação das dinâmicas necropolíticas no Brasil. Entre 2013 e 2023, a taxa de homicídios entre pessoas negras permaneceu consistentemente 2,7 vezes superior à verificada entre pessoas não negras, o que reafirma a centralidade do racismo estrutural na produção da letalidade estatal (Cerqueira; Bueno, 2025). No mesmo período, verificou-se o recrudescimento das violências direcionadas a corpos dissidentes de gênero e sexualidade: entre 2022 e 2023, as agressões contra pessoas homossexuais e bissexuais aumentaram 35%, enquanto os ataques a pessoas trans e travestis cresceram 43%. De forma particularmente alarmante, os registros de violência contra travestis apresentaram crescimento superior a 2000% entre 2014 e 2023, revelando um processo sistemático de desumanização e apagamento radical.

A lógica necropolítica manifesta-se de modo ainda mais perverso na experiência dos povos originários, alvos de processos históricos de expropriação territorial, epistemicídio e violência estrutural. A intensificação de práticas extrativistas ilegais, como o garimpo predatório e a exploração clandestina de recursos naturais, frequentemente vinculadas a circuitos de criminalidade organizada, compromete a sustentabilidade dos ecossistemas e provoca a erosão dos modos de vida tradicionais, agravando indicadores de saúde, taxas de suicídio e homicídio, e contribuindo para a desagregação dos vínculos comunitários. Nesse contexto, o necropoder articula devastação ambiental, dominação epistêmica e violência institucional como dimensões complementares de uma mesma matriz de governança letal (Fontanetto, 2013; Souza *et al.*, 2020).

Assim, emerge a ambivalência constitutiva do Estado moderno, que atua simultaneamente como gestor de precariedades socialmente produzidas e executor direto de práticas repressivas letais, revelando a centralidade de uma necropolítica interseccional, sustentada por dispositivos interligados de opressão, entre eles o racismo, o sexismo, a cisheteronormatividade e as desigualdades econômicas que, de forma sinérgica, estruturam os regimes contemporâneos de governança seletiva da vida e da morte (Collins, 2016).

Assim, superar essa racionalidade necropolítica exige mais do que reformas institucionais pontuais ou retóricas inclusivas. Requer, de forma urgente, uma reconfiguração

dos fundamentos epistêmicos, políticos e éticos da modernidade colonial, o que implica em assumir práticas insurgentes de resistência forjadas nas experiências corporificadas e nos saberes contra-hegemônicos de sujeitos histórica e sistematicamente silenciados (Quijano, 2008; Collins, 2016).

Ademais, exige a desestabilização crítica dos regimes de saber e de subjetivação que sustentam o monopólio moderno da legitimidade estatal sobre a vida, de modo que as lutas por justiça epistêmica, reparação histórica e emancipação coletiva devem ser concebidas, portanto, não como demandas periféricas, mas como vetores centrais para a constituição de uma ordem sociopolítica verdadeiramente pós-colonial, pluriversal e comprometida com a afirmação da dignidade de existências historicamente negadas (Quijano, 2008; Collins, 2016).

### 3.2 A INTERSECCIONALIDADE APLICADA A GRUPOS VULNERABILIZADOS COMO ELEMENTO DE AGRAVAMENTO OPRESSIVO

A concepção de interseccionalidade, tal como sistematizada por Kimberlé Crenshaw ao final da década de 1980, constitui um referencial metodológico direcionado à análise crítica das configurações contemporâneas das desigualdades sociais. Formulada originalmente como uma crítica à incapacidade estrutural do sistema jurídico estadunidense em reconhecer juridicamente as experiências de mulheres negras submetidas a diversas formas de opressão, a interseccionalidade desvela as limitações inerentes aos modelos analíticos convencionais frente à complexidade interrelacional dos marcadores sociais diferenciação e subalternização (Crenshaw, 1989; 2002).

Ao deslocar o eixo interpretativo das identidades individuais isoladas para os dispositivos institucionais, normativos, econômicos e socioculturais que organizam, de forma entrelaçada, as hierarquias sociais, a interseccionalidade inaugura uma abordagem de natureza estrutural, relacional e antirreducionista. Nesse escopo teórico, categorias como raça, gênero, classe, sexualidade, nacionalidade e deficiência são concebidas como dimensões co-constitutivas que se articulam de modo indissociável na constituição de regimes específicos de dominação, exclusão e resistência (Akotirene, 2019; Cho, Crenshaw; McCall, 2013).

A densidade analítica desse referencial reside, portanto, em sua capacidade de apreender os modos sincrônicos e interligados pelos quais distintos sistemas de opressão, como o racismo estrutural, o patriarcado, a cisheteronormatividade, o capacitismo e os mecanismos de exploração capitalista operam em conjunto, retroalimentando-se na produção e perpetuação das desigualdades sociais (Collins, 2000; Crenshaw, 2002). Sob esse enfoque, os marcadores identitários se apresentam como construções histórico-discursivas, continuamente produzidas, disputadas e ressignificadas em espaços atravessados por relações assimétricas de poder (Cho, Crenshaw; McCall, 2013).

A densidade crítica da interseccionalidade é intensificada quando integrada às epistemologias decoloniais, cuja função central é a de desestabilizar os alicerces ontológicos da racionalidade moderna eurocêntrica e problematizar os binarismos fundacionais que sustentam a arquitetura epistêmica ocidental, como civilizado/barbárie, razão/emoção e moderno/primitivo. Esses dualismos, historicamente instrumentalizados para a legitimação de hierarquias epistêmicas e para a produção de subjetividades subalternizadas, operam como dispositivos de controle simbólico e cognitivo, reafirmando padrões hegemônicos de validação do conhecimento (Akotirene, 2019).

Nesse horizonte, destaca-se a ideia de regime interseccional de dominação, entendida como uma formação de poder complexa e multissituada, que articula mecanismos institucionais, discursivos e simbólicos para regular, de forma desigual, o acesso a recursos materiais, direitos civis e reconhecimento social, configurando-se, assim, como um dispositivo dinâmico de reprodução de desigualdades históricas, sustentado pela interconexão entre marcadores sociais de diferença, como raça, gênero, classe, sexualidade e territorialidade (Akotirene, 2019).

A esse referencial crítico agregam-se contribuições oriundas de outras tradições teóricas, como a sociologia crítica francesa, que aprofundam a capacidade interpretativa da interseccionalidade. Nesse sentido, destaca-se o conceito de consubstancialidade, formulado por Danièle Kergoat, que aponta para a inseparabilidade estrutural, histórica e prática entre as categorias de classe, sexo e raça. Ao recusar tanto sua dissociação analítica quanto qualquer tentativa de hierarquização entre essas dimensões, Kergoat sustenta que a inteligibilidade dos processos de desigualdade requer a análise das práticas sociais concretas em que essas categorias operam de forma simultânea e mutuamente constitutiva (Kergoat, 1978; 2012).

No contexto latino-americano, a produção intelectual de Lélia Gonzalez e Heleieth Saffioti oferece importantes inflexões à teoria interseccional, ao ancorá-la nas especificidades históricas e sociopolíticas da formação social brasileira. Gonzalez (1984), ao desvelar os fundamentos ideológicos da noção de “democracia racial”, evidencia como o discurso da miscigenação é mobilizado como instrumento de apagamento das desigualdades raciais e de contenção da mobilização política da população negra, legitimando, assim, a reprodução de assimetrias materiais e simbólicas nos campos do trabalho, da educação e da cidadania. Saffioti (2004), por sua vez, formula a categoria de nó estruturante para designar a articulação orgânica entre patriarcado, capitalismo e racismo como pilares constitutivos das hierarquias sociais, permitindo compreender a posição estruturalmente subordinada das mulheres negras, historicamente inseridas em ocupações socialmente desvalorizadas e invisibilizadas, como o trabalho doméstico e as atividades de cuidado.

A solidez teórica desse referencial interseccional encontra respaldo em evidências empíricas contemporâneas que atestam sua fecundidade analítica na compreensão de fenômenos sociais complexos, tais como a segregação socioespacial, a informalidade nas relações laborais e a precarização das condições de vida (Oliveira, 2022; Colosso, 2019; Almeida Batista, 2020). Nesse sentido, a interseccionalidade consolida-se como um instrumento teórico-metodológico indispensável para desvelar os mecanismos de reprodução das desigualdades sociais, ao mesmo tempo em que oferece uma lente crítica para a análise dos processos de subjetivação e das formas sutis, porém persistentes, de exclusão, invisibilização e silenciamento de sujeitos historicamente marginalizados (Al-Rebholz, 2013; Rios; Silva, 2005).

No contexto sociopolítico brasileiro contemporâneo, as dinâmicas interseccionais tornam-se particularmente evidentes na distribuição estruturalmente desigual da violência letal, cuja incidência desproporcional sobre corpos racializados constitui expressão da persistência de processos históricos de racialização e hierarquização da vida. A realidade das mulheres negras, classificadas como pretas ou pardas, conforme os critérios do IBGE, representa uma materialização paradigmática das múltiplas opressões estruturais. Em 2023, elas corresponderam a 68,2% dos homicídios femininos no país, com uma taxa de mortalidade de 4,3 por 100 mil habitantes, e representaram 67,1% dos feminicídios com causa identificada no período entre 2013 e 2023, evidenciando a sobreposição entre violência de gênero e racialização das mortes (Cerqueira; Bueno, 2025).

Esse padrão de violência não se restringe a eventos pontuais, mas expressa um contínuo de vulnerabilização que se intensifica ao longo do ciclo de vida, como ilustrado pela condição da população negra idosa. Em 2023, homens negros com 60 anos ou mais apresentaram uma taxa de internação hospitalar por agressão de 19,3 por 100 mil habitantes, mais que o dobro do registrado entre homens não negros (9,1). No caso das mulheres negras idosas, esse índice foi de 5,9, contrastando com 2,3 entre mulheres brancas (Cerqueira; Bueno, 2025). Esses dados não apenas evidenciam um envelhecimento marcado por desigualdades estruturais, mas também denunciam a permanência de práticas institucionais racializadas que operam sob a lógica da negligência, da desproteção e da medicalização compulsória de corpos considerados menos valiosos.

A situação das mulheres negras com deficiência, por sua vez, também expõe uma das formas de interseccionalidade de opressões. Inseridas de forma marginal no sistema educacional, com apenas 0,6% de participação entre estudantes negras no ensino superior, e submetidas a uma taxa de desemprego de 13,4%, essas mulheres experienciam a articulação de múltiplas formas de exclusão que envolvem apagamento cognitivo, precarização material e deslegitimação social (IBGE, 2019).

A face mais extrema dessa lógica de desvalorização da vida manifesta-se na letalidade estatal dirigida contra corpos interseccionalmente marcados por raça, deficiência e pobreza. Exemplos dessa dinâmica incluem o assassinato de Genivaldo de Jesus Santos — homem negro com esquizofrenia, asfixiado em uma viatura da Polícia Rodoviária Federal transformada em câmara de gás improvisada —, e os homicídios de Ruan Limão do Nascimento e Thiago Aparecido Duarte de Souza — ambos jovens negros com deficiência intelectual, baleados por policiais em ações marcadas por desproporcionalidade, violência e posterior tentativa de criminalização das vítimas.

Essas ocorrências não podem ser interpretadas como anomalias ou excessos isolados, mas sim como manifestações sintomáticas de um modelo de governamentalidade fundado na exclusão. Assim, diante da complexidade e da densidade dessas articulações estruturais, impõe-se a urgência de romper com abordagens analíticas fragmentárias por meio da adoção de uma concepção ampliada, crítica e integrada de interseccionalidade para a elaboração de diagnósticos socialmente situados e para o delineamento de respostas institucionais que confrontem, de maneira eficaz, os mecanismos de dominação e exclusão social (Pacheco, 2022; Kyrillos, 2020).

### 3.3 MANIFESTAÇÕES DA DISCRIMINAÇÃO ALGORÍTMICA NA CONJUNTURA DAS MINORIAS SOCIAIS

A consolidação das tecnologias digitais tem revelado a emergência da discriminação algorítmica como uma renovação dos meios de reprodução das desigualdades estruturais, profundamente enraizadas no tecido histórico-social. Trata-se de um fenômeno cuja complexidade reside na intersecção entre técnica e poder, operando por meio de sistemas computacionais que ocultam a persistência de lógicas excludentes sob o disfarce de objetividade e racionalidade instrumental (Silva, 2020).

Esses sistemas são concebidos sob uma lógica tecnocrática orientada por parâmetros de otimização, escalabilidade e padronização, o que implica a supressão das particularidades contextuais dos sujeitos e fenômenos sociais. Nesse processo, indivíduos deixam de ser reconhecidos em sua singularidade histórico-relacional e passam a ser representados por abstrações estatísticas, como perfis de risco ou vetores preditivos em detrimento das singularidades históricas, contextuais e relacionais dos sujeitos que são objeto de processamento (Eubanks, 2018).

Esse deslocamento epistemológico e normativo implica consequências significativas para a arquitetura dos direitos fundamentais, notadamente no que tange à proteção da dignidade humana, que passa a ser reconvertida em uma métrica funcional, desprovida de densidade ética e valor substancial. O processo de metricização enseja a reconfiguração, agora sob lógicas algorítmicas, de hierarquias sociais historicamente instituídas, reproduzidas com maior legitimidade, escala e opacidade operacional (Marmelstein, 2021).

Dentre as contribuições mais relevantes nesse campo, destaca-se a sistematização proposta por Mendes, Mattiuzzo e Fujimoto (2021), que delineiam quatro dimensões interdependentes da discriminação algorítmica, articulando desde os pressupostos estatísticos que sustentam os processos de modelagem até os impactos concretos que as decisões automatizadas produzem sobre sujeitos e coletividades.

A primeira dimensão refere-se às limitações inferenciais dos modelos preditivos, que operam a partir da identificação de padrões em dados históricos. Esses modelos, concebidos com base em lógicas probabilísticas de previsão comportamental, desconsideram as distinções

fundamentais entre correlação e causalidade e, ao fazê-lo, incorrem na reprodução acrítica de regularidades empíricas contaminadas por processos de exclusão pretéritos. Os algoritmos, ao internalizarem essas distorções como parâmetros operacionais, automatizam desigualdades ao mesmo tempo em que as recobrem com a legitimidade da racionalidade técnica, esvaziando sua dimensão política e histórica.

A segunda dimensão trata da incorporação de atributos sensíveis, ou de suas representações estatisticamente correlacionadas, sem a devida mediação crítica, o que potencializa a perpetuação de hierarquias estruturais. Ainda que variáveis como raça, gênero, classe ou território não estejam explicitamente contempladas nos modelos, seu efeito discriminatório é reproduzido por meio de *proxies* que capturam tais marcadores de forma indireta. Esse processo resulta em uma discriminação latente, automatizada e opaca, cuja responsabilização jurídica e institucional é dificultada pela camuflagem técnica do viés subjacente (Silva, 2020; Marmelstein, 2021).

A terceira dimensão diz respeito à generalização empírica não contextualizada de inferências estatísticas, ou seja, à aplicação indiscriminada de padrões observados em universos amostrais específicos a populações distintas, ignorando sua diversidade sociocultural, histórica e institucional. Essa transposição acrítica de regularidades estatísticas assume a forma de normatividade algorítmica, por meio da qual a complexidade dos sujeitos sociais é reduzida à abstração quantitativa, comprometendo, assim, a equidade das decisões automatizadas e a justiça distributiva (Eubanks, 2018; Marmelstein, 2021).

A quarta e última dimensão refere-se à opacidade estrutural que caracteriza certos sistemas algorítmicos. Essa ausência de transparência impossibilita o acesso público às lógicas decisórias, fragiliza os mecanismos de auditoria e impede o contraditório, configurando um ambiente de ininteligibilidade (Mendes; Mattiuzzo; Fujimoto, 2020; Marmelstein, 2021).

Casos concretos evidenciam como a discriminação algorítmica se manifesta na contemporaneidade. Em 2009, uma falha relevante em um sistema de rastreamento facial desenvolvido pela empresa HP chamou a atenção para os riscos de viés algorítmico. O dispositivo, embutido em *webcams* comerciais, apresentou desempenho satisfatório na detecção de rostos brancos, mas falhou de maneira recorrente ao reconhecer rostos negros. Embora a empresa tenha inicialmente atribuído a falha a fatores externos, como iluminação,

investigações posteriores demonstraram que o modelo havia sido treinado com dados não representativos do espectro fenotípico humano (Wilson; Hoffman; Morgenstern, 2019).

As consequências dessa arquitetura algorítmica excludente também se estendem ao mercado de trabalho. Conforme demonstrado por Sumpter (2019), algoritmos utilizados em plataformas de anúncios de emprego tendem a favorecer homens em candidaturas a cargos de elevado capital econômico e simbólico, ao passo que mulheres, especialmente mulheres negras, são preteridas e, com frequência, alocadas em funções caracterizadas por precariedade estrutural, baixos salários e invisibilidade social, como operadoras de caixa, faxineiras ou motoristas de aplicativo (Ali *et al.*, 2019; Bogen, 2019).

Ademais, o caso emblemático da Amazon, ocorrido em 2014, ilustra os riscos éticos, sociais e institucionais implicados na implementação de sistemas algorítmicos em contextos sensíveis, como os processos seletivos. Ao desenvolver um modelo de triagem curricular fundamentado em dados passados de contratações, com a predominância de perfis masculinos em funções técnicas, a empresa concebeu um sistema que, ao invés de promover critérios meritocráticos, internalizou associações discriminatórias entre atributos de gênero e competência profissional. Mesmo após tentativas de mitigação dos vieses identificados, como a supressão de termos explícitos relacionados ao universo feminino, o algoritmo persistia em penalizar candidaturas femininas com base em inferências linguísticas implícitas, evidenciando a profundidade e a resiliência dos padrões discriminatórios incorporados ao modelo (Dastin, 2018; Reis; Graminho, 2019).

Essa problemática assume contornos ainda mais preocupantes quando essas tecnologias passam a incidir diretamente sobre decisões judiciais, como exemplifica o caso do software COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), amplamente utilizado nos Estados Unidos para a estimativa de risco de reincidência criminal. Fundamentado em variáveis de natureza sociodemográfica, comportamental e histórico-criminal, o sistema gera escores de risco que influenciam decisões como a dosimetria da pena, a aplicação de medidas cautelares e a concessão de benefícios no âmbito da execução penal (O'Neil, 2020).

Estudos empíricos de ampla repercussão pública, como o conduzido pela agência ProPublica, revelam que o COMPAS atua como vetor de reprodução e intensificação de desigualdades raciais. A análise de cerca de sete mil casos no estado da Flórida demonstrou

que indivíduos negros apresentavam uma probabilidade 45% superior de serem classificados equivocadamente como de alto risco de reincidência, quando comparados a indivíduos brancos em condições equivalentes e sem antecedentes penais relevantes (Angwin, 2016). Corroborando esses achados, a pesquisa conduzida por Caliskan-Islam, Bryson e Narayanan (2017) demonstrou que algoritmos empregados em contextos jurídicos tendem a associar nomes de origem africana a penalidades mais severas.

A culminância dessa problemática manifesta-se no caso *Loomis v. Wisconsin*. Nesse julgamento, o réu foi sentenciado com base em um relatório algorítmico cujos critérios de cálculo, lógica de ponderação e fundamentos operacionais foram integralmente vedados à sua defesa, resultando na violação substancial das garantias do devido processo legal, notadamente do contraditório e da ampla defesa, e impossibilitando a contestação técnica da classificação imposta (Karns, 2023).

A incorporação progressiva dos Sistemas de Reconhecimento Facial (SRFs) às políticas de segurança pública no Brasil tem despertado preocupações de natureza jurídica, ética e política, sobretudo no que se refere à sua aderência aos princípios constitucionais e axiológicos que estruturam o Estado Democrático de Direito. A adoção desses sistemas ocorre em um contexto histórico-social atravessado por desigualdades estruturais profundas, hierarquias raciais persistentes e práticas seletivas enraizadas nas instituições de controle social, particularmente a população negra e periférica, que passa a figurar como principal alvo da vigilância automatizada e da intervenção estatal punitiva (Silva, 2021; Lisboa, 2024).

Sob o prisma técnico-operacional, os SRFs consistem em sistemas que realizam a captação, vetorização e análise de dados biométricos faciais, os quais são posteriormente confrontados com imagens armazenadas em bancos de dados oriundos de redes de videomonitoramento. Embora demonstrem elevados índices de acurácia em ambientes laboratoriais controlados, sua eficácia decai consideravelmente quando aplicados em contextos reais, onde imperam variabilidade ambiental, diversidade fenotípica e instabilidade operacional. Elementos fisiológicos, como o envelhecimento e as alterações hormonais, e fatores ambientais, como ângulos desfavoráveis de captura, iluminação precária e baixa resolução de imagem, impactam diretamente os resultados, elevando as taxas de erro, em especial os falsos positivos, que podem culminar em abordagens policiais injustificadas e outras formas de violação de direitos (Hora, 2020).

A situação torna-se crítica quando os SRFs são empregados com finalidades preditivas, orientadas à antecipação de condutas consideradas potencialmente delitivas com base em modelos estatísticos e inferências probabilísticas. Esse paradigma representa um deslocamento epistemológico no campo penal, ao permitir que medidas de restrição de direitos sejam impostas não com base em atos concretamente verificados, mas em projeções abstratas de periculosidade de modo a subverter o princípio da presunção de inocência e comprometer o arcabouço garantista do sistema jurídico brasileiro, configurando uma violação manifesta à Constituição Federal e aos tratados internacionais de proteção dos direitos humanos aos quais o Brasil se vincula (Ferrajoli, 1995).

Ademais, a ausência de garantias mínimas de transparência, auditabilidade e responsabilização agrava esse quadro, revelando um descompasso entre a complexidade técnica desses sistemas e a capacidade institucional de regulá-los (Valente, 2022). A ausência de uma regulamentação jurídica específica, aliada à inexistência de mecanismos autônomos e efetivos de fiscalização, controle e auditoria, tem favorecido a consolidação de um modelo de governança securitária assentado em infraestruturas algorítmicas opacas, bancos de dados de procedência indefinida e mecanismos decisórios automatizados. (Silva, 2021; Lisboa, 2024). Como resultado, observa-se a proliferação de práticas coercitivas fundadas em evidências de baixa confiabilidade técnica, como abordagens discriminatórias, prisões arbitrárias e violações ao direito à liberdade (Karns, 2023; Pires *et al.*, 2021).

Conforme sustenta Silva (2021), a adoção desses dispositivos tecnológicos se ancora em três vetores estruturais que se inter-relacionam de modo sistemático. O primeiro refere-se à perpetuação de práticas repressivas de matriz colonial, que historicamente associam a negritude à periculosidade e ao desvio. O segundo eixo diz respeito à criminalização seletiva da pobreza, operada mediante procedimentos de coleta biométrica assimétricos e racializados, os quais privilegiam territórios e sujeitos já marcados pela vulnerabilidade social como alvos prioritários da vigilância algorítmica. O terceiro vetor, por sua vez, consiste na institucionalização de dispositivos tecnicamente imprecisos, cujas margens de erro, frequentemente negligenciadas, incidem de maneira desproporcional sobre indivíduos racializados, perpetuando ciclos de estigmatização, exclusão e encarceramento.

As evidências empíricas disponíveis reforçam os riscos já apontados. Dados obtidos entre março e outubro de 2019 indicam que 90,5% das pessoas identificadas por meio de tecnologias de reconhecimento facial em quatro unidades federativas eram negras,

evidenciando um padrão de seletividade racial que desmente qualquer pretensão de neutralidade algorítmica e expõe a perpetuação de dinâmicas discriminatórias sob roupagens tecnológicas (Nunes, 2019).

Nesse contexto, ganha relevo a atuação do projeto *O Panóptico*, iniciativa do Centro de Estudos de Segurança e Cidadania (CESeC), desenvolvido desde 2019, que se dedica à análise crítica das implicações ético-políticas da vigilância digital no Brasil. Segundo levantamento, há, atualmente, 337 projetos de reconhecimento facial em diferentes fases de implementação no país, com forte predominância no setor da segurança pública e foco na identificação automatizada de pessoas com pendências judiciais. (De França Netto et. al, 2025; Lisboa, 2024).

Desde o ano de 2019, o estado da Bahia passou a adotar, de forma sistemática, sistemas de reconhecimento facial em aproximadamente 80 municípios. Essa iniciativa tecnológica, frequentemente justificada sob a retórica da modernização das práticas de segurança pública, resultou, até maio de 2025, na prisão de 3.291 indivíduos. Todavia, uma análise mais detida da composição demográfica das pessoas detidas revela um dado alarmante: 90,5% dos indivíduos identificados e posteriormente encarcerados eram pessoas negras (IBGE, 2018; CNN Brasil, 2025).

Em 2024, a Polícia Militar do Estado do Rio de Janeiro alocou R\$ 18 milhões à ampliação do sistema de videomonitoramento integrado ao Centro de Operações Rio (COR), sustentando tal investimento sob a justificativa de aprimoramento da eficiência na execução de mandados judiciais. Não obstante, uma análise crítica dos dados empíricos disponíveis evidencia um descompasso substancial entre a retórica institucional de eficácia e os resultados efetivamente observados: do total aproximado de 3 milhões de imagens analisadas, apenas 0,082% resultaram em correspondências consideradas válidas. Em operações específicas, os índices de imprecisão ultrapassaram 60%, comprometendo não apenas a fidedignidade técnica do sistema, mas também sua legitimidade enquanto política pública orientada por critérios de legalidade, economicidade e respeito aos direitos fundamentais (SEPM-RJ, 2024; Nunes, 2022).

Paralelamente, na cidade de São Paulo, observa-se a consolidação de uma sofisticada infraestrutura sociotécnica de vigilância contínua, articulada principalmente por meio do programa *Smart Sampa*. A iniciativa prevê a instalação de mais de 25 mil câmeras equipadas

com tecnologias avançadas, como reconhecimento facial, leitura automatizada de placas veiculares e análise preditiva de comportamentos, todas integradas a sistemas de inteligência artificial (Dall'agnol, 2025). A intensificação dessa lógica se manifesta em iniciativas complementares, a exemplo do projeto *Muralha Paulista*, implementado nas dependências do estádio Allianz Parque, em que o acesso do público passa a ser mediado por mecanismos de verificação biométrica automatizada (PALMEIRAS ajuda..., 2023).

Casos têm exposto os efeitos concretos e potencialmente lesivos dessa política de segurança orientada por algoritmos. Em Aracaju, durante o pré-carnaval de 2023, uma mulher negra foi equivocadamente identificada como foragida da Justiça, sendo abordada e submetida, em duas ocasiões no mesmo dia, à contenção física e à exposição pública sem respaldo probatório (Nunes, 2024). De modo similar, na final do Campeonato Sergipano de 2024, um espectador foi detido de forma indevida em razão de erro de reconhecimento facial, o que culminou na suspensão emergencial da tecnologia por parte do governo estadual (SERGIPE suspende..., 2025).

A ausência de um marco regulatório específico que discipline o tratamento de dados biométricos no âmbito da segurança pública intensifica a complexidade do problema. Apesar do reconhecimento, pela Lei Geral de Proteção de Dados Pessoais (Lei nº 13.709/2018), da sensibilidade intrínseca dos dados biométricos, o artigo 4º, inciso III, da referida norma exclui de sua incidência as operações de tratamento destinadas a fins de segurança pública, remetendo sua regulamentação a instrumentos infralegais que, até o presente momento, permanecem não editados. Essa lacuna normativa configura um espaço de ampla discricionariedade estatal, no qual decisões de alto impacto sobre liberdades civis são tomadas sem observância de parâmetros fundamentais como legalidade, necessidade, proporcionalidade e finalidade (Silva, 2021).

A indefinição legislativa tem repercussões diretas sobre os modelos de governança informacional adotados por entes públicos e privados, permitindo a disseminação de práticas tecnológicas desprovidas de mecanismos de controle. Ilustra-se esse quadro pelo caso da concessionária ViaQuatro, operadora da Linha 4-Amarela do metrô de São Paulo, responsabilizada pela coleta não consentida de expressões faciais de usuários do transporte coletivo, com o intuito de aferir reações emocionais a anúncios publicitários. A ausência de consentimento informado, somada à opacidade do processo e à inexistência de respaldo legal

claro, evidencia a precariedade dos dispositivos de proteção de dados pessoais em ambientes urbanos crescentemente hipercontrolados (Neoway, 2021).

Outro caso ilustrativo ocorreu no município do Recife, com o anúncio da instalação de 108 relógios digitais equipados com câmeras de reconhecimento facial, sensores térmicos e conectividade à internet por meio de rede pública de Wi-Fi. Embora a iniciativa tenha sido justificada oficialmente pela necessidade de modernizar a infraestrutura urbana e reforçar as políticas de segurança pública, a medida foi duramente criticada por organizações da sociedade civil. Em especial, o Instituto de Pesquisa em Direito e Tecnologia do Recife (IP.rec), em articulação com outras entidades acadêmicas e jurídicas, publicou uma carta aberta denunciando possíveis violações à Lei Geral de Proteção de Dados Pessoais (LGPD). A crítica se baseava na ausência de diretrizes claras e auditáveis de segurança da informação, bem como nos riscos estruturais à fruição plena de direitos civis fundamentais, como o direito à reunião, à manifestação e à livre circulação no espaço público (Carta, 2021).

A incorporação de práticas decisórias automatizadas nas instâncias do aparato estatal, quando desvinculada de arcabouços normativos estruturados em princípios constitucionais e garantias democráticas, compromete de modo substantivo a coerência e a legitimidade da arquitetura institucional que sustenta o Estado de Direito. Nesse horizonte, Mendes e Mattiuzzo (2020) argumentam que a implementação acrítica de sistemas algorítmicos, destituída de regulação normativa robusta e transparentemente auditável, engendra um processo gradual e silencioso de erosão das salvaguardas fundamentais, de modo que converte o juízo político e ético, indispensável à função pública, em um procedimento técnico-instrumental submetido aos imperativos da eficiência e da operatividade, alheio a qualquer mediação crítica ou controle democrático.

Essa racionalidade tecnocrática, ao reconfigurar a ação estatal a partir de critérios que prescindem de deliberação moral e responsabilidade institucional, instaura modalidades de governança automatizada cujo funcionamento se dá sob a aparência de neutralidade, mas cujas implicações são estruturalmente assimétricas e normativamente opacas. Neste cenário, a reflexão de Hannah Arendt (1999) sobre a banalização do mal revela-se teórica e politicamente fecunda, ao problematizar os perigos associados à normalização de condutas administrativas privadas de densidade ética. A dissociação entre responsabilidade moral e ação governamental legitima dispositivos de controle que, embora formalmente conformes à legalidade vigente ou sustentados por racionalidades técnico-científicas, resultam

flagrantemente incompatíveis com os fundamentos normativos e axiológicos que alicerçam a democracia constitucional.

#### **4 A IMPERIOSIDADE DE ASSEGURAR OS DIREITOS FUNDAMENTAIS DAS MINORIAS NA ERA ALGORÍTMICA**

A consolidação da chamada era algorítmica configura uma ruptura paradigmática em relação aos pressupostos clássicos da dogmática jurídica, exigindo uma reavaliação crítica dos fundamentos normativos do Estado Constitucional de Direito, sobretudo diante da crescente autonomia funcional dos sistemas algorítmicos, da opacidade de suas arquiteturas e da complexidade das lógicas inferenciais que os estruturam. Sua dinâmica, muitas vezes

obscurecida pela aparência de racionalidade matemática, compromete a efetivação de compromissos constitucionais fundantes, como a igualdade substancial, a justiça distributiva e a dignidade da pessoa humana enquanto núcleo axiológico do ordenamento (Piovesan, 2008; Sarlet, 2003).

Neste contexto de crescente complexificação técnico-normativa, revela-se insuficiente qualquer tentativa de tutela dos direitos fundamentais que se restrinja a respostas hermenêuticas meramente reativas ou à aplicação mecânica de dispositivos legais concebidos para cenários tecnológicos já superados. Torna-se imperativa, portanto, a edificação de um paradigma jurídico-normativo à IA, fundado e tecnicamente apto a articular, de maneira transversal, os princípios estruturantes da democracia constitucional aos dilemas éticos, operacionais e políticos que emergem da governança algorítmica.

#### 4.1 A CORRELAÇÃO ENTRE TRANSPARÊNCIA ALGORÍTMICA E A TUTELA DOS DIREITOS FUNDAMENTAIS DAS MINORIAS

A promulgação da Constituição da República Federativa do Brasil, em 1988, representa um ponto de inflexão histórico-normativo que redefiniu os contornos do ordenamento jurídico nacional, consagrando a ruptura definitiva com o anterior regime autoritário e instituindo os fundamentos de um Estado Democrático de Direito. Trata-se de um marco normativo de matriz garantista, que inaugura uma nova racionalidade jurídico-constitucional fundada em princípios estruturantes, como a dignidade da pessoa humana, a justiça social e a igualdade material, os quais passam a exercer função hermenêutica central na interpretação e na aplicação do direito constitucional brasileiro (Sarlet, 2002; Piovesan, 2008).

Desde sua redação preambular, a Constituição de 1988 delinea um projeto político-jurídico de natureza emancipatória, comprometido com a superação das desigualdades históricas e com a reconstrução das bases institucionais da sociedade brasileira em moldes democráticos e inclusivos. Esse projeto adquire densidade jurídica nos artigos 1º, 3º e 5º, que consagram, de modo articulado, os fundamentos do Estado, os objetivos constitucionais prioritários e um extenso rol de direitos e garantias fundamentais, elevando a dignidade da pessoa humana à condição de valor jurídico absoluto, cuja força normativa

possui caráter irradiante e transversal em relação a todo o sistema constitucional (Piovesan, 2008).

No interior dessa arquitetura normativa, os direitos fundamentais assumem caráter multifuncional, operando como instrumentos simultâneos de limitação do poder estatal e de transformação social, natureza essa particularmente evidenciada pela constitucionalização dos direitos sociais, como os direitos à saúde, à educação, à moradia, ao trabalho, à segurança, entre outros direitos, os quais refletem o compromisso do constituinte originário com a justiça distributiva e a inclusão socioeconômica. Esses direitos, ademais, apresentam eficácia jurídica plena, imediata e vinculante, sendo caracterizados por atributos como universalidade, indivisibilidade, interdependência, inalienabilidade e imprescritibilidade (Diógenes Júnior, 2012; Sarlet, 2003).

Nesse contexto, o princípio da vedação ao retrocesso social emerge como cláusula de proteção normativa de matriz material, cuja finalidade precípua consiste em impedir a supressão ou mitigação de direitos fundamentais já positivados e consolidados. Esse princípio atua como salvaguarda contra eventuais regressões normativas motivadas por conjunturas adversas — sejam elas de ordem econômica, institucional ou política —, operando como mecanismo de estabilização constitucional e de garantia da perenidade dos avanços civilizatórios incorporados pelo texto constitucional (Costa, 2013).

Com o intuito de assegurar a inviolabilidade do núcleo axiológico do pacto constitucional, a Constituição de 1988 estrutura-se de forma a incorporar mecanismos internos de autodefesa normativa, os quais conferem ao texto constitucional elevado grau de rigidez formal e substancial. Nesse sentido, o §1º do artigo 5º, em consonância com as cláusulas pétreas previstas no §4º do artigo 60, estabelece limites materiais e procedimentais ao exercício do poder de reforma, resguardando a intangibilidade dos direitos e garantias fundamentais frente a eventuais tentativas de erosão normativa por meio do processo legislativo de emenda (Sarlet, 2002; Ramos, 2019).

Essa estrutura protetiva interna projeta-se para além das fronteiras estatais, sendo substancialmente reforçada pela inserção do Brasil em uma malha normativa internacional voltada à tutela dos direitos humanos com a adesão à Organização das Nações Unidas e a promulgação da Declaração Universal dos Direitos Humanos, em 1948. Esses marcos históricos inauguraram o Sistema Global de Proteção Internacional dos Direitos Humanos,

cuja lógica transnacional tem por escopo o reconhecimento da dignidade humana como valor universal e inderrogável (Leal; Lima, 2021).

No âmbito regional, essa integração normativa encontra expressão concreta na adesão do Brasil ao Sistema Interamericano de Proteção dos Direitos Humanos, consolidada pela ratificação da Convenção Americana sobre Direitos Humanos (CADH), em 1992, e pela aceitação da jurisdição contenciosa da Corte Interamericana de Direitos Humanos, formalizada pelo Decreto nº 4.463/2002 (Leal; Vargas, 2022, p. 1342). A partir dessa vinculação, o Estado brasileiro assume compromissos jurídicos de natureza internacional, os quais impõem a necessidade de conformação do ordenamento jurídico interno às disposições convencionais. O artigo 1º da CADH consagra o dever estatal de respeitar e garantir os direitos nela previstos, obrigação reafirmada pelo princípio do *pacta sunt servanda*, consagrado no artigo 26 da Convenção de Viena sobre o Direito dos Tratados.

Embora direitos fundamentais e direitos humanos compartilhem um mesmo fundamento axiológico — assentado nos valores universais da dignidade, liberdade e igualdade da pessoa humana —, distinguem-se quanto à origem normativa e à forma de positivação. Os direitos fundamentais têm gênese na ordem constitucional interna e operam como limites jurídico-institucionais ao exercício do poder estatal, enquanto os direitos humanos são consagrados em instrumentos internacionais de caráter vinculante, com vocação universalista. No entanto, essa distinção não configura oposição, mas sim uma relação de complementaridade normativa. A própria Constituição de 1988, ao estabelecer nos §§1º a 3º do artigo 5º os mecanismos de recepção e hierarquização dos tratados internacionais de direitos humanos, evidencia a abertura do ordenamento jurídico brasileiro à integração entre esferas normativas distintas. (Leal; Lima, 2021).

Nesse cenário, a dignidade da pessoa humana, consagrada no artigo 1º, inciso III, eleva-se à condição de princípio estruturante da ordem constitucional, dotado de máxima densidade normativa, plena eficácia e imediata aplicabilidade. Ancorada na tradição filosófica kantiana, essa concepção moderna de dignidade repousa na autonomia moral do indivíduo, que deve ser tratado como um fim em si mesmo e nunca como meio instrumental para fins alheios à sua subjetividade (Kant, 2003; Fraser; Honneth, 2003).

A partir dessa matriz valorativa, o princípio da igualdade é ressignificado em termos substanciais, ultrapassando a leitura estritamente formalista e abstrata oriunda do paradigma

liberal clássico. Nesse novo horizonte hermenêutico, a igualdade deixa de ser concebida como simples isonomia legal para assumir o papel de instrumento jurídico-político voltado à realização da justiça social, impondo ao Estado Democrático de Direito a adoção de condutas ativas e comprometidas com a superação das desigualdades materiais historicamente produzidas, especialmente aquelas associadas a estruturas de exclusão sistemática (Sarlet, 2002).

Nesse sentido, a leitura tridimensional do princípio da igualdade, conforme desenvolvida por Barroso (2011), aprofunda o debate ao distinguir três dimensões interdependentes: (i) a formal, que assegura o tratamento jurídico isonômico perante a lei; (ii) a material, que justifica e fundamenta ações afirmativas voltadas à eliminação de desigualdades estruturais; e (iii) a do reconhecimento, orientada à valorização das identidades, experiências e trajetórias historicamente marginalizadas. Tal concepção robustecida da igualdade impõe ao Estado obrigações normativas de natureza positiva, compatíveis com os compromissos internacionais assumidos pelo Brasil, como a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher (CEDAW, 1979) e a Convenção nº 111 da Organização Internacional do Trabalho (Kaufman; Junquillo; Reis, 2023).

Sob esse novo paradigma hermenêutico, a igualdade exerce dupla função: como cláusula antidiscriminatória de aplicação transversal a todo o ordenamento jurídico e como fundamento normativo da justiça distributiva, servindo de parâmetro para a formulação e avaliação das políticas públicas, de acordo com os princípios da proporcionalidade, da razoabilidade e da efetividade (Ávila, 2018). Essa inflexão é crucial diante da persistência de estruturas de exclusão resultantes de legados históricos como o colonialismo, a escravidão e a concentração sistêmica de poder e riqueza (Bandeira de Mello, 1998; Miranda, 2014).

Essa reconfiguração do princípio da igualdade encontra fundamento teórico na abordagem das capacidades, formulada por Amartya Sen (2010), cuja matriz crítica desloca o foco da justiça social da mera distribuição de bens para a efetiva ampliação das liberdades reais. Nessa perspectiva, a igualdade não se esgota na igualdade de oportunidades, mas pressupõe a criação de condições materiais concretas que assegurem aos indivíduos o pleno desenvolvimento de suas autonomias e a realização de seus projetos de vida em contextos sociais marcados por dignidade.

A partir desse horizonte teórico-normativo, Herrera Flores (2009) defende uma hermenêutica jurídica de cunho emancipatório, centrada na função contra-hegemônica do direito e comprometida com a transformação das estruturas opressivas do *status quo*. Tal abordagem impõe o abandono das leituras dogmáticas, reducionistas e despolitizadas do direito, demandando um engajamento epistemológico que articule teoria crítica, práxis jurídica e emancipação social. Nesse sentido, Häberle (2021) avança ao propor uma concepção ampliada da Constituição como artefato jurídico-cultural em contínuo processo de reconstrução dialógica, cuja normatividade resulta da interação plural e dinâmica entre os múltiplos sujeitos constituintes.

A despeito do avanço institucional representado pela incorporação de direitos fundamentais no texto constitucional de 1988, a persistência de desigualdades estruturais evidencia os limites materiais do projeto emancipatório constitucional. A dissonância entre a normatividade dos direitos proclamados e sua efetivação concreta no plano empírico revela a insuficiência das abordagens estritamente formalistas e impõe a adoção de uma hermenêutica constitucional crítica, fundada em coordenadas histórico-sociais e sensível às engrenagens de poder que estruturam o campo jurídico (Barbosa, 2021).

Esse descompasso é acentuado pela progressiva digitalização das esferas públicas e privadas, contexto no qual sistemas de inteligência artificial (IA) têm sido incorporados a processos decisórios em setores sensíveis, como justiça, saúde, educação e finanças. A emergência de uma “normatividade técnica”, oriunda de arquiteturas algorítmicas caracterizadas por opacidade estrutural, elevada complexidade e autonomia funcional, desafia as estruturas tradicionais de controle democrático, publicidade dos atos e garantias processuais (Barocas; Selbst, 2016).

Nesse cenário, a discriminação algorítmica desponta como uma das manifestações mais insidiosas da reprodução sistêmica de injustiças, exigindo uma reconfiguração normativa orientada por uma perspectiva interseccional, capaz de captar a sobreposição e a interação entre diferentes marcadores sociais de vulnerabilidade — como gênero, raça, classe, etnia e orientação sexual — e sustentar, a partir disso, uma teoria constitucional pluralista, comprometida com a redistribuição equitativa de recursos e o reconhecimento substantivo das identidades historicamente subalternizadas ( Piovesan, 2008; Ramos, 2019).

Conforme argumenta Escobar Roca (2018), os riscos emergentes à dignidade humana no contexto digital impõem a urgente reinterpretação e reconfiguração das categorias clássicas dos direitos fundamentais, à luz das profundas transformações introduzidas pela racionalidade algorítmica e pela arquitetura informacional contemporânea. O direito à explicabilidade algorítmica emerge, assim, como dispositivo jurídico fundamental para a viabilização do controle público, do contraditório e da responsabilização jurídica em decisões automatizadas que impactam diretamente os direitos fundamentais, o que pressupõe a constituição de um arcabouço jurídico sustentado em princípios como a transparência procedimental, a auditabilidade dos sistemas, a rastreabilidade das decisões automatizadas e a responsabilização civil, administrativa e penal pelos seus efeitos (Rodotà, 2014; Leite, 2022).

Ademais, a efetividade desse novo conjunto de direitos pressupõe uma abordagem co-responsável, na qual não apenas o Estado, mas também os entes privados — sobretudo as grandes plataformas digitais e corporações tecnológicas — sejam reconhecidos como agentes normativos relevantes na conformação das práticas sociais e da governança informacional (Barrilao, 2023).

Não obstante os avanços representados por marcos normativos como o Marco Civil da Internet (Lei nº 12.965/2014) e a Lei Geral de Proteção de Dados Pessoais (Lei nº 13.709/2018), o ordenamento jurídico brasileiro ainda carece de um regime específico e suficientemente articulado para regular os impactos da inteligência artificial sobre os direitos fundamentais. A ausência de parâmetros normativos que assegurem supervisão humana, critérios claros de governança algorítmica e mecanismos eficazes de responsabilização contribui para a perpetuação de práticas discriminatórias e para o agravamento das desigualdades estruturais (Doneda, 2019).

Nesse contexto, insere-se o Projeto de Lei nº 2.338/2023, atualmente em tramitação no Senado Federal, que propõe a formulação de um regime jurídico autônomo e especializado para a regulação da IA no Brasil. Estruturado em torno de princípios como equidade, inclusão, segurança, transparência e supervisão humana, o projeto objetiva compatibilizar a promoção do desenvolvimento tecnológico com a salvaguarda substancial dos direitos fundamentais, articulando-se ao movimento de constitucionalização da regulação digital. Além disso, estabelece diálogo com iniciativas correlatas, como o Projeto de Lei nº 21/2020, que trata do uso da IA no setor privado, e a Resolução nº 332/2020 do Conselho Nacional de Justiça, que regulamenta sua utilização no âmbito do Poder Judiciário.

A arquitetura regulatória proposta no PL 2.338/2023 adota como eixo estruturante a lógica da gestão de riscos, classificando os sistemas de IA de acordo com seu potencial impacto sobre direitos e liberdades fundamentais. Essa tipologia normativa permite a imposição de obrigações jurídicas proporcionais ao grau de risco, variando entre exigências de transparência, segurança e governança, até a proibição do uso de tecnologias consideradas desproporcionalmente danosas (Barrilao, 2023).

Com vistas à mitigação da opacidade que caracteriza os processos decisórios automatizados, o Capítulo II do projeto consagra um conjunto de garantias fundamentais voltadas à proteção dos titulares de dados. O artigo 5º estabelece, entre outros, os direitos à informação prévia sobre a utilização da IA, ao acesso a explicações claras, inteligíveis e fundamentadas acerca das decisões automatizadas, bem como à revisão dessas decisões sempre que gerarem efeitos jurídicos relevantes. De forma complementar, o artigo 7º impõe obrigações de transparência ativa aos desenvolvedores e operadores dos sistemas, exigindo a identificação da natureza automatizada das interações, a explicitação da lógica subjacente às decisões, a exposição dos riscos previsíveis e a indicação dos operadores humanos incumbidos da supervisão das operações algorítmicas.

Embora o princípio da dignidade da pessoa humana não esteja explicitamente incorporado ao corpo do texto legal, sua centralidade axiológica é reafirmada na Exposição de Motivos, que evidencia a necessidade de humanização das relações sociais mediadas por tecnologias automatizadas. A proposta normativa estrutura-se, portanto, em duas dimensões interdependentes: de um lado, a proteção jurídico-subjetiva do indivíduo frente à opacidade e à complexidade das decisões automatizadas; de outro, a conformação de um marco regulatório tecnicamente sólido, juridicamente previsível e institucionalmente confiável, apto a fomentar a inovação tecnológica sem sacrificar os postulados constitucionais.

Nessa lógica, o direito à explicação, previsto no artigo 8º, emerge como elemento normativo central da autodeterminação informacional, compreendido em duas dimensões complementares: compreensibilidade, que visa garantir a inteligibilidade das decisões algorítmicas ao titular dos dados, e contestabilidade, que assegura a possibilidade de revisão, oposição ou reparação em face dos efeitos dessas decisões. Para conferir efetividade a esse direito, o projeto estabelece critérios normativos objetivos, relativos à gratuidade, à clareza linguística e à fixação de prazos razoáveis de resposta, reconhecendo que a simples

publicização de dados técnicos não é, por si só, suficiente para assegurar uma transparência genuína.

No âmbito da governança algorítmica, os artigos 19 a 21 instituem um regime regulatório transversal aplicável a agentes públicos e privados. Entre as obrigações previstas, destacam-se: a adoção de programas internos de conformidade regulatória, a documentação sistemática das operações, o desenvolvimento de interfaces acessíveis e, no caso de sistemas considerados de alto risco, a implementação de mecanismos que viabilizem a explicabilidade dos resultados produzidos, observando-se, para tanto, os limites impostos pela proteção à propriedade intelectual.

No setor público, o artigo 21 introduz salvaguardas adicionais ao determinar que decisões automatizadas com efeitos jurídicos relevantes ou impactos substanciais sobre direitos individuais estejam sujeitas à possibilidade de revisão humana. Essa exigência incorpora os princípios de supervisão humana (*human oversight*) e centralidade do sujeito (*human centrality*), conforme delineado nas diretrizes normativas da União Europeia (European Commission, 2021). Contudo, a redação genérica do inciso IV do artigo 5º — ao se referir de forma indeterminada à “intervenção humana” — carece de precisão conceitual quanto à natureza, intensidade e alcance dessa intervenção, suscitando dificuldades interpretativas que podem comprometer sua aplicabilidade em contextos tecnicamente complexos ou institucionalmente sensíveis.

Por fim, o projeto contempla a criação de canais acessíveis, eficazes e transparentes para o exercício dos direitos informacionais pelos titulares de dados, como estratégia para mitigar as assimetrias estruturais de poder e informação que permeiam a relação entre indivíduos e sistemas algorítmicos. Entretanto, persiste uma limitação relevante: a eficácia do direito à explicação permanece condicionada à iniciativa do próprio titular, o que restringe sua operacionalização ao modelo de transparência passiva. Torna-se, portanto, imprescindível o aprimoramento do marco regulatório mediante a adoção de mecanismos informacionais proativos, capazes de densificar esse direito e assegurar sua efetividade concreta, conforme propõe Mantelero (2022), consolidando uma concepção robusta, ativa e emancipatória da autodeterminação informacional na era digital.

## 4.2 O PAPEL DO CONTROLE PREVENTIVO-REPRESSIVO NA DEFESA DOS GRUPOS MINORITÁRIOS

A concretização dos direitos fundamentais de sujeitos historicamente marginalizados impõe ao Estado o dever de instituir dispositivos normativos e arquiteturas institucionais capazes de enfrentar, de maneira estruturada e multidimensional, as múltiplas expressões da discriminação de caráter estrutural. Nesse escopo, revela-se essencial a adoção de um modelo de controle preventivo-repressivo, não apenas para repressão de práticas excludentes cristalizadas nas engrenagens institucionais, mas também para a consolidação de uma cultura organizacional pautada na valorização da diversidade, na promoção da equidade e na efetivação de uma inclusão substancial, em consonância com os preceitos e valores fundantes insculpidos na Constituição Federal de 1988.

Assim, infere-se ao controle preventivo da discriminação algorítmica a importância de conceber uma governança algorítmica que vista disciplinar, de maneira sistêmica, o ciclo de vida das tecnologias digitais, desde sua concepção e desenvolvimento até os estágios de implementação, monitoramento e descontinuação. Diferentemente de modelos regulatórios meramente procedimentais, essa abordagem articula de forma complexa os domínios técnico, jurídico, político e econômico, conformando uma infraestrutura normativa orientada à mediação entre os imperativos da inovação tecnológica e os compromissos normativos com a justiça social e a salvaguarda dos princípios democráticos (Schmidt-Assmann apud Hartmann; Silva, 2019).

A eficácia dessa governança está intrinsecamente vinculada a sua capacidade de incorporar valores ético-normativos nas estruturas decisórias automatizadas e institucionalizar mecanismos de responsabilização moral, jurídica e técnica dos diversos atores públicos e privados envolvidos, exigindo a edificação de marcos regulatórios, respaldados por instituições dotadas de autonomia decisória, expertise técnica e legitimidade democrática, aptas a enfrentar os desafios decorrentes da opacidade algorítmica e das assimetrias de poder que estruturam a economia digital contemporânea.

Localizado no escopo geopolítico do Sul Global, o Brasil apresenta um panorama marcado por profundas assimetrias de caráter racial, de gênero, territorial e colonial, de modo que a inserção acrítica de tecnologias de inteligência artificial, frequentemente projetadas com base em epistemologias eurocentradas, tendem a reproduzir e intensificar essas desigualdades,

obscurando saberes locais e consolidando estruturas hegemônicas excludentes (AI Manifesto, 2021).

Diante desse cenário, a simples transposição de modelos regulatórios exógenos, desenvolvidos em contextos sociopolíticos, econômicos e institucionais radicalmente distintos, revela-se metodologicamente inadequada e politicamente arriscada. Do contrário, a formulação de arcabouços regulatórios eficazes requer, por conseguinte, a incorporação crítica e seletiva de referências internacionais, articulada à elaboração de instrumentos normativos ancorados nas especificidades histórico-culturais do país.

Essa diretriz se manifesta na proposta legislativa do Projeto de Lei nº 2338/2023, que adota uma arquitetura regulatória estruturada em quatro eixos: (i) delimitação precisa do contexto de aplicação; (ii) avaliação sistemática dos riscos; (iii) definição de estratégias adequadas de mitigação; e (iv) consolidação de mecanismos de governança contínua, com previsão de monitoramento e revisão periódica (Brasil, 2023), propondo uma estrutura regulatória escalonada em quatro fases, preparação, identificação e análise de riscos, mitigação e monitoramento contínuo, em conformidade com os parâmetros internacionais de boas práticas regulatórias (Brasil, 2023).

No mesmo sentido, Hood *et al.* (2001) argumentam em favor de arcabouços regulatórios que sejam, simultaneamente, flexíveis, iterativos e responsivos, de modo a assegurar sua adaptabilidade diante da constante reconfiguração tecnológica que caracteriza os sistemas de IA. Essa concepção encontra ressonância nas diretrizes internacionais formuladas pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que propõe uma abordagem regulatória orientada pela distinção funcional de responsabilidades ao longo do ciclo de vida dos sistemas algorítmicos, reconhecendo estrutura, com rigor, as assimetrias operacionais e decisórias entre os agentes envolvidos, ao promover a alocação proporcional de obrigações conforme o grau de intervenção de cada ator nos processos automatizados e segundo as especificidades contextuais e institucionais de aplicação (OECD, 2023c).

Contudo, a mera incorporação formal de diretrizes supranacionais não é suficiente para garantir a efetividade regulatória em contextos locais marcados por fragilidades institucionais. A construção de um arcabouço normativo robusto exige a constituição de capacidades institucionais autônomas e tecnicamente qualificadas, aptas a enfrentar os

desafios epistemológicos impostos pela opacidade algorítmica e pela assimetria informacional que atravessa toda a cadeia de concepção, desenvolvimento e implementação dos sistemas de IA (Barfield, 2018). Nesse sentido, o princípio da precaução configura-se como fundamento jurídico e epistêmico central à utilização de sistemas inteligentes, sobretudo diante da ausência de previsibilidade plena quanto aos impactos sociais e políticos das tecnologias emergentes, e da impossibilidade de se antever, com exatidão, os danos potenciais decorrentes de sua aplicação (Sunstein, 2005).

À luz dessa perspectiva, as abordagens regulatórias baseadas em risco emergem como instrumentos estratégicos de racionalização normativa, possibilitando a calibragem da intensidade regulatória em função do nível de criticidade das aplicações tecnológicas. Esse modelo combina instrumentos *ex ante*, destinados à antecipação e prevenção de danos, com mecanismos *ex post*, voltados à reparação e responsabilização, conformando uma estrutura regulatória adaptativa, responsiva e gradativa (Kaminski, 2021).

Nesse sentido, Mantelero (2022) propõe uma abordagem tripartida, contemplando: (i) a identificação precisa dos riscos; (ii) a estimativa da probabilidade de sua materialização; e (iii) a análise da gravidade de suas consequências. O autor defende a ampliação do escopo analítico, de modo a considerar os direitos fundamentais não apenas como elementos jurídicos, mas como categorias substantivas de risco, com vistas à proteção abrangente de indivíduos, grupos vulnerabilizados e do meio ambiente.

Simultaneamente, intensifica-se a exigência de transparência e auditabilidade nos sistemas de decisão automatizada, resultando na ascensão da Inteligência Artificial Explicável (*Explainable AI – XAI*) como paradigma técnico-jurídico voltado à inteligibilidade dos modelos algorítmicos. A XAI busca compatibilizar a acurácia, enquanto expressão da robustez inferencial dos algoritmos, com a interpretabilidade, compreendida como a capacidade de tornar acessíveis e compreensíveis os critérios decisórios a atores humanos. Ao fazê-lo, contribui significativamente para o fortalecimento da legitimidade institucional e para a promoção da responsabilização democrática na governança da IA (Dignum, 2019).

Nessa direção, impõe-se a formulação de uma arquitetura regulatória ancorada em três eixos fundamentais e interdependentes: (i) a antecipação crítica dos impactos sociotécnicos, políticos e econômicos decorrentes da adoção de tecnologias baseadas em IA; (ii) a incorporação normativa de princípios éticos e garantias constitucionais nos sistemas de

decisão automatizada; e (iii) a institucionalização de práticas de responsabilização moral e jurídica por parte dos agentes técnicos e institucionais envolvidos em sua concepção e operacionalização (Dignum, 2019). Essa arquitetura, por sua vez, deve articular-se a modelos de gestão de risco que orientem a intensidade da regulação conforme o grau de criticidade da aplicação tecnológica, promovendo, assim, uma governança algorítmica sensível ao contexto, proporcional em sua intervenção e comprometida com os fundamentos do Estado Democrático de Direito (Cowgill; Dell’Acqua; Durietz, 2021).

Ademais, a Avaliação de Impacto Algorítmico (AIA) se destaca como instrumento de regulação antecipatória, orientado por uma racionalidade preventiva e prospectiva. Inspirada nos modelos consolidados de avaliação de impacto ambiental, a AIA visa identificar, mitigar e, se necessário, remediar danos potenciais decorrentes da automação de processos decisórios antes de sua implementação efetiva. Sua eficácia, contudo, está condicionada à adoção de abordagens epistemologicamente interdisciplinares, que promovam uma articulação crítica entre os domínios das ciências exatas, sociais, humanas e jurídicas, transcendendo a função técnica ao assumir um papel analítico e político na problematização das assimetrias estruturais reproduzidas por sistemas algorítmicos. Depende, ademais, da qualidade informacional, da integridade epistêmica e da transparência das bases de dados utilizadas no treinamento e na operação dos sistemas automatizados (Frazão, 2021).

Iniciativas como relatórios de banco de dados e modelos de aprendizado de máquina e inteligência artificial também tornam-se centrais na consolidação de uma cultura de documentação técnica orientada à rastreabilidade, à auditabilidade e ao escrutínio público. A construção de bases de dados eticamente orientadas exige, simultaneamente, a inclusão ativa de populações historicamente marginalizadas e a exclusão de atributos que funcionem como *proxies* de discriminação estrutural. A literatura especializada recomenda, para tanto, o emprego de metodologias baseadas em inferência causal, calibração estatística e padronização de métricas de equidade como vetores técnicos para a mitigação de vieses algorítmicos (Kulshrestha *et al.*, 2017; Landeiro; Culotta, 2016; Zagheni; Weber, 2015).

Ademais, a verificação da aderência normativa desses sistemas pressupõe a institucionalização de mecanismos de auditoria técnica conduzidos por entidades independentes, dotadas de autonomia funcional e competência interdisciplinar, capazes de aferir a compatibilidade dos sistemas automatizados com os princípios da legalidade, da equidade e da transparência. Auditorias internas, no entanto, embora operacionalmente mais

acessíveis, estão frequentemente sujeitas a conflitos de interesse que fragilizam sua legitimidade e eficácia (Pipkin, 2020).

Correlativamente, observa-se uma crescente instrumentalização do discurso ético por parte de conglomerados tecnológicos multinacionais, os quais se apropriam de narrativas de diversidade, inclusão e responsabilidade social como estratégias de legitimação simbólica, desprovidas, contudo, de correspondência com práticas concretas de redistribuição de poder informacional. Essa dissonância entre retórica corporativa e ação institucional revela um processo de esvaziamento normativo da ética, que passa a ser mobilizada como recurso de branding reputacional, em vez de orientadora substantiva de condutas e políticas organizacionais (Birhane; Prabhu, 2021).

Em contraste, auditorias externas, conduzidas por especialistas humanos com sensibilidade contextual, demonstram maior capacidade de identificar vieses epistêmicos e deficiências de modelagem, sobretudo em realidades marcadas por desigualdade social, pluralismo linguístico e diversidade cultural (Caswell *et al.*, 2020; Paullada *et al.*, 2021).

Nesse sentido, é atribuída à Autoridade Nacional de Proteção de Dados (ANPD) um papel estratégico que ultrapasse os limites da fiscalização meramente procedimental. Cabe-lhe atuar como agente catalisador de uma justiça informacional de base estrutural, promovendo políticas públicas orientadas por uma matriz normativa atenta às desigualdades sociais e territoriais. A consolidação de uma governança algorítmica ética, democrática e inclusiva, assim, requer o fortalecimento de uma infraestrutura sociotécnica resiliente, capaz de reorganizar criticamente os fluxos de poder e conhecimento que atravessam o ecossistema digital.

No que se refere ao controle repressivo, a ausência de uma regulação específica sobre inteligência artificial no Brasil revela uma postura omissiva e conivente do Legislativo nacional, cuja inércia normativa favorece grandes conglomerados tecnológicos e agrava a desproteção dos cidadãos diante de sistemas automatizados opacos e potencialmente discriminatórios. Trata-se de um cenário de regulação frouxa, em que os riscos da automação decisória se distribuem de forma desigual, penalizando sobretudo grupos vulnerabilizados socialmente.

Embora o ordenamento jurídico brasileiro ainda não disponha de uma legislação específica que discipline diretamente a discriminação algorítmica, não se pode falar em um vácuo jurídico absoluto. O sistema normativo vigente, alicerçado na Constituição Federal de 1988, especialmente no princípio da igualdade, erigido à condição de cláusula pétrea, e complementado por normas infraconstitucionais e tratados internacionais ratificados, contém instrumentos com potencial para conter práticas discriminatórias decorrentes da automação decisória. Contudo, o fato de depender de interpretações extensivas e articulações normativas fragmentadas evidencia a fragilidade dessa proteção frente à crescente complexidade tecnológica.

Ademais, a responsabilização jurídica por discriminação algorítmica, embora viável à luz da doutrina da discriminação indireta, carece de efetividade concreta. A exigência de demonstração de efeitos excludentes, sem que se imponha o ônus de auditoria aos desenvolvedores e controladores desses sistemas, favorece a impunidade técnica, uma vez que a dificuldade de acesso à lógica algorítmica, marcada por sua baixa rastreabilidade e opacidade estrutural, compromete o próprio exercício do contraditório e da ampla defesa.

Nesse cenário, a postura do Estado brasileiro em relação a compromissos multilaterais de direitos humanos revela-se profundamente paradoxal. Embora o país figure como signatário de tratados internacionais de elevada densidade normativa, como a Declaração Universal dos Direitos Humanos, a Convenção sobre a Eliminação de Todas as Formas de Discriminação contra a Mulher e a Convenção n.º 111 da OIT entre outros dispositivos, sua atuação interna tem sido marcada por um formalismo inócuo, incapaz de transformar tais obrigações em políticas públicas e dispositivos normativos compatíveis com a complexidade das novas racionalidades técnicas, de modo que a dissonância entre o discurso internacional e a prática doméstica no que tange à regulação algorítmica aponta para um déficit normativo estrutural, que compromete não apenas a eficácia dos direitos fundamentais, mas a própria legitimidade da ordem jurídica diante da transformação digital.

A promulgação da LGPD, embora representativa de um avanço normativo relevante, permanece insuficiente para enfrentar os desafios impostos pela inteligência artificial, uma vez que seu arcabouço, centrado na proteção da autodeterminação informativa e na regulação do tratamento de dados pessoais, carece de dispositivos específicos sobre os mecanismos decisórios automatizados que constituem o núcleo das tecnologias de IA. O direito à revisão de decisões automatizadas, previsto no artigo 20 da LGPD, constitui um passo inicial no

sentido da responsabilização algorítmica, mas, sem diretrizes técnicas que definam os padrões mínimos de transparência e inteligibilidade, esse direito tende à inaplicabilidade.

Ademais, o regime de responsabilidade civil objetiva, delineado no artigo 42 da LGPD, e a previsão da inversão do ônus da prova representam, sem dúvida, instrumentos jurídicos valiosos para a proteção do titular de dados, mas condicionada à possibilidade concreta de demonstrar o nexo causal entre o funcionamento do sistema automatizado e o dano experimentado, o que muitas vezes não é possível em um contexto marcado por arquiteturas opacas, decisões probabilísticas e técnicas de aprendizado profundo que escapam à lógica determinista do direito clássico.

A potencial articulação entre a LGPD e outros instrumentos normativos voltados à proteção de grupos em situação de vulnerabilidade, como a Lei nº 7.716/1989, a Lei Maria da Penha, o Estatuto da Criança e do Adolescente e o Estatuto da Pessoa com Deficiência, revela uma possibilidade normativa promissora, mas ainda não realizada. A ausência de um projeto nacional articulado de regulação tecnológica e o caráter fragmentário das políticas públicas voltadas ao enfrentamento das desigualdades digitais impedem a construção de um regime jurídico coerente e eficaz, demandando, a superação da dispersão normativa, a requalificação técnica do aparato estatal e a consolidação de uma autoridade regulatória verdadeiramente autônoma e estruturada, capaz de enfrentar os interesses econômicos que hoje moldam, de forma opaca e unilateral, a arquitetura das decisões automatizadas.

Nesse horizonte, a promulgação da Emenda Constitucional nº 115/2022 constitui um marco relevante no redimensionamento do constitucionalismo brasileiro, ao consagrar a proteção de dados pessoais como direito fundamental, mediante a introdução do inciso LXXIX ao artigo 5º da Constituição Federal. Esse reconhecimento reflete não apenas uma ampliação da esfera de proteção dos direitos fundamentais para abranger as dimensões informacionais da dignidade humana, mas também um avanço no sentido da adaptação do texto constitucional às transformações estruturais da ordem digital.

Entretanto, essa conquista normativa revela-se, por ora, insuficiente diante da ausência de mecanismos efetivos de controle repressivo e regulatório, sobretudo no que tange ao uso e desenvolvimento de sistemas de inteligência artificial, lacuna que não pode ser atribuída unicamente ao estágio atual do processo legislativo ou a meros entraves técnico-jurídicos, mas que deve ser compreendida como manifestação de uma racionalidade político-econômica que,

ao preservar o *status quo*, transfere aos indivíduos o ônus pela proteção de seus próprios direitos em um ecossistema digital caracterizado por elevada complexidade técnica e assimetrias informacionais estruturais.

Como adverte Luigi Ferrajoli (1995), a igualdade material, compreendida como limite absoluto à discricionariedade normativa e administrativa, é condição inegociável da legitimidade constitucional. A omissão regulatória diante da inteligência artificial, portanto, não constitui mero descuido técnico, mas negação prática da igualdade jurídica, incompatível com os compromissos fundantes do Estado Democrático de Direito. Em face disso, a regulação das tecnologias algorítmicas não se impõe apenas como exigência jurídica, mas como imperativo ético-político de resistência à opressão automatizada em sociedades progressivamente capturadas por regimes de racionalidade tecnológica não controlada.

## 5 CONCLUSÃO

O presente trabalho se propôs a analisar criticamente os impactos da aplicação de Sistemas Automatizados de Apoio à Decisão sobre a tutela dos direitos fundamentais de grupos historicamente subalternizados no Brasil, com especial atenção à transparência algorítmica, compreendida como um princípio jurídico, ético e político estruturante. Ao longo da investigação, buscou-se demonstrar que tais tecnologias, embora frequentemente legitimadas sob os signos da inovação, da eficiência e da neutralidade técnica, são produtos de escolhas humanas profundamente condicionadas por contextos institucionais, políticos e econômicos, as quais interferem diretamente na lógica que rege o funcionamento dos algoritmos e na seleção dos dados que os alimentam.

Ao investigar os efeitos da aplicação desses sistemas sobre aos grupos vulneráveis, este trabalho demonstrou que há riscos de tais tecnologias reproduzirem, e, por vezes, intensificarem padrões de exclusão já presentes no tecido social. Articuladas à noção de interseccionalidade, essas dinâmicas revelam o entrelaçamento estrutural de marcadores sociais como raça, gênero, classe e territorialidade nos processos algorítmicos, o que torna certos corpos e experiências mais expostos a decisões arbitrárias, desproporcionais ou discriminatórias, evidenciando a urgência de se repensar os fundamentos jurídicos que regem a governança da inteligência artificial, sobretudo em contextos historicamente desiguais, como se encontra o Brasil.

A Constituição Federal de 1988, ao instituir o Estado Democrático de Direito, alicerça-se nos princípios da dignidade da pessoa humana (artigo 1º, III), da cidadania (artigo 1º, II) e da igualdade (artigo 5º, caput), fundamentos que impõem ao ordenamento jurídico e às políticas públicas o dever de promover justiça social, inclusão e respeito à diversidade. No mesmo sentido, o artigo 3º estabelece como objetivos fundamentais da República a redução da pobreza, a redução das desigualdades sociais e regionais e a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação. Esses comandos constitucionais exigem uma atuação regulatória ativa e participativa sobre os Sistemas Automatizados de Apoio à Decisão, assegurando que o desenvolvimento e a implementação dessas tecnologias estejam subordinados aos valores democráticos e aos direitos fundamentais.

Diante da insuficiência da LGPD em assegurar garantias substantivas de revisão e explicação das decisões automatizadas, este estudo aponta para a necessidade de construção doutrinária sólida que antecipe os riscos normativos de uma aceitação acrítica da lógica algorítmica, que deve operar como contraponto à tendência tecnocrática que se insinua na gestão pública e privada, muitas vezes naturalizando práticas de exclusão, vigilância e assimetria informacional sob o manto da eficiência computacional.

Com base nas análises realizadas, evidencia-se que não há neutralidade tecnológica quando os algoritmos são aplicados em contextos atravessados por desigualdades históricas, pois mesmo sistemas baseados em lógicas matemáticas sofisticadas carregam as marcas dos ambientes em que foram concebidos e operados. Reconhecer essa realidade é condição indispensável para a formulação de uma ordem jurídica capaz de enfrentar os desafios éticos, sociais e políticos da era digital, orientando a regulação da inteligência artificial por princípios de justiça, equidade e proteção dos mais vulneráveis, para que a tecnologia contribua com um futuro mais inclusivo, transparente e democrático.

Dentre as principais contribuições desta monografia, destaca-se a construção de uma abordagem jurídica para a transparência algorítmica que ultrapassa a dimensão meramente técnica e a reposiciona como princípio jurídico estruturante, com incidência direta sobre os direitos à dignidade da pessoa humana, ao contraditório, à ampla defesa e à igualdade material. Sustenta-se que a transparência deve ser concebida como condição indispensável para inteligibilidade e a contestabilidade das decisões automatizadas, sobretudo naquelas situações em que tais decisões impactam de forma direta os direitos de indivíduos ou coletividades vulneráveis, de modo a permitir o escrutínio público, a responsabilização dos agentes envolvidos e a contestação por parte dos atingidos, sendo o acesso aos critérios decisórios, à lógica de funcionamento e à origem dos dados utilizados condição mínima para o respeito aos direitos à informação, ao contraditório, à proteção de dados e à dignidade da pessoa humana.

Em última instância, constatou-se a necessidade urgente de o Direito reconquistar sua centralidade enquanto ferramenta de resistência democrática frente à tecnorregulação silenciosa que se opera por meio de códigos opacos e sistemas ininteligíveis, de forma que a defesa da transparência algorítmica, nesse sentido, não se reduz a um imperativo técnico, mas uma exigência normativa fundamental para a preservação da dignidade humana, da igualdade material e da cidadania ativa. É necessário, portanto, reconhecer a tecnologia como arena de

disputa política e, a partir disso, instituir dispositivos legais e administrativos que garantam a auditabilidade, a compreensibilidade e a possibilidade de revisão das decisões algorítmicas, com a criação de instâncias de escuta ativa, reparação efetiva e deliberação democrática, de modo que reivindicar o direito à inteligibilidade das decisões, para, assim, afirmar a primazia do humano sobre a máquina e reafirmar o compromisso do Direito com a justiça, em tempos de governança automatizada.

## REFERÊNCIAS