

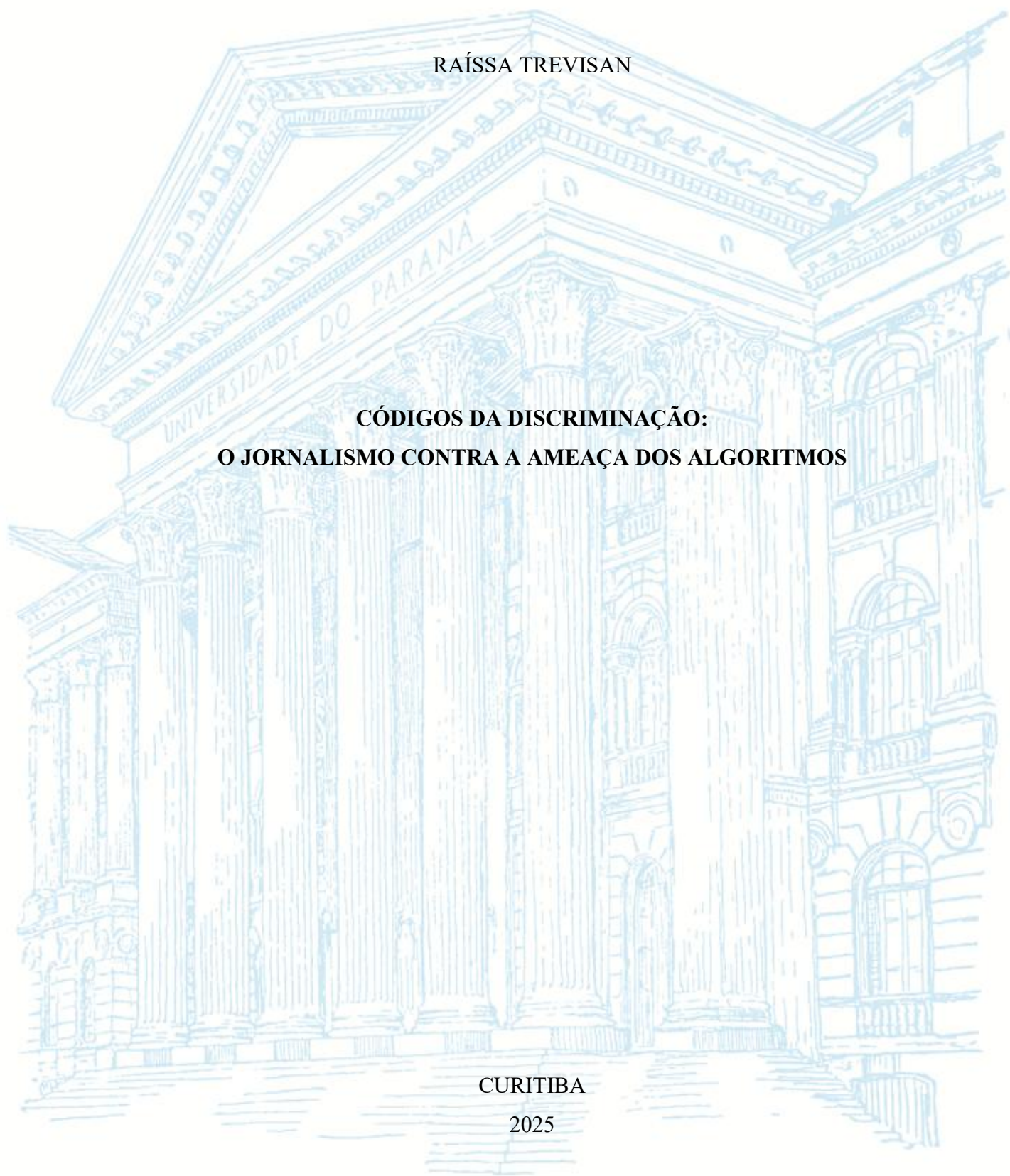
UNIVERSIDADE FEDERAL DO PARANÁ

RAÍSSA TREVISAN

**CÓDIGOS DA DISCRIMINAÇÃO:
O JORNALISMO CONTRA A AMEAÇA DOS ALGORITMOS**

CURITIBA

2025



RAÍSSA TREVISAN

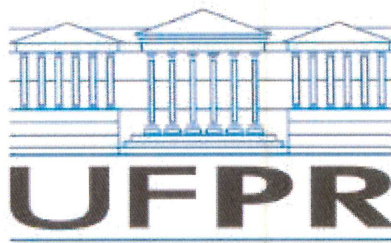
CÓDIGOS DA DISCRIMINAÇÃO:
O JORNALISMO CONTRA A AMEAÇA DOS ALGORITMOS

Monografia apresentada para a disciplina Trabalho de Conclusão de Curso II, do Curso de Jornalismo, Setor de Artes, Comunicação e Design, Universidade Federal do Paraná.

Orientador: Profº. Toni André Scharlau Vieira

CURITIBA

2025

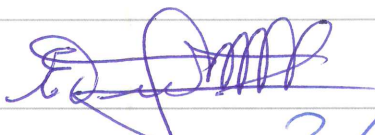
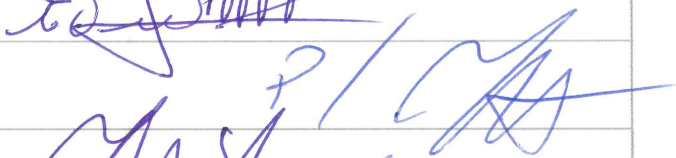
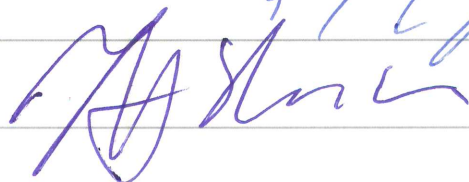


UNIVERSIDADE FEDERAL DO PARANÁ
DEPARTAMENTO DE COMUNICAÇÃO SOCIAL
Rua Bom Jesus, 650, - - Bairro Juvevê, Curitiba/PR, CEP 80035-010
Telefone: 3360-5000 - <https://ufpr.br/>

ATA DE REUNIÃO

ATA DA BANCA DE APRESENTAÇÃO DO TRABALHO DE CONCLUSÃO DO CURSO DE JORNALISMO DA UNIVERSIDADE FEDERAL DO PARANÁ

No dia 10/12/2025, às 19:30 horas, os membros da banca de avaliação reuniram-se no Departamento de Comunicação Social da UFPR, com a finalidade de avaliar a aluna **RAÍSSA PONTES TREVISAN** que apresentou o trabalho de conclusão de curso em jornalismo intitulado: **Códigos da Discriminação: o jornalismo contra a ameaça dos algoritmos**. Após informar as normas do exame de avaliação, o orientador passou a palavra para que a aluna realizasse a apresentação. Finalizada a exposição, a aluna foi arguida pelos membros da banca que atribuíram as seguintes notas:

Professora	Nota	Assinatura
ELSON FAXINA	100	
ELAINE BARCELLOS DE ARAÚJO	100	
TONI ANDRE SCHARLAU VIEIRA	100	

Sendo assim, a média aritmética atribuída à aluna na defesa de seu Trabalho de Conclusão de Curso, foi _____, nota que será lançada no SIGA pelo Professor Orientador somente após realizadas as considerações sugeridas pela banca. A aluna foi considerada aprovada na disciplina e deverá entregar o trabalho com alterações sugeridas pela banca em até 10 dias.


TONI ANDRE SCHARLAU VIEIRA
Professor Orientador



Dedico este trabalho especialmente à mamãe e Nast.

AGRADECIMENTOS

Sou muito grata à minha família, principalmente à minha mãe Helena, que batalhou para garantir que tivéssemos as melhores oportunidades. Com ela, minha guerreira e mulher admirável, aprendo todos os dias a levar a vida com leveza e bom humor, sem deixar de lado o impulso de ir atrás dos sonhos e o senso de responsabilidade. O amor, o suporte e o incentivo de mamãe e Nast, minha irmã, me dão segurança para ir, porque sei que sempre vou ter para onde voltar.

Ao melhor companheiro que eu poderia ter, Davi, muito obrigada! Dividir a vida com alguém que me faz sentir muito amada, me apoia e me ilumina tanto é uma sorte rara. Obrigada por, mais do que ninguém, ter acompanhado — com respeito, carinho e tato — esse percurso tão de perto, me dando forças para chegar até aqui.

Agradeço também aos meus amigos, em particular àqueles que fizeram companhia nessa jornada da graduação, tornando tudo mais prazeroso. Hoje vejo os bons relacionamentos que nasceram dessa trajetória e tenho a certeza de estar bem acompanhada. Um agradecimento especial à Fran, que me motiva e torce por mim há mais de uma década. De nossas conversas sempre saio inspirada, com vontade de construir infinitas coisas e evoluir como ser humano e profissional. Este trabalho, inclusive, nasceu de uma dessas trocas.

Por fim, mas não menos importante, sou extremamente grata à Universidade Federal do Paraná e seus docentes. Me graduar nessa instituição, que oferece educação pública, gratuita e de qualidade, é uma oportunidade honrável que contribuiu muito para meu amadurecimento pessoal e profissional. Faço uma menção honrosa ao melhor projeto de extensão do Decom, o NCEP - Núcleo de Comunicação e Educação Popular, que proporcionou experiências e aprendizados extremamente transformadores e marcantes na minha vida.

Sem todos os prazeres e percalços desses anos, a Raíssa de hoje não existiria.

Estamos viciados em modernidade. A maior parte das invenções é uma tentativa de nós, humanos, nos projetarmos em matéria para além de nossos corpos.

— Ailton Krenak

RESUMO

A presente monografia analisa os impactos sociopolíticos da discriminação algorítmica na esfera pública, investigando o papel do jornalismo como instância de fiscalização diante da opacidade das infraestruturas de inteligência artificial. O objetivo central é compreender como vieses raciais e sociais são convertidos em parâmetros matemáticos de decisão e discutir a viabilidade do *Algorithmic Accountability Reporting* (reportagem de auditoria algorítmica) como ferramenta de resistência. A pesquisa adota uma abordagem qualitativa e exploratória, combinando revisão bibliográfica e análise documental. A fundamentação teórica articula os estudos críticos de dados com o pensamento social brasileiro, mobilizando conceitos como racismo algorítmico e *tecnochauvinismo* para demonstrar que a tecnologia não é neutra, mas atua na reprodução de desigualdades estruturais, especialmente em contextos de segurança pública e acesso a direitos. O trabalho utiliza o caso da "Muralha Digital" de Curitiba para ilustrar a tensão entre discursos de inovação e a realidade da vigilância massiva. Os resultados apontam que a automação acrítica nas redações agrava a precarização do trabalho e amplia a dependência de plataformas proprietárias, ameaçando a autonomia jornalística. Conclui-se que a sobrevivência da profissão depende de uma reconfiguração metodológica: é imperativo que o jornalismo transite da posição de usuário para a de auditor de caixas-pretas. Defende-se a adoção de técnicas de engenharia reversa e a exigência de transparência algorítmica como compromissos essenciais para a defesa dos direitos humanos e a manutenção da democracia na era da automação.

Palavras-chave: jornalismo; algoritmos; discriminação algorítmica; racismo algorítmico; inteligência artificial; reportagem de auditoria algorítmica; *tecnochauvinismo*.

ABSTRACT

This monograph analyzes the sociopolitical impacts of algorithmic discrimination in the public sphere, investigating the role of journalism as a watchdog regarding the opacity of artificial intelligence infrastructures. The main objective is to understand how racial and social biases are converted into mathematical decision-making parameters and to discuss the viability of *Algorithmic Accountability Reporting* as a resistance tool. The research adopts a qualitative and exploratory approach, combining bibliographic review and documentary analysis. The theoretical *framework* articulates critical data studies with brazilian social thought, mobilizing concepts such as algorithmic racism and *technochauvinism* to demonstrate that technology is not neutral but acts in the reproduction of structural inequalities, especially in contexts of public security and access to rights. The work uses the case of the “Digital Wall” in Curitiba to illustrate the tension between innovation discourses and the reality of massive surveillance. The results indicate that uncritical automation in newsrooms aggravates labor precarization and expands dependence on proprietary platforms, threatening journalistic autonomy. It is concluded that the survival of the profession depends on a methodological reconfiguration: it is imperative that journalism transitions from the position of user to that of auditor of *black boxes*. The adoption of reverse engineering techniques and the demand for algorithmic transparency are defended as essential commitments for the defense of human rights and the maintenance of democracy in the automation era.

Keywords: journalism; algorithms; algorithmic discrimination; algorithmic racism; artificial intelligence; *algorithmic accountability reporting*; *technochauvinism*.

LISTA DE FIGURAS

FIGURA 1 - EVOLUÇÃO HISTÓRICA: DA INTELIGÊNCIA ARTIFICIAL À IA GENERATIVA	25
FIGURA 2 - AS QUATRO ABORDAGENS DA RELAÇÃO CIÊNCIA, TECNOLOGIA E SOCIEDADE	41
FIGURA 3 - CICLO DE RETROALIMENTAÇÃO DO VIÉS ALGORÍTMICO	51
FIGURA 4 - PRINCIPAIS DESAFIOS DE CURTO PRAZO ENFRENTADOS PELOS JORNALISTAS DO SUL GLOBAL NA ADOÇÃO DE IA NO TRABALHO	64
FIGURA 5 - CENTRO DE CONTROLE OPERACIONAL DA MURALHA DIGITAL DE CURITIBA (I)	71
FIGURA 6 - CENTRO DE CONTROLE OPERACIONAL DA MURALHA DIGITAL DE CURITIBA (II)	72

LISTA DE QUADROS

QUADRO 1 - COMPARATIVO DE ROTINAS: JORNALISMO TRADICIONAL E JORNALISMO NA ERA DA INTELIGÊNCIA ARTIFICIAL	63
QUADRO 2 - PRINCÍPIOS FATML PARA RESPONSABILIZAÇÃO ALGORÍTMICA.....	68
QUADRO 3 - ROTEIRO METODOLÓGICO: ETAPAS DA INVESTIGAÇÃO ALGORÍTMICA	68

LISTA DE SIGLAS

AAR	– <i>Algorithmic Accountability Reporting</i> (Reportagem de Auditoria Algorítmica)
ABRAJI	– Associação Brasileira de Jornalismo Investigativo
ADMs	– Armas de Destruição Matemática
AJOR	– Associação de Jornalismo Digital
API	– <i>Application Programming Interface</i> (Interface de Programação de Aplicação)
C&T	– Ciência e Tecnologia
CEP	– Código de Endereçamento Postal
CESeC	– Centro de Estudos de Segurança e Cidadania
COMPAS	– <i>Correctional Offender Management Profiling for Alternative Sanctions</i> (Perfilamento de Gestão de Infratores para Sanções Alternativas)
CTS	– Ciência, Tecnologia e Sociedade
DL	– <i>Deep learning</i> (Aprendizado Profundo)
ESCT	– Estudos Sociais da Ciência e Tecnologia
FATML	– <i>Fairness, Accountability, and Transparency in Machine Learning</i> (Justiça, Responsabilidade e Transparência no Aprendizado de Máquina)
IA	– Inteligência artificial
IJNET	– <i>International Journalists' Network</i> (Rede Internacional de Jornalistas)
INSS	– Instituto Nacional do Seguro Social
LAI	– Lei de Acesso à Informação
LGPD	– Lei Geral de Proteção de Dados
ML	– <i>Machine learning</i> (Aprendizado de Máquina)
RH	– Recursos Humanos
RSF	– Repórteres Sem Fronteiras
SEO	– <i>Search Engine Optimization</i> (Otimização para Mecanismos de Busca)
SNARC	– <i>Stochastic Neural Analog Reinforcement Calculator</i> (Calculadora Analógica Neural Estocástica de Reforço)
SUS	– Sistema Único de Saúde
TCU	– Tribunal de Contas da União
TICs	– Tecnologias de Informação e Comunicação

TRF – *Thomson Reuters Foundation*
UFPR – Universidade Federal do Paraná
UNICAMP – Universidade Estadual de Campinas

SUMÁRIO

1	INTRODUÇÃO	13
1.1	CONTEXTO E PROBLEMA.....	13
1.2	OBJETIVOS	14
1.2.1	Objetivo geral	14
1.2.2	Objetivos específicos	14
1.3	JUSTIFICATIVA	14
1.4	METODOLOGIA.....	18
1.5	ORGANIZAÇÃO TEXTUAL	19
2	A ARQUITETURA DA DECISÃO ALGORÍTMICA	21
2.1	O QUE SÃO ALGORITMOS?	21
2.2	INTELIGÊNCIA ARTIFICIAL (IA) E APRENDIZADO DE MÁQUINA.....	22
2.3	O MOTOR DA IA GENERATIVA: <i>DEEP LEARNING</i> E <i>BIG DATA</i>	23
2.4	PARADOXO DA CAIXA-PRETA: A OPACIDADE COMO CARACTERÍSTICA .	27
2.5	CAPITALISMO DE VIGILÂNCIA E ECONOMIA DOS DADOS.....	29
3	CODIFICAÇÃO DA DESIGUALDADE	32
3.1	<i>TECNOCHAUVINISMO</i> E A CRENÇA NA OBJETIVIDADE MATEMÁTICA	32
3.1.1	Raízes da crença na neutralidade digital.....	34
3.2	DETERMINISMO TECNOLÓGICO E DESPOLITIZAÇÃO DA TÉCNICA.....	36
3.2.1	O mito da neutralidade tecnológica	39
3.3	MATEMÁTICA COMO MORAL: QUANDO OS CÁLCULOS VIRAM DESTINO	41
3.4	AUTORIDADE AUTOMÁTICA E DUPLA OPACIDADE	43
3.5	A CONSTITUIÇÃO DO VIÉS ALGORÍTMICO	45
3.5.1	A desigualdade antecede o algoritmo	46
3.5.2	Colonialidade de dados e captura assimétrica	47

3.5.3	Onde nasce o viés do algoritmo?	48
3.5.4	Armas de destruição matemática: opacidade, escala e dano	51
4	DESAFIOS DO JORNALISMO NA ERA DA INTELIGÊNCIA ARTIFICIAL .	55
4.1	O PARADOXO DA INFORMAÇÃO ENTRE OS DESERTOS DE NOTÍCIAS E A SUPERPOPULAÇÃO ARTIFICIAL.....	58
4.2	DISTOPIA PROFISSIONAL SOB O ATRASO INSTITUCIONAL E PRECARIZAÇÃO	61
4.3	JORNALISMO DE AUDITORIA ALGORÍTMICA COMO CONTRAPODER E RESISTÊNCIA.....	66
4.4	A URGÊNCIA DA REINVENÇÃO JORNALÍSTICA DIANTE DA HEGEMONIA ALGORÍTMICA	75
	CONCLUSÕES.....	78
	REFERÊNCIAS.....	80

1 INTRODUÇÃO

1.1 CONTEXTO E PROBLEMA

A organização da vida social contemporânea foi profundamente reconfigurada pela infraestrutura de dados. Decisões que historicamente dependiam do julgamento humano, burocrático ou institucional - como a concessão de crédito, a triagem de currículos e o policiamento ostensivo - foram progressivamente delegadas a sistemas automatizados. Contudo, essa transição da decisão humana para a algorítmica não ocorre em um vácuo moral. Ao serem treinados com bases de dados que refletem séculos de desigualdades, esses sistemas carregam vieses intrínsecos que não apenas reproduzem, mas também escalam e amplificam preconceitos históricos, convertendo injustiças sociais em parâmetros matemáticos de eficiência.

No Brasil, essa transferência de autoridade aterrissa sobre um terreno marcado pelo racismo estrutural, onde a ‘caixa-preta’ tecnológica muitas vezes automatiza a exclusão. A natureza do problema revela-se na segurança pública: dados da Rede de Observatórios da Segurança indicam que 90,5% das prisões realizadas por reconhecimento facial no país recaem sobre pessoas negras (NUNES, 2019), enquanto o Centro de Estudos de Segurança e Cidadania - CESeC (2025) estima que mais de 87 milhões de brasileiros - quase metade da população - já estão expostos a esse monitoramento massivo. Tais números evidenciam que a tecnologia, longe de ser neutra, atua como ferramenta de validação estatística da seletividade penal.

Nesse cenário, o jornalismo, instituição historicamente responsável por fiscalizar o poder e garantir a transparência democrática, encontra-se diante de um impasse estrutural. A profissão enfrenta a dificuldade técnica de investigar processos decisórios que não deixam rastros visíveis, não emitem memorandos e são protegidos por segredos industriais. Sem o domínio de novas ferramentas de apuração e diante da precarização das redações, o campo jornalístico corre o risco de perder sua função de contrapoder, limitando-se a cobrir os efeitos da violência algorítmica sem conseguir questionar suas causas profundas.

Dessa inquietação emerge o problema central que esta monografia busca responder: de que maneira a opacidade dos sistemas algorítmicos enviesados desafia a função social do jornalismo e quais caminhos são necessários para que a profissão consiga identificar, auditar e expor as novas formas de discriminação codificada que operam na sociedade contemporânea?

1.2 OBJETIVOS

Para responder ao problema de pesquisa delimitado e enfrentar a complexidade das relações entre tecnologia, sociedade e mídia, este trabalho estabelece objetivos que partem da análise teórica das estruturas de opacidade e conduzem à discussão prática sobre a função social do jornalismo. A intenção é percorrer o caminho que vai do entendimento técnico do 'segredo' algorítmico até a proposição de estratégias profissionais de fiscalização e de transparência.

1.2.1 Objetivo geral

Analisar os mecanismos da discriminação algorítmica e seus impactos nas esferas pública e informacional, discutindo como o jornalismo pode atuar como instância de fiscalização e contrapoder diante da opacidade das infraestruturas tecnológicas de vigilância e de decisão automatizada.

1.2.2 Objetivos específicos

- I.** Descrever a arquitetura técnica e política dos sistemas de inteligência artificial, *big data* e *deep learning*, explicitando o funcionamento da 'caixa-preta' e o mito da neutralidade tecnológica;
- II.** Discutir como desigualdades estruturais são codificadas, reproduzidas e amplificadas por sistemas de decisão automatizada, especialmente em sociedades marcadas por profundas disparidades raciais e sociais, como o Brasil;
- III.** Examinar a viabilidade do jornalismo de auditoria algorítmica como resposta à precarização profissional, utilizando o cenário da "Muralha Digital" de Curitiba como contexto ilustrativo dos desafios de fiscalização.

1.3 JUSTIFICATIVA

A escolha de investigar a discriminação algorítmica e suas implicações para o jornalismo nasce de uma inquietação que atravessa tanto a realidade brasileira quanto minha própria trajetória pessoal e profissional. Vivemos em um país onde a adoção acelerada de tecnologias de vigilância

e de decisão automatizada já produz efeitos concretos sobre quem acessa direitos, quem é monitorado e quem permanece invisível. A pesquisa emerge como resposta a esse cenário de urgência - um cenário em que decisões técnicas ganham força política sem que a sociedade tenha clareza sobre seus impactos.

No plano social, a relevância do tema é incontornável. Dados do Panóptico mostram que o Brasil já possui 485 projetos oficiais de reconhecimento facial, responsáveis por potencialmente vigiar metade dos brasileiros. Apenas na região Sul há 100 projetos, alcançando mais de 11 milhões de indivíduos (CESeC, 2025). Não se trata de risco abstrato, mas de uma política de vigilância massiva que se expande sem debate público proporcional. A automatização não apenas perpetua desigualdades, mas as transforma em parâmetro de suspeição.

O problema se agrava quando decisões críticas para a vida das pessoas passam a depender de sistemas opacos. Uma auditoria do Tribunal de Contas da União (TCU) revelou que 10,9% das negativas automáticas de benefícios pelo Instituto Nacional do Seguro Social (INSS) em 2024 foram indevidas - erros cometidos por robôs que negaram direitos a cidadãos vulneráveis (CARREGOSA, 2025). Quem sofre as consequências desses ‘erros técnicos’ raramente possui meios para compreendê-los e contestá-los.

Esse deslocamento de autoridade do humano para a máquina também afeta a saúde pública. O Ministério da Saúde investe R\$ 60 milhões em projetos de inteligência artificial (IA) para otimizar diagnósticos no SUS (BRASIL, 2025a) e anunciou R\$ 4,5 bilhões para ‘hospitais inteligentes’ (BRASIL, 2025b). Num país marcado historicamente por racismo médico, em que mulheres negras recebem menos anestesia em procedimentos dolorosos do parto e morrem duas vezes mais do que mulheres brancas por complicações maternas (LEAL et al., 2017; 2023), automatizar decisões clínicas sem auditoria significa ampliar desigualdades históricas sob a aparência de eficiência. Há ainda uma dimensão geracional. A pesquisa TIC Kids Online (2025) revelou que 65% das crianças e adolescentes brasileiros já usam inteligência artificial generativa para estudar, criar conteúdo e até lidar com emoções. Uma geração inteira está aprendendo sobre o mundo por meio de sistemas proprietários que não passam por escrutínio público – o que reforça o papel do jornalismo como mediador crítico.

No plano profissional, a relevância deste estudo reside na crise de identidade do jornalismo diante de tecnologias opacas. Levantamento apresentado pela *Thomson Reuters Foundation* - TRF (2025) aponta que 57,1% dos jornalistas do Sul Global sentem preocupação

ética com a IA, e 50,6% afirmam não ter conhecimento suficiente para compreendê-la. A profissão encontra-se, portanto, simultaneamente preocupada e despreparada. É precisamente essa lacuna que este trabalho procura iluminar.

A literatura crítica sobre tecnologia também reforça a urgência da pauta. Cathy O’Neil (2020) alerta que modelos algorítmicos produzem injustiça em escala industrial. Meredith Broussard (2018) argumenta que a sociedade precisa reduzir o *tecnochauvinismo* para compreender os limites reais da técnica. Frank Pasquale (2015) descreve uma sociedade da caixa-preta na qual decisões automatizadas se tornam inquestionáveis. Todos convergem para um chamado: discutir, expor e responsabilizar os sistemas que moldam o futuro.

Quanto à dimensão pessoal, minha trajetória sempre foi marcada pelo fascínio pela tecnologia e pelo ambiente digital. Cresci imersa nesse universo, movida pela curiosidade de compreender o funcionamento das ferramentas que mediavam minhas interações. No final de 2022, quando o ChatGPT veio a público, essa relação se transformou profundamente e fiquei num misto de fascínio e preocupação. Enquanto o debate público focava em questionar ‘se a IA vai tirar nossos empregos’, comecei a ter minhas próprias *matutações* sobre o ciclo de retroalimentação: se a IA usa dados preexistentes para produzir mais dados, uma hora ela passaria a se alimentar principalmente de conteúdo criado por ela mesma, e em algum momento deixaríamos de ter clareza sobre a veracidade das informações. Foi essa angústia inicial quanto à integridade da informação que me impulsionou a investigar o tema.

A virada determinante nesta pesquisa ocorreu, no entanto, a partir de uma belíssima indicação literária: *Algoritmos de Destruição em Massa*, de Cathy O’Neil (2020). Devorei o livro em alguns dias e percebi, então, que o ‘buraco da IA’ - ou melhor, dos algoritmos - estava muito mais embaixo do que eu imaginava. Os danos e as perversidades algorítmicas não eram uma distopia futura, mas uma realidade que operava silenciosamente há anos. Me surpreendi ao constatar que, apesar da gravidade, o tema permanecia alheio ao conhecimento da maioria das pessoas com quem eu conversava.

Percebi, então, que a invisibilidade é a característica mais cruel e importante desse sistema. Somos afetados por decisões que não vemos nem conseguimos contestar. Com uma mistura de motivação e ceticismo, cada conteúdo consumido mostrava que a ameaça do viés algorítmico é profunda e extremamente desafiadora, e que uma das maiores armas para lutar contra isso é justamente a disseminação do assunto.

Essa convicção se fortaleceu em 2024, quando tive a oportunidade de participar do ciclo formativo *Antirracismo e Políticas Públicas nas Tecnologias da Informação e Comunicação (TICs) e inteligência artificial (IA)*, desenvolvido pela Fundação Mozilla e com mediação do brilhante pesquisador Tarcízio Silva - que considero um dos pioneiros e referência nos estudos do tema no Brasil. Seu livro *Racismo algorítmico: inteligência artificial e discriminação nas redes digitais*, lançado em 2022, e sua linha do tempo de Mapeamento de Danos e Discriminação Algorítmica¹, foram, com certeza, a maior inspiração para esta monografia.

Ao longo dos encontros, compreendi que mesmo especialistas e referências no assunto estão numa busca constante por respostas para tentar enfrentar esses desafios, o que reforçou a urgência de trazer esse debate para o jornalismo. Nem tudo precisa ser mediado pela tecnologia; há esferas que competem essencialmente e somente aos humanos, e jamais deveriam sair do nosso escopo. Sociedades não podem nem devem ser regidas exclusivamente por cálculos matemáticos e inferências; é extremamente ameaçador seguirmos nessa tendência de delegar, de maneira desenfreada e desregulada, todos os assuntos possíveis à tecnologia. Portanto, é dessa mistura de esperança na conscientização e desesperança diante da automação descontrolada da vida social que este trabalho nasce, com o objetivo de tornar visível o que as caixas-pretas insistem em ocultar.

Assim, esta monografia se justifica por quatro razões centrais:

I. Relevância social: o Brasil está implementando, em larga escala, tecnologias de vigilância e de decisão automatizada sem debate público proporcional. A pesquisa contribui para revelar o impacto dessas ferramentas sobre populações vulnerabilizadas e direitos fundamentais.

II. Relevância científica: o estudo articula crítica tecnológica, pensamento social brasileiro e epistemologia jornalística, oferecendo um enquadramento inovador sobre opacidade algorítmica, racismo estrutural e responsabilização.

III. Relevância jornalística: o trabalho responde a uma lacuna urgente: como jornalistas podem investigar tecnologias que não deixam rastros claros, não são auditáveis e moldam silenciosamente políticas públicas e privadas?

IV. Relevância formativa: a pesquisa consolida um percurso acadêmico marcado pelo interesse contínuo em compreender como tecnologias moldam as relações de poder, o acesso a direitos e as formas de desigualdade. O trabalho aprofunda competências analíticas essenciais para qualquer atuação crítica no ecossistema informacional contemporâneo, fortalecendo uma formação

¹ Silva, 2025. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>

voltada à investigação rigorosa, à compreensão estrutural das injustiças e ao exercício responsável da comunicação em ambientes permeados por sistemas algorítmicos.

Diante de uma realidade em que algoritmos decidem sem serem vistos - e onde o jornalismo corre o risco de perder sua função social -, compreender os mecanismos da opacidade torna-se não apenas relevante, mas indispensável. Este trabalho nasce da urgência de defender a democracia num país onde cálculos silenciosos moldam o destino de milhões de pessoas.

1.4 METODOLOGIA

Esta monografia caracteriza-se como uma pesquisa de natureza qualitativa, com fins exploratórios e descritivos. A abordagem qualitativa justifica-se pela necessidade de compreender fenômenos complexos e não quantificáveis, como a opacidade algorítmica, a precarização do trabalho jornalístico e as dinâmicas raciais e sociais embutidas em sistemas de decisão automatizada.

Quanto aos procedimentos técnicos, o trabalho adota uma estrutura mista de pesquisa bibliográfica e pesquisa documental, dividida em três eixos de apuração:

I. Pesquisa Bibliográfica: foi realizada a sistematização de literatura interdisciplinar para construir a base teórica da crítica algorítmica. O trabalho articula os Estudos Críticos de Dados e Algoritmos - com autores fundamentais como Cathy O’Neil (2020), Frank Pasquale (2015) e Shoshana Zuboff (2021) - às teorias sociais e raciais críticas, mobilizando o pensamento de intelectuais como Tarcízio Silva (2019; 2020; 2022), Renato Dagnino (2008) e Frantz Fanon (1980). Esse cruzamento visa sustentar a tese de que a tecnologia não é neutra, mas atua na reprodução de desigualdades estruturais.

II. Pesquisa Documental: foram analisados relatórios técnicos, guias éticos e documentos institucionais que contextualizam a crise do jornalismo e as respostas do campo à automação. O corpus documental inclui diretrizes internacionais de governança de IA (como a Carta de Paris e os princípios *Fairness, Accountability, and Transparency in Machine Learning*² - FATML), relatórios de entidades jornalísticas (Abraji, Ajour, *Knight Center*) e legislação pertinente (Lei Geral de Proteção de Dados e Lei de Acesso à Informação).

² Justiça, Responsabilidade e Transparência no Aprendizado de Máquina, em tradução livre

III. Contextualização Ilustrativa: o trabalho utiliza o caso da "Muralha Digital" de Curitiba como elemento ilustrativo para aplicar os conceitos de opacidade e vigilância no cenário local. A análise desse objeto baseou-se no exame do discurso institucional (materiais de divulgação da Prefeitura), na cobertura jornalística local e nas negativas de acesso a dados públicos, o que evidencia as barreiras à transparência que impedem a realização de uma auditoria técnica externa dos sistemas.

Dessa forma, o método de investigação não se propõe a realizar uma auditoria computacional direta dos códigos-fonte - protegidos por sigilo industrial e governamental -, mas sim a examinar criticamente as condições sociais, políticas e econômicas que permitem a implementação dessas ‘caixas-pretas’ e o papel do jornalismo em seu enfrentamento.

1.5 ORGANIZAÇÃO TEXTUAL

Para atingir os objetivos propostos, esta monografia está estruturada, além desta Introdução, em três capítulos de desenvolvimento e uma conclusão final:

O segundo capítulo, intitulado *A Arquitetura da Decisão Algorítmica*, dedica-se a definir as bases técnicas e políticas do objeto de estudo. Nele, são apresentados os conceitos de algoritmo, *big data*, *machine learning* e *deep learning*, não apenas como ferramentas computacionais, mas também como infraestruturas de poder. O capítulo explora o paradoxo da ‘caixa-preta’, discutindo como a opacidade técnica se converte em estratégia de mercado no contexto do capitalismo de vigilância.

O terceiro capítulo, *Codificação da Desigualdade*, aprofunda a crítica sociológica ao desconstruir o mito da neutralidade tecnológica. Apoiando-se no pensamento social brasileiro e em teorias críticas, o texto discute como o *tecnoc Chauvinismo* e o determinismo tecnológico camuflam a reprodução de hierarquias sociais. O capítulo dissecar a ‘anatomia’ do viés, demonstrando como desigualdades históricas e raciais são convertidas em parâmetros matemáticos de decisão e exclusão.

O quarto capítulo, *Desafios do Jornalismo na Era da Inteligência Artificial*, traz a discussão para o âmbito profissional. Analisa-se a crise da mediação jornalística diante da superpopulação de conteúdos artificiais e da precarização do trabalho. Como resposta, apresenta-se o *Algorithmic Accountability Reporting* (Reportagem de Auditoria Algorítmica) como método de investigação e

resistência, utilizando o contexto da "Muralha Digital" de Curitiba para ilustrar os desafios práticos dessa fiscalização.

Por fim, as Conclusões retomam os principais achados da pesquisa, apontando os limites e as possibilidades de atuação do jornalismo como contrapoder democrático em uma sociedade cada vez mais mediada por códigos opacos.

2 A ARQUITETURA DA DECISÃO ALGORÍTMICA

2.1 O QUE SÃO ALGORITMOS?

Até meados da década de 1930, as ideias de algoritmo e de computabilidade eram debatidas na comunidade matemática, sem uma definição formal e unânime. Osoba e Welser (2017) resgatam um conceito medieval de algoritmo, em que o matemático persa Al-Khwarizmi, cujo nome a palavra deriva, o descrevia como ‘procedimentos confiáveis, com um passo a passo, para computar soluções de equações’.

Embora a noção de um ‘passo a passo’ fosse intuitiva, foi somente em 1937 que a definição de algoritmo foi formalizada no campo científico, graças aos trabalhos de Alonzo Church e Alan Turing. Os matemáticos e cientistas da computação introduziram os conceitos de ‘computabilidade’ e de ‘funções computáveis’, que buscam definir os limites do que pode ser resolvido e calculado por um processo mecânico (OSOBA; WELSER, 2017).

Ao estabelecer o que era teoricamente ‘computável’, eles forneceram a base para conceituar um algoritmo como uma sequência finita de instruções/comandos bem definidos para produzir um resultado final (*output*) a partir de um ponto de partida (*input*), visando resolver um problema ou executar uma tarefa (ALMEIDA; DONEDA, 2016).

Apesar da palavra ‘algoritmo’ costumar remeter primeiro às redes sociais, sua influência hoje é muito mais vasta. Com crescente valor econômico e político, os algoritmos se tornaram onipresentes e estratégicos ao serem incorporados em muitos serviços, produtos e processos do cotidiano: recursos humanos e contratação; seguros de vida e saúde; policiamento e área jurídica; empréstimos financeiros e linhas de crédito; escolas e universidades, são áreas da sociedade que já possuem histórico de se basear em previsões algorítmicas para tomar decisões. E muitas vezes nem percebemos.

Esse poder de aconselhar ou de tomar decisões sobre como dados e informações devem ser interpretados e utilizados ocorre sobretudo ao atrelar os algoritmos à inteligência artificial (IA),

área em que Minsky (1961) e outros pioneiros já concebiam o algoritmo como uma ferramenta capaz de capacitar sistemas computacionais com a inteligência humana.

2.2 INTELIGÊNCIA ARTIFICIAL (IA) E APRENDIZADO DE MÁQUINA

De acordo com Barr e Feigenbaum (1981, p. 3, tradução nossa), a inteligência artificial (IA) é um ramo da ciência da computação focado em desenvolver sistemas de computador “que exibem características que associamos à inteligência no comportamento humano - compreensão da linguagem, aprendizado, raciocínio, resolução de problemas e assim por diante”³.

O que hoje entendemos como IA remonta a 1943, quando Warren McCulloch e Walter Pitts propuseram um modelo de neurônios artificiais inspirado no cérebro humano (RUSSEL; NORVIG, 2013). Já em 1951, Minsky e Edmonds construíram o primeiro computador de rede neural, o SNARC (Stochastic Neural Analog Reinforcement Calculator), que navegava por um labirinto virtual, tentando aprender a saída de forma autônoma (BARBOSA; BEZERRA; 2020). Só em 1956, porém, o termo inteligência artificial foi cunhado pela primeira vez pelo cientista da computação John McCarthy (STRYKER; KAVLAKOGLU, 2024).

Entendendo que uma das características da inteligência humana é a capacidade de se adaptar, otimizar ou aprender a partir de uma experiência, os esforços dos pioneiros da IA resultaram nos ‘algoritmos de aprendizado’, que treinam sistemas para aprender, criar ou melhorar modelos com base em dados coletados. Um modelo, por sua vez, é uma representação abstrata de um processo que utiliza informações existentes para gerar respostas por meio de análises preditivas em situações diversas.

Um exemplo cotidiano de modelo é preparar uma refeição. Antes de cozinhar, o responsável pode considerar as preferências alimentares, os ingredientes disponíveis, o tempo e a própria disposição. Essas informações servem como entradas (*input*) do modelo, enquanto a decisão sobre o prato a ser preparado representa sua saída (*output*). A observação do resultado, se a refeição foi bem aceita, se todos comeram o suficiente e se o preparo foi adequado, permite ajustar mentalmente o modelo para as próximas vezes. Esse processo de observação e ajuste constante

³ “systems that exhibit the characteristics we associate with intelligence in human behavior — understanding language, learning, reasoning, solving problems, and so on.”

caracteriza os modelos dinâmicos - eles não permanecem fixos, mas se transformam conforme surgem novos dados:

(...) os algoritmos não estão apenas processando números através de modelos matemáticos estáticos, mas sim atualizando seu comportamento repetidamente com base em modelos ajustados em resposta à sua experiência (dados de entrada/*input*) e métricas de desempenho.⁴ (OSOBA; WELSER, 2017, p. 5, tradução nossa)

O problema desses modelos é que não são capazes de representar todas as nuances da comunicação, psiquê ou complexidade dos seres humanos e do ‘mundo real’ para as máquinas, obrigando a pessoa ou organização que o desenvolve a fazer escolhas sobre o que é importante o suficiente para ser incluído no modelo, “simplificando o mundo numa versão de brinquedo que possa ser facilmente entendida” (O’NEIL, 2020, p. 21).

Ainda sobre os algoritmos de aprendizado, boa parte do trabalho dos pioneiros da IA sustenta a maioria dos sistemas automatizados atuais. Esses algoritmos nada mais são do que sofisticados métodos de estimativa estatística, que usam modelos matemáticos para prever resultados (OSOBA; WELSER, 2017). O aprendizado de máquina (*machine learning* ou ML) é um campo da IA que permite que os sistemas reconheçam padrões em uma determinada base de dados e, depois, apliquem esse aprendizado para reconhecer variáveis em outros conjuntos de dados. Com o ML, os sistemas passam a executar tarefas cognitivas, como a identificação de rostos e objetos nas imagens ou a tradução de linguagem natural (STRYKER; KAVLAKOGLU, 2024).

2.3 O MOTOR DA IA GENERATIVA: *DEEP LEARNING* E *BIG DATA*

Apesar do termo inteligência artificial ter sido cunhado em 1956 e ser objeto de estudo desde a década de 1940, recentemente as utilizações de IA generativa, como *chatbots*, ganharam atenção do público ‘leigo’ - em especial por causa do lançamento do ChatGPT, em novembro de 2022. A título de ilustração, três meses após o lançamento, em janeiro de 2023, o serviço atingiu

⁴ “the algorithms are not just crunching numbers through static mathematical models but update their behavior iteratively based on models tuned in response to their experience (input data) and performance metrics.”

100 milhões de usuários, tornando-se a ferramenta digital com maior adoção em massa na história (HU, 2023).

Encabeçada por esse avanço e relevância, uma inteligência artificial generativa é aquela capaz de usar o reconhecimento de padrões do ML para ir além e criar conteúdos ‘originais’: textos, imagens, códigos, áudios e vídeos surgem a partir de comandos simples do usuário - o *prompt*. Tal capacidade revolucionária sustenta-se em dois pilares tecnológicos que se aliam para aprimorar os modelos de algoritmos generativos: o aprendizado profundo (*deep learning* ou DL) - uma subcategoria do ML - e o *big data* - área dedicada ao “tratamento, análise e obtenção de informações a partir de conjuntos de dados grandes demais para serem analisados” (KREMER, 2023, p. 13) por humanos individualmente.

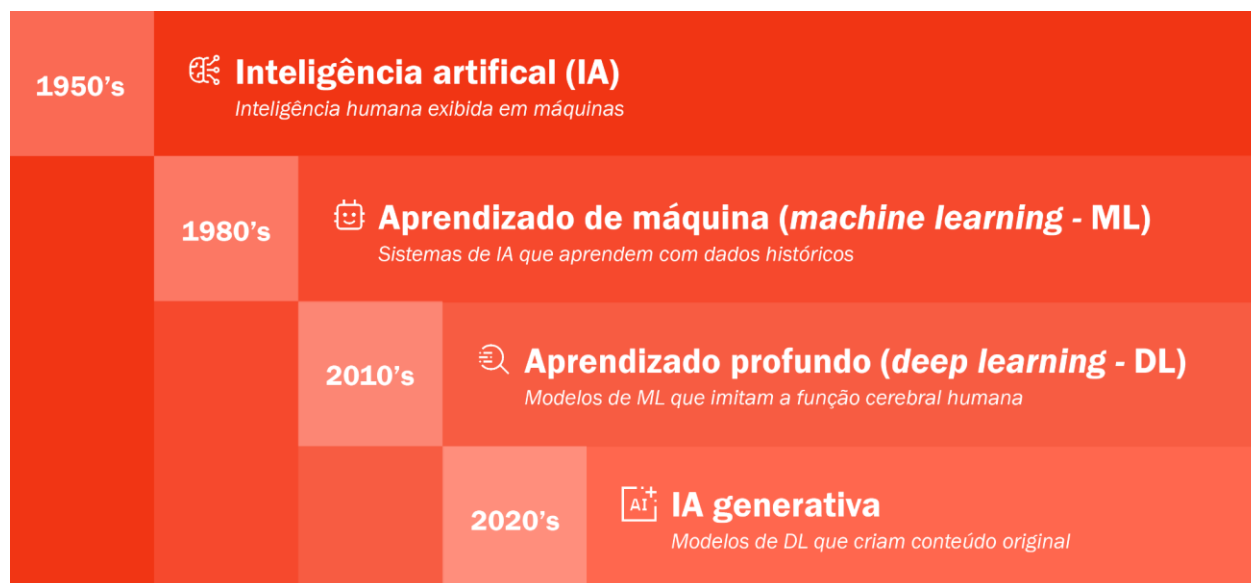
É nesse ponto que se torna necessário distinguir com mais precisão o que diferencia o ML ‘clássico’ do DL. No aprendizado de máquina tradicional, o trabalho humano desempenha um papel central, pois os pesquisadores precisam decidir previamente quais características dos dados são relevantes para o modelo aprender, processo conhecido como engenharia de atributos. A lógica é relativamente direta: o ser humano seleciona, organiza e transforma o mundo complexo em um conjunto limitado de variáveis e, a partir delas, o algoritmo aprende a mapear entradas (*inputs*) em saídas (*outputs*). Trata-se de um método de aprendizado baseado em instruções explícitas sobre ‘o que olhar’ e ‘como olhar’.

O aprendizado de máquina profundo, ao contrário, abandona esse mecanismo manual e simula o poder de tomada de decisão do cérebro humano. Enquanto os modelos tradicionais utilizam redes neurais simples com uma ou duas camadas computacionais, os modelos de *deep learning* se apoiam em redes neurais profundas, normalmente com centenas ou milhares dessas camadas, capazes de aprender autonomamente representações internas dos dados, sem que alguém precise dizer quais características buscar (HOLDSWORTH; SCAPICCHIO, 2024):

Em vez de trabalhar com dados previamente rotulados e filtrados, o DL recebe grandes volumes de dados relativamente brutos — pixels, trechos de texto, multimídia — e, por meio de sucessivas camadas ocultas, constrói compreensões internas de padrões cada vez mais abstratos. A complexidade desse processo é o que permite ao DL alcançar resultados notáveis em áreas como a visão computacional e o processamento de linguagem natural. O *deep learning* é particularmente útil em domínios com dados em larga escala e dimensionalidade, ruidosos e não estruturados, razão pela qual “superam os algoritmos de aprendizado de máquina superficiais na maioria das aplicações

em que dados de texto, imagem, vídeo, fala e áudio precisam ser processados” (LECUN et al.⁵, 2015 apud JANIESCH; ZSCHECH; HEINRICH, 2021, p. 688). A FIGURA 1 situa cronologicamente o surgimento da IA Generativa como o desdobramento mais recente dessa trajetória evolutiva:

FIGURA 1 - EVOLUÇÃO HISTÓRICA: DA INTELIGÊNCIA ARTIFICIAL À IA GENERATIVA



FONTE: Adaptado de Stryker e Kavlakoglu (2024).

Tal mudança de paradigma entre ML e DL, contudo, não é apenas técnica, mas também implica profundas implicações epistemológicas e éticas. De maneira oposta aos modelos lineares, em que decisões podem ser inspecionadas e justificadas, os modelos de redes profundas (e o que ocorre entre o *input* e o *output*) que operam com milhões ou bilhões de parâmetros ajustados automaticamente tornam-se, à medida em que aprendem por meio de camadas cada vez mais densas e opacas, mais indecifráveis e incompreensíveis pela capacidade humana (JANIESCH; ZSCHECH; HEINRICH, 2021). O resultado é que o modelo não apenas aprende padrões, mas passa a construir estruturas internas que nem mesmo seus criadores conseguem explicar; a própria

⁵ LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>.

noção de ‘entender’ como uma decisão foi tomada torna-se tecnicamente inviável. A opacidade não é um acidente, mas uma consequência direta da profundidade e da escala computacional.

O desafio interpretativo se intensifica quando consideramos o segundo pilar dos sistemas generativos: o *big data*, que representa um regime de produção e circulação de informação baseado em três elementos: volume massivo, variedade heterogênea e velocidade contínua. Esse modelo dos ‘3Vs’ foi formulado por Doug Laney (2001) para descrever os desafios de gestão de dados corporativos e, embora não utilizasse o termo *big data* na época, tornou-se posteriormente a base conceitual para a caracterização do fenômeno. O *big data* não é apenas uma descrição quantitativa, mas uma mudança cultural naquilo que consideramos conhecimento válido e verificável (BOYD; CRAWFORD, 2012). A lógica deixa de ser direcionada pelo que é teoricamente relevante e passa a ser guiada pelo que pode ser medido, acumulado e correlacionado; e o dado deixa de ser apenas registro, tornando-se matéria-prima da própria governança técnica.

Os benefícios desse regime são frequentemente destacados: maior capacidade preditiva, automação de processos, identificação rápida de padrões, otimização de decisões em larga escala. Contudo, esses mesmos fatores implicam riscos profundos. Grandes volumes de dados carregam consigo vieses históricos, sociais e raciais. Bases extensas não são neutras; são retratos do mundo como ele é, e o mundo, como sabemos, é estruturalmente desigual. Quando modelos profundos são treinados com esses dados, eles amplificam e automatizam essas desigualdades em velocidade e escala.

Além disso, o *big data* facilita o surgimento de uma falsa sensação de objetividade: quanto maior a base de dados, maior tende a ser a confiança pública nos resultados, ainda que a qualidade e a origem desses dados sejam problemáticas. Ao mesmo tempo, o próprio ato de coletar dados em escala massiva, muitas vezes sem transparência ou consentimento efetivo, constitui uma dinâmica de poder que extrapola a dimensão técnica. Para Barocas e Selbst (2016), é impossível desenvolver um algoritmo a partir do *big data* que resulte em comportamentos justos ou neutros.

Trata-se de um ambiente propício à extração sistemática de comportamentos humanos, preferências, rotinas e padrões sociais. Isso prepara o terreno para formas sofisticadas de vigilância e controle, estabelecendo as bases para o que autores posteriores chamaram de ‘capitalismo de vigilância’. Assim, a combinação entre *deep learning* e *big data* não apenas impulsiona o potencial das IAs generativas, mas também produz um novo modo de organizar conhecimento, controlar

populações e extrair valor econômico a partir de comportamentos humanos. É uma combinação poderosa, mas profundamente ambígua, que exige análise crítica constante.

2.4 PARADOXO DA CAIXA-PRETA: A OPACIDADE COMO CARACTERÍSTICA

Dessa aliança fundamental entre o *deep learning* e o *big data* como motores dos algoritmos generativos, emerge o paradoxo central mencionado: a mesma complexidade que permite que esses sistemas atinjam um desempenho inédito é também a fonte de sua principal vulnerabilidade, o efeito da ‘caixa-preta’ (*black box*). Embora possamos ver a entrada (*input*, o *prompt*) e a saída (*output*, o conteúdo gerado), o ‘como’ e o ‘porquê’ da resposta permanecem um mistério praticamente impenetrável até mesmo para os desenvolvedores (JANIESCH; ZSCHECH; HEINRICH, 2021).

A opacidade não é um defeito, mas uma característica inerente ao método. Isso levanta questões críticas sobre confiabilidade, detecção de vieses e responsabilidade por erros, posicionando o desafio da interpretabilidade como uma das fronteiras mais importantes para o futuro ético e seguro da inteligência artificial.

Contudo, a opacidade vai muito além de uma limitação técnica e configura uma dimensão política e epistemológica. O jurista estadunidense Frank Pasquale (2015), ao cunhar o termo sociedade da caixa-preta, destaca que sistemas automatizados não se tornam obscuros apenas porque são complexos, mas também porque as instituições públicas e privadas têm interesse em preservar essa obscuridade. A falta de transparência passa a operar como um recurso de poder: quanto menos compreensíveis são os critérios de decisão, menos contestáveis se tornam as decisões automatizadas.

Do ponto de vista epistêmico, a caixa-preta representa a ruptura entre o raciocínio humano e o algorítmico. Modelos profundos criam representações matemáticas internas impossíveis de traduzir diretamente em termos humanos: tensores, pesos, vetores de atenção e arquiteturas de camadas se combinam de maneiras que ultrapassam a capacidade cognitiva humana de rastrear causalidades. Como observa Burrell (2015), existem três formas fundamentais de opacidade:

I. Opacidade intencional: quando empresas mantêm o funcionamento de seus modelos sob sigilo corporativo;

II. Opacidade técnica: quando a arquitetura do modelo é tão complexa que se torna indecifrável;

III. Opacidade decorrente da escala: quando a quantidade de dados e parâmetros gera um volume de interações impossível de explicar individualmente.

Nos modelos generativos contemporâneos, as três formas coexistem. Além disso, há o problema central do aprendizado profundo de que os parâmetros internos não representam diretamente conceitos humanos. Uma rede neural não ‘entende’ o mundo, mas aproxima relações numéricas que funcionam bem no seu espaço estatístico. Quando observamos um resultado discriminatório, a falta de interpretabilidade impede que possamos demonstrar causalidade, identificar responsabilidades ou corrigir o erro com precisão.

Para agravar a situação, a geração contínua de dados alimenta a própria caixa-preta. À medida que modelos são treinados com dados produzidos por outros modelos, o risco de erros acumularem-se e tornarem-se invisíveis aumenta, fenômeno identificado por Shumailov et al. (2024) como *model collapse* (colapso do modelo, em tradução livre). Em outras palavras, a caixa-preta não é estática, mas, sim, cresce de maneira exponencial, tornando-se mais densa, mais opaca e menos auditável.

A opacidade estrutural também limita os mecanismos jurídicos e regulatórios tradicionais. Como responsabilizar uma empresa por uma decisão algorítmica se nem a própria empresa consegue explicá-la? Como exigir transparência se a transparência é tecnicamente impossível? Na prática, essa tensão gera um déficit democrático. Sem a capacidade de compreender e questionar decisões algorítmicas, usuários, cidadãos e até instituições públicas ficam à mercê de sistemas que operam segundo lógicas opacas e potencialmente discriminatórias. A caixa-preta cria assimetrias profundas: quem desenvolve e controla os algoritmos detém vantagem informacional, enquanto aqueles afetados por eles, que raramente têm conhecimento disso, permanecem sem meios de contestação (O’NEIL, 2020).

Em um cenário em que decisões automatizadas influenciam desde diagnósticos médicos até regras de crédito, contratações, vigilância pública e segurança, essa assimetria torna-se um risco sistêmico para direitos fundamentais. Com isso, é evidente que a opacidade não pode ser tratada apenas como uma limitação científica, mas como uma característica central da governança algorítmica contemporânea, com implicações diretas nas dinâmicas de poder e nas formas de discriminação digital.

Assim, compreender a opacidade algorítmica implica reconhecer que ela integra uma arquitetura mais ampla, sendo instrumentalizada em modelos de extração e controle baseados em

dados: o capitalismo de vigilância, estrutura econômica que transforma o *big data* e os próprios sistemas algorítmicos em mecanismos de acumulação de poder e lucro (ZUBOFF, 2021). É nesse encontro entre ocultamento técnico e exploração informacional que sistemas automatizados ganham força institucional e passam a moldar, silenciosamente, práticas sociais e decisões que afetam direitos, oportunidades e formas de vida.

2.5 CAPITALISMO DE VIGILÂNCIA E ECONOMIA DOS DADOS

A era do *big data* transformou radicalmente a forma como as experiências humanas são capturadas, interpretadas e convertidas em valor econômico. Assim que empresas de tecnologia perceberam o potencial lucrativo dos rastros digitais - cliques, deslocamentos, preferências, interações sociais e até expressões faciais - a coleta massiva de dados tornou-se infraestrutura da vida cotidiana por estratégia. Segundo relatório da consultoria Burton-Taylor (MARKETS MEDIA, 2024), dados e notícias do setor financeiro movimentaram cerca de 42 bilhões de dólares apenas em 2023. O número não revela apenas uma área aquecida; indica que a informação deixou de desempenhar um papel auxiliar para assumir posição central nas formas contemporâneas de acumulação e de poder.

É nesse cenário que a economista norte-americana Shoshana Zuboff (2021) apresenta o conceito de capitalismo de vigilância na obra *A Era do Capitalismo de Vigilância*. Para a socióloga, trata-se de um regime econômico que “reivindica a experiência humana como matéria-prima gratuita para práticas comerciais dissimuladas de extração, previsão e vendas” (ZUBOFF, 2021, p. 14). Aqui, o objetivo é reorganizar a vida de modo a torná-la permanentemente legível e útil para modelos que buscam antecipar (e intervir) comportamentos. O dado se torna um recurso estruturado, contínuo e explorável, transformando relações cotidianas em material preditivo.

Tal expansão é possível porque, como argumenta o sociólogo David Lyon (2007), vivemos hoje em um ‘clima de vigilância’, no qual o monitoramento não é mais uma exceção, mas um estado permanente. Smartphones, redes sociais, câmeras urbanas, plataformas de trabalho, aplicativos bancários e assistentes virtuais produzem, em tempo real, um fluxo contínuo de dados que reconfigura a relação entre indivíduos e instituições, gerando uma vigilância distribuída, quase imperceptível, mas constante.

Mark Andrejevic (2019) descreve esse fenômeno como a transição para uma cultura de ‘captura total’: tudo o que pode ser convertido em dado será. O processo não ocorre apenas porque há interesse econômico, mas porque a própria lógica das plataformas é orientada à extração contínua. Cada ação humana se converte em insumo para sistemas automáticos que interpretam, comparam e classificam indivíduos sem que eles tenham consciência disso. O resultado é um desequilíbrio informacional profundo em que pessoas tornam-se plenamente visíveis para plataformas e governos, enquanto permanecemos incapazes de ver e compreender como somos analisados.

Em outra esfera, esse processo também afeta a própria noção de identidade. Cheney-Lippold (2017) demonstra como algoritmos fabricam identidades probabilísticas, classificando pessoas com base em padrões correlacionais sugeridos pelos dados, e não por autoidentificação. Ao sermos transformados em perfis, *scores* e categorias preditivas, as classificações moldam oportunidades de emprego, acesso a crédito, anúncios recebidos, ofertas de consumo e até políticas de segurança. A subjetividade humana é substituída e as etiquetas algorítmicas ganham prioridade sobre as narrativas pessoais.

A concentração informacional só é viabilizada naquilo que Kai-Fu Lee (2018) denomina ‘Era dos Dados’, momento em que o sucesso dos algoritmos depende de três fatores: volume de dados, poder computacional e engenheiros capacitados. Dentre esses elementos, Lee destaca o primeiro como decisivo. À medida que os sistemas se expandem, informações realmente úteis tornam-se mais raras e, portanto, mais valiosas; quanto mais dados proveitosos uma empresa acumula, maior sua capacidade de prever, influenciar e moldar comportamentos. Esse modelo de acumulação cria desigualdades políticas e sociais porque as plataformas que concentram grandes volumes de dados ganham poder para influenciar comportamentos individuais, mercados, instituições públicas e até mesmo estruturas democráticas. A concentração torna-se um elemento central da própria discriminação algorítmica ao viabilizar a substituição das formas democráticas de avaliação e mediação social por categorias estatísticas.

Desse modo, o capitalismo de vigilância não é apenas um fenômeno tecnológico, mas também um modo de organização social que converte experiências humanas em fontes de lucro e em mecanismos de controle. Ele estrutura desigualdades ao:

- transformar grupos historicamente marginalizados em objetos preferenciais de vigilância;

- criar categorias algorítmicas que reforçam estereótipos;
- alimentar sistemas preditivos que reproduzem ciclos de exclusão;
- concentrar poder informacional nas mãos de empresas capazes de moldar comportamentos;
- obscurecer processos decisórios por meio de opacidade técnica e jurídica.

Em síntese, a extração de dados é o alicerce que sustenta tanto o capitalismo de vigilância quanto a discriminação algorítmica estrutural. A coleta massiva, quando combinada à opacidade dos modelos e às desigualdades históricas, produz sistemas que refletem, ampliam, automatizam e legitimam as injustiças sociais sob a aparência de neutralidade e de autoridade técnica, raramente questionada, consolidando um sistema no qual as assimetrias do passado são reescritas como previsões confiáveis.

3 CODIFICAÇÃO DA DESIGUALDADE

A incorporação de sistemas algorítmicos em diversas esferas da vida social produz uma reorganização silenciosa dos modos de classificar, avaliar e distinguir indivíduos. A aparência técnica dessas operações costuma ocultar que se apoiam em estruturas profundamente marcadas pela desigualdade. Mais do que reproduzir padrões existentes, os algoritmos estabilizam interpretações sobre o social e as transformam em critérios operacionais, convertendo disputas históricas em cálculos que parecem neutros.

Desse modo, práticas antes reconhecidas como escolhas políticas passam a ser apresentadas como simples consequência de modelos estatísticos. É nessa transição que a desigualdade se adensa: quando deixa de ser percebida como construção social e passa a circular como efeito lógico de um sistema que afirma apenas ‘processar dados’. O prestígio atribuído a esse processamento - e à suposta racionalidade que ele carrega - funciona como uma camada adicional de proteção a esses sistemas, dificultando o questionamento público e reforçando a naturalização dos resultados.

Entre essas esferas, o próprio campo da comunicação e do jornalismo passa a ser organizado por lógicas de pontuação, ranqueamento e filtragem automática, em que critérios de noticiabilidade, disputas editoriais e decisões históricas de cobertura são reescritas como cálculos de relevância e engajamento definidos por sistemas opacos.

3.1 TECNOCHAUVINISMO E A CRENÇA NA OBJETIVIDADE MATEMÁTICA

A noção de que sistemas computacionais oferecem uma visão mais precisa do mundo sustenta uma parte importante das expectativas depositadas na inteligência artificial. Essa confiança repousa sobre uma leitura particular da matemática, entendida como linguagem desinteressada e, portanto, moralmente superior às avaliações humanas. A pesquisadora e jornalista de dados estadunidense Meredith Broussard (2018) observa que esse prestígio decorre de uma crença generalizada na capacidade do cálculo de revelar ‘verdades objetivas’, como se números pudessem eliminar ambiguidades e conflitos presentes em qualquer processo social. O resultado é uma valorização automática de soluções computacionais, tratadas como se carregassem, por si mesmas, uma imparcialidade intrínseca.

Essa expectativa sustenta o que a autora denomina *tecnochauvinismo*: a tendência a considerar o digital como alternativa sempre melhor, mais eficiente e mais justa. O problema não

está apenas no otimismo tecnológico cego, mas também na falta de cautela na sua aplicação e na forma como esse otimismo reorganiza hierarquias de autoridade. Quando se assume que o cálculo é melhor do que o julgamento, as decisões produzidas por modelos ganham um peso desproporcional e passam a ser tratadas como conclusões inevitáveis e à prova de falhas. Essa inevitabilidade é perigosa e enganosa. Broussard (2018) argumenta que os dados são socialmente construídos e, mesmo ao interagir com um computador, lidamos com um programa escrito por um ser humano - alguém com seus próprios contextos, pressupostos e vieses. Algoritmos não observam o mundo; apenas recombina registros produzidos por instituições cuja história é marcada por desigualdades profundas.

Há aqui uma dimensão crítica que não pode ser ignorada. A crença na objetividade matemática deixa de operar apenas como suposição epistemológica e passa a atuar como mecanismo de redistribuição de responsabilidades. Quando o modelo produz um resultado, a tendência é deslocar a discussão do campo das escolhas sociais para o das supostas limitações técnicas. O cálculo, ao ser apresentado como mero processamento, ganha uma autoridade que não deveria ter, justamente porque não foi concebido para interpretar a complexidade da vida social, mas para simplificar. Essa simplificação produz conforto político: torna mais fácil aceitar assimetrias como efeitos naturais de um sistema que apenas ‘funciona’. O cálculo não se torna moral por conta própria, mas passa a ser tratado como tal sempre que sua saída é recebida como verdade. É nesse ponto que a técnica deixa de parecer uma ferramenta e passa a operar como referência normativa.

O *tecnoc Chauvinismo* contemporâneo emerge desse cenário, herdando a crença na superioridade moral, cognitiva e política da tecnologia e transformando a inovação digital em solução genérica para problemas sociais, sustentada pelo imaginário que desloca conflitos sociais para fora do âmbito da discussão pública e os apresenta como superáveis pela inovação técnica. Em vez de promover inclusão, isso estrutura uma cultura que traduz desigualdades em questões de design e eficiência, restringindo o alcance das discussões democráticas. Nas redações jornalísticas, essa cultura se manifesta quando métricas automatizadas de audiência, SEO (*Search Engine Optimization*, ou Otimização para Mecanismos de Busca, em tradução livre) e engajamento são tratados como diagnósticos incontestáveis do interesse do público, transpondo os critérios editoriais para o terreno de dashboards e relatórios algorítmicos, em vez de decisões jornalísticas abertamente discutidas.

3.1.1 Raízes da crença na neutralidade digital

A crença contemporânea de que a tecnologia é capaz de operar acima das tensões sociais tem origens que antecedem a própria inteligência artificial. Durante os anos 1990, momento em que a internet comercial se consolidava e o Vale do Silício se afirmava como centro simbólico da inovação, formou-se um imaginário no qual o digital se apresentava como um espaço destinado a superar as limitações do mundo material. A ideia de que o ciberespaço permitiria interações livres de corpo, raça, gênero ou classe tornou-se um elemento estruturante desse entusiasmo. Surgia, assim, a promessa da descorporificação: um ambiente em que as diferenças históricas perderiam relevância e as identidades poderiam se reorganizar em torno da vontade individual, e não das marcas sociais que moldam a vida offline.

O manifesto mais emblemático desse imaginário é a Declaração de Independência do Ciberespaço, publicada por John Perry Barlow em 1996. Ali, a internet é apresentada como “uma nova casa da Mente”, livre da legislação e das hierarquias do “mundo da carne e do aço” (BARLOW, 1996, s.p.). A formulação não apenas rejeita a interferência estatal, mas também sugere que o ambiente digital seria capaz de sustentar uma forma superior de convivência, amparada na racionalidade e na autonomia dos indivíduos. Sob essa visão, conflitos raciais, desigualdades de gênero e disputas por poder seriam fenômenos restritos ao mundo físico; o ciberespaço surgia como um território onde essas estruturas perderiam sua força.

Essa concepção está no cerne do que Richard Barbrook e Andy Cameron (1996) denominaram *Ideologia Californiana*. Para os autores, a visão de mundo que orientou o ecossistema tecnológico nascente resulta da fusão entre o individualismo da contracultura dos anos 1960 e o neoliberalismo pró-mercado da Nova Direita. A criatividade do programador, o espírito empreendedor e a confiança no mercado aparecem como motores legítimos do progresso, enquanto a intervenção estatal, a regulação democrática e a crítica social são tratadas como entraves. Assim, a tecnologia é elevada à condição de força civilizatória, capaz de impulsionar transformações sociais por seu mérito intrínseco.

A utopia da descorporificação encontrou limites rapidamente. Como analisou Lisa Nakamura (2002), a promessa de interação ‘sem corpo’ não eliminou as desigualdades; apenas deslocou suas formas. No ambiente digital, prevaleceu a figura implícita do ‘usuário universal’: um sujeito branco, anglófono, masculino e ocidental, cuja perspectiva passou a orientar interfaces,

linguagens e modos de participação. Para a professora, esse enquadramento produziu os *cybertypes*, padrões de representação racial e cultural que se expressam tanto em conteúdos quanto nas expectativas de comportamento, mostrando que a internet não se desvinculou das hierarquias do mundo físico - ao contrário, reconfigurou-as.

Pesquisadores brasileiros ampliam essa crítica ao evidenciar como o mito da neutralidade digital se articulou ao contexto socioeconômico local. Pesquisador focado em governança e no estudo dos impactos sociais de tecnologias digitais, como a internet e a inteligência artificial, Tarcízio Silva (2022) aponta que a ausência de diversidade entre desenvolvedores, pesquisadores e gestores do setor tecnológico reforçou a ideia de que as plataformas digitais seriam naturalmente neutras. Essa ausência faz parte das condições que permitiram a consolidação de um ambiente em que o corpo parece desaparecer, mas apenas para reforçar como norma aqueles que nunca foram marcados pela diferença. Assim, a promessa de uma rede ‘sem raça’ ou ‘sem gênero’ tornou-se conveniente para negligenciar desigualdades de acesso, violência simbólica e padrões discriminatórios que atravessam a vida online.

Essas narrativas se tornaram tão influentes que moldaram a forma como o setor tecnológico passou a apresentar seus produtos. Sob a ideia de que dados e códigos operam acima do conflito social, a cultura do Vale do Silício consolidou a convicção de que problemas coletivos poderiam ser resolvidos por meio de soluções computacionais. A crítica de Morozov (2013) ao solucionismo tecnológico conecta-se a esse ponto: a crença de que processos complexos podem ser tratados como questões de design reforça a autoridade de engenheiros e empreendedores, ao mesmo tempo em que desloca disputas sobre desigualdade, poder e justiça para o domínio exclusivo da técnica.

O que emerge desse processo é um conjunto de expectativas que naturalizam a autoridade da tecnologia. A promessa da descorporificação legitima a ideia de que sistemas digitais nascem desvinculados das estruturas sociais; a Ideologia Californiana projeta a tecnologia como caminho inevitável do progresso; os *cybertypes* evidenciam que esse ideal nunca se realizou; e o solucionismo reforça a convicção de que conflitos podem ser resolvidos pela lógica computacional. A crença contemporânea na objetividade matemática alimenta-se exatamente desse legado. Antes de ser um traço técnico, ela é uma construção cultural que preparou terreno para que a inteligência artificial aparecesse como árbitro neutro de decisões que são, desde a origem, profundamente sociais.

3.2 DETERMINISMO TECNOLÓGICO E DESPOLITIZAÇÃO DA TÉCNICA

As expectativas de neutralidade que moldaram o imaginário digital também influenciaram a forma como a tecnologia passou a ser compreendida como instância racional, capaz de operar acima das disputas sociais. A ideia de que sistemas computacionais funcionam segundo uma lógica própria, imunes a condicionamentos históricos e institucionais, está atrelada ao determinismo tecnológico e tem desempenhado um papel central na forma como as decisões algorítmicas são justificadas. O efeito disso é que decisões humanas - o desenho de modelos: escolhas sobre quais dados coletar, como classificá-los ou que resultados considerar desejáveis - são apresentadas como simples exigências técnicas, desprovidas de alternativa.

Conforme esse enquadramento se consolida, conflitos que deveriam ser debatidos coletivamente são deslocados para dentro da arquitetura dos sistemas, protegidos por uma linguagem que parece neutra, mas que carrega prioridades e valores definidos por quem tem poder de projetar e implementar essas ferramentas. Dagnino (2010) observa que tal deslocamento não decorre de ingenuidade metodológica, mas de um processo sistemático cirurgicamente intencionado de neutralização das tensões sociais, que passam a ser tratadas como questões de eficiência ou de aprimoramento técnico.

A confiança excessiva na tecnologia também produz uma sensação de inevitabilidade. Nesse ambiente, quando sistemas computacionais são apresentados como soluções ‘ótimas’ ou ‘precisas’, seus resultados passam a ser tomados como interpretações mais fiéis da realidade. Esse prestígio se sustenta na forma como a técnica incorpora valores específicos sem reconhecê-los como tais, reduzindo escolhas políticas a parâmetros de modelagem, como se debates sobre o modo de governar populações pudessem ser traduzidos em ajustes de performance (DAGNINO, 2010). No debate sobre a automatização do jornalismo, o determinismo tecnológico aparece justamente na promessa de que as redações precisam adotar sistemas de inteligência artificial (IA) para sobreviver, como se não existisse alternativa política senão adaptar rotinas e critérios editoriais às exigências de plataformas e fornecedores tecnológicos.

Com isso, decisões políticas codificadas em linguagem técnica passam a ser envolvidas pelo vocabulário da eficiência, e seus efeitos sociais são percebidos como consequências naturais do funcionamento de sistemas supostamente racionais. Pasquale (2015) mostra como esses sistemas opacos oferecem respaldo e consolo institucionais ao fornecer justificativas técnicas para

escolhas que, antes, exigiam defesa explícita e debate público. Em vez de assumir a responsabilidade por decisões controversas, organizações e empresas passam a invocar as limitações ou as ‘exigências’ do modelo.

A automatização produz conforto porque desloca a responsabilidade: se o algoritmo é visto como instância objetiva, o julgamento humano deixa de ser questionado e a própria noção de erro se transforma, passando a ser enquadrada como um problema de otimização. O que deveria ser reconhecido como escolha passa a ser apresentado como cálculo. A partir desse enquadramento, disputas por interpretação sobre risco, mérito, prioridade ou perigo são transformadas em resultados aparentemente evidentes, ainda que produzidos por modelos que carregam pressupostos muito específicos.

Esse esvaziamento político é premeditado e cria um ambiente em que discriminações estruturais podem ser reconfiguradas como falhas de dados ou ruído estatístico, desconsiderando a origem dessas assimetrias e enfraquecendo a crítica. Contudo, quando desigualdades persistentes são tratadas apenas como desvios ou anomalias a serem corrigidas pelo desempenho algorítmico, perde-se de vista a estrutura que as produz. É nesse ponto que a leitura crítica se torna indispensável. A técnica nunca atua sozinha. Sempre há um conjunto de forças - econômicas, políticas, epistemológicas - que definem seus objetivos e efeitos.

A crítica ao determinismo tecnológico, portanto, não rejeita a tecnologia, mas sim o mito de que ela seja neutra. Recuperar a dimensão política da técnica significa reconhecer que sistemas automatizados não são construídos em ambientes neutros, mas sim em estruturas sociais discriminatórias. A crença no caráter autossuficiente da tecnologia funciona como dispositivo que desvia o olhar do que realmente está em disputa, permitindo que conflitos sociais sejam traduzidos em parâmetros e métricas que escapam à deliberação pública.

Essa operação de neutralizar tensões ganha contornos ainda mais claros em situações concretas em que os limites técnicos são tratados como detalhes menores. A discussão sobre veículos autônomos (autodirigíveis) é um desses casos. Broussard (2018) investiga como as falhas reais da tecnologia são tratadas na pauta pública, mostrando que a narrativa do ‘progresso inevitável’ convive com limitações frequentemente ignoradas pela indústria e pela cobertura jornalística. Nesses casos, os sistemas não fracassam por imprevisibilidade humana ou comportamento inadequado dos usuários, mas por causa da incapacidade dos modelos de antecipar a complexidade das situações cotidianas:

A vida é, por definição, imprevisível. É impossível para programadores antecipar todas as situações problemáticas ou surpreendentes que podem surgir, o que significa que os sistemas de *machine learning* existentes continuam suscetíveis a falhas quando se deparam com as irregularidades e a imprevisibilidade das circunstâncias do mundo real. [...] Essa abordagem simplesmente não é escalável.⁶ (DARPA, 2017, s.p., tradução nossa)

Essa evidência técnica contrasta com a forma como a discussão costuma ser conduzida. Muitas empresas e entusiastas insistem em atribuir falhas a indivíduos - motoristas que não compreenderam o sistema, pedestres imprevisíveis, usuários que não seguiram as instruções. Esse deslocamento simplifica o problema, evita discutir limitações estruturais e protege a crença de que a tecnologia está sempre prestes a funcionar plenamente, mesmo quando seus próprios desenvolvedores reconhecem barreiras técnicas difíceis de contornar.

O debate público também se apoia repetidamente nos mesmos indicadores estatísticos de taxas de acidentes, estimativas de precisão e projeções de adoção, que parecem comprovar a superioridade dos sistemas automatizados. Entretanto, é essencial o alerta de que números que surgem sem explicação de origem e aparecem incessantemente merecem atenção redobrada, pois geralmente vêm de fontes escassas com interesses muito bem definidos. Essa recorrência artificial ecoa o alerta clássico de Darrell Huff (2019), em *Como Mentir com Estatística*, de que as estatísticas podem funcionar como instrumentos de persuasão, manipulação e camuflagem quando descoladas do contexto. Broussard (2018, p. 149, tradução nossa) sintetiza a falácia em uma frase irônica, após analisar o uso distorcido de comparações entre erros humanos e falhas técnicas: “É claro que os humanos são responsáveis pelos erros de direção. Os humanos são os únicos que dirigem carros!”⁷. A observação desmonta a falsa equivalência que sustenta grande parte da retórica pró-automação.

A crítica final da escritora articula uma alternativa que contraria o imaginário dominante da automação total. Em vez de insistir em projetos que buscam substituir o julgamento humano (quando os próprios especialistas reconhecem que os sistemas não conseguem operar de forma autônoma), a autora sustenta que “deveríamos realmente focar em criar sistemas de assistência

⁶ “Life is by definition unpredictable. It is impossible for programmers to anticipate every problematic or surprising situation that might arise, which means existing ML systems remain susceptible to failures as they encounter the irregularities and unpredictability of real-world circumstances. [...] This approach is just not scalable.”

⁷ “Of course humans are responsible for driving errors. Humans are the only ones driving cars!”

humana, e não sistemas de substituição humana”⁸ (BROUSSARD, 2018, p. 159, tradução nossa). Não se trata de negar avanços tecnológicos, mas de questionar para quais finalidades são mobilizados, quando são realmente necessários e quem se beneficia dessa narrativa.

Sendo assim, é possível compreender que o determinismo tecnológico não se sustenta apenas em teorias ou discursos entusiasmados, mas também em uma operação contínua. A despolitização da técnica opera como uma forma específica de organização do poder. Ao substituir o debate pela otimização; a escolha pelo cálculo; a política pela tecnologia; e a responsabilidade pela figura abstrata do sistema, a autoridade simbólica da tecnologia é preservada mesmo quando seu funcionamento real evidencia fragilidades profundas e essa despolitização consolida-se, produzindo seus efeitos mais nefastos e duradouros.

3.2.1 O mito da neutralidade tecnológica

As interpretações sobre determinismo tecnológico e a relação entre ciência, tecnologia e sociedade tornam-se mais nítidas com as ideias apresentadas pelo professor do Departamento de Política Científica e Tecnológica da Unicamp Renato Dagnino (2010) em sua obra *Neutralidade da ciência e determinismo tecnológico*, em que propõe uma classificação de como os Estudos Sociais da Ciência e Tecnologia (ESCT) abordam esse relacionamento. Tal sistematização ajuda a compreender por que certos discursos ainda insistem em apresentar tecnologias como esferas separadas do conflito social.

Segundo o autor, é possível distinguir duas grandes categorias, cada uma subdividida em duas variantes. A primeira categoria, centrada na própria produção científica e tecnológica, entende que esse campo se desenvolve de maneira autônoma e linear em busca da verdade, numa lógica evolucionista em que sobrevivem as soluções mais aptas, eficientes e produtivas. Nesse cenário, o desenvolvimento técnico é visto como dotado de uma barreira que protegeria “o ambiente de produção científico-tecnológica do contexto social, político e econômico, que impediria que os interesses dos atores sociais envolvidos no desenvolvimento da C&T pudessem determinar a trajetória de inovação” (DAGNINO, 2010, p. 35).

⁸ “We should really focus on making human-assistance systems instead of on making human-replacement systems.”

Dentro dessa perspectiva, a primeira subdivisão, da neutralidade, concebe essa barreira como impermeável em ambos os sentidos: a tecnologia não influenciaria nem seria influenciada pela sociedade, permanecendo completamente independente. Já a segunda subdivisão, justamente a do determinismo tecnológico, entende essa barreira como “uma espécie de membrana impermeável no sentido da sociedade para a C&T, mas não no sentido contrário” (DAGNINO, 2010, p. 36). Aqui, o avanço técnico é visto como essencial para o desenvolvimento econômico, enquanto a tecnologia assume o papel de força condutora e determinante da estrutura social.

A segunda grande categoria, por sua vez, desloca o foco para a sociedade. Nela, distinguem-se duas abordagens que negam a neutralidade da tecnologia: a Tese Fraca e a Tese Forte. A primeira defende que o desenvolvimento científico-tecnológico é moldado pela sociedade, refletindo suas relações sociais, econômicas e políticas. Consequentemente, o conhecimento produzido incorpora características do contexto em que surge e adquire uma funcionalidade que contribui para a reprodução e continuidade desse mesmo contexto.

Para compreender de forma mais profunda como a tecnologia atua na manutenção das estruturas de poder, a Tese Forte da Não-Neutralidade dá um passo além: a C&T, uma vez moldada para servir a ordem social em que foi criada, permanece intrinsecamente comprometida com a preservação dessa ordem. Assim, tornam-se instrumentos de reprodução do sistema capitalista, assegurando a manutenção e a replicação de privilégios e desigualdades historicamente estabelecidos (DAGNINO, 2010).

Como síntese visual dessas distinções teóricas, a FIGURA 2 estrutura as quatro abordagens fundamentais descritas pelo autor, destacando visualmente a transição da crença na neutralidade (onde há uma barreira ou um fluxo unilateral) para a perspectiva crítica da não-neutralidade, na qual a tecnologia é compreendida não apenas como reflexo do meio, mas como elemento ativo na reprodução de privilégios e na manutenção da ordem social:

FIGURA 2 - AS QUATRO ABORDAGENS DA RELAÇÃO CIÊNCIA, TECNOLOGIA E SOCIEDADE

Neutralidade

ênfase na ciência e tecnologia

Não-neutralidade

ênfase na sociedade

Neutralidade da C&T

- autonomia
- barreira impermeável
- sem influência social



Tese fraca

- contexto social influencia
- C&T moldada pela sociedade
- reprodução das relações



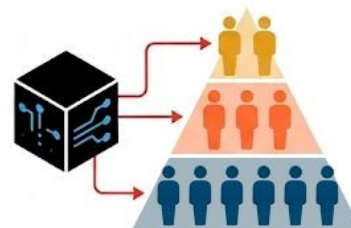
Determinismo tecnológico

- C&T > sociedade
- desenvolvimento econômico
- progresso linear



Tese forte

- C&T a serviço do capital
- manutenção da ordem
- reprodução de privilégios



FONTE: Adaptada de Dagnino (2010).

Esse enquadramento teórico permite compreender com mais precisão como a discriminação algorítmica se vincula a processos sociais mais amplos. Quando modelos automatizados incorporam padrões de desigualdade e os apresentam como resultados estatísticos legítimos, operam exatamente no sentido descrito pela Tese Forte. A tecnologia deixa de ser apenas um reflexo da sociedade e passa a atuar ativamente na reprodução da sociedade. Reconhecer essa não-neutralidade significa entender que a técnica não flutua acima do conflito social: ela participa, consolida e estabiliza lógicas que antecedem o algoritmo. Nesse contexto, a construção de uma sociedade mais justa requer não apenas transformações sociais, mas também uma reconfiguração profunda da própria base tecnológica que sustenta, legitima e reproduz essas relações de poder.

3.3 MATEMÁTICA COMO MORAL: QUANDO OS CÁLCULOS VIRAM DESTINO

A incorporação de sistemas algorítmicos no cotidiano transformou categorias numéricas em categorias sociais. À medida que métricas passam a classificar indivíduos em termos de risco, mérito ou relevância, a matemática deixa de ser apenas um método de análise e se torna um modo

de produzir interpretações sobre quem as pessoas ‘são’. Esse movimento pode ser descrito como a formação de tipos interativos, grupos que surgem quando classificações passam a moldar a forma como indivíduos são tratados por instituições. Não se trata apenas de descrever populações, mas de reorganizar formas de vida com base em parâmetros estatísticos.

Uma pontuação numérica passa, então, a circular como critério moral que influencia decisões cotidianas, pois os números carregam um prestígio que lhes permite definir quem deve ser monitorado, quem merece oportunidade ou quem representa ameaça. Cheney-Lippold (2017) chama esse fenômeno de identidades algorítmicas, categorias formadas não por experiências vividas, mas por padrões inferidos a partir de dados.

O efeito institucional desse processo se manifesta em diversas políticas públicas. Na educação, modelos que estimam a probabilidade de evasão orientam as prioridades de intervenção. Na segurança pública, sistemas de análise histórica ampliam o patrulhamento em determinadas regiões. No crédito, as métricas de perfil financeiro definem o acesso a financiamentos e moldam as expectativas sobre quem é considerado elegível. Em programas sociais, sistemas automatizados segmentam beneficiários com base em padrões estatísticos e orientam decisões administrativas. Em todos esses casos, as métricas se tornam critérios estruturantes de ação. No ecossistema informacional, lógicas semelhantes estruturam o que é considerado notícia relevante: algoritmos de recomendação, *trending topics* e sistemas de ranqueamento em buscadores distribuem visibilidade com base em pontuações que combinam previsões de engajamento e interesses comerciais, reconfigurando silenciosamente o que chega ao público e o que permanece invisível.

Esses mecanismos ajudam a entender como padrões históricos se transformam em previsões que orientam o futuro de indivíduos e grupos. Quando desigualdades acumuladas são tratadas como variáveis que antecipam comportamentos, o cálculo converte esse passado em previsões individuais, como se fossem características intrínsecas das pessoas. Bowker e Star (1999) observam que as classificações moldam rotinas e práticas institucionais ao organizarem o que deve ser visto, registrado e priorizado.

A matemática, nesse contexto, participa da fabricação do destino ao orientar decisões futuras. Quando modelos traduzem vidas em categorias calculadas e essas categorias retornam para moldar comportamentos, políticas e oportunidades, o número ganha uma autoridade que reorganiza o mundo social. Um exemplo simples é o de “um desempregado de longa data, que descobre que

poucos o contratarão porque está sem emprego há tempo demais. É um ciclo de *feedback* em espiral, frustrante para os azarados que ficam presos nele” (O’NEIL, 2020, p. 139-140).

A crítica, portanto, não está na quantificação em si, mas na forma como ela transforma diferenças históricas em argumentos técnicos e expectativas morais. Reconhecer esse processo é fundamental para compreender por que decisões algorítmicas tendem a estabilizar desigualdades: elas tratam padrões herdados como se fossem destinos inevitáveis, quando são apenas escolhas convertidas em cálculos.

3.4 AUTORIDADE AUTOMÁTICA E DUPLA OPACIDADE

À medida que sistemas algorítmicos passam a estruturar setores inteiros da administração pública e privada, consolidam uma forma particular de autoridade sustentada pela pouca transparência de seus mecanismos internos. Em vez de gerar cautela, a dificuldade de compreender como um modelo opera abre espaço para novas hierarquias institucionais baseadas na deferência técnica. A opacidade deixa de ser percebida como limitação e passa a funcionar como selo de especialização, sugerindo que quanto menos acessível for o funcionamento do algoritmo, mais legítimo ele tende a parecer.

Em ambientes organizacionais, essa dinâmica altera rotinas e hierarquias. Relatórios e classificações produzidos por sistemas automatizados passam a orientar decisões, mesmo quando seus critérios não são conhecidos ou compreendidos pela maioria dos profissionais (BRAYNE, 2021). A especialização exigida para interpretar modelos complexos restringe quem pode participar das deliberações, redefinindo o que conta como justificativa legítima. Quando poucos dominam a linguagem do modelo, a discussão se desloca do campo da negociação para o domínio da interpretação técnica, o que transforma o debate institucional em uma consulta a um resultado. No jornalismo, isso se traduz em profissionais que passam a receber listas prontas de pautas ‘quentes’, títulos ‘otimizados’ e perfis ‘promissores’ de audiência gerados por softwares proprietários, de modo que decisões editoriais vão sendo pouco a pouco deslocadas para interfaces de sistema cujo funcionamento raramente é transparente para a própria redação.

Além disso, à medida que essa reconfiguração interna avança, as escolhas inscritas na modelagem ganham aparência de objetividade. O volume de dados e a velocidade de processamento produzem uma falsa impressão de rigor. A questão é que essa suposta precisão

depende de escolhas feitas muito antes da execução do modelo que raramente se tornam públicas, reordenando silenciosamente o que é possível contestar (O'NEIL, 2020).

Ao negar sua própria materialidade racializada e parcial, as tecnologias digitais instauram aquilo que Tarcízio Silva (2019; 2020) denomina dupla opacidade. A primeira, técnica, decorre da complexidade dos sistemas - códigos proprietários, múltiplas camadas de processamento, modelos importados sem documentação adequada. A segunda, política, silencia as condições sociais que deram origem aos dados utilizados. Ambas se reforçam mutuamente: enquanto a opacidade técnica impede o acompanhamento do funcionamento interno, a opacidade política impede questionar por que determinados padrões aparecem nos dados e como foram produzidos nas instituições que os alimentam.

Essa combinação se torna evidente ao analisarmos processos que atravessam o cotidiano brasileiro. Nos sistemas de crédito, por exemplo, modelos que classificam a capacidade de pagamento raramente tornam visíveis os critérios que definem as escalas de risco. No Cadastro Único, ajustes em pesos e cruzamentos entre bases de dados redefinem perfis de elegibilidade sem que a maior parte dos usuários compreenda as razões por trás da mudança. Nas plataformas financeiras, pequenas variações no histórico bancário, endereço ou ocupação podem alterar pontuações de forma decisiva, embora permaneçam inacessíveis para quem é afetado. Trata-se de uma estrutura que impede que indivíduos acompanhem o percurso que leva ao resultado e, simultaneamente, dilui a responsabilidade institucional.

Em paralelo, a autoridade automática reorganiza a forma como as instituições percebem risco, mérito e prioridade. Quando resultados algorítmicos são tratados como leituras 'superiores' do mundo social, deixam de funcionar como projeções e passam a operar como normas. A ausência de visibilidade sobre os fundamentos desses cálculos dificulta questionamentos mais amplos: quem definiu o que conta como risco? Que tipo de comportamento é premiado? Quais grupos se tornam foco de vigilância ou de intervenção?

Assim, a autoridade dos algoritmos não deriva apenas de sua sofisticação técnica, mas também da combinação entre obscuridade, dependência institucional, interesses econômicos e estruturas históricas que moldam os dados utilizados. A convergência entre opacidade técnica e opacidade política produz um cenário em que decisões automatizadas circulam com legitimidade ampliada, mesmo quando se apoiam em critérios que permanecem invisíveis para aqueles que deveriam poder debatê-los. É nesse terreno, opaco, onde decisões se naturalizam e

responsabilidades se dispersam, que os mecanismos estruturais do viés algorítmico entram em cena.

3.5 A CONSTITUIÇÃO DO VIÉS ALGORÍTMICO

A consolidação de sistemas algorítmicos nos campos da gestão pública, da segurança, do crédito, da educação e da organização do trabalho ampliou a transformação das desigualdades históricas em decisões automatizadas. Esse processo não se inicia com a tecnologia, mas encontra nela um ambiente de amplificação eficiente. Em sociedades marcadas por assimetrias persistentes, como o Brasil, os modelos reproduzem padrões acumulados e contribuem para sua reorganização e endurecimento. O viés algorítmico, portanto, não corresponde a um defeito técnico, mas a uma modernização da discriminação estrutural mediada por modelos matemáticos e executada por instituições que já operavam seletivamente muito antes da inteligência artificial.

Logo no início da automação moderna, tornou-se possível observar como injustiças podem ser reinterpretadas como diagnósticos neutros. Um exemplo marcante ocorreu no final dos anos 1970, quando a Escola de Medicina de St. George's, no Reino Unido, implementou um algoritmo para agilizar a triagem de candidatos. Após constatarem que a ferramenta reproduzia com precisão as decisões dos avaliadores humanos, a ferramenta foi incorporada sem resistência. Anos depois, uma auditoria interna revelou que o sistema aplicava penalidades automáticas e descontava pontos a candidatos com nomes não europeus e a mulheres (SCHWARTZ, 2019). O modelo não inventou critérios discriminatórios; apenas codificou escolhas historicamente presentes na instituição e lhes conferiu regularidade estatística.

O episódio evidencia o funcionamento central da lógica algorítmica em contextos desiguais, em que injustiças preexistentes tornam-se instruções de cálculo. Ao operar sobre bases produzidas por instituições marcadas por seletividades acumuladas, os modelos transformam padrões de exclusão em parâmetros de previsão e de tratamento. A estética da precisão - exibida na forma de porcentagens, métricas e classificações - legitima práticas que, em sua origem, não são neutras.

Quando diferenças estruturais passam a ser traduzidas em categorias como 'risco', 'perfil', 'probabilidade' ou 'adequação', desigualdades históricas ganham a aparência de propriedades numéricas da população, o que antes dependia de decisão política ou de disputa institucional passa a assumir o formato de um resultado técnico. Essa conversão prepara o terreno para compreender

como hierarquias sociais precedem o algoritmo e continuam a operar nele, moldando tanto seus dados quanto suas interpretações.

3.5.1 A desigualdade antecede o algoritmo

A conversão de desigualdades históricas em parâmetros matemáticos torna-se possível porque o terreno sobre o qual os algoritmos operam já é profundamente moldado por disparidades. Antes de qualquer linha de código, existem estruturas que orientam a classificação, documentação e administração de populações. No Brasil, esse pano de fundo consolidado ao longo de séculos influencia diretamente a forma como sistemas automatizados reproduzem padrões de exclusão.

A violência racial é um processo que permanece e se atualiza sempre que encontra novos meios materiais para operar:

Estudar as relações entre o racismo e a cultura é levantar a questão da sua ação recíproca. Se a cultura é o conjunto dos comportamentos motores e mentais nascido do encontro do homem com a natureza e com o seu semelhante, devemos dizer que o racismo é sem sombra de dúvida um elemento cultural. (...) Contudo, este elemento cultural preciso não se enquistou. O racismo não pôde esclerosar-se. Teve de se renovar, de se matizar, de mudar de fisionomia. Teve de sofrer a sorte do conjunto cultural que o informava. (FANON, 1980, p. 36)

Essa capacidade de renovação do racismo alertada pelo autor encontra na contemporaneidade um terreno fértil nos sistemas automatizados. É nesse sentido que a análise de Paixão et al. (2010) se torna uma chave de leitura fundamental para compreender a inteligência artificial, mesmo antecedendo o debate tecnológico atual. Ao definir a discriminação como um mecanismo institucional que altera as probabilidades de acesso a ativos fundamentais - como crédito, emprego, educação e proteção jurídica -, ele descreve, com precisão cirúrgica, a lógica operacional que viria a ser adotada pelos algoritmos. O viés algorítmico, sob essa ótica, atua como a ferramenta moderna que automatiza a exclusão mapeada por Paixão et al., processando o passado discriminatório para ‘decidir’ quem terá maior probabilidade de acessar o futuro.

Carneiro (2005) reforça a dinâmica ao mostrar que o racismo brasileiro opera por meio da naturalização. Diferenças produzidas historicamente passam a ser tratadas como características intrínsecas do grupo racializado, apagando a violência estrutural por trás de explicações aparentemente neutras. Quando tais diferenças surgem nos dados como tendências objetivas, e não

como efeitos de desigualdade, modelos matemáticos reproduzem essa naturalização com a regularidade da estatística.

Sendo assim, as infraestruturas digitais carregam camadas de poder sedimentadas nos fluxos de dados, resultantes de décadas de processos de categorização e controle que antecedem a tecnologia (SILVA, 2022). A coleta, organização e aplicação de dados refletem assimetrias sociais produzidas por instituições que definem quem é registrado, como é registrado e para que finalidade, passando a compor a lógica de decisão do próprio modelo.

O conjunto dessas reflexões revela que o algoritmo não cria desigualdade; apenas reproduz estruturas já consolidadas por instituições cuja atuação antecede as tecnologias: “o racismo algorítmico não se apresenta apenas como uma falha no sistema operacional, mas como parte estruturante dele” (KREMER et al., 2023, p. 19-20). Nessa passagem do social para o numérico, abre-se o caminho para compreender como as relações de poder moldam a própria produção das bases informacionais que alimentam sistemas computacionais.

3.5.2 Colonialidade de dados e captura assimétrica

No Brasil, a produção de dados nunca ocorreu em terreno neutro. As bases que hoje alimentam sistemas algorítmicos são herdeiras de práticas estatais e privadas que sempre registraram populações de forma desigual - com maior vigilância sobre determinados grupos, menor detalhamento sobre outros e lacunas produzidas por séculos de exclusão. Por isso, os bancos de dados usados em modelos computacionais não apenas descrevem a realidade: reproduzem a forma como instituições historicamente enxergaram e administraram diferentes segmentos da população.

O conceito de colonialidade de dados, formulado por Couldry e Mejias (2019), ajuda a iluminar esse cenário. Para os autores, a extração contínua de informações sobre populações vulnerabilizadas opera como forma contemporânea de colonialismo, em que o dado assume o papel de recurso natural, tornando-se matéria-prima de valor econômico e instrumento de governo. Não por acaso, a coleta massiva recai com maior intensidade sobre grupos já marcados pela desigualdade.

Silva (2022) aprofunda essa crítica ao demonstrar que a coleta de dados no país segue padrões de captura distintos. Populações negras, periféricas e pobres aparecem nas bases sobretudo por meio de mecanismos de vigilância - boletins de ocorrência, registros de abordagens policiais,

cadastros assistenciais, sistemas de reconhecimento facial -, enquanto grupos privilegiados surgem majoritariamente em registros voluntários e positivos - histórico educacional, consumo formal ou documentação civil consolidada. A assimetria não está apenas no volume de dados, mas também no tipo de relação institucional que define como cada grupo é visto, registrado e monitorado.

As desigualdades raciais no Brasil não se distribuem ao acaso, mas decorrem de processos institucionais que moldam trajetórias coletivas. Quando bases de dados são construídas a partir dessas mesmas instituições, acabam carregando consigo os efeitos acumulados da seletividade policial, das omissões no atendimento público e médico, das lacunas no registro civil, das desigualdades territoriais e da organização racializada do trabalho. O resultado é que os dados apresentados como neutros são, na verdade, produtos históricos de um Estado que sempre administrou populações de forma desigual.

Esse processo se completa no momento em que essas bases de dados entram no fluxo computacional. Uma vez incorporadas ao ciclo de treinamento, validação e aplicação, passam a operar como a base que define padrões, pesos e categorias. O algoritmo é tecnicamente incapaz de distinguir desigualdade histórica de tendência estatística e acaba interpretando ambos como sinais relevantes para o cálculo.

Tudo isso revela que o problema não se resume à qualidade dos dados, tampouco à sua quantidade. O centro da questão está na forma como as relações estruturais de poder moldam a própria produção das bases que alimentam os sistemas de decisão. Se os dados já chegam ao algoritmo com desigualdade, a modelagem não pode corrigi-la automaticamente - ao contrário, tende a amplificá-la. Esse movimento evidencia a necessidade de observar como o percurso técnico incorpora e legitima as discriminações.

3.5.3 Onde nasce o viés do algoritmo?

As desigualdades inscritas nos dados não são a única razão pela qual os sistemas algorítmicos passam a reproduzir padrões discriminatórios. Mesmo quando a base informacional parece limpa ou tecnicamente adequada, o viés surge em diferentes momentos do ciclo de modelagem, pois cada etapa da construção do algoritmo envolve escolhas normativas, hipóteses estatísticas e decisões institucionais que moldam o resultado final. Como é impossível reproduzir a complexidade e as nuances humanas para uma máquina, é preciso decidir o que é ‘mais

importante’. O’Neil (2020, p. 22) sintetiza bem ao afirmar que “modelos são opiniões embutidas em matemática”.

É aqui que a limitação epistemológica da tecnologia se revela significativa. Sistemas computacionais têm dificuldade de compreender conceitos humanos mais abstratos como justiça, moralidade, amor ou dignidade:

Eles “entendem” beleza apenas como uma palavra associada ao Grand Canyon, pores do sol no mar e dicas de depilação da revista Vogue. Eles tentam em vão medir “amizade” contando curtidas e conexões no Facebook. E o conceito de justiça inteiramente os escapa. Programadores não sabem como programá-lo, e poucos de seus patrões os pedem para fazê-lo. (O’NEIL, 2020, p. 91)

O ponto inicial da modelagem é a definição do problema, e decidir o que deve ser previsto não é uma operação neutra. Optar por antecipar risco, evasão, produtividade, suspeição ou propensão ao crédito implica atribuir sentido social a comportamentos específicos. Como observam Barocas e Selbst (2016), todo modelo nasce orientado por uma finalidade que favorece certos grupos e desfavorece outros, e essa orientação antecede qualquer decisão técnica posterior. O algoritmo já carrega um direcionamento normativo antes mesmo de os dados serem coletados.

Em seguida, a curadoria das bases adiciona novas camadas de seletividade. Escolher quais registros entram ou são excluídos, como variáveis são consolidadas, que categorias são criadas ou mescladas, que lacunas são tratadas como ausências ou como zeros - tudo isso forma um arranjo que traduz desigualdades históricas em estrutura numérica. Kremer (2023) demonstra que as bases policiais, educacionais, assistenciais e financeiras no Brasil são atravessadas por décadas de omissões, seletividades e inconsistências. Quando essas distorções entram no fluxo técnico, tornam-se parte integrante das relações que o modelo aprenderá.

No nível seguinte, o próprio desenho estatístico introduz vieses. Seleção de atributos, definição de limites, calibração de pesos, métricas de otimização e parâmetros de regularização são momentos em que opções subjetivas se transformam em fórmulas matemáticas. Modelos que buscam maximizar a acurácia tendem a penalizar grupos minoritários justamente porque a acurácia privilegia o padrão e o comportamento dominantes da base, ou seja, as regularidades produzidas por uma sociedade desigual.

Com redes neurais profundas, no *deep learning*, essa dinâmica se intensifica porque os algoritmos de aprendizado costumam ser vulneráveis às características dos dados de treinamento:

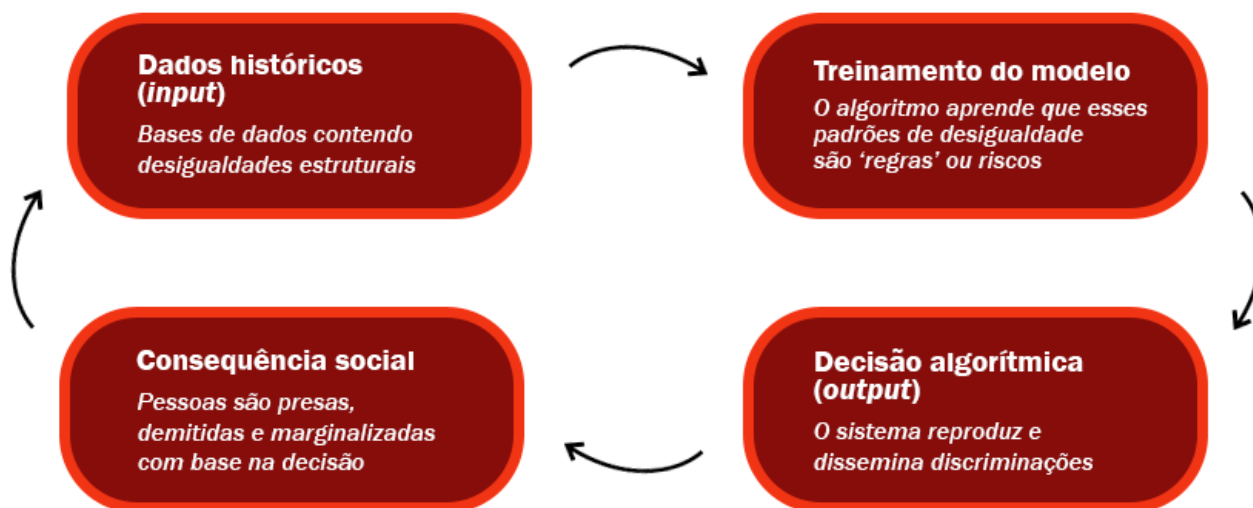
a mesma ‘vantagem’ de se adaptar de maneira autônoma, sem supervisão humana contínua, diante das mudanças dos dados de *input* também representa uma ameaça; tanto relacionado à incorporação de vieses e discriminações quanto a ataques de usuários mal-intencionados. Osoba e Welser (2017) descrevem esse fenômeno como um acoplamento entre a complexidade e a reprodução de vieses. Quanto maior a flexibilidade do modelo, maior a probabilidade de que hierarquias sociais se convertam em parâmetros internos de decisão, estáveis e invisíveis.

A busca pela otimização a qualquer custo cria, inclusive, barreiras à eficácia da regulação estatal. Enquanto leis de proteção de dados, como a Lei Geral de Proteção de Dados - LGPD (BRASIL, 2018), tentam impedir a discriminação vetando o uso explícito de informações sensíveis (como raça ou religião), os algoritmos contornam essa proibição por meio de *proxies* - variáveis substitutas, como o CEP, o modelo do celular ou padrões de linguagem, que permitem ao sistema inferir, por correlação, exatamente o dado protegido que a lei pretendia ocultar. Estabelece-se, assim, um descompasso estrutural. Enquanto a legislação opera em ritmo moroso, incapaz de acompanhar a velocidade de atualização dos modelos, as *big techs* utilizam seu poder político para frear regulações mais rígidas (O’NEIL, 2020).

E, mesmo após o treinamento, o viés pode emergir e se intensificar na etapa de aplicação. Quando sistemas automatizados passam a interferir diretamente em trajetórias individuais, negando crédito, priorizando vigilância policial, classificando estudantes e ranqueando candidatos a emprego, esses resultados se tornam novos dados de entrada (*input*). Essa retroalimentação transforma consequências institucionais em supostas evidências empíricas. Ao confundir seus resultados com a realidade, o algoritmo cria ciclos destrutivos de *feedback* e reforça o mundo que ajudou a produzir. Modelos de sentença que perfilam indivíduos pela circunstância ajudam a moldar o comportamento social, reforçando as premissas injustas que justificaram essas decisões inicialmente, num ciclo de injustiça recursiva (FIGURA 3):

FIGURA 3 - CICLO DE RETROALIMENTAÇÃO DO VIÉS ALGORÍTMICO

Ciclo vicioso do viés algorítmico



Fonte: Adaptada de O'Neil (2020).

Esse processo evidencia que o viés não reside apenas no dado, nem pode ser tratado como contaminação externa. Surge do encontro entre definições normativas, escolhas técnicas e estruturas institucionais que orientam a modelagem. Cada etapa adiciona uma camada própria de assimetria, e todas se acumulam no cálculo final. Quando essas camadas convergem, a desigualdade histórica se converte em tendência estatística, a tendência estatística se converte em probabilidade e a probabilidade se converte em justificativa técnica. Esse movimento estabelece o terreno no qual os sistemas de classificação passam a funcionar como mecanismos de administração social.

Tal lógica prepara o fundamento para compreender por que certos modelos se tornam particularmente perigosos. Quando a opacidade técnica se combina com decisões de larga escala e produz efeitos concentrados sobre grupos vulnerabilizados, o algoritmo deixa de ser apenas um instrumento de previsão e passa a operar como uma estrutura que formaliza desigualdades sob a aparência de neutralidade.

3.5.4 Armas de destruição matemática: opacidade, escala e dano

Modelos nocivos e enviesados aplicados a áreas da sociedade tornam-se especialmente destrutivos quando combinam três elementos estruturantes: opacidade, escala e dano. Essa tríade compõe as Armas de Destruição Matemáticas (ADMs), termo criado pela cientista de dados estadunidense Cathy O’Neil (2020) para batizar tais modelos. Após ver de perto como a matemática pode ser nociva, a autora lançou o livro *Weapons of Math Destruction: how big data increases inequality and threatens democracy* (Algoritmos de destruição em massa: como o *big data* aumenta a desigualdade e ameaça a democracia).

Esses três elementos não operam como um conjunto de falhas ocasionais, mas como uma arquitetura sociotécnica capaz de transformar desigualdades antigas em decisões automatizadas com aparência de precisão. No Brasil, essa dinâmica assume contornos mais graves porque se acopla a instituições historicamente marcadas pela seletividade racial e por formas persistentes de desigualdade.

A opacidade refere-se à falta de transparência do modelo, de seus objetivos e de seus critérios e funciona como primeiro eixo desse funcionamento. Parte dela deriva da complexidade técnica dos sistemas, especialmente os de *deep learning*, que tomam decisões por meio de camadas de processamento inacessíveis até mesmo a especialistas. Outra parte decorre de escolhas institucionais: códigos proprietários, ausência de documentação, impossibilidade de auditoria e sigilos corporativos. Pasquale (2015) argumenta que mecanismos desse tipo convertem decisões políticas em cálculos invisíveis, blindando o processo decisório contra escrutínio público.

No Brasil, essa invisibilidade se intensifica quando decisões automatizadas operam em sistemas cujas próprias instituições já funcionam como caixas-pretas; de delegacias a secretarias de educação, de órgãos de crédito a plataformas privadas. O resultado é que populações historicamente alvo de vigilância e controle também são aquelas que menos conseguem compreender, contestar ou revisar decisões algorítmicas que afetam sua vida cotidiana. Essa é a natureza de uma ADM - as máquinas falam por si e, por sua vez, são incontestáveis; suas sentenças são inflexíveis:

Mas não se pode recorrer contra uma ADM. Isso é parte de seu poder assustador. Elas não ouvem. Tampouco vergam. São surdas não somente a charme, ameaça e adulação mas também à lógica — mesmo quando há boas razões para se questionar os dados que alimentam seus resultados. Sim, caso fique claro que sistemas automatizados estão errando de modo vergonhoso e sistemático, os programadores irão voltar atrás e ajustar os algoritmos. Mas na maior parte das vezes os programas entregam sentenças inflexíveis, e

os seres humanos os utilizando dão de ombros como se dissessem, “bem, fazer o quê?”. (O’NEIL, 2020, p. 13)

O segundo eixo, a escala, refere-se ao amplo alcance das ADMs, que, ao operarem em grande proporção, transformam distorções locais em estruturas de exclusão amplificadas. Enquanto decisões humanas tendem a ser fragmentadas e contraditórias, sistemas automatizados podem aplicar o mesmo padrão enviesado a milhões de pessoas simultaneamente. O perigo não é apenas cometer erros, mas repetir o mesmo erro em larga escala com velocidade industrial. Kremer et al. (2023) observam que, quando essas métricas são integradas a políticas públicas, criam circuitos permanentes de exclusão, nos quais populações vulnerabilizadas são reiteradamente classificadas como risco.

O terceiro eixo é o dano - não apenas a consequência final, mas o próprio modo de operação desses sistemas, facilmente representado pela discriminação algorítmica. É o atributo de reforçar desigualdades e perpetuar discriminações já existentes, prejudicando grupos vulneráveis ao tomar decisões enviesadas em áreas críticas da sociedade, reproduzindo um ciclo de exclusão difícil de romper. Os modelos não se limitam a descrever a realidade, mas também a moldar. Silva (2022) argumenta que infraestruturas digitais funcionam como extensões de regimes de controle que já operavam por linhas raciais, transformando desigualdade acumulada em cálculo de risco.

Quando opacidade, escala e dano se somam, os sistemas deixam de operar como ferramentas de análise e passam a funcionar como mecanismos de administração social. A matemática é apresentada como critério de neutralidade, mas sua função prática é estabilizar hierarquias que nasceram fora do ambiente digital. A autoridade técnica, discutida anteriormente, reforça esse processo ao se apoiar no prestígio da estatística. Aquilo que historicamente foi produzido ressurge como probabilidade numérica, e a probabilidade passa a ser tratada como verdade.

Não se trata de corrigir uma distorção pontual, mas de reconhecer que muitos sistemas automatizados funcionam como engrenagens que formalizam desigualdades sociais em formato computável. Essa formalização é o que confere às ADMs seu poder normativo: ao transformar desigualdade em cálculo, o modelo altera o próprio horizonte de possibilidades de quem é avaliado. No Brasil, ADMs já se materializam em casos concretos, especialmente quando modelos importados de outros contextos são aplicados a populações cuja dinâmica histórica e social não corresponde às bases em que esses sistemas foram originalmente treinados. É também sobre essa

infraestrutura de cálculo que se assentam muitos dos sistemas que filtram, bloqueiam e priorizam conteúdos informativos, definindo, de forma invisível, a circulação do jornalismo e preparando o terreno para a crise de mediação analisada no capítulo seguinte.

4 DESAFIOS DO JORNALISMO NA ERA DA INTELIGÊNCIA ARTIFICIAL

A incorporação de sistemas algorítmicos na mediação da vida pública reorganiza silenciosamente as condições de possibilidade da própria comunicação social. A mesma lógica que estrutura a vigilância, o crédito, a assistência estatal e a classificação populacional passa a definir também os fluxos de visibilidade e noticiabilidade, os critérios de relevância e as formas pelas quais sociedades interpretam acontecimentos. Quando algoritmos se tornam responsáveis por ordenar o que aparece e o que desaparece, a esfera pública deixa de funcionar como espaço de disputa política e passa a operar segundo parâmetros de cálculo. A filtragem automática torna-se o novo regime de circulação e, com isso, as assimetrias herdadas da desigualdade estrutural ganham forma digital, reproduzidas pela opacidade técnica que impede o escrutínio e pela autoridade automática que confere legitimidade a decisões não supervisionadas.

Nesse ambiente, a presença do jornalismo se fragiliza: sua função histórica de produzir verificabilidade e contestação converte-se em mais um elemento subordinado a sistemas que respondem prioritariamente à lógica do engajamento. Essa dinâmica revela um alinhamento perverso entre a arquitetura algorítmica das *big techs* e os espectros políticos radicais. Como os sistemas de recomendação são calibrados para maximizar a retenção do usuário por meio da intensidade emocional, acabam por privilegiar discursos de ódio e narrativas polarizadas que mobilizam afetos extremos. Nesse cenário, materiais sensacionalistas e desinformativos - frequentemente associados à extrema-direita - ganham tração artificial justamente por gerarem taxas de interação mais elevadas do que a sobriedade necessária à apuração jornalística. O algoritmo, ao operar estritamente sob a métrica do lucro e sem a devida responsabilização editorial, converte-se em um vetor de amplificação de conteúdos que corroem o debate democrático, recompensando a radicalização em detrimento da verificação factual (ABRAJI, 2025).

Relatórios recentes mostram um ecossistema informacional em rápida transformação. A Associação Brasileira de Jornalismo Investigativo (Abraji) alerta para o avanço da plataformação imposta pela inteligência artificial (IA), processo em que redações passam a operar em ecossistemas fechados, dependentes de soluções proprietárias, o que as leva a perder autonomia tecnológica e editorial (2025). Segundo a entidade, esse rearranjo altera as prioridades das empresas e empurra jornalistas para um ambiente em que a produção é acelerada pela automação enquanto a verificação se torna mais difícil.

Na análise da Associação de Jornalismo Digital (Ajour), que representa veículos digitais independentes, o avanço das IAs generativas aprofunda desigualdades históricas entre grandes conglomerados e pequenas redações, sobretudo no Sul Global, onde a infraestrutura tecnológica é insuficiente para competir com ferramentas que centralizam recursos, dados e capacidade computacional (2024). Sem políticas de proteção e de investimento, os veículos locais correm o risco de desaparecer à mesma velocidade com que a automação se consolida.

Esse cenário se agrava porque, simultaneamente, cresce a percepção pública de que a mediação humana se tornou dispensável. Plataformas de busca já não funcionam como diretórios de informação, mas sim como mecanismos de resposta. A substituição do link pelo resumo, visível nas demonstrações recentes de sistemas integrados de inteligência artificial generativa no Google e no Bing, cria uma sensação de soberania algorítmica, em que uma resposta sintética parece mais confiável do que o trabalho de uma redação. Tal mudança desloca a centralidade da apuração profissional, pois a IA não tem compromisso com a verdade, mas com a verossimilhança:

Quando os humanos saem fora da produção da notícia, eles comprometem a interpretação, deixam de lado a crítica ou o debate e desaparece o contraditório, que são componentes-chave para você construir a notícia. Sem os humanos, riscos éticos se tornam elevados. A IA não possui empatia, senso crítico, capacidade investigativa. São qualidades essenciais do bom jornalismo. (ARBIX, s.p., 2025)

O resultado imediato é a erosão do valor público da reportagem. Uma investigação complexa, que exige tempo, deslocamento e entrevistas, pode ser soterrada por conteúdos automáticos cujo alcance depende apenas da otimização das plataformas. O jornalismo perde a batalha pela visibilidade antes mesmo de entrar em campo.

O deslocamento ocorre ao mesmo tempo em que há regiões inteiras do Brasil onde a ausência de redações cria verdadeiros desertos de notícias. A superpopulação artificial de conteúdo gerado por modelos de linguagem convive com a desertificação informacional em cidades pequenas, periferias e territórios onde os veículos locais fecharam as portas ou reduziram drasticamente o número de equipes. Duchiaide (2025) registra que a IA amplia essa assimetria porque acelera a produção em grandes redações, mas não fortalece a apuração local. O risco é que comunidades periféricas fiquem ainda mais dependentes de plataformas, justamente onde a desinformação circula com maior intensidade.

O paradoxo se torna evidente: nunca houve tanto conteúdo, mas nunca houve tão pouca informação cuja origem seja verificável, contextualizada e responsável. Andamos às cegas em um oceano de textos produzidos por sistemas que não são responsabilizados pelos danos que causam. O impacto dessa transformação recai diretamente sobre o campo profissional. Há um entusiasmo crescente em empresas de comunicação que veem na IA a promessa de produtividade ilimitada, enquanto jornalistas enfrentam cortes sucessivos, o acúmulo de funções e o desaparecimento de postos de trabalho. Ao automatizar etapas inteiras da produção, as empresas passam a tratar o conteúdo como *commodity*, reduzindo a complexidade do trabalho jornalístico a um produto replicável (ESTADÃO, 2025).

A precarização não é acidental e decorre da substituição sistemática de tarefas, da falta de regulação e do atraso sindical diante do ritmo acelerado da inovação. A Rede de Jornalistas Internacionais (IJNET) cita casos de redações latino-americanas em que jornalistas foram pressionados a usar IA sem treinamento adequado, assumindo para si a responsabilidade ética e legal por conteúdos gerados por sistemas que não entendem (RADCLIFFE, 2025). Esse deslocamento de responsabilidade ilustra o risco de apagamento, em que a figura humana se torna o único elo responsável numa cadeia de produção automatizada que ninguém controla.

À medida que aumenta a dependência das redações de ferramentas opacas, cresce também a dificuldade de fiscalizar instituições que já operam sob lógicas algorítmicas. Sistemas que modulam publicidade estatal, organizam cronogramas de publicação, filtram comentários e distribuem conteúdo entre plataformas aplicam critérios invisíveis que podem amplificar ou suprimir temas sensíveis. A autoridade automática deixa de atuar apenas sobre o público; atua também sobre o próprio jornalismo, modulando o que os profissionais veem, o que conseguem verificar e o que obtém circulação. A lógica descrita por pesquisadores da área - algoritmos como caixas-pretas que reproduzem assimetrias estruturais e se alimentam de métricas de performance - alcança aqui seu ponto mais crítico: um jornalismo que depende desses sistemas perde a capacidade de funcionar como contrapoder. Uma redação que não consegue ver os critérios que organizam sua própria visibilidade não consegue fiscalizar governos, empresas nem instituições que operam com as mesmas tecnologias.

Esse conjunto de transformações configura uma ameaça existencial. A IA reestrutura as prioridades profissionais, redefine a autoridade sobre o que é verdadeiro e desloca a função social do jornalismo para as margens de um ecossistema governado por sistemas que não respondem ao

controle democrático. A celebração acrítica desses mecanismos por parte de empresas de mídia revela o risco de um campo profissional que ‘joga confete’ no agente, o que pode se tornar seu próprio ceifador. A adoção entusiasmada ocorre num ambiente de regulação inexistente, com sindicatos fragilizados, tribunais despreparados e legisladores atrasados no debate técnico. A combinação entre desigualdade estrutural, opacidade matemática e concentração econômica avança sobre o jornalismo como força modeladora, capaz de redefinir seu futuro institucional.

A urgência de compreender essa conjuntura não reside em rejeitar a tecnologia, mas sim em reconhecer a profundidade da ameaça. A informação continua sendo o fundamento mínimo da democracia, e o jornalismo permanece a única instituição dedicada à produção pública de informações verificáveis. Quando algoritmos passam a filtrar, ordenar e sintetizar essa verdade, o risco não é apenas perder empregos, mas também perder a própria possibilidade de uma esfera pública plural. A seguir, torna-se necessário examinar como essa crise se manifesta materialmente: na produção, no trabalho e na estrutura democrática. O futuro do jornalismo depende da capacidade de enfrentar essa reorganização antes que ela se complete de forma irreversível.

4.1 O PARADOXO DA INFORMAÇÃO ENTRE OS DESERTOS DE NOTÍCIAS E A SUPERPOPULAÇÃO ARTIFICIAL

O ecossistema informacional contemporâneo vive sob uma contradição estrutural. A inteligência artificial multiplicou o volume de conteúdo em proporções inéditas ao mesmo tempo em que as condições materiais do jornalismo profissional se deterioram. Redações fecham, equipes encolhem, veículos locais se extinguem, enquanto fluxos automáticos de textos fabricados inundam as plataformas. Pequenas redações brasileiras operam num cenário de competição desigual, pressionadas por uma massa de conteúdo automatizado que parece oferecer informação, mas não entrega verificação, responsabilidade ou presença territorial (AJOR, 2024). A superpopulação artificial encobre o jornalismo. Quanto mais textos são gerados, menos visível se torna a ausência de reportagem e mais natural parece a ideia de que a apuração humana é dispensável.

Em várias regiões fora dos grandes centros urbanos, a saturação algorítmica substitui o jornalismo local como fonte primária de sentido. Relatórios como o *Digital News Report* (Relatório de Notícias Digitais, em tradução livre) registram que muitas comunidades do Sul Global dependem quase exclusivamente de *WhatsApp*, *TikTok* e *Instagram* para receber informações, não

por preferência, mas por falta de alternativas estruturais (NEWMAN et al., 2024). Essa dependência apresenta um recorte racial nítido: o relatório Desigualdades Informativas (ALÁFIA LAB, 2024) revela que 45,8% das pessoas negras utilizam o Instagram como fonte prioritária de informação - contra 37,2% entre brancos -, o que expõe essa população de forma desproporcional à curadoria opaca dos algoritmos, em detrimento do jornalismo profissional. Não por acaso, uma pesquisa realizada em 2024 pela Data Favela revelou que o impacto das *fake news* é maior nas periferias.

A desertificação informacional não se apresenta como silêncio, mas como uma avalanche. O excesso permanente de vídeos, resumos automáticos e conteúdos virais encobre o fato de que não há repórteres acompanhando políticas municipais, monitorando obras, fiscalizando recursos públicos ou narrando disputas locais. A sensação de hiperconectividade funciona como cortina para o colapso da presença jornalística. Um território pode estar saturado de mensagens e, ainda assim, politicamente invisível.

A abundância artificial transforma a própria lógica de produção do conhecimento. Boyd e Crawford (2012) demonstram que grandes volumes de dados produzem mais correlação do que compreensão, o que estimula interpretações superficiais que parecem neutras, mas decorrem de métodos opacos e enviesados. A IA generativa intensifica esse quadro ao produzir textos que soam explicativos, apesar de não apresentarem compromisso causal. A correlação se disfarça de conclusão. Zuboff (2021) descreve esse ambiente como um de saturação informacional, no qual a sensação de saber é fabricada por repetições automatizadas e não pela mediação crítica. A proliferação de resumos, respostas instantâneas e explicações genéricas estabelece um ‘mundo sem memória’, em que sistemas reproduzem padrões linguísticos sem relação direta com acontecimentos, dissolvendo o conceito de fato e substituindo-o por uma estética de plausibilidade. A forma da explicação se sobrepõe ao vínculo com a realidade.

O mito da autonomia informacional emerge desse contexto. A crença de que o usuário pode buscar por conta própria no Google ou solicitar ao ChatGPT uma síntese imediata sugere independência, quando, na verdade, intensifica a dependência de sistemas privados que filtram e priorizam conteúdos segundo métricas invisíveis. IAs generativas oferecem ‘aparência de compreensão’, produzindo textos que dispensam esforço cognitivo e fomentam a ilusão de domínio intelectual (ARBIX, 2025). A autonomia aparente converte-se em heteronomia profunda. A liberdade de consulta esconde o fato de que as plataformas decidem o que aparece, como aparece

e em que ordem. A desigualdade informacional se disfarça de escolha. O usuário sente que pensa quando, na verdade, apenas consome.

No cenário digital, a atenção deixou de ser espontânea para se tornar arquitetada. Plataformas modulam o comportamento por meio de sistemas de recomendação que priorizam a previsibilidade e a intensidade emocional. Esse modelo favorece conteúdos polarizados, sensacionalistas e desinformativos, especialmente os associados à extrema-direita, que costumam gerar elevada interação (ABRAJI, 2025). A audiência é moldada pelas engrenagens que a mantêm em rotação. A esfera pública se transforma em um ambiente de modulação, no qual a atenção, os afetos e as crenças são ajustados segundo modelos matemáticos treinados para maximizar a permanência. O debate público se dobra ao imperativo da viralidade.

Essa modulação não impacta todos os territórios da mesma forma. A lógica algorítmica privilegia conteúdos universais, visuais e rápidos, o que empurra pautas hiperlocais para posições de irrelevância estrutural. O algoritmo age como editor oculto do mundo, determinando o que tem chance de existir publicamente. Diferentemente do editor humano, que tem responsabilidade pública e ética, o algoritmo opera sob a lógica do lucro e do engajamento, sem responsabilização (BUCCI, 2023).

A política municipal, a educação pública, o transporte, a saúde básica, a segurança cotidiana - tudo o que tradicionalmente é coberto por redações locais - torna-se invisível quando não se encaixa no fluxo viral. A morte do jornalismo local se manifesta como um apagamento progressivo: primeiro perde alcance, depois perde financiamento, depois perde relevância algorítmica e, finalmente, desaparece. E quando desaparece, a comunidade some com ele.

A consequência de longo prazo é um ciclo de retroalimentação negativa que reforça a própria degradação do ecossistema informacional. Menos jornalismo significa menos dados verificáveis circulando na esfera pública. Menos dados circulando significam modelos generativos treinados sobre conteúdos cada vez mais frágeis, imprecisos ou fabricados. Modelos mais frágeis geram explicações mais pobres. Explicações pobres reforçam a desconfiança no jornalismo profissional. O ciclo se fecha e recomeça. O'Neil (2020) descreve esse processo como 'cascatas de erro', em que sistemas opacos transformam suas próprias saídas em novas entradas, perpetuando distorções que ganham aparência estatisticamente normal.

A crise não se explica apenas por processos tecnológicos, mas pelo cruzamento entre automação, desigualdade estrutural e abandono institucional. A precarização das redações,

acelerada pelo uso indiscriminado de IA, empobrece a capacidade investigativa, amplifica as dependências de plataformas e transforma jornalistas em operadores de fluxo. A profissão perde autonomia enquanto o sistema que a substitui ganha poder. A lógica econômica que orienta empresas de mídia empurra o campo para uma competição desigual com conteúdos automáticos produzidos a custo quase zero, deslocando o jornalismo da esfera do interesse público para a da sobrevivência mercantil.

À medida que a informação se torna um produto replicável por máquinas, a própria noção de verdade pública se enfraquece. A superpopulação informacional produz desorientação. Os desertos de notícias produzem vulnerabilidade. A disputa pela atenção substitui a disputa pela verificação. A estética da resposta instantânea destrói a cultura da apuração demorada. E a autoridade automática que confere legitimidade a textos gerados por IA torna-se mais influente do que o trabalho do repórter que investiga, entrevista, contextualiza e responde por suas conclusões.

O paradoxo da informação é, portanto, o de uma democracia submetida a sistemas de cálculo. Há informação demais e jornalismo de menos. Há voz demais e testemunho de menos. Há explicação demais e compreensão de menos. A IA consegue organizar o fluxo, mas não estabelece sentido. Nessa organização, o jornalismo perde visibilidade, perde autoridade e perde a batalha mais fundamental: o direito de definir o que é notícia.

4.2 DISTOPIA PROFISSIONAL SOB O ATRASO INSTITUCIONAL E PRECARIZAÇÃO

A incorporação da inteligência artificial ao jornalismo vem sendo anunciada como sinal de modernização e como prova de que as redações estão ‘atualizadas’ diante da revolução tecnológica. Ferramentas generativas são apresentadas como aliadas da produtividade, capazes de automatizar tarefas repetitivas e de liberar tempo para investigações mais complexas. A retórica da inovação, porém, encobre um movimento mais profundo: a tecnologia é utilizada prioritariamente para reduzir custos, reduzir equipes e transformar trabalho intelectual em conteúdo barato e escalável. Sem debate estruturado com jornalistas, a IA tende a se consolidar como instrumento de ‘plataformização imposta’, deslocando o poder de decisão das redações para empresas de tecnologia e gestores financeiros (AJOR, 2024). O que se vende como futuro da profissão funciona, na prática, como mecanismo de reestruturação empresarial que trata o jornalista como variável descartável.

Esse processo corrói a cultura profissional por dentro. Habilidades que constituíram a identidade ocupacional do jornalista - apurar com base em conflito e dúvida, contextualizar, entrevistar, editar com responsabilidade - passam a disputar espaço com sistemas que oferecem textos prontos, leads padronizados, títulos otimizados e até perguntas sugeridas para entrevistas. Beckett e Yaseen (2023) descrevem a transformação como uma mudança de paradigma, em que redações deixam de ser espaços em que o julgamento humano organiza o fluxo informativo para se tornarem ambientes em que esse julgamento precisa se adaptar às sugestões de ferramentas automatizadas, sob pena de ser visto como lento ou ineficiente.

A lógica do ‘conteúdo como produto’ se impõe com violência. Em vez de reportagens pensadas a partir de problemas públicos, prolifera um regime de entrega contínua de textos, posts, vídeos curtos e *newsletters*, calibrados para alimentar plataformas em tempo real. A pressão por volume torna-se um critério central de avaliação. Muitos jornalistas relatam jornadas fragmentadas, acúmulo de funções e uma sensação permanente de inadequação diante de métricas que medem cliques, tempo de permanência e ‘performance’ individual, mas ignoram a complexidade da pauta, o risco pessoal e a relevância social (RADCLIFFE, 2025). A precarização passa a ser também afetiva e cognitiva: não se trata apenas de ganhar pouco, mas de trabalhar sob a impressão contínua de estar devendo algo a máquinas que nunca se cansam.

Enquanto a automação avança, as instituições que deveriam proteger a categoria permanecem atrasadas. A Abraji (2025) reconhece que o debate trabalhista e ético sobre IA chega atrasado e com pouca força diante do ritmo de implementação das empresas. Sindicatos ainda discutem pautas clássicas, ao passo que redações já reorganizaram rotinas inteiras por meio de automação, sem consulta prévia aos profissionais. O legislativo não oferece respostas mínimas sobre limites para a substituição de equipes por modelos generativos, sobre transparência no uso de IA em conteúdos noticiosos ou sobre proteção contra vigilância algorítmica nas redações. O descompasso é estrutural. Decisões que definem o futuro do trabalho jornalístico são tomadas em conselhos de administração, enquanto instrumentos de defesa coletiva seguem presos a um mundo pré-algoritmo. O QUADRO 1 sistematiza as diferenças impostas pela IA sob o campo jornalístico, evidenciando o contraste entre o paradigma da apuração humana e a dinâmica da produção automatizada:

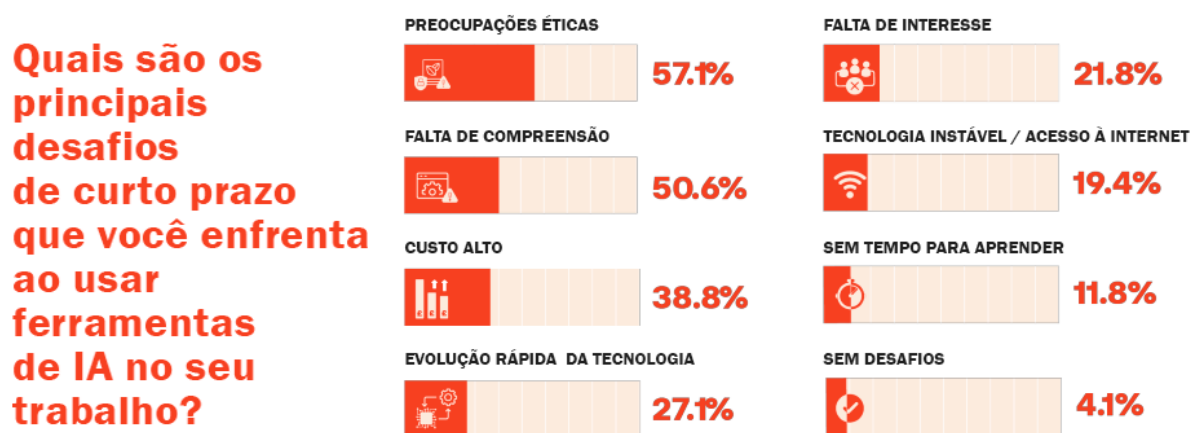
QUADRO 1 - COMPARATIVO DE ROTINAS: JORNALISMO TRADICIONAL E JORNALISMO NA ERA DA INTELIGÊNCIA ARTIFICIAL

Dimensão	Jornalismo tradicional (lógica da apuração)	Jornalismo na era da IA (lógica da precarização)
Matéria-prima	Fatos, entrevistas, documentos e observação empírica	Bases de dados pré-existentes e raspagem de conteúdo (<i>scraping</i>)
Tempo de produção	Determinado pela complexidade da investigação e verificação (<i>slow news</i>)	Determinado pela velocidade de processamento e meta de cliques
Função do jornalista	Autor e investigador (agente intelectual)	Operador de <i>prompts</i> e editor de fluxo (agente supervisor)
Critério de valor e noticiabilidade	Relevância pública, ineditismo e impacto social	Engajamento, SEO e escalabilidade (volume)
Risco principal	Erro humano na apuração ou viés editorial consciente	Alucinação da IA, reprodução de vieses estruturais opacos e perda de autoria
Responsabilização	Individual (repórter) e institucional (veículo)	Diluída (atribuída à "falha do sistema" ou à ferramenta proprietária)

FONTE: A autora (2025).

Essa virada de chave claramente não ocorre sem atritos. A imposição de novas rotinas sem o devido letramento digital gera incerteza entre os profissionais. Dados levantados pela *Thomson Reuters Foundation* (2025) confirmam que, para além das barreiras técnicas ou de custo, o principal obstáculo para jornalistas no Sul Global é de ordem moral e cognitiva. Conforme demonstra a FIGURA 4, o medo de violações éticas e a falta de compreensão sobre como as ferramentas funcionam superam até mesmo as dificuldades de infraestrutura:

FIGURA 4 - PRINCIPAIS DESAFIOS DE CURTO PRAZO ENFRENTADOS PELOS JORNALISTAS DO SUL GLOBAL NA ADOÇÃO DE IA NO TRABALHO



FONTE: Adaptado de Radcliffe (2025).

Os impactos variam conforme o formato, mas convergem para o mesmo ponto. No telejornalismo, experiências com apresentadores sintéticos e roteiros parcialmente gerados por IA colocam em risco as funções de repórteres, editores e produtores. No rádio, vozes artificiais substituem locutores e operadores em boletins e chamadas. No impresso e no digital, scripts de IA redigem notas, reescrevem textos, reorganizam notícias segundo métricas de cliques e sugerem manchetes já otimizadas para SEO. Estudos recentes sobre a adoção de IA em redações mostram que a automação foi incorporada primeiro às tarefas de entrada, justamente aquelas que historicamente atuavam como porta de entrada para novos profissionais, comprometendo a renovação geracional e a transmissão do conhecimento interno (BECKETT, YASEEN, 2023).

A desigualdade competitiva entre o jornalismo profissional e a desinformação torna-se ainda mais agressiva nesse contexto. Produzir reportagem custa caro, exige tempo, deslocamento, acesso a documentos, editores, advogados e uma estrutura mínima de redação. Produzir boatos custa muito pouco e, com a IA generativa, passa a custar quase nada; barateando a mentira e encarecendo a verdade. Enquanto o jornalismo investe recursos escassos para comprovar fatos, atores mal-intencionados usam ferramentas automáticas para produzir, adaptar e traduzir conteúdos enganosos em escala industrial. A competição entre verdade e falsidade se torna uma corrida em que um dos lados carrega um peso nas costas.

O risco de que a combinação entre IA e plataforma aprofunde a hegemonia de narrativas produzidas a partir de grandes centros urbanos, brancos e de classe média, enquanto

periferias, quilombos, territórios indígenas e cidades do interior desaparecem da cobertura regular revela o quanto a distopia é também geográfica. Comunidades perdem a voz pública e a experiência do país é ‘universalizada’ a partir de um recorte estreito. A colonialidade de dados converte-se em colonialidade territorial. O apagamento de regiões inteiras é um produto direto de uma infraestrutura que trata o que não gera engajamento como irrelevante.

A distopia profissional também se instala nas rotinas de formação. Escolas de jornalismo e cursos de comunicação têm incorporado a IA de maneira desigual. Enquanto alguns programas ainda tratam o tema como uma curiosidade tecnológica, entidades como o Knight Center e a Abraji já oferecem cursos específicos sobre cobertura crítica de inteligência artificial, riscos democráticos e implicações para a liberdade de expressão (KNIGHT CENTER, 2025; ABRAJI, 2025). A oferta de oficinas sobre ‘como usar IA na redação’ cresce rapidamente; em contrapartida, as formações sobre ‘como auditar sistemas’, ‘como investigar algoritmos’ e ‘como enfrentar opacidades técnicas’ avançam em ritmo bem mais lento. Sem uma atualização curricular profunda, há risco de formar não repórteres capazes de disputar o poder com sistemas automatizados, mas operadores dóceis de ferramentas proprietárias.

No plano ético, multiplicam-se iniciativas para orientar o uso responsável da IA, mas a distância entre princípios e prática permanece grande. Em 2023, Repórteres sem Fronteiras (RSF) coordenou a Carta de Paris sobre IA e jornalismo, que estabelece dez princípios para o uso responsável da tecnologia por veículos de comunicação, defendendo a transparência, a responsabilidade e a proteção do direito à informação (REPÓRTERES SEM FRONTEIRAS, 2023). O Conselho da Europa (2023) aprovou diretrizes para a implementação responsável de IA no jornalismo, com recomendações às empresas de mídia, aos Estados e às plataformas sobre transparência, supervisão humana e preservação da liberdade de expressão.

Em paralelo, organizações como o Poynter Institute (2025) vêm desenvolvendo kits práticos e *templates* para que redações criem seus próprios códigos internos de uso de IA, definindo o que é aceitável e o que viola valores jornalísticos. O *AI ethics guidelines* (Diretrizes de ética em IA, em tradução livre) é um exemplo de tentativa de transformar princípios gerais em políticas aplicáveis no cotidiano. Na América Latina, associações de mídia endossaram princípios para o desenvolvimento ético e justo da IA generativa, enfatizando a transparência, a responsabilidade e a integridade na relação entre veículos e tecnologia (LINARES, 2023). Em comum, essas

iniciativas reconhecem que, sem limites claros, a automação ameaça não só empregos, mas também a própria integridade do ofício.

Apesar desse movimento, a ética aplicada permanece frágil. Diretrizes internacionais têm força recomendatória, mas não se traduzem automaticamente em cláusulas contratuais, normas de redação ou dispositivos legais. Em muitas empresas, não há protocolos públicos de transparência no uso de IA; o público não sabe quando interage com conteúdo automatizado; as fontes não são informadas se estão falando para um sistema que será processado por modelos generativos; jornalistas não têm garantias sobre os limites de uso de seus textos para treinar ferramentas internas. O contraste entre a existência de cartas, manifestos e *guidelines* e a realidade de redações precarizadas e silenciosas quanto a isso expõe o núcleo da distopia profissional: a tecnologia avança com disciplina e velocidade, enquanto a proteção da profissão avança com timidez e atraso.

Esse conjunto de forças - cultura profissional corroída, identidade ocupacional diluída, custo desproporcional entre verdade e mentira, apagamento territorial, formação insuficiente e ética aplicada frágil - compõe um cenário em que a perda de controle humano sobre a produção de informação deixa de ser hipótese teórica para se tornar um processo em curso. Celebrar a IA como inovação neutra, sem enfrentar essas tensões, equivale a comemorar um mecanismo que pode atuar como ceifador da própria profissão. Se o jornalismo abdicar do direito de decidir como, quando e por que usar sistemas automatizados, corre o risco de sobreviver apenas como marca respeitável colada em conteúdos decididos, filtrados e escritos por algoritmos que ninguém elegeu e que não respondem a nenhum compromisso com a democracia.

4.3 JORNALISMO DE AUDITORIA ALGORÍTMICA COMO CONTRAPODER E RESISTÊNCIA

A expansão de sistemas algorítmicos em diversas esferas da sociedade desloca o eixo clássico da fiscalização jornalística. Poder não é apenas aquilo que se decide em gabinetes, plenários ou conselhos de administração; é também o que se codifica em modelos matemáticos que definem quem é vigiado, quem recebe crédito, quem tem acesso a direitos, o que aparece na tela e o que desaparece do campo de visão coletivo. A técnica deixou de ser apenas um instrumento de mediação para se tornar uma instância de decisão. Em um mundo mediado por algoritmos, investigar como esses sistemas funcionam torna-se parte da própria definição de jornalismo. Ignorar essa camada é aceitar que uma parte decisiva do poder opere fora do alcance da apuração.

Nicholas Diakopoulos (2014) propõe o conceito de *Algorithmic Accountability Reporting* (reportagem de auditoria algorítmica, em tradução livre) justamente para descrever o jornalismo que se volta a esses sistemas como objeto central de apuração. Em um artigo clássico, o autor define a prática como a investigação sistemática de estruturas de poder computacionais, com o objetivo de tornar visíveis seus vieses, impactos e assimetrias. O jornalista detalha métodos de amostragem de algoritmos, a reconstrução da relação entre entradas e saídas e o uso de engenharia reversa como estratégias para expor como decisões automatizadas influenciam a vida social. A proposta é direta: o jornalismo precisa deslocar a inteligência artificial do lugar de ferramenta para o de objeto a ser interrogado. A mesma curiosidade que, historicamente, se volta a orçamentos, licitações e bastidores institucionais deve ser dirigida a modelos de recomendação, sistemas de pontuação e infraestruturas de vigilância.

A discussão sobre responsabilidade algorítmica reforça essa exigência. O relatório *Algorithmic Accountability: A Primer* (Responsabilização Algorítmica: Uma Introdução, em tradução livre), do instituto de pesquisa *Data & Society*, define auditoria algorítmica como o processo de atribuir responsabilidade por danos quando decisões automatizadas produzem resultados discriminatórios (CAPLAN et al., 2018). As políticas de responsabilização ainda são frágeis, especialmente no setor público, o que evidencia a necessidade de mecanismos independentes capazes de auditar os algoritmos utilizados em decisões estatais (AI NOW INSTITUTE, 2018). Esses documentos dialogam com a função jornalística: se algoritmos podem limitar oportunidades, restringir serviços e distribuir risco de maneira desigual, o repórter que ignora essas engrenagens abdica da tarefa de expor quem ganha e quem perde com a automação.

Para operacionalizar essa vigilância, é necessário estruturar a investigação em duas dimensões complementares - a definição dos critérios éticos de avaliação e o estabelecimento de um rito metodológico de apuração. Os quadros a seguir sistematizam essa abordagem. O QUADRO 2 estabelece os pilares fundamentais da responsabilidade algorítmica, com base nas diretrizes da organização *Fairness, Accountability, and Transparency in Machine Learning* (FATML, 2016). O QUADRO 3 propõe um roteiro prático de execução, adaptando as técnicas de engenharia reversa de Diakopoulos (2014) ao fluxo de produção jornalística:

QUADRO 2 - PRINCÍPIOS FATML PARA RESPONSABILIZAÇÃO ALGORÍTMICA

Princípio	Descrição
Justiça (<i>fairness</i>)	Garantir que decisões algorítmicas não gerem impactos discriminatórios ou injustos entre diferentes grupos demográficos.
Explicabilidade (<i>explainability</i>)	Possibilitar que decisões e dados que as sustentam sejam explicados em termos não técnicos para usuários e demais interessados.
Auditabilidade (<i>auditability</i>)	Permitir que terceiros interessados possam inspecionar e revisar o comportamento do sistema, com documentação, APIs e termos de uso que viabilizem auditoria.
Responsabilidade (<i>responsibility</i>)	Definir responsáveis claros, canais de reclamação e mecanismos de reparação para efeitos adversos individuais ou sociais causados por sistemas algorítmicos.
Acurácia (<i>accuracy</i>)	Identificar e registrar fontes de erro e incerteza, de forma que impactos esperados e de “pior caso” possam ser compreendidos e mitigados.

FONTE: Adaptado de FATML (2016).

QUADRO 3 - ROTEIRO METODOLÓGICO: ETAPAS DA INVESTIGAÇÃO ALGORÍTMICA

(continua)

Etapa	Objetivo central	Perguntas orientadoras	Procedimentos / técnicas de reportagem
1. Identificação do algoritmo-alvo	Escolher sistemas com alto impacto social e risco de dano	Em que decisões o algoritmo intervém? Quem é afetado (em escala, vulnerabilidade, direitos)?	Monitorar políticas públicas e serviços digitais; acompanhar denúncias, processos judiciais, reportagens e reclamações; focar em setores sensíveis.
2. Formulação da hipótese de poder e risco	Tornar explícito que tipo de poder algorítmico será investigado	O algoritmo prioriza, classifica, associa ou filtra o quê? Onde podem surgir vieses, exclusões ou censura?	Usar a tipologia de decisões algorítmicas (prioritização, classificação, associação, filtragem) para estruturar hipóteses sobre possíveis injustiças.
3. Mapeamento de contexto e atores	Entender o sistema sociotécnico em torno do algoritmo	Quem projeta, contrata, opera, regula e é afetado pelo sistema? Que leis e políticas se aplicam?	Leitura de relatórios técnicos, contratos, leis e políticas; entrevistas com desenvolvedores, gestores públicos/empresas, reguladores e pessoas atingidas; identificar dependência de fornecedores “caixa-preta”.
4. Estratégia de amostragem (inputs/outputs)	Coletar dados para observar o comportamento do algoritmo	Que entradas posso observar ou simular? Que saídas consigo registrar de forma sistemática?	Uso de APIs e <i>scraping</i> quando possível; pedidos de acesso à informação; coleta colaborativa de decisões (<i>crowdsourcing</i> de prints, cartas, telas); definição de perfis de teste relevantes para a hipótese.

QUADRO 4 - ROTEIRO METODOLÓGICO: ETAPAS DA INVESTIGAÇÃO ALGORÍTMICA
(conclusão)

Etapa	Objetivo central	Perguntas orientadoras	Procedimentos / técnicas de reportagem
5. Testes e “engenharia reversa”	Revelar padrões, vieses e erros sistemáticos na relação entrada-saída	Como a saída muda ao variar perfis, atributos sensíveis ou contexto? Há diferenças entre grupos comparáveis?	Experimentos controlados variando uma variável por vez; comparação de taxas de erro, aprovação/negação, ranqueamento ou recomendação entre grupos; uso de estatísticas básicas para evidenciar disparidades.
6. Interpretação normativa dos achados	Traduzir resultados em termos de direitos, justiça e princípios de <i>accountability</i>	Os padrões observados violam princípios de justiça, explicabilidade ou proteção de dados? Quem é mais prejudicado?	Confrontar evidências com princípios FATML e com marcos legais (antidiscriminação, proteção de dados, direitos fundamentais); diálogo com especialistas e grupos afetados para validar a interpretação.
7. Reportagem humana e contraditório	Mostrar o impacto concreto na vida das pessoas e ouvir os responsáveis	Como o algoritmo aparece na experiência de indivíduos e comunidades? O que dizem empresas e governos?	Entrevistas em profundidade com pessoas atingidas; uso de casos emblemáticos; solicitação de posicionamentos oficiais e planos de correção; inclusão de respostas e justificativas no texto.
8. Transparência metodológica e debate público	Permitir escrutínio, replicação e uso dos achados no debate democrático	O público entende o método, seus limites e implicações? Outros podem replicar ou expandir a investigação?	Descrição clara de dados, amostragem, testes e limitações; visualizações acessíveis; quando possível, publicação de <i>scripts</i> /código; conexão dos achados com debates de regulação, justiça algorítmica e políticas públicas.

FONTE: Adaptado de Diakopoulos (2014).

A prática já mostrou, em escala internacional, que esse tipo de investigação é possível e necessário. A série *Machine Bias*, realizada em 2016 pela organização de jornalismo investigativo sem fins lucrativos dos ProPublica, analisou o algoritmo COMPAS⁹, usado para estimar risco de reincidência criminal nos Estados Unidos, e demonstrou que o sistema atribuía *scores* mais altos e falsos positivos de alto risco com muito mais frequência a réus negros do que a réus brancos em condições semelhantes ou até piores (ANGWIN et al., 2016).

Em outro caso, reportagens da mesma organização revelaram que o sistema de anúncios do Facebook permitia a anunciantes de moradia e de emprego excluir usuários com determinadas ‘afinidades étnicas’, contribuindo para práticas discriminatórias que levaram a ações judiciais e a acordos com o Departamento de Justiça dos EUA (ANGWIN; TOBIN; VARNER, 2017). Em

⁹ *Correctional Offender Management Profiling for Alternative Sanctions* (Perfilamento de Gestão de Infratores para Sanções Alternativas, em tradução livre)

ambas as situações, a combinação de jornalismo de dados, pedidos de acesso à informação e testes sistemáticos expôs aquilo que empresas e autoridades tratavam como detalhe técnico. A fiscalização do poder algorítmico não foi obra de auditorias internas, mas de redações que optaram por colocar sistemas automatizados no centro da apuração.

No cenário nacional, a viabilidade desse modelo de investigação foi demonstrada pela reportagem *Reprovados por robôs: Como plataformas de inteligência artificial podem discriminar mulheres, idosos e faculdades populares em processos seletivos*, apurada pela jornalista Ianaira Neves e publicada em 2022 no *The Intercept Brasil*. Ao investigar suspeitas de que a plataforma de recursos humanos Gupy reproduzia preconceitos de gênero e idade, a autora não se limitou a coletar denúncias, mas também realizou testes práticos de auditoria para confrontar a lógica do sistema. O método revelou inconsistências graves: uma desenvolvedora sênior foi reprovada na triagem algorítmica, enquanto sua irmã, utilizando o mesmo perfil profissional mas preenchendo o teste com respostas aleatórias, foi aprovada (NEVES, 2022).

Esse caso serve como prova de conceito de que a opacidade das ‘caixas-pretas’ corporativas brasileiras pode ser perfurada pelo jornalismo, o que transforma relatos individuais sobre discriminação em evidências estruturais. Esse tipo de investigação exige aprendizado, mas não exige que jornalistas se tornem engenheiros de software. O jornalismo se torna uma ferramenta central de desmontagem do sistema. A pauta deixa de ser ‘uso de IA na cidade’ para se tornar ‘como este sistema específico redistribui risco, direitos e violência’.

A necessidade de auditoria torna-se ainda mais premente quando o objeto de análise deixa de ser apenas uma plataforma privada, como redes sociais ou sistemas de RH, e passa a ser a própria infraestrutura urbana. A transição para o modelo de ‘cidades inteligentes’ (*smart cities*) desloca a opacidade algorítmica dos servidores fechados para o espaço público, instalando mecanismos de coleta de dados e vigilância em locais de convivência coletiva. Nesse cenário, o cidadão é monitorado pelo simples ato de transitar pela cidade.

Para o jornalismo, cobrir essa transformação exige ir além da reprodução dos discursos oficiais de modernização e eficiência. É preciso confrontar o fato de que, sob o verniz da inovação tecnológica, muitas administrações públicas estão implementando sistemas de controle social cujos critérios de funcionamento permanecem inacessíveis. A ‘caixa-preta’ se materializa no mobiliário urbano, financiada por recursos públicos, mas protegida por segredos industriais que dificultam o escrutínio sobre quem é vigiado e como as decisões de segurança são automatizadas.

É dentro dessa conjuntura que a capital paranaense se apresenta como um caso emblemático para a aplicação da auditoria algorítmica. O projeto da 'Muralha Digital' em Curitiba ilustra com nitidez essa tensão entre a promessa de uma gestão inteligente e os riscos de uma vigilância massiva e opaca. Ao investigar esse sistema, busca-se compreender não apenas a tecnologia em si, mas como ela reconfigura as relações de poder no território local, servindo de laboratório para os desafios que a reportagem enfrenta diante da governança digital das cidades.

A Muralha Digital, inaugurada em 2021, é apresentada pela prefeitura como um cerco digital de segurança que integra milhares de câmeras com reconhecimento facial em pontos estratégicos, radares de trânsito, câmeras corporais de agentes de segurança e, mais recentemente, imagens de sistemas privados de monitoramento de empresas e condomínios (CURITIBA, 2025; GAZETA DO POVO, 2025). A expansão do programa Conecta Muralha promete integrar até 6 mil câmeras públicas e privadas, sob a justificativa de reforçar a segurança pública e modernizar o controle urbano. O discurso oficial mobiliza a linguagem da eficiência, da prevenção e da tecnologia a serviço do cidadão.

FIGURA 5 - CENTRO DE CONTROLE OPERACIONAL DA MURALHA DIGITAL DE CURITIBA (I)



FONTE: Rodrigo Felix Leal (2021).

FIGURA 6 - CENTRO DE CONTROLE OPERACIONAL DA MURALHA DIGITAL DE CURITIBA (II)



FONTE: José Fernando Ogura / AEN (2021).

Análises críticas lembram, porém, que o mesmo arranjo técnico que permite localizar um veículo roubado ou identificar um suspeito também possibilita monitorar deslocamentos cotidianos, mapear rotinas de grupos específicos e consolidar um regime de vigilância em massa. Matéria recente descreve a Muralha Digital como símbolo de como projetos de ‘cidades inteligentes’ têm se convertido em cidades vigiadas, combinando crença na eficiência da IA com o medo da violência urbana para legitimar a expansão contínua de dispositivos de controle (ABAD et al., 2025). Enquanto textos institucionais e reportagens de viés promocional celebram supostas reduções de criminalidade e a promessa de uma cidade completamente assegurada por câmeras, pesquisadores associados à UFPR e outros espaços acadêmicos alertam para riscos de violação de privacidade, falta de transparência, possibilidade de discriminação e ausência de estudos independentes que comprovem a eficácia real do sistema.

É necessário reconhecer, também, que o jornalismo local enfrenta um dilema direto. Redações fragilizadas, dependentes de publicidade oficial e pressionadas por métricas de engajamento têm dificuldade em deslocar o foco de inaugurações, coletivas de imprensa e narrativas de ‘modernização da segurança’ para perguntas incômodas: quais bairros concentram câmeras, quais populações são mais vigiadas, que critérios definem a instalação dos equipamentos,

quem tem acesso às imagens e por quanto tempo, que mecanismos de controle externo existem, que erros já ocorreram, quantas pessoas foram injustamente abordadas com base em identificações falhas.

Sem capacidade técnica, tempo e segurança institucional para enfrentar essas questões, o jornalismo local corre o risco de se limitar a reproduzir o discurso de ‘cidade inteligente’ enquanto a cidade real é submetida a uma vigilância cuja lógica permanece invisível. Essa opacidade foi confirmada em episódios sucessivos de negativa de transparência. Em 2023, a administração municipal da capital paranaense recusou-se a detalhar o conteúdo das bases de dados que alimentam o reconhecimento facial, alegando motivos de segurança (COLOMBO, 2023). A restrição ao acesso intensificou-se quando o Comitê Gestor de Proteção de Dados Pessoais de Curitiba negou todos os pedidos, feitos via Lei de Acesso à Informação - LAI (BRASIL, 2011), de acesso às imagens captadas pela Muralha Digital feitas pelo jornal Plural em 2024 (FREITAS, 2025). O episódio revela uma ironia perversa: a administração utilizou a LGPD (BRASIL, 2018) - norma criada justamente para proteger a privacidade dos indivíduos contra abusos - como fundamentação jurídica para sonegar informações públicas sobre o aparato que os monitora, convertendo uma legislação de direitos civis em escudo para o segredo estatal.

A assimetria de visibilidade é estrutural. Sistemas como a Muralha Digital veem tudo, ou quase tudo, que interessa ao seu desenho; cidadãos veem apenas o que lhes é mostrado em peças de comunicação oficial; jornalistas, quando muito, têm acesso a tours guiados em centros de controle e a números selecionados. Políticas de responsabilização algorítmica no setor público exigem abertura de informações, documentação acessível e mecanismos que permitam a grupos externos testar e contestar sistemas (AI NOW INSTITUTE, 2018). Sem essa abertura, o jornalismo permanece, na melhor das hipóteses, na superfície do problema. A combinação entre segredo comercial, contratos pouco transparentes e linguagem tecnicista transforma algoritmos em muralhas jurídicas que protegem o poder, ao mesmo tempo em que a muralha física e digital expande o olhar do Estado sobre os corpos e territórios mais vulneráveis.

Essa disputa não se restringe à segurança pública. Algoritmos ordenam filas de acesso a políticas sociais, definem quem recebe ou não determinado benefício, ranqueiam notícias em buscadores, modulam a entrega de conteúdos em redes sociais e distribuem anúncios de emprego, crédito e moradia. Cada um desses domínios já foi palco de escândalos em que jornalismo e sociedade civil expuseram discriminações profundas. O caso COMPAS mostrou como modelos de

risco criminal podem amplificar desigualdades raciais sob a aparência de neutralidade estatística; investigações sobre o sistema de anúncios do Facebook revelaram que ferramentas de segmentação permitiam excluir grupos protegidos de ofertas essenciais, a ponto de órgãos reguladores classificarem a prática como forma contemporânea de discriminação (ANGWIN et al., 2016). Em todos esses casos, foi o trabalho de jornalistas e pesquisadores engajados na responsabilização algorítmica que transformou problemas aparentemente técnicos em temas de debate público e de disputa jurídica.

Organizações ligadas à defesa da liberdade de imprensa reconhecem esse papel. A Carta de Paris sobre IA e jornalismo afirma que a profissão continua a ser um pilar de confiança na sociedade e defende que qualquer uso de sistemas de IA com impacto significativo na produção ou distribuição de conteúdos deve ser transparente, com a preservação do julgamento humano como instância final de decisão (RSF, 2023). Para proteger o direito à informação, é necessário que jornalistas e meios de comunicação não apenas usem a tecnologia de forma responsável, mas também participem ativamente de sua governança, denunciando abusos, resistindo a opacidades e recusando-se a delegar decisões editoriais a sistemas que não respondem publicamente por seus efeitos.

A bifurcação colocada à profissão aparece com nitidez. De um lado, está a tentação de incorporar a IA como mera ferramenta produtiva, aceitando sua presença em todas as etapas do trabalho sem questionar o desenho, o objetivo e os interesses embutidos em cada sistema. De outro lado, está a responsabilidade de assumir que os algoritmos se tornaram uma das principais formas de exercício de poder na contemporaneidade e que, portanto, precisam ser vigiados com a mesma intensidade com que se vigiam governos, empresas e instituições tradicionais. A fórmula é simples e brutal. O jornalismo precisa vigiar os algoritmos, mas antes precisa sobreviver a eles. Sobrevivência significa garantir condições materiais, políticas e éticas para que existam redações capazes de investigar caixas-pretas, enfrentar a retórica tecnosolucionista e expor as engrenagens de cidades vigiadas e de plataformas opacas. Se essa tarefa for abandonada em nome da eficiência ou do deslumbramento, a profissão continuará existindo apenas como fachada honrosa de um ecossistema informacional governado por sistemas que ninguém elegeu e que, no limite, dispensam completamente a presença do jornalista.

4.4 A URGÊNCIA DA REINVENÇÃO JORNALÍSTICA DIANTE DA HEGEMONIA ALGORÍTMICA

O quadro que se desenha é o de uma profissão encurralada entre a abundância artificial e o esvaziamento estrutural. Enquanto a superpopulação de conteúdos gerados por IA e modulados por plataformas produz a sensação de que a informação está disponível em excesso, acessível em qualquer tela, a qualquer hora, os desertos de notícias se aprofundam, as redações encolhem, o jornalismo local perde presença e a mediação humana na produção de sentido público se fragiliza. A combinação entre autoridade automática, opacidade algorítmica e precarização trabalhista redefine o lugar do jornalismo na esfera pública. A profissão corre o risco de ser rebaixada à função de ilustrar um mundo já narrado por sistemas que não respondem a critérios democráticos, apenas a métricas internas de desempenho.

Ao longo deste capítulo, emerge com força a ideia de que a ameaça não é apenas econômica ou tecnológica, mas também existencial. A automação é incorporada pelas empresas como estratégia de corte de custos e não como ampliação da capacidade investigativa. O atraso sindical, a inércia legislativa e o vácuo regulatório abrem caminho para um cenário em que a IA se torna infraestrutura invisível da produção de informação. A lógica de engajamento das plataformas empurra o jornalismo para uma posição periférica, enquanto a desinformação automatizada se torna cada vez mais barata e eficiente. A consequência é um deslocamento profundo, em que o controle humano sobre a produção e a circulação da informação deixa de ser um pressuposto democrático e passa a ser uma exceção disputada.

Nesse contexto, a sobrevivência do jornalismo como instituição democrática depende de assumir explicitamente sua função de contrapoder algorítmico. Não basta usar ferramentas de IA para ganhar tempo ou produzir mais conteúdo. Não basta cobrir inteligência artificial como tema, reproduzindo promessas de inovação e eficiência. É necessário investigar a tecnologia. A proposta de reportagem de auditoria algorítmica formulada por Diakopoulos (2014) aponta com clareza esse caminho ao definir o jornalismo como instância de apuração das estruturas de poder computacionais, responsável por tornar visíveis os vieses, as decisões e os interesses embutidos nos algoritmos que organizam a vida social.

Assumir essa tarefa implica incorporar, como tese mínima, alguns gestos concretos. O primeiro é exigir dados e documentação sempre que governos e empresas adotarem sistemas automatizados para decisões que afetam direitos, o acesso a políticas públicas ou a circulação de

informação. A Lei de Acesso à Informação (BRASIL, 2011) não foi concebida para o mundo algorítmico, mas pode - em teoria - ser usada para solicitar contratos, relatórios de impacto, documentação técnica simplificada e avaliações de risco de sistemas aplicados em áreas como segurança pública, filas de saúde, cadastro de benefícios sociais ou triagem educacional. A responsabilização algorítmica passa justamente pela possibilidade de conhecer os contextos em que esses sistemas são usados, os dados em que se baseiam e os mecanismos de reparação em caso de dano.

O segundo gesto é testar, simular, experimentar. Um dos métodos que jornalistas usam é a criação de perfis distintos para observar como um sistema de recomendação, ranqueamento ou pontuação trata pessoas e grupos de forma desigual (DIAKOPOULOS, 2014). Esse tipo de experimento já foi utilizado em reportagens sobre anúncios discriminatórios, buscas personalizadas e algoritmos de risco criminal. No contexto brasileiro, a mesma lógica pode ser aplicada a sistemas que ranqueiam candidatos a vagas de emprego, definem oferta de crédito, priorizam certos bairros em políticas de policiamento ou modulam a entrega de notícias em plataformas. A técnica deixa de ser fetiche e passa a ser um método de denúncia.

O terceiro gesto é abandonar a fantasia de atuar de forma solitária e trabalhar em ecossistema. Jornalistas não precisam, nem conseguirão, decifrar sozinhos cada linha de código de sistemas complexos. A literatura sobre responsabilização algorítmica insiste na importância de parcerias entre redações, pesquisadores em ciência de dados, organizações de direitos digitais e coletivos de periferia para construir investigações robustas e socialmente situadas. Contudo, a viabilidade material dessa auditoria algorítmica não pode depender apenas da vontade individual ou de parcerias espontâneas.

É imperativo reconhecer que investigações dessa natureza exigem recursos financeiros, tempo e capacidade técnica que a maioria das redações locais, asfixiadas pela precarização, não dispõe. A figura do 'jornalista herói solitário' torna-se obsoleta e ineficaz diante da complexidade e da escala das caixas-pretas computacionais. A única saída realista reside na formação de consórcios investigativos e na criação de fundos de financiamento independentes, especificamente desenhados para o letramento e a vigilância tecnológica. Entretanto, a obtenção desse suporte enfrenta barreiras estruturais severas: como a hegemonia econômica das *big techs* controla não apenas a infraestrutura digital, mas também grande parte dos fluxos de capital que sustentam a inovação, torna-se politicamente árduo financiar iniciativas cujo objetivo primário é justamente

auditar e confrontar os interesses dessas corporações, que possuem - até mais do que deviam - grande influência na estrutura social.

Relatórios recentes da TRF (2025) indicam que jornalistas do Sul Global já usam IA em larga escala, mas pedem orientação ética e técnica para não se tornarem reféns de ferramentas que compreendem mal. O desafio é apropriar-se criticamente desses recursos, orientando-os para a vigilância do poder e não apenas para a produtividade. A ponte com o jornalismo de dados é evidente. Não se trata de usar planilhas para adornar reportagens com gráficos, mas de compreender os dados como campo de disputa política e técnica. O mesmo arsenal metodológico que permite mapear orçamentos, redes de influência ou padrões de violência também pode ser utilizado para desmontar caixas-pretas algorítmicas. Bases públicas, bancos de dados produzidos por movimentos sociais, registros de chamadas, registros de abordagens policiais e documentos administrativos podem ser articulados para mostrar quem é mais vigiado, quem é mais punido, quem é mais excluído por sistemas apresentados como neutros. Quando a colonialidade de dados encontra a colonialidade territorial, como no caso da Muralha Digital de Curitiba, cabe ao jornalismo recusar o papel de vitrine tecnológica e assumir o de cartógrafo crítico da vigilância.

A frase que atravessa o capítulo se sustenta como conclusão: o jornalismo precisa vigiar os algoritmos, mas antes precisa sobreviver a eles. Entre a adesão acrítica que celebra a IA e a escolha de assumir a função de contrapoder algorítmico, não há espaço neutro. Um campo que aceita reduzir-se a operador de ferramenta contribui para a própria irrelevância. Um campo que assume o risco de enfrentar a opacidade, exigir dados, testar sistemas e expor os danos da automação pode recuperar parte da legitimidade que vem perdendo. A encruzilhada é profissional, mas também democrática. Se o jornalismo abdicar desse lugar, a tarefa de vigiar o poder algorítmico ficará a cargo de poucos pesquisadores e organizações isoladas, enquanto o grosso da esfera pública será organizado por sistemas que dispensam a mediação humana e, no limite, a própria existência do jornalista.

CONCLUSÕES

A trajetória percorrida por esta monografia permite afirmar que a ‘caixa-preta’ algorítmica não constitui apenas um obstáculo técnico, mas um dispositivo político que reorganiza as dinâmicas de poder na sociedade contemporânea. Ao longo da pesquisa, a análise da arquitetura da inteligência artificial demonstrou que a opacidade é uma característica intrínseca aos modelos de aprendizado profundo, mas também uma estratégia de mercado que protege a propriedade intelectual em detrimento do escrutínio público.

A discussão teórica confirmou que a crença na neutralidade da tecnologia - o *tecnochauvinismo* - atua como uma cortina de fumaça para a manutenção de hierarquias sociais preexistentes. Em contextos marcados por desigualdades estruturais como o Brasil, observou-se que sistemas de decisão automatizada tendem a codificar e amplificar o racismo e a exclusão social. Seja na segurança pública, onde o reconhecimento facial vigia desproporcionalmente corpos negros, ou na gestão de direitos, onde a eficiência administrativa se sobrepõe à proteção social, a tecnologia opera validando estatisticamente preconceitos históricos sob a aparência de objetividade matemática.

Diante desse cenário, os resultados indicam que o jornalismo enfrenta uma ameaça existencial dupla. Por um lado, a incorporação acrítica de ferramentas de automação nas redações agrava a precarização do trabalho e aprofunda a dependência de plataformas proprietárias. Por outro lado, a perda do monopólio da verdade por sistemas geradores de conteúdo sintético corrói a relevância social da imprensa. Constatou-se que a falta de letramento técnico e o atraso institucional deixam a profissão vulnerável, limitando sua capacidade de fiscalizar as novas infraestruturas de poder.

No entanto, a pesquisa aponta um caminho de resistência viável: o *Algorithmic Accountability Reporting* (Reportagem de Auditoria Algorítmica). A investigação demonstrou ser possível - e urgente - que jornalistas assumam a função de auditores do poder computacional, utilizando métodos de engenharia reversa e análise de dados para desvendar os vieses ocultos nos códigos. O contexto da "Muralha Digital" de Curitiba ilustra o desafio prático dessa missão, exigindo que o repórter confronte a opacidade de contratos e de discursos de ‘cidade inteligente’ que muitas vezes ocultam regimes de vigilância.

Em suma, este trabalho sustenta que a sobrevivência do jornalismo depende de sua capacidade de evoluir de mero usuário de ferramentas digitais a fiscalizador de seus códigos. Se a

democracia morre no escuro, na era da inteligência artificial, corre o risco de morrer na opacidade. Cabe ao jornalismo, portanto, a tarefa inadiável de acender a luz dentro da caixa-preta, reafirmando seu compromisso ético com a defesa dos direitos humanos contra a automação da injustiça.

REFERÊNCIAS

- ABAD, Livia et al. Muralha Digital de Curitiba é um símbolo de como ‘cidades inteligentes’ viraram ‘cidades vigiadas’. **The Conversation**, 11 ago. 2025. Disponível em: <https://theconversation.com/muralha-digital-de-curitiba-e-um-simbolo-de-como-cidades-inteligentes-viraram-cidades-vigiadas-262003>. Acesso em: 22 set. 2025.
- ABRAJI. Desafios do jornalismo em 2025: plataforma imposta pela IA. **Abraji**, São Paulo, 2025. Disponível em: <https://www.abraji.org.br/noticias/desafios-do-jornalismo-em-2025-plataformizacao-imposta-pela-ia>. Acesso em: 24 jun. 2025.
- AI NOW INSTITUTE. Algorithmic Accountability Policy Toolkit. Nova York: AI Now Institute, 2018. Disponível em: <https://ainowinstitute.org/publications/algorithmic-accountability-policy-toolkit>. Acesso em: 25 jun. 2025.
- AJOR. A era da inteligência artificial chegou ao jornalismo. **Associação de Jornalismo Digital**, 2024. Disponível em: <https://ajor.org.br>. Acesso em: 25 jun. 2025.
- ALÁFIA LAB. **Desigualdades Informativas**: Entendendo os caminhos informativos dos brasileiros na internet 2024. Disponível em: <https://alafialab.org/wp-content/uploads/2025/05/Desigualdades-informativas-2024.pdf>. Acesso em: 15 nov. 2025.
- ALMEIDA, Virgilio A. F.; DONEDA, Danilo. O que é a governança de algoritmos? **Revista poliTICs**, n. 24, p. 59-63, out. 2016. Disponível em: <https://politics.org.br/pt-br/privacidade-e-vigilancia-news/o-que-e-governanca-de-algoritmos>. Acesso em: 13 jun. 2024.
- ANDREJEVIC, Mark. **Automated Media**. Abingdon: Routledge, 2019.
- ANGWIN, Julia et al. **Machine Bias**: There’s software used across the country to predict future criminals. And it’s biased against blacks. **ProPublica**, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 10 jun. 2024.
- ANGWIN, Julia; TOBIN, Ariana; VARNER, Madeleine. Facebook (Still) Letting Housing Advertisers Exclude Users by Race. **ProPublica**, New York, 21 nov. 2017. Disponível em: <https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin>. Acesso em: 30 nov. 2025.
- ARBIX, Glauco. Como fica o jornalismo com o avanço da IA? **Jornal da USP**, São Paulo, 24 jun. 2025. Disponível em: <https://jornal.usp.br/radio-usp/como-fica-o-jornalismo-com-o-avanco-da-ia/>. Acesso em: 30 nov. 2025.
- BARBROOK, Richard; CAMERON, Andy. The Californian Ideology. **Science as Culture**, v. 6, n. 1, p. 44-72, 1996. Disponível em: https://monoskop.org/images/d/dc/Barbrook_Richard_Cameron_Andy_1996_The_Californian_Ideology.pdf. Acesso em: 01 set. 2025.

BARBOSA, Xênia de Castro; BEZERRA, Ruth Ferreira. Breve introdução à história da inteligência artificial. **Jamaxi**, Rio Branco, v. 4, n. 1, p. 90-97, 2020. Disponível em: <https://periodicos.ufac.br/index.php/jamaxi/article/view/4730>. Acesso em: 04 mar. 2025.

BARLOW, John Perry. **A Declaration of the Independence of Cyberspace**. Davos: Electronic Frontier Foundation, 8 fev. 1996. Disponível em: <https://www.eff.org/cyberspace-independence>. Acesso em: 21 set. 2025.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **California Law Review**, Berkeley, v. 104, n. 3, p. 671-732, 2016. Disponível em: <https://doi.org/10.15779/Z38BG31>. Acesso em: 01 ago. 2024.

BARR, Avron; FEIGENBAUM, Edward A. (org.). **The Handbook of Artificial Intelligence**. v. 1. Los Altos: William Kaufmann, 1981.

BECKETT, Charlie; YASEEN, Mira. **Generating Change: A global survey of what news organisations are doing with AI**. London: London School of Economics and Political Science (LSE), 2023. Disponível em: <https://blogs.lse.ac.uk/polis/2023/06/26/how-newsrooms-around-the-world-use-ai-a-journalism-2023-global-survey/>. Acesso em: 01 ago. 2025.

BOWKER, Geoffrey C.; STAR, Susan Leigh. **Sorting Things Out: Classification and Its Consequences**. Cambridge: The MIT Press, 1999. Disponível em: <https://doi.org/10.7551/mitpress/6352.001.0001>. Acesso em: 11 set. 2025.

BOYD, Danah; CRAWFORD, Kate. Critical questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon. **Information, Communication & Society**, v. 15, n. 5, p. 662-679, 2012. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2012.678878>. Acesso em: 06 dez. 2024.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações (LAI). **Diário Oficial da União**, Brasília, DF, 18 nov. 2011.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais (LGPD). **Diário Oficial da União**, Brasília, DF, 15 ago. 2018.

BRASIL. Ministério da Saúde. **Regulação da inteligência artificial no setor saúde: benefícios e riscos**. Apresentação em Audiência Pública na Comissão de Seguridade Social e Família da Câmara dos Deputados. Brasília, 5 ago. 2025. Disponível em: <https://www2.camara.leg.br/atividade-legislativa/comissoes/comissoes-permanentes/cssf/apresentacoes-em-eventos/eventos-2025/05-08-2025-audiencia-publica-regulacao-da-inteligencia-artificial-no-setor-saude-beneficios-e-riscos-1/apresentacao-ministerio-da-saude/view>. Acesso em: 25 nov. 2025.

BRASIL. Ministério da Saúde. **Ministério da Saúde vai investir R\$ 4,5 bilhões em rede de hospitais e serviços inteligentes do SUS**. Brasília, nov. 2025. Disponível em:

<https://www.gov.br/saude/pt-br/assuntos/noticias/2025/novembro/ministerio-da-saude-vai-investir-r-4-5-bilhoes-em-rede-de-hospitais-e-servicos-inteligentes-do-sus>. Acesso em: 25 nov. 2025.

BRAYNE, Sarah. **Predict and Surveil**: Data, Discretion, and the Future of Policing. Oxford: Oxford University Press, 2021.

BROUSSARD, Meredith. **Artificial Unintelligence**: How Computers Misunderstand the World. Cambridge: MIT Press, 2018.

BUCCI, Eugênio. **Incerteza, um ensaio**: como pensamos a ideia que nos desorienta (e orienta o mundo digital). Belo Horizonte: Autêntica, 2023.

BURRELL, Jenna. How the machine ‘thinks’: understanding opacity in machine learning algorithms. **Big Data & Society**, v. 3, n. 1, 2016. Disponível em: <https://doi.org/10.1177/2053951715622512>. Acesso em: 18 nov. 2024.

CAPLAN, Robyn et al. **Algorithmic Accountability**: A Primer. New York: Data & Society Research Institute, 2018. Disponível em: https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL-4.pdf. Acesso em: 31 jul. 2025.

CARREGOSA, Lais. TCU estima que 10,9% das negativas automáticas para concessão de benefícios do INSS em 2024 foram indevidas. **G1**, Brasília, 26 mar. 2025. Disponível em: <https://g1.globo.com/politica/noticia/2025/03/26/tcu-estima-que-109percent-negativas-automaticas-para-concessao-de-beneficios-do-inss-em-2024-foram-indevidas.ghml>. Acesso em: 25 nov. 2025.

CARNEIRO, Sueli. **Racismo, sexismo e desigualdade no Brasil**. São Paulo: Selo Negro, 2011.

CESec. **O Panóptico**: Monitor do reconhecimento facial no Brasil. Rio de Janeiro: Centro de Estudos de Segurança e Cidadania, 2025. Disponível em: <https://www.opanoptico.com.br/>. Acesso em: 01 dez. 2025.

CETIC.BR. **TIC Kids Online Brasil**: 65% das crianças e dos adolescentes usam IA generativa para estudar, criar conteúdo e lidar com emoções. São Paulo: Cetic.br, 22 out. 2025. Disponível em: <https://cetic.br/pt/noticia/tic-kids-online-brasil-65-das-criancas-e-dos-adolescentes-usam-ia-generativa-para-estudar-criar-conteudo-e-lidar-com-emocoes/>. Acesso em: 25 nov. 2025.

CHENEY-LIPPOLD, John. **We Are Data**: Algorithms and the Making of Our Digital Selves. New York: New York University Press, 2017.

COLOMBO, Mariah. Curitiba usa reconhecimento facial como ferramenta na segurança pública sem ter regulação específica; 'sociedade de controle', critica especialista. **G1 Paraná**, Curitiba, 02 out. 2023. Disponível em: <https://g1.globo.com/pr/parana/noticia/2023/10/02/curitiba-usa-reconhecimento-facial-como-ferramenta-na-seguranca-publica-sem-ter-regulacao-especifica-sociedade-de-controle-critica-especialista.ghml>. Acesso em: 17 nov. 2025.

COUNCIL OF EUROPE. **Guidelines on the responsible implementation of artificial intelligence systems in journalism**. Strasbourg: Council of Europe, 2023. Disponível em: <https://rm.coe.int/cdmsi-2023-014-guidelines-on-the-responsible-implementation-of-artific/1680adb4c6>. Acesso em: 08 fev. 2025.

COULDRY, Nick; MEJIAS, Ulises A. **The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism**. Stanford: Stanford University Press, 2019.

CURITIBA. Prefeitura Municipal. **Conecta Muralha Curitiba: Integrar Câmera Privada à Segurança Pública**. Curitiba: Prefeitura Municipal, 2025. Disponível em: <https://www.curitiba.pr.gov.br/servicos/conecta-muralha-curitiba-integrar-camera-privada-a-seguranca-publica/1001>. Acesso em: 31 out. 2025.

DAGNINO, Renato. **Neutralidade da ciência e determinismo tecnológico: um debate sobre a tecnociência**. Campinas, SP: Editora da Unicamp, 2008.

DARPA. **Toward Machines that Improve with Experience**. Defense Advanced Research Projects Agency, 16 mar. 2017. Disponível em: <https://www.darpa.mil/news/2017/machines-improve-experience>. Acesso em: 24 ago. 2025.

DATA FAVELA. **89% de moradores da periferia no Brasil já foram vítimas de fake news, aponta pesquisa**. São Paulo, 5 jul. 2024. Disponível em: <https://g1.globo.com/sp/sao-paulo/noticia/2024/07/05/89percent-de-moradores-da-periferia-no-brasil-ja-foram-vitimas-de-fake-news-aponta-pesquisa.ghtml>. Acesso em: 24 nov. 2025.

DIAKOPOULOS, Nicholas. **Algorithmic Accountability Reporting: On the Investigation of Black Boxes**. New York: Tow Center for Digital Journalism, Columbia University, 2014. Disponível em: <https://doi.org/10.7916/D8ZK5TW2>. Acesso em: 29 mai. 2025.

DUCHIADE, André. Jornalistas brasileiras adotam IA em suas rotinas, mas setor vê ameaça econômica, mostra pesquisa. **LatAm Journalism Review**, Austin, 14 fev. 2025. Disponível em: <https://latamjournalismreview.org/pt-br/articles/jornalistas-brasileiras-adotam-ia-em-suas-rotinas-mas-setor-ve-ameaca-economica-mostra-pesquisa/>. Acesso em: 24 nov. 2025.

ESTADÃO. **Estadão 150 anos: qual o futuro do jornalismo?** | Eugênio Bucci e Paula Miraglia | Dois Pontos #60. 22 jan. 2025. 1 vídeo (58 min). Publicado pelo canal Estadão. Disponível em: <https://www.youtube.com/watch?v=pVW2WGclnBQ>. Acesso em: 20 set. 2025.

FANON, Frantz. **Em defesa da revolução africana**. Lisboa: Livraria Sá da Costa, 1980.

FATML. **Principles for Accountable Algorithms and a Social Impact Statement for Algorithms**. [S.l.]: Fairness, Accountability, and Transparency in Machine Learning, 2016. Disponível em: <https://www.fatml.org/resources/principles-for-accountable-algorithms>. Acesso em: 29 mai. 2025.

FREITAS, Rosiane Correia de. Comitê rejeitou todos os pedidos sobre a Muralha Digital. **Plural**, Curitiba, 15 jan. 2025. Disponível em: <https://www.plural.jor.br/comite-rejeitou-pedidos-muralha-digital/>. Acesso em: 17 nov. 2025.

GAZETA DO POVO. Muralha Digital: tecnologia e inteligência na segurança de Curitiba. **Gazeta do Povo**, Curitiba, 2025. Especial 332 Anos de Curitiba. Disponível em: <https://www.gazetadopovo.com.br/conteudo-publicitario/332-anos-curitiba/muralha-digital-curitiba/>. Acesso em: 30 nov. 2025.

HOLDSWORTH, J.; SCAPICCHIO, M. **O que é o Deep Learning?** IBM, 2024. Disponível em: <https://www.ibm.com/br-pt/topics/deep-learning>. Acesso em: 26 nov. 2024.

HU, Krystal. ChatGPT sets record for fastest-growing user base - analyst note. **Reuters**, 1 fev. 2023. Disponível em: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>. Acesso em: 15 jan. 2025.

HUFF, Darrell. **Como mentir com estatística**. Rio de Janeiro: Intrínseca, 2019.

JANIESCH, Christian; ZSCHECH, Patrick; HEINRICH, Kai. Machine learning and deep learning. **Electronic Markets**, v. 31, n. 3, p. 685-695, set. 2021. Disponível em: <https://doi.org/10.1007/s12525-021-00475-2>. Acesso em: 26 nov. 2024.

KNIGHT CENTER FOR JOURNALISM IN THE AMERICAS. **Generative AI for journalists**: discovering what data can do. Austin: University of Texas at Austin, 2025. Disponível em: <https://journalismcourses.org/course/generative-ai-for-journalists-boc/>. Acesso em: 26 nov. 2025.

KREMER, Bianca. **Racismo algorítmico**. Organização de Thallita G. L. Lima e Pablo Nunes. Rio de Janeiro: CESeC, 2023. (Coleção Panorama). Disponível em: https://www.academia.edu/114040800/Racismo_Algorítmico. Acesso em: 04 jun. 2024.

LANEY, Doug. **3D Data Management**: Controlling Data Volume, Velocity and Variety. Stamford: META Group, 2001. (META Delta, File 949). Disponível em: <https://diegonogare.net/wp-content/uploads/2020/08/3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. Acesso em: 26 nov. 2024.

LEAL, Maria do Carmo et al. A cor da dor: iniquidades raciais na atenção pré-natal e ao parto no Brasil. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 33, sup. 1, e00078816, 2017. Disponível em: <https://doi.org/10.1590/0102-311X00078816>. Acesso em: 26 nov. 2024.

LEAL, Maria do Carmo et al. **Nascer no Brasil II**: pesquisa nacional sobre aborto, parto e nascimento 2022-2023: dados preliminares da pesquisa para oficina "Morte Materna de Mulheres Negras no Contexto do SUS". Rio de Janeiro: ENSP/Fiocruz, 2023. Disponível em: <https://nascernobrasil.ensp.fiocruz.br/wp-content/uploads/2023/11/Dados-preliminares-da-pesquisa-Nascer-no-Brasil-2.pdf>. Acesso em: 26 nov. 2024.

LEAL, Rodrigo Felix. [Fotografia]. In: PARANÁ. Agência Estadual de Notícias. **Novo centro de controle em Curitiba contribui com monitoramento do Estado**. Curitiba, 5 jan. 2021.

Disponível em: <https://www.parana.pr.gov.br/aen/Noticia/Novo-centro-de-controle-em-Curitiba-contribui-com-monitoramento-do-Estado>. Acesso em: 04 mai. 2024.

LEE, Kai-Fu. **AI Superpowers: China, Silicon Valley, and the New World Order**. Boston: Houghton Mifflin Harcourt, 2018.

LINARES, César. Associações de mídia da América Latina endossam princípios para desenvolvimento ético e justo da IA generativa. **LatAm Journalism Review**, Austin, 6 nov. 2023. Disponível em: <https://latamjournalismreview.org/pt-br/articles/associacoes-de-midia-da-america-latina-endossam-principios-para-desenvolvimento-etico-e-justo-da-ia-generativa/>. Acesso em: 27 set. 2025.

LYON, David. **Surveillance Studies: An Overview**. Cambridge: Polity Press, 2007.

MARKETS MEDIA. **Financial Market Data Spending Reaches Record \$42bn**. New York, 29 abr. 2024. Disponível em: <https://www.marketsmedia.com/financial-market-data-spending-reaches-record-42bn/>. Acesso em: 16 jun. 2024.

MINSKY, Marvin. Steps toward Artificial Intelligence. **Proceedings of the IRE**, v. 49, n. 1, p. 8-30, jan. 1961. Disponível em: <https://courses.csail.mit.edu/6.803/pdf/steps.pdf>. Acesso em: 26 nov. 2024.

MOROZOV, Evgeny. **To Save Everything, Click Here: The Folly of Technological Solutionism**. Nova York: PublicAffairs, 2013.

NAKAMURA, Lisa. **Cybertypes: Race, Ethnicity, and Identity on the Internet**. Nova York: Routledge, 2002.

NEVES, Ianaira. Reprovados por robôs: como plataformas de inteligência artificial podem discriminar mulheres, idosos e faculdades populares em processos seletivos. **The Intercept Brasil**, 24 nov. 2022. Disponível em: <https://www.intercept.com.br/2022/11/24/como-plataformas-de-inteligencia-artificial-podem-discriminar-mulheres-idosos-e-faculdades-populares-em-processos-seletivos/>. Acesso em: 11 fev. 2024.

NEWMAN, Nic et al. **Reuters Institute Digital News Report 2024**. Oxford: Reuters Institute for the Study of Journalism, 2024. Disponível em: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2024-06/RISJ_DNR_2024_Digital_v10%20lr.pdf. Acesso em: 01 dez. 2025.

NUNES, Pablo. Exclusivo: levantamento revela que 90,5% dos presos por monitoramento facial no Brasil são negros. **The Intercept Brasil**, 21 nov. 2019. Disponível em: <https://www.intercept.com.br/2019/11/21/presos-monitoramento-facial-brasil-negros/>. Acesso em: 11 fev. 2024.

O'NEIL, Cathy. **Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia**. Tradução de Rafael Abraham. Santo André: Editora Rua do Sabão, 2020.

OGURA, José Fernando. [Fotografia]. In: PARANÁ. Agência Estadual de Notícias. **Novo centro de controle em Curitiba contribui com monitoramento do Estado**. Curitiba, 5 jan. 2021. Disponível em: <https://www.parana.pr.gov.br/aen/Noticia/Novo-centro-de-controle-em-Curitiba-contribui-com-monitoramento-do-Estado>. Acesso em: 04 mai. 2024.

OSOBA, Osonde; WELSER IV, William. **An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence**. Santa Monica: RAND Corporation, 2017. Disponível em: <https://www.rand.org/t/RR1744>. Acesso em: 29 out. 2023.

PAIXÃO, Marcelo et al. (org.). **Relatório anual das desigualdades raciais no Brasil; 2009-2010**. Rio de Janeiro: Garamond, 2010.

PASQUALE, Frank. **The Black Box Society: The Secret Algorithms That Control Money and Information**. Cambridge: Harvard University Press, 2015.

POYNTER INSTITUTE. **AI ethics guidelines: a starter kit for newsrooms**. St. Petersburg: The Poynter Institute, 2025. Disponível em: <https://www.poynter.org/ai-ethics-journalism/ai-ethics-guidelines/>. Acesso em: 30 nov. 2025.

RADCLIFFE, Damian. Como a IA está mudando o jornalismo no Sul Global. **IJNet – Rede de Jornalistas Internacionais**, 6 mar. 2025. Disponível em: <https://ijnet.org/pt-br/story/como-ia-está-mudando-o-jornalismo-no-sul-global>. Acesso em: 01 nov. 2025.

REPÓRTERES SEM FRONTEIRAS (RSF). **Carta de Paris sobre IA e Jornalismo**. Paris: Repórteres Sem Fronteiras, 2023. Disponível em: https://rsf.org/sites/default/files/medias/file/2023/11/Carte%20de%20Paris%20sobre%20IA%20e%20jornalismo_0.pdf. Acesso em: 01 nov. 2025.

RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência artificial**. Tradução de Regina Célia Simille. 3. ed. Rio de Janeiro: Elsevier, 2013.

SCHWARTZ, Oscar. Untold History of AI: The Birth of Machine Bias. **IEEE Spectrum**, 16 abr. 2019. Disponível em: <https://spectrum.ieee.org/untold-history-of-ai-the-birth-of-machine-bias>. Acesso em: 11 fev. 2024.

SHUMAILOV, Ilia et al. AI models collapse when trained on recursively generated data. **Nature**, v. 631, p. 455–459, jul. 2024.

SILVA, Tarcízio (ed.) et al. **Mapeamento de Danos e Discriminação Algorítmica**. Desvelar, 2025. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>. Acesso em: 31 jun. 2025.

SILVA, Tarcízio (Org.). **Racismo Algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Edições Sesc São Paulo, 2022.

SILVA, Tarcízio. **Teoria Racial Crítica e Comunicação Digital: conexões contra a dupla opacidade**. Anais do 42º Congresso Brasileiro de Ciências da Comunicação. Belém, Pará, 2019.

SILVA, Tarcízio. Visão Computacional e Racismo Algorítmico: Branquitude e Opacidade no Aprendizado de Máquina. **Revista da ABPN**, v. 12, n. 31, p. 428-448, fev./jul. 2020. Disponível em: <https://abpnrevista.org.br/site/article/view/744>. Acesso em: 31 out. 2024

STRYKER, Cole; KAVLAKOGLU, Eda. What is Artificial Intelligence (AI)? **IBM**, 2024. Disponível em: <https://www.ibm.com/br-pt/think/topics/artificial-intelligence>. Acesso em: 26 nov. 2024.

THOMSON REUTERS FOUNDATION. **Journalism in the AI Era: A TRF Insights Survey**. Londres, 2025. Disponível em: <https://www.trust.org/wp-content/uploads/2025/01/TRF-Insights-Journalism-in-the-AI-Era.pdf>. Acesso em: 01 dez. 2025.

ZUBOFF, Shoshana. **A era do capitalismo de vigilância**: A disputa por um futuro humano na nova fronteira do poder. Rio de Janeiro: Intrínseca, 2021.