

Fabiano Couto Corrêa da Silva

QUEM CONTROLA SEUS DADOS?

Ciência Aberta, Colonialismo de Dados e Soberania
na era da Inteligência Artificial e do Big Data



Fabiano Couto Corrêa da Silva

QUEM CONTROLA SEUS DADOS?

Ciência Aberta, Colonialismo de Dados e Soberania
na era da Inteligência Artificial e do Big Data



Programa de Pós-Graduação
em Ciência da Informação

| São Paulo | 2025 |



DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

S586q

Silva, Fabiano Couto Corrêa da -

Quem controla seus dados? Ciência Aberta, Colonialismo de Dados e Soberania na era da Inteligência Artificial e do Big Data / Fabiano Couto Corrêa da Silva. – São Paulo: Pimenta Cultural, 2025.

Livro em PDF

ISBN 978-85-7221-474-2

DOI 10.31560/pimentacultural/978-85-7221-474-2

1. Ciência Aberta. 2. Colonialismo de dados. 3. Soberania informacional. 4. Governança e preservação de dados. 5. Justiça epistêmica. I. Silva, Fabiano Couto Corrêa da. II. Título.

CDD 004.21

Índice para catálogo sistemático:

I. Governança e preservação de dados

Simone Sales - Bibliotecária - CRB ES-000814/0

Copyright © Pimenta Cultural, alguns direitos reservados.

Copyright do texto © 2025 o autor.

Copyright da edição © 2025 Pimenta Cultural.

Esta obra é licenciada por uma Licença Creative Commons:
Atribuição-NãoComercial-SemDerivações 4.0 Internacional - (CC BY-NC-ND 4.0).
Os termos desta licença estão disponíveis em:
<<https://creativecommons.org/licenses/>>.
Direitos para esta edição cedidos à Pimenta Cultural.
O conteúdo publicado não representa a posição oficial da Pimenta Cultural.

Direção editorial	Patricia Bieging Raul Inácio Busarello
Editora executiva	Patricia Bieging
Gerente editorial	Landressa Rita Schiefelbein
Assistente editorial	Júlia Marra Torres
Estagiária editorial	Ana Flávia Pivisan Kobata
Diretor de criação	Raul Inácio Busarello
Assistente de arte	Naiara Von Groll
Editoração eletrônica	Andressa Karina Voltolini
Estagiária em editoração	Stela Tiemi Hashimoto Kanada
Imagens da capa	cuong, titima037 - Freepik.com
Tipografias	Acumin, Obviously
Revisão	O autor
Autor	Fabiano Couto Corrêa da Silva

PIMENTA CULTURAL
São Paulo • SP
+55 (11) 96766 2200
livro@pimentacultural.com
www.pimentacultural.com



CONSELHO EDITORIAL CIENTÍFICO

Doutores e Doutoradas

Adilson Cristiano Habowski <i>Universidade La Salle, Brasil</i>	Bernadette Beber <i>Universidade Federal de Santa Catarina, Brasil</i>
Adriana Flávia Neu <i>Universidade Federal de Santa Maria, Brasil</i>	Bruna Carolina de Lima Siqueira dos Santos <i>Universidade do Vale do Itajaí, Brasil</i>
Adriana Regina Vettorazzi Schmitt <i>Instituto Federal de Santa Catarina, Brasil</i>	Bruno Rafael Silva Nogueira Barbosa <i>Universidade Federal da Paraíba, Brasil</i>
Aguimario Pimentel Silva <i>Instituto Federal de Alagoas, Brasil</i>	Caio Cesar Portella Santos <i>Instituto Municipal de Ensino Superior de São Manuel, Brasil</i>
Alaim Passos Bispo <i>Universidade Federal do Rio Grande do Norte, Brasil</i>	Carla Wanessa do Amaral Caffagni <i>Universidade de São Paulo, Brasil</i>
Alaim Souza Neto <i>Universidade Federal de Santa Catarina, Brasil</i>	Carlos Adriano Martins <i>Universidade Cruzeiro do Sul, Brasil</i>
Alessandra Knoll <i>Universidade Federal de Santa Catarina, Brasil</i>	Carlos Jordan Lapa Alves <i>Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil</i>
Alessandra Regina Müller Germani <i>Universidade Federal de Santa Maria, Brasil</i>	Caroline Chioquetta Lorenset <i>Universidade Federal de Santa Catarina, Brasil</i>
Aline Corso <i>Universidade do Vale do Rio dos Sinos, Brasil</i>	Cassia Cordeiro Furtado <i>Universidade Federal do Maranhão, Brasil</i>
Aline Wendpap Nunes de Siqueira <i>Universidade Federal de Mato Grosso, Brasil</i>	Cássio Michel dos Santos Camargo <i>Universidade Federal do Rio Grande do Sul, Brasil</i>
Ana Rosangela Colares Lavand <i>Universidade Estadual do Norte do Paraná, Brasil</i>	Cecília Machado Henriques <i>Universidade Federal de Santa Catarina, Brasil</i>
André Gobbo <i>Universidade Federal da Paraíba, Brasil</i>	Christiano Martino Otero Avila <i>Universidade Federal de Pelotas, Brasil</i>
André Tanus Cesário de Souza <i>Faculdade Anhanguera, Brasil</i>	Cláudia Samuel Kessler <i>Universidade Federal do Rio Grande do Sul, Brasil</i>
Andressa Antunes <i>Universidade Federal de Ouro Preto, Brasil</i>	Cristiana Barcelos da Silva <i>Universidade do Estado de Minas Gerais, Brasil</i>
Andressa Wiebusch <i>Universidade Federal de Santa Maria, Brasil</i>	Cristiane Silva Fontes <i>Universidade Federal de Minas Gerais, Brasil</i>
Andreza Regina Lopes da Silva <i>Universidade Federal de Santa Catarina, Brasil</i>	Daniela Susana Segre Guertzenstein <i>Universidade de São Paulo, Brasil</i>
Angela Maria Farah <i>Universidade de São Paulo, Brasil</i>	Daniele Cristine Rodrigues <i>Universidade de São Paulo, Brasil</i>
Anísio Batista Pereira <i>Universidade do Estado do Amapá, Brasil</i>	Dayse Centurion da Silva <i>Universidade Anhanguera, Brasil</i>
Antonio Edson Alves da Silva <i>Universidade Estadual do Ceará, Brasil</i>	Dayse Sampaio Lopes Borges <i>Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil</i>
Antonio Henrique Coutelo de Moraes <i>Universidade Federal de Rondonópolis, Brasil</i>	Deilson do Carmo Trindade <i>Instituto Federal de Educação, Ciência e Tecnologia do Amazonas, Brasil</i>
Arthur Vianna Ferreira <i>Universidade do Estado do Rio de Janeiro, Brasil</i>	Diego Pizarro <i>Instituto Federal de Brasília, Brasil</i>
Ary Albuquerque Cavalcanti Junior <i>Universidade Federal de Mato Grosso, Brasil</i>	Dorama de Miranda Carvalho <i>Escola Superior de Propaganda e Marketing, Brasil</i>
Asterlindo Bandeira de Oliveira Júnior <i>Universidade Federal da Bahia, Brasil</i>	Edilson de Araújo dos Santos <i>Universidade de São Paulo, Brasil</i>
Bárbara Amaral da Silva <i>Universidade Federal de Minas Gerais, Brasil</i>	Edson da Silva <i>Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brasil</i>

Elena Maria Mallmann

Universidade Federal de Santa Maria, Brasil

Eleonora das Neves Simões

Universidade Federal do Rio Grande do Sul, Brasil

Eliane Silva Souza

Universidade do Estado da Bahia, Brasil

Elvira Rodrigues de Santana

Universidade Federal da Bahia, Brasil

Estevão Schultz Campos

Centro Universitário Adventista de São Paulo, Brasil

Éverly Pegoraro

Universidade Federal do Rio de Janeiro, Brasil

Fábio Santos de Andrade

Universidade Federal de Mato Grosso, Brasil

Fabrícia Lopes Pinheiro

Universidade Federal do Estado do Rio de Janeiro, Brasil

Fauston Negreiros

Faustidade de Brasília, Brasil

Felipe Henrique Monteiro Oliveira

Universidade Federal da Bahia, Brasil

Fernando Vieira da Cruz

Universidade Estadual de Campinas, Brasil

Flávia Fernanda Santos Silva

Universidade Federal do Amazonas, Brasil

Gabriela Moysés Pereira

Universidade Federal do Rio de Janeiro, Brasil

Gabriella Eldereti Machado

Universidade Federal de Santa Maria, Brasil

Germano Ehler Pollnow

Universidade Federal de Pelotas, Brasil

Geuciane Felipe Guerim Fernandes

Universidade Federal do Pará, Brasil

Geymeesson Brito da Silva

Universidade Federal de Pernambuco, Brasil

Giovanna Ofretorio de Oliveira Martin Franchi

Universidade Federal de Santa Catarina, Brasil

Handherson Leylton Costa Damasceno

Universidade Federal da Bahia, Brasil

Hebert Elias Lobo Sosa

Universidad de Los Andes, Venezuela

Helciclever Barros da Silva Sales

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasil

Helena Azevedo Paulo de Almeida

Universidade Federal de Ouro Preto, Brasil

Hendy Barbosa Santos

Faculdade de Artes do Paraná, Brasil

Humberto Costa

Universidade Federal do Paraná, Brasil

Igor Alexandre Barcelos Graciano Borges

Universidade de Brasília, Brasil

Inara Antunes Vieira Willerding

Universidade Federal de Santa Catarina, Brasil

Jaziel Vasconcelos Dorneles

Universidade de Coimbra, Portugal

Jean Carlos Gonçalves

Universidade Federal do Paraná, Brasil

Joao Adalberto Campato Junior

Universidade Brasil, Brasil

Jocimara Rodrigues de Sousa

Universidade de São Paulo, Brasil

Joelson Alves Onofre

Universidade Estadual de Santa Cruz, Brasil

Jónata Ferreira de Moura

Universidade São Francisco, Brasil

Jonathan Machado Domingues

Universidade Federal de São Paulo, Brasil

Jorge Eschriqui Vieira Pinto

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Jorge Luís de Oliveira Pinto Filho

Universidade Federal do Rio Grande do Norte, Brasil

Juliana de Oliveira Vicentini

Universidade de São Paulo, Brasil

Juliano Milton Kruger

Instituto Federal do Amazonas, Brasil

Juliano Pizzano Ayoub

Universidade Estadual de Ponta Grossa, Brasil

Julierme Sebastião Morais Souza

Universidade Federal de Uberlândia, Brasil

Junior César Ferreira de Castro

Universidade de Brasília, Brasil

Katia Bruginski Mulik

Universidade de São Paulo, Brasil

Laionel Vieira da Silva

Universidade Federal da Paraíba, Brasil

Lauro Sérgio Machado Pereira

Instituto Federal do Norte de Minas Gerais, Brasil

Leonardo Freire Marino

Universidade do Estado do Rio de Janeiro, Brasil

Leonardo Pinheiro Mozdzenski

Universidade Federal de Pernambuco, Brasil

Letícia Cristina Alcântara Rodrigues

Faculdade de Artes do Paraná, Brasil

Lucila Romano Tragtenberg

Pontifícia Universidade Católica de São Paulo, Brasil

Lucimara Rett

Universidade Metodista de São Paulo, Brasil

Luiz Eduardo Neves dos Santos

Universidade Federal do Maranhão, Brasil

Maikel Pons Giralt

Universidade de Santa Cruz do Sul, Brasil

Manoel Augusto Polastrelli Barbosa

Universidade Federal do Espírito Santo, Brasil

Marcelo Nicomedes dos Reis Silva Filho

Universidade Estadual do Oeste do Paraná, Brasil

Márcia Alves da Silva

Universidade Federal de Pelotas, Brasil

Marcio Bernardino Sirino

Universidade Federal do Estado do Rio de Janeiro, Brasil

Marcos Pereira dos Santos

Universidad Internacional Iberoamericana del Mexico, México

Marcos Uzel Pereira da Silva

Universidade Federal da Bahia, Brasil

Marcus Fernando da Silva Praxedes

Universidade Federal do Recôncavo da Bahia, Brasil

Maria Aparecida da Silva Santandel

Universidade Federal de Mato Grosso do Sul, Brasil

Maria Cristina Giorgi

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brasil

Maria Edith Maroca de Avelar

Universidade Federal de Ouro Preto, Brasil

Marina Bezerra da Silva

Instituto Federal do Piauí, Brasil

Marines Rute de Oliveira

Universidade Estadual do Oeste do Paraná, Brasil

Mauricio José de Souza Neto

Universidade Federal da Bahia, Brasil

Michele Marcelo Silva Bortolai

Universidade de São Paulo, Brasil

Mônica Tavares Orsini

Universidade Federal do Rio de Janeiro, Brasil

Nara Oliveira Salles

Universidade do Estado do Rio de Janeiro, Brasil

Neide Araujo Castilho Teno

Universidade Estadual de Mato Grosso do Sul, Brasil

Neli Maria Mengalli

Pontifícia Universidade Católica de São Paulo, Brasil

Patrícia Bieging

Universidade de São Paulo, Brasil

Patricia Flavia Mota

Universidade do Estado do Rio de Janeiro, Brasil

Patrícia Helena dos Santos Carneiro

Universidade Federal de Rondônia, Brasil

Rainei Rodrigues Jadejiski

Universidade Federal do Espírito Santo, Brasil

Raul Inácio Busarello

Universidade Federal de Santa Catarina, Brasil

Raymundo Carlos Machado Ferreira Filho

Universidade Federal do Rio Grande do Sul, Brasil

Ricardo Luiz de Bittencourt

Universidade do Extremo Sul Catarinense, Brasil

Roberta Rodrigues Ponciano

Universidade Federal de Uberlândia, Brasil

Robson Teles Gomes

Universidade Católica de Pernambuco, Brasil

Rodiney Marcelo Braga dos Santos

Universidade Federal de Roraima, Brasil

Rodrigo Amancio de Assis

Universidade Federal de Mato Grosso, Brasil

Rodrigo Sarruge Molina

Universidade Federal do Espírito Santo, Brasil

Rogério Rauber

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Rosane de Fatima Antunes Obregon

Universidade Federal do Maranhão, Brasil

Samuel André Pompeo

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Sebastião Silva Soares

Universidade Federal do Tocantins, Brasil

Silmar José Spinardi Franchi

Universidade Federal de Santa Catarina, Brasil

Simone Alves de Carvalho

Universidade de São Paulo, Brasil

Simoni Urnau Bonfiglio

Universidade Federal da Paraíba, Brasil

Stela Maris Vaucher Farias

Universidade Federal do Rio Grande do Sul, Brasil

Tadeu João Ribeiro Baptista

Universidade Federal do Rio Grande do Norte, Brasil

Taiane Aparecida Ribeiro Nepomoceno

Universidade Estadual do Oeste do Paraná, Brasil

Taíza da Silva Gama

Universidade de São Paulo, Brasil

Tania Micheline Miorando

Universidade Federal de Santa Maria, Brasil

Tarcísio Vanzin

Universidade Federal de Santa Catarina, Brasil

Tascieli Feltrin

Universidade Federal de Santa Maria, Brasil

Tatiana da Costa Jansen

Serviço Nacional de Aprendizagem Comercial, Brasil

Tayson Ribeiro Teles

Universidade Federal do Acre, Brasil

Thiago Barbosa Soares

Universidade Federal do Tocantins, Brasil

Thiago Camargo Iwamoto

Universidade Estadual de Goiás, Brasil

Thiago Medeiros Barros

Universidade Federal do Rio Grande do Norte, Brasil

Tiago Mendes de Oliveira

Universidade Federal do Rio Grande do Norte, Brasil

Vanessa de Sales Marruche

Universidade Federal do Amazonas, Brasil

Vanessa Elisabete Raue Rodrigues

Universidade Estadual do Centro Oeste, Brasil

Vania Ribas Ulbricht
Universidade Federal de Santa Catarina, Brasil

Vinicius da Silva Freitas
Centro Universitário Vale do Cricaré, Brasil

Wellington Furtado Ramos
Universidade Federal de Mato Grosso do Sul, Brasil

Wellton da Silva de Fatima
Instituto Federal de Alagoas, Brasil

Wenis Vargas de Carvalho
Universidade Federal da Grande Dourados, Brasil

Yan Masetto Nicolai
Universidade Federal de São Carlos, Brasil

PARECERISTAS E REVISORES(AS) POR PARES

Avaliadores e avaliadoras Ad-Hoc

Alcidinei Dias Alves
Logos University International, Estados Unidos

Alessandra Figueiró Thornton
Universidade Luterana do Brasil, Brasil

Alexandre João Appio
Universidade do Vale do Rio dos Sinos, Brasil

Artur Pires de Camargos Júnior
Universidade do Vale do Sapucaí, Brasil

Bianka de Abreu Severo
Universidade Federal de Santa Maria, Brasil

Carlos Eduardo B. Alves
Universidade Federal do Agreste de Pernambuco, Brasil

Carlos Eduardo Damian Leite
Universidade de São Paulo, Brasil

Catarina Prestes de Carvalho
Instituto Federal Sul-Rio-Grandense, Brasil

Davi Fernandes Costa
Secretaria Municipal de Educação de São Paulo, Brasil

Denilson Marques dos Santos
Universidade do Estado do Pará, Brasil

Domingos Aparecido dos Reis
Must University, Estados Unidos

Edson Vieira da Silva de Camargos
Logos University International, Estados Unidos

Edwins de Moura Ramires
Serviço Nacional de Aprendizagem Comercial, Brasil

Elisiene Borges Leal
Universidade Federal do Piauí, Brasil

Elizabete de Paula Pacheco
Universidade Federal de Ubertândia, Brasil

Elton Simomukay
Universidade Estadual de Ponta Grossa, Brasil

Francisco Geová Goveia Silva Júnior
Universidade Potiguar, Brasil

Indiamaris Pereira
Universidade do Vale do Itajaí, Brasil

Jacqueline de Castro Rimá
Universidade Federal da Paraíba, Brasil

Jonas Lacchini
Pontifícia Universidade Católica de Minas Gerais, Brasil

Lucimar Romeu Fernandes
Instituto Politécnico de Bragança, Brasil

Marcos de Souza Machado
Universidade Federal da Bahia, Brasil

Michele de Oliveira Sampaio
Universidade Federal do Espírito Santo, Brasil

Nívea Consuêlo Carvalho dos Santos
Serviço Nacional de Aprendizagem Comercial, Brasil

Pedro Augusto Paula do Carmo
Universidade Paulista, Brasil

Rayner do Nascimento Souza
Serviço Nacional de Aprendizagem Comercial, Brasil

Samara Castro da Silva
Universidade de Caxias do Sul, Brasil

Sidney Pereira Da Silva
Stockholm University, Suécia

Suêlen Rodrigues de Freitas Costa
Universidade Federal do Espírito Santo, Brasil

Thais Karina Souza do Nascimento
Instituto de Ciências das Artes, Brasil

Viviane Gil da Silva Oliveira
Universidade Federal do Amazonas, Brasil

Walmir Fernandes Pereira
Miami University of Science and Technology, Estados Unidos

Weyber Rodrigues de Souza
Pontifícia Universidade Católica de Goiás, Brasil

William Roslindo Paranhos
Universidade Federal de Santa Catarina, Brasil

Parecer e revisão por pares

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Pimenta Cultural, bem como revisados por pares, sendo indicados para a publicação.

*Os dados que não governamos
nos governam.*

— *Fabiano Couto Corrêa da Silva*

SUMÁRIO

Prefácio	17

PARTE I

FUNDAMENTOS TEÓRICOS E CONCEITUAIS	28

CAPÍTULO 1

Introdução ao colonialismo de dados.....	29
1.1 A economia dos dados e as grandes corporações (<i>Big Tech</i>)	40
1.2 Assimetrias de poder e vigilância na era do Big Data	45
1.3 Marcos teóricos: dos estudos pós-coloniais à teoria crítica da informação	50
1.4 Paradoxos e contradições do colonialismo de dados.....	55
1.5 Resistências e alternativas emergentes	58
1.6 Dimensões psicológicas e subjetivas.....	60
1.7 Impactos ambientais e sustentabilidade.....	61
1.8 Perspectivas futuras e tendências emergentes.....	62

CAPÍTULO 2

Ciência aberta: tensões, contradições e paradoxos	64
2.1 A evolução do conceito de ciência aberta	66

2.2 Princípios FAIR e democratização do conhecimento.....	68
2.3 Modelos de acesso aberto e sustentabilidade	74
2.4 Tensões entre abertura e controle.....	78
2.5 Algoritmos e curadoria: o poder invisível da descoberta	84
2.6 Ciência cidadã e participação democrática	88
2.7 Paradoxos da democratização científica.....	91
2.8 Alternativas e resistências à ciência aberta hegemônica	92
2.9 Perspectivas futuras	94

PARTE II

ANÁLISE CRÍTICA E IMPACTOS 96

CAPÍTULO 3

O impacto do colonialismo de dados na produção científica e no conhecimento 97

3.1. Desigualdades na produção científica global (Norte global vs. Sul global)	98
3.2 Apropriação de dados científicos e parachute science	102
3.3. Epistemologias marginalizadas e colonialidade do saber na ciência.....	106
3.4. Relação entre dados abertos e assimetria de poder	113
3.5 Algoritmos, vieses e reprodução de desigualdades.....	115

CAPÍTULO 4

**Fraudes científicas, manipulação
de dados e práticas antiéticas..... 118**

**4.1. Fundamentos históricos e teóricos
da integridade científica..... 123**

4.1.1 Evolução histórica dos conceitos
de integridade científica..... 123

4.1.2 Marcos normativos contemporâneos
e codificação internacional 126

4.1.3 Dimensões neurobiológicas
e psicológicas da ética científica..... 128

4.1.4 Teorias contemporâneas
de integridade científica..... 130

4.1.5 Colonialismo científico
e estruturas de poder global 132

**4.2. A crise contemporânea:
dimensões e magnitude do problema 134**

4.2.1. Análise estatística da escalada
de retrações científicas..... 134

4.2.2 Paper Mills: a industrialização
sistemática da fraude..... 136

4.2.3 Impactos da pandemia COVID-19
na integridade científica..... 138

4.2.4 Dimensões econômicas
e desperdício de recursos sistêmicos..... 140

**4.3 Inteligência artificial:
a transformação fundamental da má conduta científica..... 142**

4.3.1 A dupla face da revolução
tecnológica: promessa e ameaça simultâneas..... 143

4.3.2 Tipologia expandida da má conduta
científica digitalmente amplificada 146

4.3.3 Geração de conteúdo fraudulento: análise multimodal.....	147

4.3.4 Implicações epistemológicas: redefinindo a natureza da evidência.....	150

4.3.5 Características identificáveis, padrões de detecção e evolução estratégica: uma análise integrada da sofisticação operacional.....	152

4.4 Ferramentas de detecção e contramedidas tecnológicas: a evolução da resposta científica à fraude digital	161

4.4.1 Evolução sistêmica dos paradigmas de detecção: da análise manual à inteligência artificial interpretável.....	161

4.4.2 Detecção avançada de manipulação visual: forense digital e análise de autenticidade.....	164

4.4.3 Análise linguística avançada e detecção de conteúdo gerado por ia: além da detecção de plágio tradicional.....	166

4.4.4 Tecnologias emergentes e direções futuras: blockchain, computação quântica e realidade aumentada.....	170

4.5 Dimensões éticas e sociais: impactos sistêmicos na confiança pública e coesão epistêmica	172

4.5.1 Erosão da confiança pública e fragmentação do consenso científico: uma análise multidimensional	174

4.5.2 Colonialismo científico e vulnerabilidades sistêmicas: reprodução de desigualdades globais	177

4.5.3 Dimensões neurobiológicas e psicológicas: estresse sistêmico e degradação do julgamento ético	180

4.5.4 Responsabilidade institucional e transformação cultural: além da responsabilização individual	182

4.6 Perspectivas futuras e recomendações estratégicas: construindo um sistema científico íntegro e equitativo	185

4.6.1 Visão integrada para governança da integridade científica: coordenação multi-escalar e abordagem sistêmica	187

4.6.2 Recomendações tecnológicas: inovação responsável, interpretabilidade e democratização	189

4.6.3 Reformas institucionais: transformação de incentivos, culturas, e estruturas de poder	191

4.6.4 Cooperação internacional e soberania científica: equidade global e justiça epistêmica	193

4.6.5 <i>Roadmap</i> de implementação: estratégias integradas de curto, médio e longo prazo para transformação sistêmica da integridade científica	195

PARTE III

ESTUDOS DE CASO E EXPERIÊNCIAS PRÁTICAS.	199

CAPÍTULO 5

Estudos de caso internacionais	200

5.1. Caso 1: Dados genéticos indígenas e desrespeito ao consentimento (o caso Havasupai)	201

5.2. Caso 2: SciELO e a democratização da comunicação científica na América Latina.....	208
5.3. Caso 3: universidades e Big Tech - a dependência de plataformas privadas na educação pública.....	214
5.4. Caso 4: Governança de dados em saúde durante a pandemia de COVID-19	221

CAPÍTULO 6

Experiências brasileiras e latino-americanas.....	228
6.1 Políticas de ciência aberta no Brasil	229
6.2 Repositórios institucionais e iniciativas regionais	231
6.3 Desafios da implementação da LGPD na pesquisa.....	233
6.4 Movimentos de software livre e tecnologias sociais.....	234
6.5 Ciência cidadã e participação comunitária.....	236
6.6 Redes de colaboração Sul-Sul	238

PARTE IV

PROPOSTAS E PERSPECTIVAS FUTURAS	239

CAPÍTULO 7

Propostas e soluções para uma governança de dados justa e ética	240
7.1 Soberania de dados e marcos regulatórios.....	241
7.2 Infraestruturas abertas e tecnologias decoloniais.....	243

7.3 Princípios éticos e inclusão	245
7.4 Recomendações de políticas públicas e iniciativas comunitárias	248
7.5 Modelos alternativos de financiamento e sustentabilidade	249
7.6 Educação e capacitação para soberania digital	251

CAPÍTULO 8

Perspectivas futuras e agenda de pesquisa	253
8.1 As novas fronteiras tecnológicas da pesquisa científica	255
8.2 inteligência artificial (IA) como catalisadora e desafio	255
8.3 Ciência descentralizada (DeSci) e a promessa da Blockchain	256
8.4 O dilema central: abertura global, soberania digital e o espectro do colonialismo digital	258
8.5 Agenda de pesquisa para uma ciência aberta, equitativa e soberana	260
8.6 Inclusão e Ciência Cidadã: democratizando a produção do conhecimento além da academia	269
Considerações finais: rumo a uma transformação sistêmica da ciência global	304
Referências Bibliográficas	307
Sobre o Autor	341
Índice remissivo	342

PREFÁCIO

A sociedade contemporânea atravessa uma transformação silenciosa, mas profundamente estruturante: os dados se tornaram o fundamento sobre o qual se constroem sistemas econômicos, dinâmicas políticas, relações sociais e estruturas epistêmicas. Já não são apenas artefatos auxiliares da ciência ou da tecnologia, mas ativos estratégicos disputados por corporações, governos e instituições em escala global. A centralidade dos dados na era digital não é neutra nem equânime: ela está ancorada em relações assimétricas de poder, que operam tanto no nível material quanto simbólico, ecoando padrões históricos de dominação que remontam aos primórdios da modernidade colonial.

Quando Cristóvão Colombo registrou em seu diário, em 12 de outubro de 1492, as primeiras observações sobre os povos das Américas (COLUMBUS, 1493), além de documentar encontros culturais, também inaugurou uma prática sistemática de catalogação, classificação e apropriação de conhecimentos que se tornaria central ao projeto colonial europeu. Os mapas, censos, registros etnográficos e inventários de recursos naturais produzidos durante os séculos de colonialismo não eram meros instrumentos técnicos, mas tecnologias de poder que permitiam administrar territórios, controlar populações e extrair riquezas. Como observa Benedict Anderson em *Comunidades Imaginadas*, o censo colonial não se limitava a contar pessoas. Ele as criava como categorias administrativas, definindo quem existia e como existia para o Estado colonial (ANDERSON, 2008).

Essa genealogia histórica é fundamental para compreender o que Couldry e Mejias (2019) denominaram “colonialismo de dados”. Não como metáfora, mas como continuidade estrutural de práticas

coloniais adaptadas à era digital (COULDRY). Assim como as potências coloniais dos séculos XVI ao XX se apropriaram de territórios, recursos e saberes, hoje grandes corporações tecnológicas e Estados vigilantes apropriam-se dos dados gerados pela vida humana em sociedade, convertendo experiências cotidianas em matéria-prima para acumulação de capital e exercício de poder.

A analogia não é superficial. O colonialismo histórico operava através de quatro mecanismos fundamentais que encontram paralelos diretos no colonialismo de dados contemporâneo: a exploração inicial de novos territórios (hoje, novos domínios da vida social digitalizável), a exploração sistemática de recursos (dados pessoais e comportamentais), a expansão territorial e populacional (plataformização global), e a exterminação ou assimilação de formas alternativas de organização social (homogeneização algorítmica e epistemicídio digital) (GROSFOGUEL, 2007).

Diante dessa nova paisagem, emergem questionamentos que ecoam as resistências anticoloniais do passado: quem controla os dados? A quem pertencem? Quais populações são mais expostas à extração digital e quais se beneficiam dessa exploração? Tais perguntas não são meramente técnicas, mas éticas, políticas e históricas. Elas nos remetem às lutas de Toussaint Louverture no Haiti (1791–1804), de Mahatma Gandhi na Índia (1869–1948), de Amílcar Cabral na Guiné-Bissau (1924–1973) e de tantos outros que questionaram não só a dominação política, como também os próprios fundamentos epistemológicos e ontológicos do projeto colonial (JAMES, 1989; GANDHI, 1997; CABRAL, 1973).

Este livro parte da hipótese de que estamos testemunhando uma nova forma de colonialismo, não centrada na posse de territórios físicos, mas na apropriação contínua da vida, convertida em dado. Esse fenômeno articula práticas de extração, classificação e comercialização de informações pessoais e coletivas, muitas vezes sem o consentimento ou o benefício das comunidades envolvidas.

Mais profundamente, ele envolve a padronização de epistemologias, a imposição de lógicas algorítmicas e o silenciamento de formas plurais de conhecimento.

A dimensão epistêmica dessa nova colonialidade é particularmente relevante em relação às tensões com a Ciência Aberta. Quando Frantz Fanon escreveu “Os Condenados da Terra” em 1961, ele alertava para como o colonialismo, além de explorar recursos materiais, destruía sistematicamente as formas de conhecimento dos povos colonizados, substituindo-as por epistemologias europeias apresentadas como universais (FANON, 1961). Hoje, algoritmos de classificação e sistemas de inteligência artificial reproduzem essa violência epistêmica em escala global, invisibilizando línguas, práticas culturais e modos de vida que não se encaixam nos padrões dominantes.

Os efeitos do colonialismo de dados são sentidos nos sistemas computacionais, como também nos modos de existir, resistir e conhecer. Comunidades marginalizadas veem-se alijadas da governança dos próprios dados, enquanto atores hegemônicos concentram capital, inovação e infraestrutura digital. O conhecimento deixa de ser um bem comum para tornar-se ativo financeiro; a ciência, frequentemente apresentada como neutra, torna-se coadjuvante na reprodução dessas assimetrias.

A crise contemporânea da soberania não se limita ao campo territorial ou militar; ela se manifesta com particular intensidade no domínio informacional. A soberania informacional refere-se à capacidade de indivíduos, comunidades e nações de controlar os dados que produzem, os fluxos informacionais aos quais estão submetidos e as decisões derivadas desses dados. Este conceito ecoa as lutas por autodeterminação que marcaram os processos de descolonização do século XX, mas ganha contornos específicos na era digital.

Quando Kwame Nkrumah cunhou o termo “neocolonialismo” em 1965, ele se referia à perpetuação da dominação através de mecanismos econômicos e culturais após a independência política formal (NKRUMAH, 1965). Hoje, países do Sul Global enfrentam uma situação análoga no domínio digital: formalmente soberanos, mas dependentes de plataformas cujo funcionamento interno é opaco, cujos termos de uso são impostos unilateralmente, e cujos interesses frequentemente divergem das necessidades de suas populações.

Essa assimetria digital mina a capacidade de formular políticas públicas baseadas em dados confiáveis e soberanos, compromete a autonomia científica nacional e expõe governos a interferências externas. A dependência de infraestruturas digitais controladas por corporações estrangeiras cria vulnerabilidades que vão além do econômico, afetando a própria capacidade de autodeterminação democrática.

Ao lado da soberania informacional, cresce a preocupação com a soberania tecnológica, conceito que extrapola o domínio da infraestrutura digital para abranger o desenvolvimento autônomo de soluções tecnológicas, a produção local de conhecimento e a capacidade de decidir, eticamente, sobre o que é adotado ou rejeitado em determinado contexto. Em tempos de inteligência artificial, vigilância ubíqua e colonialismo de plataforma, torna-se imperativo discutir quem projeta, treina, mantém e audita os sistemas que moldam as sociedades digitais.

A história da transferência de tecnologia durante o período colonial oferece lições importantes para compreender esses desafios contemporâneos. As ferrovias construídas pelos britânicos na Índia entre 1853 e 1947, por exemplo, não foram projetadas para promover o desenvolvimento local, mas para facilitar a extração de recursos e o controle territorial (KERR, 2007). Similarmente, muitas das tecnologias digitais transferidas hoje para países do Sul Global são projetadas para facilitar a extração de dados e a integração em cadeias globais de valor controladas por corporações do Norte Global.

A dependência de *softwares* proprietários, servidores externos, modelos preditivos não auditáveis e algoritmos opacos impede que comunidades tenham controle sobre os efeitos das tecnologias sobre suas vidas. Portanto, a soberania tecnológica exige não apenas infraestrutura, mas também política pública, formação técnica e uma visão crítica sobre a transferência de tecnologia que, muitas vezes, reproduz relações neocoloniais sob o discurso da inovação.

Por fim, a soberania epistêmica emerge como dimensão incontornável da disputa pelo sentido e pelo valor dos conhecimentos. Ela diz respeito à capacidade das comunidades de produzir, validar e utilizar seus próprios saberes, em seus próprios termos. Historicamente, as epistemologias do Sul foram deslegitimadas pelas estruturas científicas ocidentais, com seus critérios de objetividade, replicabilidade e formalização universal.

Quando Boaventura de Sousa Santos propõe uma “ecologia de saberes” que reconheça a diversidade epistemológica do mundo, ele está respondendo a séculos de epistemicídio - a destruição sistemática de formas de conhecimento não-ocidentais (SANTOS, 2007). Na era dos dados, essa marginalização ganha novas formas: os sistemas de classificação automatizados invisibilizam línguas, práticas culturais e modos de vida que não se encaixam nos padrões dominantes; bases de dados globais padronizadas exigem traduções que achatam as especificidades locais; e políticas de Ciência Aberta, se mal desenhadas, podem promover uma nova extração de saberes sem retorno ou reconhecimento.

A luta por soberania epistêmica não é apenas uma reação ao colonialismo informacional, mas um projeto afirmativo de pluralização das formas de conhecer e de proteger o direito coletivo de decidir o que deve e o que não deve ser compartilhado. Ela se inspira nas tradições de resistência epistêmica que vão desde a preservação de conhecimentos indígenas durante a colonização até os movimentos contemporâneos de descolonização das universidades.

Este livro propõe-se a explorar o ponto de tensão entre dois movimentos aparentemente contraditórios: o avanço do colonialismo de dados e a emergência da Ciência Aberta. Enquanto o primeiro aprofunda estruturas de dominação e vigilância, o segundo se apresenta como promessa de democratização do conhecimento, da transparência e da colaboração científica. No entanto, a Ciência Aberta, para cumprir esse papel emancipador, precisa confrontar seus próprios limites e contradições.

A abertura de dados e saberes, se feita sem critérios éticos e políticos claros, pode ser apropriada pelas mesmas forças que alimentam a extração predatória de informações. A chave para esse dilema reside em uma pergunta fundamental que ecoa os debates anticoloniais do passado: abrir para quem, em quais condições, e sob quais termos?

A leitura crítica da Ciência Aberta, neste contexto, exige que se vá além da retórica da inovação e da eficiência. É preciso compreender que a abertura de dados não é apenas uma questão técnica, mas profundamente política. As estruturas que sustentam a produção, o acesso e o uso de dados são marcadas por desigualdades históricas, geopolíticas e epistêmicas que remontam ao período colonial e se perpetuam através de novos mecanismos.

A ideia de ciência como patrimônio universal, embora nobre em seus ideais, frequentemente mascara o fato de que grande parte do conhecimento foi (e continua sendo) produzido a partir da invisibilização de saberes locais e da exploração de comunidades periféricas. Quando Alexander von Humboldt realizou suas expedições científicas pelas Américas entre 1799 e 1804, ele dependeu fundamentalmente do conhecimento de guias indígenas, mas esses saberes foram sistematicamente apagados dos registros científicos oficiais, contribuindo para a construção de uma ciência supostamente universal, mas de fato profundamente eurocêntrica (PRATT, 1992).

Hoje, práticas similares persistem sob novas formas. Dados sobre biodiversidade coletados em países tropicais são processados em laboratórios do Norte Global e patenteados por corporações farmacêuticas, sem retorno para as comunidades que preservaram esses conhecimentos por gerações. Algoritmos de inteligência artificial são treinados com dados extraídos de populações do Sul Global, mas os benefícios econômicos e tecnológicos se concentram em empresas sediadas no Norte Global.

A análise proposta nesta obra está ancorada em marcos teóricos que incluem os estudos pós-coloniais, a teoria crítica da informação, a ética da ciência e as epistemologias do Sul. Com base nesses referenciais, os capítulos aqui reunidos examinam casos concretos, revisões conceituais e propostas de ação transformadora. São discutidas as consequências da captura de dados biométricos por Estados autoritários; os impactos da plataformação da pesquisa científica; os riscos da inteligência artificial treinada sobre bases enviesadas; os desafios da governança de dados em contextos indígenas e o papel das infraestruturas tecnológicas como arenas de disputa geopolítica.

A relevância deste livro também reside em sua capacidade de articular teoria e prática, conectando reflexões conceituais com experiências concretas de resistência e reinvenção. As análises aqui apresentadas dialogam com iniciativas como os projetos de ciência cidadã que emergem das favelas do Rio de Janeiro, os repositórios comunitários desenvolvidos por organizações indígenas na Amazônia, os protocolos de governança ética de dados como os princípios CARE (Collective benefit, Authority to control, Responsibility, Ethics) e FAIR (Findable, Accessible, Interoperable, Reusable), e as iniciativas de soberania digital em comunidades tradicionais da África e Ásia.

Esses exemplos demonstram que não faltam alternativas viáveis e justas à lógica extrativista dominante. Eles se inscrevem

em uma longa tradição de resistência e criatividade que vai desde os quilombos e palenques das Américas até os movimentos contemporâneos de software livre e tecnologias sociais. O que está em jogo é a capacidade de reconhecer essas alternativas, fortalecê-las e inseri-las nos circuitos formais da ciência e da política pública.

No centro desta obra está o reconhecimento de que dados são mais do que números: são representações da vida, da memória, da cultura e da identidade dos povos. Tratar os dados como recurso exige responsabilidade. Tratar os dados como direito exige soberania. E tratar os dados como conhecimento compartilhado exige equidade e cuidado. Por isso, a Ciência Aberta precisa se articular com uma agenda de justiça epistêmica e decolonial, que leve em consideração tanto o acesso aos dados quanto os regimes de poder que definem quem pode produzir, interpretar e se beneficiar desses dados.

Num momento histórico em que a integridade científica é posta à prova por crises de confiabilidade, pressões por produtividade e proliferação de desinformação, o compromisso com uma ciência crítica, ética e comprometida com o bem comum nunca foi tão necessário. A pandemia de COVID-19 revelou tanto o potencial transformador da colaboração científica global quanto os riscos da dependência de infraestruturas digitais controladas por poucos atores (ZUBOFF, 2019). A crise climática exige respostas científicas urgentes, mas também levanta questões sobre quem define as prioridades de pesquisa e quem se beneficia das soluções tecnológicas.

A reconfiguração dos sistemas de conhecimento na era digital precisa ser acompanhada por uma reconfiguração ética e política que aprenda com as lutas anticoloniais do passado e se inspire nas resistências contemporâneas. Não se trata apenas de imaginar novas tecnologias, mas de reinventar os fundamentos que orientam a produção e a circulação do conhecimento, questionando hierarquias epistêmicas estabelecidas e criando espaços para a emergência de formas plurais de saber.

A tensão entre Ciência Aberta e colonialismo de dados representa um dos desafios mais complexos e urgentes da era digital. Como demonstrado ao longo deste livro, essa tensão não é meramente técnica ou regulatória, mas reflete questões fundamentais sobre poder, justiça, e o futuro da produção de conhecimento humano.

A Ciência Aberta, em seus ideais mais elevados, promete democratizar o acesso ao conhecimento, acelerar descobertas científicas, e promover transparência e *accountability* na pesquisa. No entanto, como vimos através dos casos e análises apresentados, a implementação da Ciência Aberta em contextos de desigualdade estrutural pode inadvertidamente reproduzir ou amplificar essas desigualdades.

O colonialismo de dados opera através de mecanismos sutis, mas poderosos, que transformam a vida humana em recurso extraível. Na ciência, isso se manifesta através da apropriação de dados científicos, da imposição de padrões epistemológicos dominantes, e da concentração de infraestruturas digitais em poucas mãos. Esses processos não apenas reproduzem desigualdades históricas, mas criam novas formas de dependência e exclusão. Busca-se demonstrar que alternativas são possíveis. Iniciativas como o SciELO na América Latina, os princípios CARE para dados indígenas, e movimentos de *software* livre globalmente mostram que é possível desenvolver abordagens para Ciência Aberta que promovem justiça e equidade. Essas iniciativas oferecem modelos importantes para transformações mais amplas.

A construção de um futuro onde a Ciência Aberta cumpra suas promessas democratizantes requer transformações fundamentais em como pensamos sobre dados, conhecimento e poder. Isso inclui o desenvolvimento de marcos regulatórios que protejam direitos individuais e coletivos, o investimento em infraestruturas abertas e sustentáveis, e a criação de incentivos que promovam colaboração e compartilhamento equitativos.

Mais fundamentalmente, requer um reconhecimento de que dados não são apenas recursos técnicos, mas representações da vida, cultura, e identidade humanas. Tratar dados com o respeito e cuidado que merecem exige ir além de considerações técnicas para incorporar dimensões éticas, culturais e políticas.

O caminho à frente não é simples nem direto. Requer navegação cuidadosa de tensões entre abertura e proteção, entre eficiência e equidade, entre inovação e tradição. Requer colaboração entre múltiplos atores, incluindo pesquisadores, instituições, governos, organizações da sociedade civil e comunidades, cada um contribuindo com suas perspectivas e capacidades únicas.

Mais importante, requer um compromisso contínuo com os valores de justiça, equidade, e dignidade humana que devem estar no centro de qualquer sistema de conhecimento verdadeiramente democrático. A Ciência Aberta não é um fim em si mesma, mas um meio para criar um mundo onde o conhecimento serve ao bem comum e onde todos têm oportunidade de participar na grande conversa humana sobre como compreender e melhorar nossa condição compartilhada.

O futuro da ciência na era digital ainda está sendo escrito. As escolhas que fazemos hoje sobre como governar dados, como estruturar sistemas de conhecimento, e como equilibrar diferentes valores e interesses determinarão se esse futuro será caracterizado por maior justiça e inclusão ou por novas formas de dominação e exclusão.

A tarefa não é fácil, mas é urgente e necessária. O conhecimento é um dos recursos mais preciosos da humanidade. A forma como escolhemos produzir, compartilhar e utilizar esse conhecimento determinarão futuro da ciência e, inseparavelmente, o futuro de nossa sociedade como um todo.

Este livro é, portanto, uma contribuição a esse esforço coletivo de reimaginação. Longe de oferecer respostas prontas ou soluções técnicas simples, ele oferece chaves de leitura crítica, caminhos metodológicos rigorosos e propostas concretas para aqueles que desejam construir uma Ciência Aberta verdadeiramente inclusiva, situada e transformadora. Uma ciência que compartilhe resultados e, ao mesmo tempo, distribua poder; que armazene dados, mas também escute vozes; que inove, sem deixar de reparar injustiças históricas.

A obra se dirige a pesquisadores, estudantes, ativistas, gestores públicos e todos aqueles que reconhecem a urgência de pensar criticamente os rumos da ciência na era dos dados. Ela busca contribuir para a desconstrução de mitos de neutralidade e universalidade que ainda permeiam o discurso científico, e para a construção de pontes entre saberes diversos que foram historicamente separados por fronteiras disciplinares, geográficas e epistêmicas.

Além de ser um diagnóstico dos problemas contemporâneos, este livro é um convite à imaginação política e epistêmica necessária para sonhar e realizar futuros pós-coloniais, nos quais o conhecimento seja verdadeiramente um bem comum da humanidade. Futuros nos quais a Ciência Aberta não seja apenas uma questão de acesso técnico, mas de justiça social; nos quais a democratização do conhecimento não reproduza as exclusões do passado, mas as transforme em inclusões do presente.

Parte

FUNDAMENTOS TEÓRICOS E CONCEITUAIS

1

INTRODUÇÃO AO COLONIALISMO DE DADOS

O colonialismo de dados, termo cunhado por Couldry e Mejias (2019) para capturar a essência predatória da economia digital, não é uma mera metáfora. Trata-se de uma estrutura de poder que converte a vida humana em matéria-prima, inaugurando uma nova era de colonialidade organizada em quatro estágios: exploração, extração, expansão e extermínio. No ambiente digital, essa lógica manifesta-se quando mercados, populações e experiências pessoais são quantificados e governados por dados, reconfigurando a própria existência social. E ainda que os modos e intensidades difiram do colonialismo histórico, a sua função essencial de desapropriação permanece.

Para compreender a profundidade dessa analogia, é necessário examinar como cada estágio do colonialismo de dados se manifesta no contexto digital contemporâneo. A exploração inicial envolve a identificação de novos mercados, populações e domínios da vida humana que podem ser convertidos em fontes de dados, abrangendo desde a expansão geográfica para países e regiões ainda não explorados até a penetração em aspectos da experiência humana que antes permaneciam fora do alcance da quantificação digital. Essa transformação, porém, vai além das analogias superficiais, revelando complexidades estruturais que definem nossa era: vivemos um momento histórico singular, em que a própria natureza da existência humana está sendo reconfigurada por processos de dataficação que alcançam os aspectos mais íntimos da vida cotidiana.

Em outras palavras, o colonialismo de dados envolve a extração sistemática de dados pessoais e comportamentais de indivíduos em escala global, que são então gerenciados e monetizados por poucos atores hegemônicos, criando uma ordem social baseada no rastreamento constante e na vigilância ubíqua (ZUBOFF, 2019). Para compreender essa nova lógica, é necessário traçar um paralelo com o colonialismo histórico. Desde o período colonial, o controle da informação foi um instrumento de dominação: mapas, censos e registros etnográficos serviram para administrar territórios e populações (GROSSI; LEPRATTI, 2021).

Contudo, a escala e a intensidade da extração contemporânea superam qualquer precedente histórico. Enquanto os censos coloniais eram realizados periodicamente e capturavam apenas informações básicas sobre populações, os sistemas de coleta de dados digitais operam continuamente, 24 horas por dia, 7 dias por semana, capturando informações que vão muito além dos dados demográficos, abrangendo também padrões comportamentais, preferências, emoções, relacionamentos sociais, movimentos físicos, hábitos de consumo, e até mesmo dados biométricos como batimentos cardíacos e padrões de sono.

Hoje, com a explosão da coleta digital, cada clique, cada movimento e cada transação geram trilhas de dados que podem ser capturadas, armazenadas e analisadas. Os indivíduos tornam-se fontes de matéria-prima digital, frequentemente sem saber como seus dados são utilizados. Assim como os colonizados historicamente foram privados dos frutos das riquezas exploradas, os usuários hoje raramente controlam os benefícios extraídos de seus próprios dados (RICAURTE, 2019).

Cada interação em mídias sociais, cada busca *online* ou deslocamento registrado por smartphones contribui para um fluxo de dados explorado comercialmente. Esse processo de transformação da vida cotidiana em insumos informacionais levou Couldry e Mejias (2019) a propor que estamos diante de uma nova forma de colonialismo, no qual a apropriação da vida humana como recurso torna-se central.

O paradoxo fundamental do colonialismo de dados é sua sutileza: ele não se impõe pela força, como seu predecessor histórico, mas opera pela sedução tecnológica e por um consentimento aparente. Os usuários voluntariamente se inscrevem em plataformas digitais, baixam aplicativos, utilizam serviços “gratuitos” e compartilham informações pessoais em troca de conveniência, entretenimento e conectividade social. Esta dimensão consensual cria uma

ilusão de agência e escolha que mascara as relações de poder assimétricas que caracterizam o sistema.

Contudo, essas dependências transcendem o tecnológico: são profundamente sociais, econômicas e psicológicas. A arquitetura das plataformas digitais é deliberadamente projetada para maximizar o engajamento e a coleta de dados, utilizando técnicas derivadas da psicologia comportamental e da neurociência para criar padrões de uso compulsivo. Notificações *push*, *feeds* infinitos, sistemas de recompensa variável, e algoritmos de recomendação são algumas das estratégias utilizadas para manter os usuários constantemente conectados e gerando dados.

Considere, por exemplo, como a pandemia de COVID-19 acelerou dramaticamente a digitalização de atividades que anteriormente eram realizadas *offline*. O trabalho remoto, a educação à distância, as consultas médicas virtuais, as compras *online* e até mesmo as interações sociais foram massivamente transferidas para plataformas digitais, criando novas oportunidades de coleta de dados em escala sem precedentes. O que antes eram conversas privadas em escritórios tornaram-se reuniões do Zoom passíveis de gravação e análise; o que antes eram aulas presenciais tornaram-se sessões *online* que podem ser monitoradas e quantificadas; o que antes eram consultas médicas confidenciais tornaram-se interações digitais que geram rastros de dados sobre condições de saúde.

Esta nova forma de colonialismo opera através das chamadas “relações de dados”, que garantem a conversão contínua da experiência cotidiana em fluxo informacional monetizável. Trata-se de uma nova ordem social baseada na vigilância, que normaliza a exploração do ser humano por meio de seus dados. Como observa Chinmayi Arun, em sua análise crítica da obra de Mejjias e Couldry (2019), “o colonialismo de dados tanto reforça danos antigos quanto inflige novas lesões” (ARUN, 2025).

O paralelismo histórico é fundamental para compreender a natureza estrutural dessa nova forma de dominação. Durante o período colonial clássico, as potências europeias estabeleceram sistemas de extração de recursos naturais e exploração de mão de obra que seguiam padrões específicos: primeiro, a exploração de territórios desconhecidos para identificar recursos; segundo, a extração desses recursos; terceiro, a expansão do sistema para novos territórios; e finalmente, o extermínio ou subjugação das populações locais que resistiam ao processo.

Mas há uma diferença fundamental que torna o colonialismo de dados potencialmente mais pervasivo e duradouro que seu predecessor histórico: enquanto o colonialismo tradicional dependia da ocupação física de territórios e do controle direto sobre populações, o colonialismo de dados opera através de infraestruturas digitais que transcendem fronteiras geográficas e penetram diretamente na vida cotidiana dos indivíduos. Um smartphone carregado no bolso de uma pessoa em Lagos, Nigéria, pode estar coletando dados que são processados em servidores na Califórnia, analisados por algoritmos desenvolvidos em Tel Aviv, e utilizados para tomar decisões que afetam oportunidades de emprego em Porto Alegre.

No colonialismo de dados contemporâneo, observamos padrões similares aplicados ao domínio informacional. As grandes corporações tecnológicas primeiro exploram novos mercados e populações para identificar fontes de dados. Em seguida, estabelecem sistemas de extração sistemática através de plataformas digitais, aplicativos e dispositivos conectados. A expansão ocorre através da penetração em novos mercados geográficos e demográficos, enquanto o “extermínio” manifesta-se na eliminação de alternativas tecnológicas locais e na imposição de padrões globais de classificação e processamento de dados.

Um exemplo paradigmático dessa dinâmica é a expansão do Facebook (agora Meta) para mercados emergentes através de

iniciativas como o Internet.org (posteriormente renomeado Free Basics). Sob o pretexto de democratizar o acesso à internet, o Facebook ofereceu acesso “gratuito” a uma versão limitada da internet que incluía principalmente seus próprios serviços. Esta estratégia permitiu à empresa estabelecer uma posição dominante em mercados onde a infraestrutura de internet era limitada, criando dependências estruturais que são difíceis de reverter uma vez estabelecidas.

A eficiência de extração, característica fundamental do colonialismo histórico, também se manifesta no colonialismo de dados. Como destacam Mejias e Couldry (2019), o colonialismo histórico utilizou métodos produtivos de ponta do século XIX em contabilidade, financiamento e gestão ‘racional’ de corpos humanos para produzir resultado econômico na maior escala possível e ao menor custo possível. Hoje, algoritmos de aprendizado de máquina e sistemas de inteligência artificial permitem a extração e processamento de dados em escalas sem precedentes, com custos marginais próximos de zero.

Esta eficiência é amplificada pela natureza automatizada dos sistemas de coleta de dados digitais. Enquanto o colonialismo histórico requeria vastos aparatos administrativos e militares para manter o controle sobre territórios e populações, o colonialismo de dados opera através de sistemas algorítmicos que podem processar bilhões de transações de dados simultaneamente, sem intervenção humana direta. Um único algoritmo pode analisar os padrões de comportamento de centenas de milhões de usuários, identificar tendências, fazer previsões e tomar decisões em tempo real.

A dimensão temporal do colonialismo de dados também merece atenção especial. Enquanto o colonialismo histórico operava através de ciclos de conquista, exploração e eventual descolonização, o colonialismo de dados estabelece relações de dependência contínua e auto-reprodutiva. Os usuários deixam de ser meros fornecedores pontuais de dados para se tornarem alvos de monitoramento constante, gerando fluxos perpétuos de valor que alimentam os sistemas de acumulação digital.

Essa perpetuidade é garantida pela natureza viciante e necessária das tecnologias digitais contemporâneas. Diferentemente dos recursos naturais, que podem ser esgotados, os dados humanos são constantemente regenerados através das atividades cotidianas. Cada respiração monitorada por um smartwatch, cada passo registrado por um smartphone, cada interação social mediada por plataformas digitais contribui para esse fluxo contínuo de extração.

Mais intrigante ainda é como essa extração se intensifica através do que podemos chamar de “efeito rede” do colonialismo de dados. Quanto mais pessoas utilizam uma plataforma digital, mais valiosa ela se torna para todos os usuários, criando incentivos para a adesão universal. Este efeito de rede, combinado com estratégias de preços predatórios e aquisições agressivas de concorrentes, permite que as grandes corporações tecnológicas estabeleçam posições monopolísticas que são extremamente difíceis de contestar.

O caso do WhatsApp ilustra perfeitamente essa dinâmica. Quando o Facebook adquiriu o WhatsApp por 19 bilhões de dólares em 2014, muitos observadores questionaram o valor astronômico pago por uma empresa que gerava relativamente pouca receita (Constine, 2014). No entanto, a aquisição fazia sentido estratégico perfeito do ponto de vista do colonialismo de dados: o WhatsApp havia se tornado a plataforma de comunicação dominante em muitos países do Sul Global, criando uma infraestrutura de coleta de dados que seria extremamente difícil de replicar. Ao adquirir o WhatsApp, o Facebook não estava apenas comprando uma empresa, mas sim colonizando digitalmente bilhões de relacionamentos sociais e padrões de comunicação (Helmond *et al.*, 2019).

Publicamente, a justificativa de Mark Zuckerberg para o preço exorbitante focava-se no crescimento e no envolvimento dos utilizadores, métricas centrais para a economia da atenção. Ele afirmou: “O WhatsApp está a caminho de conectar 1 bilhão de pessoas. Os serviços que atingem esse marco são todos incrivelmente valiosos”.

Numa entrevista posterior, descreveu o WhatsApp como “a aplicação mais envolvente que existe no telemóvel que já vimos”, com 70% dos seus utilizadores ativos diariamente. Esta narrativa desloca deliberadamente a avaliação do valor da receita para os “utilizadores ativos”, o recurso primário na economia das plataformas (Constine, 2014; *Reuters*, 2014).

Esta lógica contrasta fortemente com a situação financeira do WhatsApp. Na primeira metade de 2014, a empresa gerou apenas 15,95 milhões de dólares em receita, ao mesmo tempo que registrou um prejuízo líquido de 233 milhões de dólares (TIME, 2014). O seu modelo de negócio consistia numa simples subscrição anual de 0,99 dólares após um primeiro ano gratuito, um modelo que estava apenas a começar a ser implementado em alguns países (Constine, 2014). Este modelo era explicitamente anti-publicidade, alinhado com o lema dos fundadores: “Sem anúncios, sem jogos, sem truques”. A desconexão profunda entre o desempenho financeiro do WhatsApp e o seu preço de aquisição é a prova mais clara de que o Facebook não estava adquirindo uma empresa no sentido tradicional. Um ator econômico racional não paga mais de mil vezes a receita anual de uma empresa. A transação foi motivada pela aquisição de um ativo estratégico cujo valor era medido em utilizadores, conexões sociais e fluxos de dados, não em lucros. Esta anomalia aponta para uma fonte alternativa de valor: a matéria-prima da própria vida social.

O verdadeiro valor do WhatsApp para o Facebook não residia na sua tecnologia ou no seu modesto fluxo de receitas, mas sim na sua base de utilizadores, especificamente na sua escala, no seu nível de envolvimento e, mais crucialmente, na sua localização geográfica. A aquisição foi uma “apropriação de terras de dados” (*data landgrab*) calculada para garantir uma posição dominante nos mercados do Sul Global, onde o Facebook, enquanto plataforma, não era o principal modo de interação digital (Koum, 2014; Backlinko, 2024).

No momento do anúncio da aquisição, em fevereiro de 2014, o WhatsApp já contava com mais de 450 milhões de utilizadores ativos mensais. Este número cresceu a um ritmo vertiginoso, atingindo 500 milhões em abril e 600 milhões em agosto do mesmo ano. A plataforma adicionava mais de 1 milhão de novos utilizadores registados por dia. Mais importante do que a escala era o nível de dependência que a plataforma gerava. Com 70% dos seus utilizadores ativos diariamente, o WhatsApp exibia um nível de envolvimento que superava em muito as aplicações de redes sociais típicas, sinalizando que, para muitos, não era apenas uma aplicação, mas uma utilidade essencial (Koum, 2014; Backlinko, 2024).

O valor estratégico desta base de utilizadores era amplificado pela sua distribuição. O WhatsApp tinha-se tornado o principal, e em muitos casos o padrão, meio de comunicação em vastas regiões do Sul Global. Um relatório de 2014 da GlobalWebIndex afirmava explicitamente que o uso do WhatsApp era maior na América Latina, no Oriente Médio e na África (Ugur, 2014).

A penetração em mercados-chave era imensa. Na América Latina, países como o Brasil e o México registraram uma adoção massiva. O Brasil viria a tornar-se o segundo maior mercado da aplicação, a seguir à Índia, que por si só representa a maior base de utilizadores do mundo. Em África, países como o Quênia, a África do Sul e a Nigéria apresentam taxas de penetração superiores a 95% entre os utilizadores de internet (Askazi, 2025).

Esta dominância foi facilitada por fatores socioeconómicos, nomeadamente a prática de *"zero-rating"*, através da qual as operadoras de telecomunicações ofereciam acesso de dados gratuito ao WhatsApp. Uma investigação na América Latina revelou que, de 15 países que ofereciam planos de *"zero-rating"*, 14 davam acesso gratuito ao WhatsApp ou ao Facebook (CENTRE FOR INTERNATIONAL GOVERNANCE INNOVATION, 2016). Esta prática tornou o WhatsApp sinónimo de *"internet"* para muitos utilizadores de baixos

rendimentos, criando uma profunda dependência infraestrutural e uma audiência cativa. Este mecanismo funciona como uma forma moderna de construção de infraestrutura colonial. Assim como as potências coloniais construíram ferrovias para extrair matérias-primas de forma mais eficiente, as plataformas tecnológicas usam o "zero-rating" para extrair dados de forma mais eficiente. Ao tornar o acesso ao WhatsApp gratuito, criam um "jardim murado" que desincentiva os utilizadores a explorar a web aberta, limitando a sua autonomia e prendendo-os num ecossistema onde as suas interações sociais se tornam matéria-prima para o Facebook.

A aquisição foi, portanto, um ataque preventivo contra a irrelevância da plataforma num mundo cada vez mais móvel. Nascido na web de desktop, o Facebook corria o risco de ser ultrapassado em mercados emergentes onde a principal "rampa de acesso" à internet era uma aplicação de mensagens móveis, não uma rede social. Ao comprar o WhatsApp, o Facebook não comprou apenas utilizadores; comprou a própria infraestrutura através da qual centenas de milhões de pessoas vivenciavam a vida digital, garantindo a sua própria relevância contínua. Além disso, a aquisição não visava apenas o número de utilizadores, mas a *natureza* dos dados. O WhatsApp detinha o grafo social íntimo, de um para um e de pequenos grupos, isto é, um conjunto de dados indiscutivelmente mais valioso para compreender a influência e o comportamento social do que os dados mais públicos e performativos da plataforma principal do Facebook.

Para interpretar os fatos apresentados, é essencial recorrer a enquadramentos teóricos críticos que fornecem a linguagem analítica necessária. As teorias do "capitalismo de vigilância" de Shoshana Zuboff (2019) e do "colonialismo de dados" de Nick Couldry e Ulises Mejias oferecem lentes através das quais a aquisição do WhatsApp se revela um caso de estudo paradigmático. Enquanto o capitalismo de vigilância descreve o motor económico da extração de dados, o colonialismo de dados elucida seu modo de expansão geopolítica e a estrutura de poder que o sustenta.

Este enquadramento postula uma nova lógica económica, pioneira da Google e aperfeiçoada pelo Facebook, que trata a experiência humana como uma fonte gratuita de matéria-prima. Esta matéria-prima é designada por “excedente comportamental” (*behavioral surplus*), ou seja, os dados de exaustão deixados pelas nossas atividades *online* (cliques, gostos, pesquisas, localização) que não são necessários para a melhoria do serviço. Este excedente é processado por “inteligência de máquina” para criar “produtos de previsão” que antecipam o nosso comportamento, sendo estes depois vendidos a clientes empresariais em “mercados de futuros comportamentais”, como a publicidade direcionada.

A aquisição do WhatsApp proporcionou ao Facebook (agora Meta) uma nova e colossal fonte de excedente comportamental. Embora o conteúdo das mensagens fosse encriptado de ponta a ponta, os *metadados* não o eram. Estes metadados incluem:

- Com quem o utilizador fala e quando.
- A sua lista de contatos.
- O seu número de telefone, nome de perfil e fotografia.
- Informações do dispositivo (modelo, sistema operativo, nível da bateria, força do sinal).
- Informações de ligação (endereço IP, operadora móvel, fuso horário).
- Informações de localização (derivadas de endereços IP e códigos de área telefónicos).

Estes metadados são estratégicos para construir e refinar produtos de previsão, permitindo ao Facebook “ver redes e inferir conteúdo” mesmo sem ler as mensagens. O debate público foi estrategicamente desviado para a privacidade do “conteúdo da mensagem”, obscurecendo o imenso valor dos metadados não encriptados que

estavam a ser extraídos. A verdadeira mais-valia para um capitalista de vigilância reside nos padrões, não necessariamente no conteúdo.

Esta teoria enquadra o momento atual como uma nova fase histórica do colonialismo, onde o recurso apropriado não é a terra ou o trabalho no sentido tradicional, mas a própria vida humana, transformada em dados. Trata-se de uma “apropriação de terras de dados” que cerca os “bens comuns de dados” (*data commons*), isto é, a nossa existência social coletiva, relações e comunicações, e os transforma num ativo proprietário para lucro corporativo.

As duas teorias não são mutuamente exclusivas, mas descrevem facetas diferentes do mesmo processo. O capitalismo de vigilância é o *motor económico* da nova ordem, enquanto o colonialismo de dados é o seu *modo de expansão e estrutura geopolítica*. A aquisição do WhatsApp foi o ato colonial; as subsequentes alterações à política de privacidade para permitir a partilha de dados foram a implementação da lógica capitalista de vigilância.

As ações tomadas pelo Facebook após a aquisição servem como prova concreta da tese do colonialismo de dados. Estas ações revelam a verdadeira intenção por trás da compra, transitando de promessas de autonomia para práticas de integração e extração de dados.

1.1 A ECONOMIA DOS DADOS E AS GRANDES CORPORAÇÕES (*BIG TECH*)

A economia contemporânea dos dados é dominada por um pequeno número de corporações transnacionais, frequentemente denominadas “*Big Tech*”, que incluem empresas como Google (Alphabet), Meta (Facebook), Amazon, Apple e Microsoft.

Essas corporações construíram impérios econômicos baseados na coleta, processamento e monetização de dados pessoais em escala global, estabelecendo o que pode ser caracterizado como um novo modo de acumulação capitalista.

Para compreender a magnitude dessa transformação, considere que em 2023, as cinco maiores empresas de tecnologia dos Estados Unidos possuíam uma capitalização de mercado combinada de aproximadamente 9 trilhões de dólares - um valor que supera o PIB de qualquer país individual, exceto os próprios Estados Unidos e a China. (Reuters, 2025; Sibilis Research, 2025). Esta concentração de valor econômico em empresas cujo principal ativo são dados e algoritmos representa uma mudança fundamental na natureza do capitalismo contemporâneo.

O modelo de negócios dessas empresas fundamenta-se na transformação de atividades cotidianas em dados comercializáveis. Cada busca no Google, cada interação no Facebook, cada compra na Amazon, cada localização registrada pelo smartphone contribui para vastos bancos de dados que são posteriormente utilizados para direcionamento publicitário, desenvolvimento de produtos e serviços, e tomada de decisões algorítmicas que afetam a vida de bilhões de pessoas.

Mas a verdadeira genialidade, e também a perversidade, desse modelo reside na sua capacidade de transformar custos em receitas. Tradicionalmente, as empresas precisavam pagar pelos insumos necessários para produzir seus produtos. No modelo de negócios baseado em dados, os próprios consumidores fornecem gratuitamente a matéria-prima (seus dados pessoais) e ainda pagam pelo privilégio de fazê-lo, seja por meio da compra de dispositivos, da assinatura de serviços de internet ou do tempo gasto nas plataformas, que poderia ser utilizado em outras atividades produtivas.

A concentração de poder nessas corporações é impressionante. Segundo dados recentes, as cinco maiores empresas de tecnologia dos Estados Unidos possuem uma capitalização de mercado combinada que supera o PIB de muitos países desenvolvidos. Essa concentração econômica traduz-se em poder político e social, permitindo que essas empresas influenciem regulamentações, moldem comportamentos sociais e determinem os termos sob os quais bilhões de pessoas interagem com a tecnologia digital.

Esta influência política manifesta-se de múltiplas formas. As grandes corporações tecnológicas gastam centenas de milhões de dólares anualmente em lobbying para influenciar legislações que afetam seus interesses. Elas empregam ex-funcionários governamentais em posições de liderança e frequentemente “emprestam” executivos para ocupar posições em agências regulatórias (Arun, 2025). O fenômeno da “porta giratória” entre o setor privado e o governo é particularmente pronunciado na indústria tecnológica, criando conflitos de interesse estruturais que dificultam a regulamentação efetiva.

O conceito de “capitalismo de vigilância”, desenvolvido por Shoshana Zuboff, oferece uma perspectiva complementar ao colonialismo de dados para compreender essa dinâmica. Zuboff (2019) argumenta que as empresas de tecnologia criaram um novo tipo de capitalismo baseado na extração de “excedente comportamental” - dados sobre comportamentos humanos que são coletados além do necessário para melhorar produtos e serviços, sendo posteriormente utilizados para prever e influenciar comportamentos futuros.

O excedente comportamental representa uma inovação fundamental na lógica capitalista. Tradicionalmente, o capitalismo operava através da extração de mais-valia do trabalho humano - a diferença entre o valor produzido pelos trabalhadores e os salários que eles recebiam. No capitalismo de vigilância, uma nova forma de extração emerge: a captura de dados comportamentais que são utilizados para criar produtos preditivos vendidos em mercados de comportamento futuro.

No entanto, como observa Arun, a análise de Zuboff não engaja adequadamente com “quão longe suas características prejudiciais vêm do capitalismo, seu impacto desigual em diferentes grupos de pessoas, ou sua dependência da lei para seu poder” (ARUN, 2025). É precisamente nesse ponto que a perspectiva do colonialismo de dados oferece novas perspectivas ao destacar as dimensões geopolíticas e as desigualdades estruturais que caracterizam a economia digital contemporânea.

A dimensão colonial torna-se evidente quando examinamos os fluxos globais de dados. Países do Sul Global frequentemente servem como fontes de dados brutos, que são processados e monetizados por empresas sediadas no Norte Global. Essa dinâmica reproduz padrões históricos de extração de recursos, onde matérias-primas são extraídas da periferia para serem processadas e comercializadas no centro do sistema mundial.

Um exemplo revelador dessa dinâmica é o mercado global de trabalho digital. Plataformas como Amazon Mechanical Turk, Clickworker e Figure Eight conectam empresas que precisam de tarefas simples realizadas (como rotulagem de imagens, transcrição de áudio, ou moderação de conteúdo) com trabalhadores dispostos a realizá-las por pagamentos muito baixos. A maioria desses trabalhadores está localizada em países do Sul Global, onde os salários locais tornam atrativo aceitar pagamentos que seriam considerados exploratórios em países desenvolvidos (data brokers, 2025).

Esses trabalhadores digitais, além de fornecerem mão de obra barata, também geram dados importantes sobre seus próprios padrões de trabalho, eficiência, e qualidade de output. Esses dados são posteriormente utilizados para desenvolver algoritmos de inteligência artificial que podem eventualmente automatizar as próprias tarefas que esses trabalhadores realizam, criando uma forma particularmente perversa de exploração onde os trabalhadores contribuem involuntariamente para sua própria obsolescência.

Além disso, a infraestrutura digital global está concentrada em poucos países e regiões. A maioria dos servidores, cabos submarinos, centros de dados e tecnologias fundamentais da internet são controlados por empresas e governos do Norte Global. Isso cria dependências estruturais que limitam a autonomia tecnológica de países e regiões periféricas, reproduzindo relações de dominação características do colonialismo histórico.

A geografia dos cabos submarinos que transportam a maior parte do tráfego de internet global ilustra perfeitamente essa dinâmica colonial. Esses cabos conectam principalmente centros financeiros e tecnológicos do Norte Global, com conexões limitadas e frequentemente inadequadas para países do Sul Global. Isso significa que dados gerados em Lagos podem precisar viajar através de Londres para chegar a Joanesburgo, criando latências, custos e vulnerabilidades que colocam usuários e empresas do Sul Global em desvantagem estrutural.

A financeirização dos dados representa outro aspecto importante da economia digital contemporânea. Dados pessoais tornaram-se ativos financeiros que podem ser comprados, vendidos, emprestados e utilizados como garantia em transações complexas. Mercados de dados pessoais operam em escalas globais, com brokers especializados intermediando transações entre coletores e compradores de informações pessoais.

Esta financeirização cria novos tipos de desigualdade e vulnerabilidade. Populações com maior poder aquisitivo podem pagar por serviços que protegem sua privacidade, enquanto populações mais pobres são forçadas a “pagar” com seus dados pessoais para acessar serviços digitais básicos. Isso cria uma economia de dados estratificada, onde a privacidade torna-se um luxo disponível apenas para aqueles que podem pagar por ela.

O fenômeno dos “data brokers” - empresas especializadas na coleta, agregação e venda de dados pessoais - ilustra como essa financeirização opera na prática. Empresas como Acxiom, Experian e LexisNexis coletam dados de centenas de fontes diferentes (registros públicos, transações comerciais, atividade online, dados de localização, etc.) para criar perfis detalhados de indivíduos que são posteriormente vendidos para empresas interessadas em marketing direcionado, avaliação de risco, ou tomada de decisões automatizada.

Esses perfis podem incluir informações extremamente sensíveis sobre saúde, situação financeira, orientação política, preferências sexuais, e outros aspectos íntimos da vida pessoal. Mais preocupante ainda, esses dados são frequentemente imprecisos, desatualizados ou baseados em inferências algorítmicas questionáveis, mas são utilizados para tomar decisões importantes sobre acesso a crédito, emprego, seguro de saúde, e outros aspectos fundamentais da vida.

1.2 ASSIMETRIAS DE PODER E VIGILÂNCIA NA ERA DO BIG DATA

A era do Big Data caracteriza-se por assimetrias de poder que permeiam todos os aspectos da sociedade digital. Essas assimetrias manifestam-se em múltiplas dimensões: entre coletores e fornecedores de dados, entre países centrais e periféricos no sistema mundial, entre populações com diferentes níveis de literacia digital, e entre comunidades com distintos graus de proteção legal e institucional.

Para compreender a magnitude dessas assimetrias, considere que uma única empresa como o Google processa dezenas de milhões de consultas, algo em média de 15 bilhões de buscas por dia em 2024-2025, o que evidencia o poder informacional concentrado em poucos intermediários digitais (Cardillo, 2025; Lahey, 2025).

Cada uma dessas buscas revela informações sobre interesses, preocupações, desejos e necessidades dos usuários. Agregadas e analisadas, essas informações fornecem um retrato em tempo real da consciência coletiva humana que é de valor inestimável para governos, empresas e outros atores interessados em prever e influenciar comportamentos.

A vigilância ubíqua emerge como uma das características mais preocupantes dessa nova configuração social. Diferentemente dos sistemas de vigilância tradicionais, que eram limitados por restrições tecnológicas e de recursos, a vigilância digital contemporânea opera de forma contínua, automatizada e em escala global. Cada dispositivo conectado à internet torna-se um potencial ponto de coleta de dados, criando uma rede de vigilância que penetra nos espaços mais íntimos da vida privada (Foucault, 1975).

O conceito de “panóptico digital” oferece uma metáfora útil para compreender essa transformação. O panóptico, originalmente concebido pelo filósofo Jeremy Bentham como um modelo de prisão onde um único guarda poderia observar todos os prisioneiros sem ser visto, foi utilizado por Michel Foucault como metáfora para os mecanismos de poder disciplinar na sociedade moderna. No panóptico digital, não há um único observador, mas sim uma rede distribuída de sensores, algoritmos e sistemas de análise que monitoram continuamente o comportamento humano.

Essa vigilância não é meramente passiva. Os sistemas de Big Data vão além da simples coleta de informações sobre comportamentos passados, utilizando algoritmos preditivos para antecipar e influenciar comportamentos futuros. Isso cria o que pode ser caracterizado como uma forma de “poder preditivo”, onde decisões sobre acesso a crédito, emprego, educação, saúde e outros aspectos fundamentais da vida são tomadas com base em análises algorítmicas de dados pessoais.

O poder preditivo representa uma evolução qualitativa nos mecanismos de controle social. Tradicionalmente, o poder operava principalmente através da punição após a transgressão, em que você era punido depois de cometer um crime, demitido depois de um desempenho inadequado, ou negado crédito depois de demonstrar incapacidade de pagamento. No regime do poder preditivo, as decisões são tomadas com base em probabilidades calculadas algoritmicamente sobre comportamentos futuros, criando formas de discriminação e exclusão que operam antes mesmo que qualquer comportamento específico tenha ocorrido.

As assimetrias de poder tornam-se particularmente evidentes quando consideramos quem controla os algoritmos e sistemas de inteligência artificial que processam esses dados. A maioria dos algoritmos mais avançados é desenvolvida por um pequeno número de empresas e instituições de pesquisa, principalmente localizadas nos Estados Unidos e na China. Isso significa que as lógicas, valores e vieses incorporados nesses sistemas refletem as perspectivas e interesses de grupos específicos, frequentemente em detrimento de populações marginalizadas.

A questão dos vieses algorítmicos ilustra como as assimetrias de poder se reproduzem e amplificam através da tecnologia. Sistemas de reconhecimento facial demonstram taxas de erro significativamente maiores para pessoas negras, especialmente mulheres negras. Algoritmos de contratação discriminam contra mulheres e minorias étnicas (Grother *et al.*, 2019). Sistemas de pontuação de crédito penalizam populações de baixa renda. Essas não são falhas técnicas isoladas, mas reflexos sistemáticos das desigualdades estruturais da sociedade, amplificadas pela escala e velocidade dos sistemas digitais.

Um caso particularmente revelador é o do algoritmo COMPAS (Correctional Offender Management Profiling for Alternative Sanctions), utilizado pelo sistema de justiça criminal americano

para avaliar o risco de reincidência de réus. Uma investigação da ProPublica revelou que o algoritmo demonstrava viés racial sistemático, classificando incorretamente réus negros como de alto risco de reincidência quase duas vezes mais frequentemente que réus brancos, enquanto classificava incorretamente réus brancos como de baixo risco mais frequentemente que réus negros (Angwin *et al.*, 2016).

O que torna esse caso particularmente preocupante é que o algoritmo COMPAS é utilizado por juízes para tomar decisões sobre liberdade condicional, sentenças e outras questões que afetam fundamentalmente a vida das pessoas. O viés algorítmico é muito mais que uma questão técnica abstrata, pois tem consequências concretas e devastadoras para indivíduos e comunidades já marginalizadas pelo sistema de justiça criminal.

A dimensão geopolítica das assimetrias de poder no Big Data também merece atenção especial. O controle sobre dados tornou-se uma questão de segurança nacional e soberania para muitos países. A dependência de plataformas digitais estrangeiras para serviços essenciais cria vulnerabilidades estratégicas que podem ser exploradas para fins de espionagem, influência política ou coerção econômica.

O caso da empresa chinesa TikTok ilustra essas tensões geopolíticas. Preocupações sobre o acesso do governo chinês aos dados de usuários americanos levaram a tentativas de banir ou forçar a venda da plataforma nos Estados Unidos. Similarmente, empresas americanas como Google e Facebook enfrentam restrições em vários países devido a preocupações sobre privacidade, segurança e soberania digital (New York Times, 2024). Em 2024, os EUA aprovaram uma lei exigindo a venda da TikTok por sua controladora chinesa ou a saída do mercado; em janeiro de 2025, a Suprema Corte manteve a validade da medida, elevando o patamar da disputa entre segurança nacional e liberdade de expressão no ambiente digital (Shepardson, 2024; SUPREME COURT OF THE UNITED STATES, 2025).

Essas dinâmicas revelam como o colonialismo de dados opera não apenas no nível individual, mas também no nível de Estados e nações. Países com capacidades tecnológicas limitadas encontram-se em posições de dependência estrutural em relação às potências digitais, reproduzindo padrões de dominação característicos do sistema colonial histórico.

A “balcanização” da internet global, ou seja, a fragmentação em redes nacionais ou regionais com diferentes regras, padrões e controles, representa uma resposta a essas assimetrias de poder, mas também cria novos desafios e contradições. Enquanto a soberania digital pode proteger países de formas de exploração colonial, ela também pode ser utilizada por regimes autoritários para controlar informação e suprimir dissidência.

A resistência a essas assimetrias de poder tem emergido em várias formas. Regulamentações como o Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia representam tentativas de reequilibrar as relações de poder entre indivíduos e corporações. Iniciativas de soberania digital em países como Brasil, Índia e França buscam reduzir a dependência de tecnologias estrangeiras. Movimentos sociais e organizações da sociedade civil mobilizam-se para denunciar abusos e promover alternativas mais justas e democráticas (European Commission, 2020). No plano normativo, a GDPR e a LGPD disciplinam tratamento, transferência internacional e direitos dos titulares, incluindo limites à perfilação; na UE, o AI Act introduz exigências graduais para sistemas de alto risco e modelos de propósito geral, o que incide diretamente sobre práticas de análise preditiva em larga escala (EUROPEAN COMMISSION, 2025).

No entanto, essas resistências enfrentam desafios significativos. O poder econômico e político das grandes corporações tecnológicas permite que elas influenciem processos regulatórios e políticos. A complexidade técnica dos sistemas digitais cria barreiras para a participação democrática na governança tecnológica.

A dependência crescente de tecnologias digitais para atividades essenciais torna difícil imaginar alternativas viáveis ao sistema atual.

Mais fundamentalmente, a resistência ao colonialismo de dados exige mais do que mudanças regulatórias ou tecnológicas, e requer uma transformação profunda nas estruturas econômicas e sociais que sustentam o sistema atual. Isso inclui questionar o modelo de capitalismo baseado em crescimento infinito, repensar as relações entre trabalho e valor na era digital, e imaginar formas alternativas de organização social que priorizem o bem-estar humano e a sustentabilidade ambiental sobre a acumulação de capital (EFF, 2023).

1.3 MARCOS TEÓRICOS: DOS ESTUDOS PÓS-COLONIAIS À TEORIA CRÍTICA DA INFORMAÇÃO

A análise do colonialismo de dados requer uma abordagem teórica multidisciplinar que combine contribuições dos estudos pós-coloniais, da teoria crítica da informação, da economia política, da comunicação e das epistemologias do Sul. Essa convergência teórica permite uma compreensão mais nuançada das dinâmicas de poder que caracterizam a era digital, evitando tanto o determinismo tecnológico quanto o reducionismo econômico.

Os estudos pós-coloniais oferecem ferramentas conceituais fundamentais para compreender como as relações de dominação se perpetuam e se transformam ao longo do tempo. O conceito de “colonialidade”, desenvolvido por Aníbal Quijano, é particularmente relevante para entender como estruturas coloniais persistem mesmo após o fim formal do colonialismo político (QUIJANO, 2000). A colonialidade opera em três dimensões interconectadas: a colonialidade

do poder (estruturas políticas e econômicas), a colonialidade do saber (sistemas de conhecimento e epistemologias), e a colonialidade do ser (subjetividades e identidades).

No contexto do colonialismo de dados, essas três dimensões da colonialidade manifestam-se de formas específicas e interconectadas. A colonialidade do poder digital manifesta-se através da concentração de controle sobre infraestruturas digitais, algoritmos e dados nas mãos de um pequeno número de corporações e Estados, principalmente localizados no Norte Global. Esta concentração reproduz e amplifica assimetrias de poder preexistentes, criando novas formas de dependência e subordinação.

A colonialidade do saber digital opera através da imposição de formas específicas de classificação, categorização e processamento de informação que refletem epistemologias dominantes, frequentemente marginalizando ou invisibilizando formas alternativas de conhecimento. Algoritmos de busca, sistemas de recomendação e ferramentas de tradução automática incorporam vieses culturais e linguísticos que privilegiam perspectivas hegemônicas.

A colonialidade do ser digital manifesta-se através da transformação de subjetividades e identidades através de processos de dataficação. Os indivíduos são reduzidos a perfis algorítmicos, suas identidades são fragmentadas em categorias comercializáveis, e suas possibilidades de auto-determinação são limitadas por sistemas preditivos que antecipam e moldam comportamentos futuros.

O conceito de “epistemicídio”, desenvolvido por Boaventura de Sousa Santos, oferece outra perspectiva para compreender as dimensões cognitivas do colonialismo de dados (SANTOS, 2007). O epistemicídio refere-se à destruição sistemática de formas de conhecimento não-hegemônicas, um processo que foi central ao projeto colonial histórico e que continua operando através de novos mecanismos na era digital.

No contexto digital, o epistemicídio manifesta-se através de várias formas. Algoritmos de busca e sistemas de recomendação podem invisibilizar conteúdos que não se conformam aos padrões dominantes de relevância e autoridade. Sistemas de tradução automática podem distorcer ou simplificar conceitos que não têm equivalentes diretos em línguas hegemônicas. Plataformas de mídia social podem suprimir conteúdos que são considerados “problemáticos” por algoritmos treinados com base em valores e normas específicas.

A teoria crítica da informação, desenvolvida por autores como Christian Fuchs e Tiziana Terranova, oferece ferramentas adicionais para compreender as dimensões econômicas e políticas da sociedade digital (FUCHS, 2014; TERRANOVA, 2000). Esta perspectiva enfatiza como as tecnologias de informação e comunicação estão inseridas em relações capitalistas de produção e como elas reproduzem e transformam estruturas de classe, poder e exploração. Nesta obra, usamos “teoria crítica da informação” para designar um campo que interroga a própria noção de informação, seus regimes documentais e implicações éticas (DAY; FROHMANN; CAPURRO), articulando-se à economia política da comunicação (SMYTHE; SCHILLER; FUCHS) quando examinamos propriedade, trabalho e mercantilização nas plataformas.

O conceito de “trabalho digital gratuito”, desenvolvido por Terranova, é particularmente relevante para compreender como o valor é extraído dos usuários de plataformas digitais. Quando os usuários criam conteúdo, interagem com outros usuários, ou simplesmente navegam em plataformas digitais, eles estão realizando trabalho que gera valor para as empresas proprietárias dessas plataformas. Este trabalho não é remunerado, mas é essencial para o funcionamento e a lucratividade do sistema (Terranova, 2000).

A economia política da comunicação, representada por autores como Dallas Smythe, Herbert Schiller e Dan Schiller, oferece indicações sobre como as indústrias de comunicação e informação

operam dentro do sistema capitalista global. O conceito de “audiência como mercadoria”, desenvolvido por Smythe, antecipou muitas das dinâmicas que caracterizam a economia digital contemporânea, onde a atenção dos usuários é capturada e vendida para anunciantes (Smythe, 1981).

As epistemologias do Sul, desenvolvidas por autores como Boaventura de Sousa Santos, Aníbal Quijano, Walter Dignolo e Ramón Grosfoguel, oferecem perspectivas alternativas para compreender e resistir às formas de dominação características da modernidade colonial. Essas perspectivas enfatizam a importância de reconhecer e valorizar formas de conhecimento que foram marginalizadas ou suprimidas pelo projeto colonial (Dignolo, 2011; Grosfoguel, 2016; Santos, 2009).

No contexto do colonialismo de dados, as epistemologias do Sul oferecem recursos para imaginar e construir alternativas ao sistema atual. Isso inclui o desenvolvimento de tecnologias que incorporem valores e princípios diferentes daqueles que dominam a indústria tecnológica atual, a criação de sistemas de governança de dados que respeitem a autonomia e autodeterminação das comunidades, e a promoção de formas de literacia digital que capacitem os indivíduos a compreender e resistir às formas de exploração digital.

A convergência dessas perspectivas teóricas permite uma análise mais rica e nuançada do colonialismo de dados, que vai além das abordagens puramente técnicas ou econômicas para incluir dimensões culturais, epistemológicas e ontológicas. Esta abordagem multidisciplinar é essencial para compreender a complexidade dos desafios que enfrentamos na era digital e para desenvolver estratégias efetivas de resistência e transformação.

Além disso, essa convergência teórica permite identificar conexões entre lutas aparentemente distintas - desde movimentos por justiça racial e de gênero até lutas por soberania indígena

e sustentabilidade ambiental - revelando como todas essas lutas estão conectadas por estruturas mais amplas de dominação que se manifestam tanto no mundo físico quanto no digital.

A análise do colonialismo de dados também se beneficia das contribuições da teoria feminista, particularmente do trabalho de autoras como Silvia Rivera Cusicanqui, María Lugones e Ochy Curiel, que desenvolveram perspectivas interseccionais sobre como diferentes formas de opressão se articulam e se reforçam mutuamente. No contexto digital, isso significa compreender como o colonialismo de dados afeta diferentemente mulheres e homens, pessoas de diferentes raças e etnias, classes sociais, orientações sexuais e identidades de gênero (Cusicanqui, 2010; Lugones, 2008; Curiel, 2007).

Por exemplo, mulheres são frequentemente alvo de formas específicas de vigilância e controle digital, desde o monitoramento de atividade reprodutiva através de aplicativos de saúde até a exposição a violência de gênero *online*. Comunidades LGBTQ+ enfrentam riscos particulares relacionados à coleta e uso de dados sobre orientação sexual e identidade de gênero. Populações racializadas são desproporcionalmente afetadas por vieses algorítmicos em sistemas de justiça criminal, contratação e acesso a serviços (Noble, 2018; Eubanks, 2018).

Esta análise interseccional é fundamental para desenvolver estratégias de resistência que sejam verdadeiramente inclusivas e que não reproduzam as exclusões e marginalizações que caracterizam o sistema atual. Ela também revela como a luta contra o colonialismo de dados está intrinsecamente conectada a lutas mais amplas por justiça social, igualdade e autodeterminação.

1.4 PARADOXOS E CONTRADIÇÕES DO COLONIALISMO DE DADOS

O colonialismo de dados está repleto de paradoxos e contradições que revelam a complexidade das relações de poder na era digital. Esses paradoxos não são meras curiosidades intelectuais, mas tensões fundamentais que moldam as experiências cotidianas de bilhões de pessoas e que oferecem pontos de entrada para resistência e transformação.

O primeiro e talvez mais fundamental paradoxo é o da “participação voluntária na própria exploração”. Diferentemente do colonialismo histórico, que era imposto através da força militar e da coerção política explícita, o colonialismo de dados opera através do que aparenta ser consentimento livre e informado. Os usuários voluntariamente se inscrevem em plataformas digitais, aceitam termos de serviço, e compartilham informações pessoais em troca de conveniência, entretenimento e conectividade social (Zuboff, 2019).

Esta aparência de voluntariedade, no entanto, mascara coerções mais sutis mas não menos efetivas. A crescente digitalização de serviços essenciais - desde educação e saúde até trabalho e participação cívica - torna a participação em plataformas digitais cada vez menos opcional. Durante a pandemia de COVID-19, por exemplo, milhões de estudantes foram forçados a utilizar plataformas educacionais digitais para continuar seus estudos, trabalhadores foram obrigados a utilizar ferramentas de videoconferência para manter seus empregos, e cidadãos precisaram utilizar aplicativos de rastreamento de contato para acessar espaços públicos (UNESCO, 2020; ONU, 2020).

O segundo paradoxo é o da “transparência opaca”. As grandes corporações tecnológicas frequentemente se apresentam como campeãs da transparência e da abertura, publicando relatórios

detalhados sobre suas práticas de coleta de dados, oferecendo controles de privacidade aos usuários, e defendendo princípios de “dados abertos” e “algoritmos explicáveis”. No entanto, essa transparência superficial frequentemente obscurece opacidades mais profundas.

Os termos de serviço e políticas de privacidade, embora tecnicamente disponíveis para todos os usuários, são escritos em linguagem jurídica complexa que é incompreensível para a maioria das pessoas. Um estudo estimou que seria necessário aproximadamente 76 dias de trabalho por ano para ler todas as políticas de privacidade dos serviços digitais que um usuário médio utiliza. Além disso, essas políticas frequentemente mudam sem aviso prévio, e os usuários são considerados como tendo “consentido” às mudanças simplesmente por continuar utilizando os serviços (McDonald; Cranor, 2008).

O terceiro paradoxo é o da “personalização despersonalizante”. As tecnologias digitais prometem experiências cada vez mais personalizadas e adaptadas às necessidades e preferências individuais (Eubanks, 2018; Noble, 2018). Algoritmos de recomendação prometem mostrar exatamente o conteúdo que cada usuário quer ver, assistentes virtuais prometem aprender os hábitos e preferências de cada pessoa, e sistemas de inteligência artificial prometem fornecer respostas e soluções customizadas para cada situação específica.

Paradoxalmente, essa personalização extrema frequentemente resulta em uma forma de despersonalização, onde os indivíduos são reduzidos a conjuntos de dados e padrões algorítmicos. A “personalização” é baseada não em uma compreensão genuína da complexidade e singularidade de cada pessoa, mas em correlações estatísticas derivadas de vastos bancos de dados comportamentais. O resultado é frequentemente uma forma de determinismo algorítmico onde as pessoas são tratadas como versões previsíveis de categorias pré-definidas.

O quarto paradoxo é o da “conectividade isolante”. As tecnologias digitais prometem conectar pessoas através de distâncias geográficas e diferenças culturais, criando uma “aldeia global” onde todos podem se comunicar e colaborar (APA, 2017). As mídias sociais prometem fortalecer relacionamentos e criar novas comunidades, as plataformas de colaboração prometem facilitar o trabalho em equipe, e os serviços de comunicação prometem manter famílias e amigos conectados independentemente da distância.

No entanto, pesquisas crescentes sugerem que o uso intensivo de tecnologias digitais está associado a aumentos em solidão, ansiedade, depressão e outros problemas de saúde mental (Pariser, 2011; Pew Research Center, 2020). A “conectividade” digital frequentemente substitui formas mais profundas de conexão humana, criando uma ilusão de intimidade que pode ser mais isolante que o isolamento físico. Além disso, algoritmos de recomendação frequentemente criam “bolhas de filtro” que limitam a exposição a perspectivas diversas, potencialmente aumentando polarização e fragmentação social.

O quinto paradoxo é o da “democratização elitista”. As tecnologias digitais são frequentemente apresentadas como ferramentas de democratização que podem nivelar campos de jogo desiguais, dar voz aos sem voz, e criar oportunidades para todos (ITU, 2022; ONU, 2021). A internet é celebrada como um espaço onde qualquer pessoa pode publicar conteúdo, criar um negócio, ou mobilizar apoio para causas importantes.

Embora existam exemplos genuínos de democratização através da tecnologia, a realidade é que as tecnologias digitais frequentemente reproduzem e amplificam desigualdades existentes. O acesso a tecnologias avançadas requer recursos financeiros, educacionais e culturais que não estão igualmente distribuídos. A capacidade de utilizar efetivamente essas tecnologias requer formas de literacia digital que são desenvolvidas através de privilégios educacionais e sociais.

E a capacidade de se beneficiar economicamente dessas tecnologias frequentemente requer acesso a redes sociais e formas de capital que são concentradas em grupos já privilegiados.

1.5 RESISTÊNCIAS E ALTERNATIVAS EMERGENTES

Apesar da aparente hegemonia do colonialismo de dados, resistências e alternativas emergem constantemente em diferentes escalas e contextos. Essas resistências vão desde ações individuais de proteção da privacidade até movimentos sociais organizados, desde iniciativas tecnológicas alternativas até políticas públicas regulatórias, desde práticas artísticas e culturais até experimentos econômicos cooperativos.

No nível individual, práticas de “higiene digital” estão se tornando cada vez mais comuns. Isso inclui o uso de navegadores que bloqueiam rastreamento, a utilização de redes privadas virtuais (VPNs) para mascarar localização, a escolha de serviços de email e mensagens criptografadas, e a adoção de práticas de compartilhamento seletivo de informações pessoais (EFF, 2023; PrivacyTools, 2022). Embora essas práticas sejam importantes para conscientização e proteção individual, elas têm limitações estruturais significativas, pois colocam o ônus da proteção nos indivíduos em vez de abordar as causas sistêmicas do problema.

No nível coletivo, movimentos sociais estão se organizando para contestar o poder das grandes corporações tecnológicas e promover alternativas mais justas. O movimento “Tech Workers Coalition” organiza trabalhadores da indústria tecnológica para resistir a projetos que consideram prejudiciais, como contratos militares ou sistemas de vigilância (Tech Workers Coalition, 2023).

O movimento “Algorithmic Justice League” trabalha para identificar e combater vieses em sistemas de inteligência artificial (AJL, 2022). Organizações como a “Electronic Frontier Foundation” e a “Privacy International” advogam por políticas que protegem direitos digitais (EFF, 2023; Privacy International, 2022).

No nível tecnológico, iniciativas de “tecnologia apropriada” ou “tecnologia decolonial” buscam desenvolver alternativas que incorporem valores diferentes daqueles que dominam a indústria tecnológica mainstream. Isso inclui projetos de *software* livre e código aberto que priorizam transparência e controle comunitário, redes descentralizadas que reduzem a dependência de infraestruturas corporativas, e sistemas de governança de dados que respeitam autonomia e autodeterminação das comunidades (Costanza-Chock, 2020).

Um exemplo fascinante é o projeto “Ushahidi”, desenvolvido no Quênia para mapear violência pós-eleitoral em 2008. A plataforma permite que cidadãos relatem incidentes através de SMS, email ou web, criando mapas em tempo real de eventos importantes (Ushahidi, 2008). O projeto foi posteriormente utilizado em contextos diversos, desde monitoramento de eleições até resposta a desastres naturais, sempre com foco em empoderamento comunitário e transparência.

No nível regulatório, governos ao redor do mundo estão experimentando com diferentes abordagens para regular a economia digital. O GDPR da União Europeia estabeleceu novos padrões globais para proteção de dados pessoais (European Union, 2018; Brasil, 2018; Mehta, 2020). A Lei Geral de Proteção de Dados (LGPD) do Brasil adaptou muitos princípios do GDPR para o contexto latino-americano. Países como a Índia estão desenvolvendo frameworks regulatórios que enfatizam soberania digital e localização de dados.

No nível econômico, experimentos com modelos alternativos de organização econômica estão emergindo. Cooperativas de plataforma buscam criar alternativas de propriedade coletiva a

plataformas corporativas. Projetos de “commons digitais” exploram formas de gestão coletiva de recursos digitais. Iniciativas de “economia solidária digital” experimentam com moedas alternativas, sistemas de troca não-monetários, e formas cooperativas de produção e distribuição (Scholz, 2016; Platform Coop, 2023).

1.6 DIMENSÕES PSICOLÓGICAS E SUBJETIVAS

O colonialismo de dados não opera apenas através de mecanismos econômicos e políticos externos, mas também através da colonização de subjetividades e da transformação de formas de consciência e identidade. Esta dimensão psicológica e subjetiva é frequentemente negligenciada em análises que se focam apenas em aspectos estruturais, mas é fundamental para compreender como o sistema se reproduz e como pode ser contestado.

A “gamificação” da vida cotidiana representa uma das formas mais pervasivas através das quais o colonialismo de dados penetra na subjetividade. Aplicativos de fitness transformam exercício físico em jogos com pontos, níveis e recompensas. Plataformas de mídia social utilizam “likes”, “shares” e “comments” como formas de validação social (Whitson, 2013; Zichermann; Linder, 2010). Aplicativos de produtividade transformam tarefas cotidianas em conquistas a serem desbloqueadas.

Essa gamificação não é neutra; ela incorpora lógicas específicas de competição, quantificação e otimização que moldam como as pessoas compreendem a si mesmas e suas relações com os outros (Whitson, 2013; Fuchs, 2017). A vida torna-se uma série de métricas a serem otimizadas, relacionamentos tornam-se redes a serem expandidas, e experiências tornam-se conteúdo a ser compartilhado e monetizado.

A “ansiedade algorítmica” emerge como uma nova forma de sofrimento psicológico característico da era digital (Binns *et al.*, 2018). As pessoas desenvolvem ansiedade sobre como algoritmos podem estar interpretando seus comportamentos, preocupação sobre como suas ações *online* podem afetar oportunidades futuras, e estresse relacionado à necessidade constante de gerenciar sua “presença digital”. Esta ansiedade é particularmente aguda para populações que já enfrentam discriminação e marginalização, que sabem que algoritmos podem amplificar vieses existentes.

A “síndrome FOMO” (Fear of Missing Out - medo de estar perdendo algo) ilustra como as tecnologias digitais podem criar formas de dependência psicológica (Przybylski *et al.*, 2013; APA, 2017). A exposição constante a feeds de mídia social que mostram as experiências (frequentemente curadas e idealizadas) de outros pode criar sentimentos de inadequação e uma compulsão para estar constantemente conectado e atualizado.

1.7 IMPACTOS AMBIENTAIS E SUSTENTABILIDADE

Uma dimensão frequentemente invisibilizada do colonialismo de dados são seus impactos ambientais massivos. A infraestrutura digital global - incluindo centros de dados, redes de telecomunicações, e bilhões de dispositivos conectados - consome quantidades enormes de energia e recursos naturais, contribuindo significativamente para mudanças climáticas e degradação ambiental.

Estima-se que a indústria digital seja responsável por aproximadamente 4% das emissões globais de gases de efeito estufa, um número que está crescendo rapidamente (The Shift Project, 2019; IEA, 2022). Centros de dados sozinhos consomem aproximadamente

1% da eletricidade global, e esse consumo está dobrando a cada quatro anos (IEA, 2022; Hintemann, 2020). O treinamento de um único modelo de inteligência artificial pode gerar emissões de carbono equivalentes a cinco carros durante suas vidas úteis inteiras (Strubell *et al.*, 2019).

Além do consumo de energia, a produção de dispositivos digitais requer a extração de minerais raros, frequentemente em condições que causam devastação ambiental e exploração de trabalhadores (UNEP, 2020; WEF, 2021). O “lixo eletrônico” tornou-se uma das formas de resíduo que mais cresce globalmente, com apenas uma pequena fração sendo adequadamente reciclada (ITU, 2022; Baldé *et al.*, 2020).

Esses impactos ambientais não são distribuídos igualmente. Os benefícios econômicos da economia digital concentram-se principalmente no Norte Global, enquanto muitos dos custos ambientais - desde a extração de minerais até o descarte de lixo eletrônico - são externalizados para o Sul Global. Isso representa uma forma de “colonialismo ambiental” que complementa e reforça o colonialismo de dados (Parikka, 2015; Nixon, 2011).

1.8 PERSPECTIVAS FUTURAS E TENDÊNCIAS EMERGENTES

À medida que olhamos para o futuro, várias tendências emergentes prometem intensificar tanto as oportunidades quanto os riscos associados ao colonialismo de dados. A proliferação da “Internet das Coisas” (IoT) está expandindo a coleta de dados para praticamente todos os aspectos da vida cotidiana, desde geladeiras “inteligentes” que monitoram hábitos alimentares até cidades “inteligentes” que rastreiam movimentos de todos os cidadãos (Zuboff, 2019; WEF, 2020).

O desenvolvimento de tecnologias de inteligência artificial cada vez mais sofisticadas está criando novas possibilidades para análise e manipulação de dados comportamentais. Sistemas de “deep learning” podem identificar padrões em dados que são invisíveis para análise humana, potencialmente revelando informações sensíveis sobre saúde mental, orientação sexual, afiliação política e outras características pessoais mesmo quando essas informações não foram explicitamente compartilhadas (Kosinski *et al.*, 2013; Raji; Buolamwini, 2020).

A emergência de tecnologias de realidade virtual e aumentada promete criar novas fronteiras para coleta de dados, incluindo dados sobre movimentos corporais, expressões faciais, padrões de atenção visual, e até mesmo respostas fisiológicas como batimentos cardíacos e condutância da pele (Meta, 2022; Slater *et al.*, 2020). Esses dados podem fornecer inovações sem precedentes sobre estados psicológicos e emocionais, criando novas possibilidades tanto para manipulação quanto para compreensão humana.

Ao mesmo tempo, tecnologias emergentes também oferecem possibilidades para resistência e alternativas. Tecnologias de *blockchain* podem permitir formas descentralizadas de organização que reduzem a dependência de intermediários corporativos (Tapscott; Tapscott, 2016). Avanços em criptografia podem fortalecer proteções de privacidade. Desenvolvimentos em inteligência artificial podem democratizar acesso a ferramentas analíticas poderosas (EFF, 2023; Cloudflare, 2022).

O futuro do colonialismo de dados não está predeterminado. Ele será moldado pelas escolhas que fazemos hoje sobre como desenvolver, regular e utilizar tecnologias digitais. Compreender as dinâmicas atuais é o primeiro passo para imaginar e construir futuros alternativos que priorizem justiça, sustentabilidade e bem-estar humano sobre acumulação de capital e controle social.

2

CIÊNCIA ABERTA:
TENSÕES, CONTRADIÇÕES
E PARADOXOS

A Ciência Aberta emerge como um dos movimentos mais significativos na transformação contemporânea dos sistemas de conhecimento, prometendo democratizar o acesso à informação científica, acelerar a inovação através da colaboração global, e tornar a pesquisa mais transparente, reproduzível e socialmente relevante (Opus Project, 2023). No entanto, como veremos ao longo deste capítulo, a Ciência Aberta está longe de ser um conceito unívoco ou uma prática neutra. Ela é atravessada por tensões, contradições e paradoxos que refletem as complexas relações de poder que caracterizam tanto o sistema científico quanto a sociedade digital mais ampla.

O movimento pela Ciência Aberta não surgiu no vácuo, mas como resposta a crises múltiplas que afetam o sistema científico contemporâneo: a crise de reprodutibilidade, onde uma proporção significativa de resultados científicos publicados não pode ser replicada (Baker, 2016); a crise de acesso, onde o conhecimento científico permanece trancado atrás de paywalls que impedem seu acesso por pesquisadores, estudantes e cidadãos, especialmente no Sul Global; a crise de relevância, onde a pesquisa científica frequentemente parece desconectada das necessidades e prioridades sociais mais urgentes; e a crise de confiança, onde escândalos de má conduta científica e a politização de questões científicas minam a autoridade epistêmica da ciência.

Ao mesmo tempo, a Ciência Aberta é impulsionada por oportunidades tecnológicas sem precedentes. A digitalização massiva da produção científica, o desenvolvimento de infraestruturas de comunicação global, e a emergência de ferramentas colaborativas sofisticadas criam possibilidades inéditas para compartilhamento, colaboração e participação na produção de conhecimento (Open Science Cycle, 2023). Pela primeira vez na história, é tecnicamente possível imaginar um sistema científico verdadeiramente global, onde qualquer pessoa, em qualquer lugar, pode acessar o conhecimento científico acumulado pela humanidade e contribuir para sua expansão.

No entanto, a realização dessas possibilidades não é automática nem neutra. A Ciência Aberta está sendo moldada por forças assimétricas, como a concentração de poder em plataformas corporativas (Elsevier, Springer Nature, Google Scholar), a reprodução de desigualdades geopolíticas e epistêmicas, e a subordinação de valores científicos a lógicas comerciais e métricas quantitativas (Albornoz *et al.*, 2020). Compreender essas tensões é fundamental para desenvolver formas de Ciência Aberta que sejam verdadeiramente emancipatórias e transformadoras.

2.1. A EVOLUÇÃO DO CONCEITO DE CIÊNCIA ABERTA

Para compreender adequadamente os desafios e possibilidades da Ciência Aberta contemporânea, é necessário situá-la em uma perspectiva histórica mais ampla que reconheça tanto as continuidades quanto as rupturas em relação a formas anteriores de organização científica. A ciência sempre teve dimensões “abertas” e “fechadas”, e as tensões entre compartilhamento e apropriação, colaboração e competição, universalismo e particularismo têm sido constantes ao longo da história da ciência moderna.

Os ideais normativos da ciência moderna, articulados classicamente por Robert Merton em 1942, incluíam o “comunalismo” – a ideia de que o conhecimento científico deveria ser propriedade comum da humanidade, disponível para escrutínio e utilização por todos (Merton, 1942; Antunes *et al.*, 2021). Este ideal estava enraizado tanto em considerações epistemológicas (a ciência progride através do escrutínio crítico coletivo) quanto em considerações éticas (o conhecimento financiado publicamente deveria beneficiar o público).

Embora muitas vezes vinculada à ideia de abertura, a ciência moderna carregou, desde o início, práticas de fechamento e apropriação. O sistema de patentes, desenvolvido no século XVIII, permitiu a apropriação privada de conhecimentos científicos e tecnológicos (WIPO, 2020). O desenvolvimento de revistas científicas comerciais, especialmente a partir do século XX, criou barreiras econômicas ao acesso ao conhecimento. A militarização da ciência durante as duas guerras mundiais e a Guerra Fria estabeleceu regimes de segredo e classificação que limitaram drasticamente o compartilhamento de conhecimento em áreas consideradas estratégicas (Morandin *et al.*, 2023).

A emergência da internet e das tecnologias digitais nas décadas de 1990 e 2000 criou novas possibilidades para o compartilhamento de conhecimento científico, mas também novos desafios. Por um lado, a digitalização reduziu drasticamente os custos de reprodução e distribuição de informação, tornando tecnicamente possível disponibilizar todo o conhecimento científico para qualquer pessoa com acesso à internet. Por outro lado, a digitalização também criou novas oportunidades para controle e monetização da informação, levando a formas ainda mais restritivas de apropriação do conhecimento.

O movimento contemporâneo pela Ciência Aberta pode ser entendido como uma resposta a essas tensões, buscando aproveitar as possibilidades tecnológicas da era digital para realizar os ideais normativos da ciência moderna. No entanto, como veremos, este movimento está longe de ser homogêneo, abrangendo uma diversidade de perspectivas, práticas e interesses que frequentemente entram em conflito.

A Declaração de Budapeste sobre Acesso Aberto, publicada em 2002, é frequentemente citada como um marco fundacional do movimento contemporâneo pela Ciência Aberta (BUDAPEST OPEN ACCESS INITIATIVE, 2002). A declaração articulou uma visão onde “literatura científica revisada por pares deveria estar disponível

online, gratuitamente e sem restrições, para todos os pesquisadores, educadores, estudantes e membros curiosos do público.” Esta visão foi posteriormente expandida para incluir não apenas publicações, mas também dados científicos, *software*, materiais educacionais e outras formas de conhecimento científico.

Desde então, o movimento pela Ciência Aberta tem se expandido e diversificado enormemente. Iniciativas como o PLOS (Public Library of Science), arXiv e PubMed Central criaram infraestruturas para publicação e acesso aberto (PLOS, 2025). Políticas de mandato de acesso aberto foram adotadas por agências de financiamento e instituições de pesquisa ao redor do mundo. Ferramentas e plataformas para compartilhamento de dados, código e materiais de pesquisa proliferaram. Movimentos para ciência cidadã e participativa expandiram as fronteiras de quem pode participar na produção de conhecimento científico.

No entanto, essa expansão também revelou tensões e contradições fundamentais. O “acesso aberto” pode significar coisas muito diferentes dependendo de quem o define e implementa. Para alguns, significa simplesmente tornar publicações disponíveis gratuitamente *online*, mantendo todos os outros aspectos do sistema científico inalterados. Para outros, significa uma transformação fundamental nas relações de poder que governam a produção e distribuição de conhecimento (Albornoz *et al.*, 2020).

2.2 PRINCÍPIOS FAIR E DEMOCRATIZAÇÃO DO CONHECIMENTO

Os princípios FAIR (Findable, Accessible, Interoperable, Reusable – Encontrável, Acessível, Interoperável, Reutilizável) emergiram como um framework influente para orientar práticas de Ciência Aberta,

especialmente no que se refere ao compartilhamento de dados científicos (Wilkinson *et al.*, 2016). Desenvolvidos através de um processo colaborativo envolvendo pesquisadores, bibliotecários, editores e outros stakeholders, os princípios FAIR buscam estabelecer padrões técnicos e metodológicos que facilitem o descobrimento, acesso e reutilização de dados científicos.

O princípio da “Encontrabilidade” (Findability) enfatiza a importância de tornar dados e metadados facilmente descobríveis através de mecanismos de busca e catálogos. Isso inclui a atribuição de identificadores únicos e persistentes, a utilização de vocabulários controlados e ontologias padronizadas, e a implementação de sistemas de indexação robustos. A encontrabilidade é fundamental para realizar o potencial da Ciência Aberta, pois dados que não podem ser descobertos não podem ser reutilizados.

O princípio da “Acessibilidade” (Accessibility) foca na remoção de barreiras técnicas, legais e econômicas ao acesso a dados e metadados. Isso inclui a utilização de protocolos de comunicação padronizados e abertos, a implementação de sistemas de autenticação e autorização apropriados, e a disponibilização de dados em formatos que não requerem software proprietário para acesso. A acessibilidade é fundamental para democratizar o acesso ao conhecimento científico.

O princípio da ‘Interoperabilidade’ (Interoperability) enfatiza a importância de utilizar padrões, vocabulários e formatos que permitam a integração e análise conjunta de dados de diferentes fontes. Isso é particularmente importante na era do Big Data, onde descobertas científicas frequentemente emergem da combinação de datasets diversos. A interoperabilidade técnica é complementada pela interoperabilidade semântica, que requer o uso de ontologias e vocabulários compartilhados.

O princípio da “Reutilização” (Reusability) foca em garantir que dados sejam adequadamente documentados e licenciados para facilitar sua reutilização por outros pesquisadores. Isso inclui a provisão de metadados ricos que descrevem como os dados foram coletados e processados, a especificação clara de licenças de uso, e a aderência a padrões comunitários relevantes. A reutilização é fundamental para maximizar o valor científico e social dos investimentos em pesquisa.

Embora os princípios FAIR tenham sido amplamente adotados e tenham contribuído significativamente para o avanço de práticas de Ciência Aberta, eles também têm sido objeto de críticas importantes. Uma crítica central é que os princípios FAIR são principalmente técnicos e procedimentais, não abordando adequadamente questões de poder, justiça e equidade que são fundamentais para uma verdadeira democratização do conhecimento (Bender *et al.*, 2021; Leonelli, 2020).

Por exemplo, tornar dados “encontráveis” e “acessíveis” segundo padrões técnicos não garante que eles sejam realmente acessíveis para pesquisadores e comunidades que carecem de recursos tecnológicos, capacidades técnicas ou literacia digital necessárias para utilizá-los efetivamente. A “interoperabilidade” pode privilegiar padrões e vocabulários desenvolvidos por e para comunidades científicas dominantes, marginalizando formas alternativas de organização e classificação do conhecimento. A “reutilização” pode facilitar a apropriação de conhecimentos tradicionais e comunitários por pesquisadores e instituições externas, sem retorno adequado para as comunidades que originalmente desenvolveram esses conhecimentos.

Essas limitações levaram ao desenvolvimento de frameworks complementares que buscam abordar dimensões éticas e políticas da Ciência Aberta. Os princípios CARE (Collective Benefit, Authority to Control, Responsibility, Ethics – Benefício Coletivo, Autoridade

para Controlar, Responsabilidade, Ética) foram desenvolvidos especificamente para orientar o compartilhamento de dados relacionados a povos indígenas, enfatizando a importância de soberania de dados, benefício comunitário e responsabilidade ética (Carroll *et al.*, 2020).

O princípio do “Benefício Coletivo” enfatiza que o compartilhamento de dados deve resultar em benefícios para as comunidades das quais os dados se originam, e não exclusivamente para pesquisadores e instituições externas. Isso requer mecanismos que garantam que pesquisas baseadas em dados comunitários contribuam para prioridades e necessidades definidas pelas próprias comunidades.

O princípio da “Autoridade para Controlar” reconhece o direito das comunidades de manter controle sobre dados que se relacionam a elas, incluindo o direito de determinar como esses dados são coletados, armazenados, acessados e utilizados. Isso desafia modelos tradicionais de pesquisa onde pesquisadores externos têm controle total sobre dados coletados em comunidades.

O princípio da “Responsabilidade” enfatiza que pesquisadores e instituições têm responsabilidades contínuas para com as comunidades das quais coletam dados, incluindo responsabilidades de prestação de contas, transparência e reciprocidade. Isso vai além de considerações éticas tradicionais para incluir compromissos de longo prazo com bem-estar comunitário.

O princípio da “Ética” enfatiza que o compartilhamento de dados deve ser guiado por considerações éticas que vão além de conformidade com regulamentações formais para incluir respeito por valores, culturas e sistemas de conhecimento comunitários. Isso requer sensibilidade cultural e disposição para adaptar práticas de pesquisa a contextos locais específicos.

A implementação conjunta de ambos os conjuntos de princípios enfrenta desafios práticos significativos. Para o FAIR, obstáculos comuns incluem sistemas de dados fragmentados, formatos

heterogêneos e a falta de metadados padronizados ou ontologias. Para o CARE, os desafios são mais complexos, envolvendo a necessidade de superar desigualdades históricas, construir capacidade de dados em comunidades indígenas, e garantir que as estruturas de governança de dados respeitem os direitos e interesses indígenas, que muitas vezes foram marginalizados ou explorados. A questão da soberania de dados indígenas é particularmente sensível, dada a história de coleta e uso de dados por governos e instituições colonizadoras sem o consentimento ou benefício das comunidades.

Exemplos de tensão em contextos reais são evidentes. Em dados biomédicos, a busca por compartilhamento amplo (FAIR) pode colidir com a necessidade de controle e proteção de dados de populações indígenas, dadas as histórias de exploração e falta de benefício para as comunidades de origem. A governança de grandes bases de dados, que frequentemente adotam modelos de compartilhamento amplos, contrasta com mecanismos mais rigorosos de proteção e relações de confiança que facilitaram a coleta original. Na pesquisa em saúde, a priorização da “justiça” na distribuição de benefícios e riscos pode levar a escolhas difíceis sobre alocação de recursos, onde a eficiência (FAIR) pode ser secundária à equidade (CARE). Em arqueologia, a publicação de dados abertos (FAIR) deve ser cuidadosamente balanceada com a soberania de dados culturais e a necessidade de compartilhar benefícios com comunidades descendentes, garantindo que os dados sensíveis sejam gerenciados de forma responsável, mesmo que isso signifique não serem totalmente abertos.

A tensão entre FAIR e CARE não é meramente uma questão de prioridades, mas uma indicação de que o conceito de “abertura” na ciência deve ser reavaliado. A “soberania de dados” emerge como um conceito fundamental que transcende a simples “acessibilidade” do FAIR. Ela exige que o controle e a governança dos dados residam nas comunidades de onde os dados se originam, especialmente em contextos de conhecimento tradicional e dados indígenas.

Isso implica que a verdadeira democratização não é apenas sobre tornar dados tecnicamente disponíveis, mas sobre empoderar as comunidades para que decidam

como e *se* seus dados são compartilhados e reutilizados, e para *quem* os benefícios se destinam. A coexistência e a integração bem-sucedida de FAIR e CARE exigem uma mudança cultural profunda no ecossistema científico, indo além das métricas de eficiência para abraçar valores de justiça e autodeterminação.

A tensão entre princípios FAIR e CARE ilustra contradições mais amplas no movimento pela Ciência Aberta. Enquanto os princípios FAIR enfatizam eficiência, padronização e maximização da reutilização de dados, os princípios CARE enfatizam justiça, autodeterminação e respeito por diversidade cultural e epistêmica. Essas diferentes ênfases refletem diferentes visões sobre os propósitos e valores que deveriam orientar a Ciência Aberta (Lovett *et al.*, 2022).

Tabela 1 - Comparativo FAIR vs. CARE: foco e implicações

Característica	Princípios FAIR	Princípios CARE
Orientação Principal	Dados e tecnologia	Pessoas e propósito
Objetivo Primário	Eficiência, padronização, maximização da reutilização de dados	Justiça, autodeterminação, benefício comunitário, respeito cultural
Desafios Abordados	Fragmentação de dados, baixa reutilização, falta de padronização	Desigualdades de poder, apropriação de conhecimento, falta de controle comunitário
Implicações Éticas/Sociais	Risco de marginalização, injustiça epistêmica, colonialismo de dados (sem CARE)	Empoderamento, soberania de dados, respeito cultural, equidade

Fonte: Elaborado pelo autor (2025).

A adoção isolada de FAIR pode reforçar assimetrias globais ao privilegiar atores com maior capacidade tecnológica, enquanto o alinhamento com CARE exige diálogo com comunidades, construção de confiança e reconhecimento de saberes locais e tradicionais.

Nesse sentido, práticas de governança de dados que conciliem os dois conjuntos de princípios podem favorecer uma Ciência Aberta mais inclusiva e responsável, na qual a eficiência tecnológica esteja acompanhada de justiça social e respeito à diversidade epistêmica.

2.3 MODELOS DE ACESSO ABERTO E SUSTENTABILIDADE

O movimento pelo acesso aberto desenvolveu múltiplos modelos para tornar publicações científicas livremente disponíveis, cada um com suas próprias implicações para sustentabilidade econômica, qualidade científica e equidade global. Compreender esses diferentes modelos e suas consequências é fundamental para avaliar criticamente as promessas e limitações da Ciência Aberta contemporânea (Suber, 2012; SPARC, 2020).

O modelo “Green Open Access” (via verde) permite que autores depositem versões de seus artigos em repositórios institucionais ou disciplinares, frequentemente após um período de embargo determinado pelos editores. Este modelo preserva o sistema tradicional de publicação científica enquanto cria vias alternativas de acesso. Embora seja relativamente fácil de implementar e tenha baixos custos para autores, o modelo verde frequentemente resulta em acesso a versões não-finais de artigos e pode ser limitado por políticas restritivas de editores (Piwowar *et al.*, 2018).

O modelo “Gold Open Access” (via dourada) torna artigos imediatamente disponíveis nos sites dos editores, frequentemente através do pagamento de “Article Processing Charges” (APCs) pelos autores ou suas instituições. Este modelo pode garantir acesso a versões finais e editadas de artigos, mas cria novas barreiras econômicas que podem excluir pesquisadores e instituições

com recursos limitados. O custo médio de APCs varia enormemente, de algumas centenas a vários milhares de dólares por artigo (Solomon; Björk, 2016).

O modelo “Diamond Open Access” (via diamante) disponibiliza artigos gratuitamente tanto para leitores quanto para autores, sendo sustentado através de financiamento institucional, governamental ou comunitário. Este modelo evita tanto barreiras de acesso quanto barreiras de publicação, mas requer formas alternativas de sustentabilidade econômica que podem ser difíceis de estabelecer e manter (Fuchs; Sandoval, 2013; UNESCO, 2021).

Tabela 2 - Modelos de acesso aberto: vantagens, desvantagens e sustentabilidade

Característica	Acesso Aberto Verde (Green)	Acesso Aberto Dourado (Gold)	Acesso Aberto Diamante (Diamond)
Mecanismo	Autoarquivamento do manuscrito pelo autor em um repositório.	Publicação imediata e livre no site da editora.	Publicação imediata e livre no site da editora/ plataforma.
Custo para Autores	Geralmente nenhum.	Pagamento de Taxa de Processamento de Artigo (APC).	Nenhum.
Custo para Leitores	Nenhum.	Nenhum.	Nenhum.
Barreiras	Períodos de embargo; acesso a versões não-finais (pré-print).	Barreira financeira para autores/instituições com menos recursos.	Desafios de sustentabilidade do modelo de financiamento.
Modelo de Financiamento	Mantido por instituições (repositórios).	Financiado por APCs pagas pelos autores ou agências de fomento.	Financiado por instituições, consórcios, voluntariado, filantropia.
Implicação para Equidade	Mais inclusivo para publicação, mas o acesso pode ser atrasado.	Excludente para autores do Sul Global e instituições sem verba para APCs.	Modelo mais equitativo, mas com desafios de escala e sustentabilidade.

Fonte: Elaborado pelo autor (2025).

Cada um desses modelos tem implicações diferentes para equidade global e justiça epistêmica. O modelo dourado, embora promova acesso amplo a publicações, pode criar novas formas de exclusão baseadas na capacidade de pagamento. Pesquisadores em países de baixa e média renda, que já enfrentam desvantagens significativas no sistema científico global, podem ser ainda mais marginalizados se não puderem pagar APCs. Isso pode resultar em uma Ciência Aberta que é “aberta” para leitura, mas “fechada” para contribuição por parte de pesquisadores com recursos limitados (Chan *et al.*, 2011).

A emergência de “predatory publishing” – editores que cobram APCs mas não fornecem serviços editoriais adequados – ilustra como modelos de acesso aberto podem ser explorados para fins comerciais que prejudicam a qualidade e integridade científica. Esses editores frequentemente targetam pesquisadores em países em desenvolvimento, explorando pressões por publicação e falta de familiaridade com padrões editoriais internacionais (Beall, 2012; Grudniewicz *et al.*, 2019).

Ao mesmo tempo, iniciativas como o “Plan S”, lançado pela cOAlition S na Europa, é um exemplo notável. O plano exige que pesquisas financiadas publicamente sejam disponibilizadas em Acesso Aberto imediatamente após a publicação. Seus requisitos incluem a publicação sob licenças Creative Commons Attribution (CC BY) e a cobertura dos APCs pelos financiadores ou instituições, não pelos pesquisadores individuais. O Plan S visa, em última instância, transformar o panorama da publicação científica, garantindo que o conhecimento financiado publicamente seja um bem comum.

No entanto, o Plan S gerou controvérsias e críticas significativas. Uma das principais é a alegação de que ele pode violar a liberdade acadêmica ao impor restrições sobre onde e como os pesquisadores podem publicar. A imposição de uma licença específica (CC BY) e a proibição de publicação em periódicos por assinatura

ou híbridos não transformadores são vistas por alguns como uma limitação indevida da autonomia dos pesquisadores na escolha de seus veículos de disseminação. Essa imposição de um modelo de licenciamento específico e a restrição de opções de publicação podem ser percebidas como uma preferência pelos direitos do sistema científico como um todo em detrimento dos direitos individuais dos pesquisadores.

Embora bem-intencionadas, as iniciativas do Plan S também podem ter consequências não intencionais. Há preocupações de que elas possam aumentar os custos para as instituições de pesquisa, especialmente aquelas que precisam arcar com os APCs ou com a transição para novos modelos. Além disso, se o resto do mundo não aderir a requisitos semelhantes, o Plan S poderia excluir pesquisadores que não podem cumprir com os requisitos, potencialmente prejudicando a internacionalização de estudantes de pós-graduação e pós-doutorandos, e desencorajando colaborações globais. Também há temores de que a pressão para publicar em periódicos OA compatíveis possa levar a uma sobrecarga de artigos de baixa qualidade ou limitar as oportunidades de publicação em periódicos de alto impacto, caso estes não se adaptem aos requisitos do plano.

O Plan S representa uma tentativa ambiciosa de uma mudança sistêmica *top-down* no cenário da publicação científica. Seu objetivo é nobre – tornar a pesquisa financiada publicamente imediatamente acessível. No entanto, ao impor mandatos rigorosos, como a licença CC BY e a proibição de periódicos por assinatura, ele entra em conflito direto com o princípio da “liberdade acadêmica” e a autonomia dos pesquisadores na escolha de seus veículos de publicação. Esse paradoxo ilustra a dificuldade de implementar políticas de “abertura” em larga escala sem gerar efeitos colaterais não intencionais, como a potencial marginalização de certos periódicos ou a criação de novas barreiras para pesquisadores que não se encaixam nos requisitos do plano. A tensão entre a necessidade de coordenação global para a Ciência Aberta e a preservação da autonomia individual e institucional é um desafio persistente.

A sustentabilidade econômica do acesso aberto permanece um desafio fundamental. Estimativas sugerem que a transição completa para acesso aberto poderia resultar em economias significativas para o sistema científico global, eliminando custos de assinatura e reduzindo barreiras ao acesso (Schimmer *et al.*, 2015). No entanto, a transição requer coordenação global e pode resultar em custos de curto prazo que são difíceis de absorver para muitas instituições.

Modelos alternativos de sustentabilidade estão sendo explorados, incluindo consórcios de bibliotecas que financiam coletivamente publicações de acesso aberto, fundos governamentais dedicados a APCs, e sistemas de “freemium” onde serviços básicos são gratuitos mas serviços premium são pagos. Cada modelo tem suas próprias vantagens e limitações, e é provável que o futuro do acesso aberto envolva uma combinação de diferentes abordagens. Contudo, para que esses modelos de abertura alcancem seu potencial democratizante, é imperativo confrontar as tensões inerentes entre abertura e controle que permeiam o ecossistema científico atual.

2.4 TENSÕES ENTRE ABERTURA E CONTROLE

Uma das contradições mais fundamentais da Ciência Aberta contemporânea é a tensão entre ideais de abertura, compartilhamento, democratização por um lado, e realidades de controle, apropriação, concentração de poder por outro. Esta tensão manifesta-se em múltiplas dimensões e escalas, desde questões técnicas sobre padrões e protocolos até questões geopolíticas sobre soberania de dados e dependência tecnológica.

No nível técnico, a implementação de sistemas de Ciência Aberta frequentemente requer a adoção de padrões, protocolos,

plataformas específicas, o que pode criar novas formas de dependência e controle. Por exemplo, a migração para plataformas de publicação digital pode aumentar a dependência de empresas de tecnologia que controlam essas plataformas. A adoção de padrões específicos de metadados pode privilegiar certas formas de organização do conhecimento em detrimento de outras. A utilização de identificadores únicos pode criar pontos centralizados de controle sobre a identidade científica.

A concentração crescente do mercado de publicação científica em um pequeno número de grandes corporações ilustra como a abertura pode coexistir ou até facilitar formas extremas de concentração de poder (Piwowar *et al.*, 2018; van Bellen *et al.*, 2024). Empresas como Elsevier, Springer Nature e Wiley controlam uma proporção significativa das publicações científicas globais, e têm expandido suas operações para incluir não apenas publicação, mas também análise de dados, avaliação de pesquisa, e outros serviços científicos. Essa integração vertical permite que essas empresas capturem valor em múltiplos pontos da cadeia de produção científica (Scholarly Kitchen, 2023; FT, 2024).

Paradoxalmente, algumas dessas mesmas empresas se posicionam como líderes no movimento pela Ciência Aberta, oferecendo serviços de acesso aberto e desenvolvendo plataformas para compartilhamento de dados. No entanto, críticos argumentam que isso representa uma forma de “open washing,” ou seja, a apropriação da retórica da abertura para legitimar práticas que na verdade consolidam controle corporativo sobre o sistema científico (Kember; Brand, 2023).

A Elsevier, por exemplo, promove ativamente iniciativas de acesso aberto através de sua plataforma ScienceDirect Open Access, oferece ferramentas para gestão de dados de pesquisa através do Mendeley Data, e participa de consórcios internacionais para desenvolvimento de padrões abertos. No entanto, críticos argumentam que

essas iniciativas servem principalmente para manter a empresa no centro do ecossistema científico enquanto a transição para modelos mais abertos ocorre (Broersma, 2024).

O “open washing” é particularmente problemático porque pode cooptar a energia e legitimidade do movimento pela Ciência Aberta sem alterar fundamentalmente as estruturas de poder subjacentes. Quando grandes editoras oferecem opções de acesso aberto, elas frequentemente fazem isso através de modelos que mantêm ou até aumentam suas receitas, como as taxas de processamento de artigos (APCs) que podem custar milhares de dólares por publicação.

No nível geopolítico, a Ciência Aberta pode reproduzir ou amplificar desigualdades entre países e regiões. Embora a abertura possa, teoricamente, beneficiar pesquisadores em todo lugar, na prática, a capacidade de se beneficiar da Ciência Aberta depende de recursos tecnológicos, capacidades institucionais, e formas de capital cultural que não estão igualmente distribuídos. Países com infraestruturas digitais avançadas, instituições bem financiadas, e populações com alta literacia digital estão melhor posicionados para aproveitar as oportunidades da Ciência Aberta. Países com limitações nessas áreas podem encontrar-se em posição de dependência, consumindo conhecimento produzido em outros lugares, mas contribuindo relativamente pouco para sua produção (van Bellen *et al.*, 2024).

A questão da língua ilustra bem essas dinâmicas. Embora a Ciência Aberta prometa democratizar o acesso ao conhecimento, a dominância do inglês como língua franca da ciência significa que muito do conhecimento disponibilizado permanece inacessível para populações que não dominam o idioma. Ao mesmo tempo, conhecimentos produzidos em outras línguas podem ser marginalizados ou invisibilizados em sistemas de Ciência Aberta que privilegiam conteúdo em inglês (van Bellen *et al.*, 2024).

A dominância do inglês na comunicação científica representa uma das manifestações mais pervasivas e problemáticas da tensão entre abertura e controle na Ciência Aberta. Embora frequentemente naturalizada como uma necessidade prática para comunicação global, a hegemonia linguística do inglês constitui uma forma de violência epistêmica que marginaliza conhecimentos, perspectivas e comunidades inteiras de pesquisa. Esta seção examina as dimensões quantitativas, qualitativas e estruturais desta questão, bem como suas implicações para a democratização genuína do conhecimento científico.

Os dados sobre a dominância do inglês na publicação científica são impressionantes em sua magnitude. Pesquisa recente indica que aproximadamente 98% da pesquisa científica mundial é publicada em inglês, enquanto apenas cerca de 18% da população mundial fala este idioma (Ramírez-Castañeda, 2020). Esta disparidade revela uma desconexão fundamental entre a diversidade linguística global e a representação na literatura científica.

Análise de citações em avaliações internacionais de biodiversidade do Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) demonstra que 96,6% das citações são em inglês, apesar de avaliações nacionais dependerem significativamente de artigos científicos em idiomas não-ingleses (The Conversation, 2024). Esta discrepância sugere que tomadores de decisão internacionais podem estar sistematicamente ignorando conhecimento científico importante simplesmente devido a barreiras linguísticas.

O impacto da barreira linguística na revisão por pares é particularmente bem documentado. Pesquisa empírica demonstra que autores cuja primeira língua não é o inglês experimentam resultados significativamente piores na revisão por pares. Crucialmente, quando a identidade do autor é ocultada dos revisores, essa diferença desaparece, fornecendo evidência clara de viés linguístico no processo de avaliação científica (Nature Human Behaviour, 2023).

A emergência de plataformas de Ciência Aberta controladas por grandes corporações tecnológicas cria novas formas de dependência e vulnerabilidade. Quando pesquisadores depositam dados em plataformas comerciais, podem estar inadvertidamente contribuindo para a criação de recursos posteriormente monetizados por essas empresas. Quando instituições adotam ferramentas e serviços corporativos, criam dependências difíceis de reverter (FT, 2024).

A questão da soberania de dados torna-se particularmente complexa no contexto da Ciência Aberta. Por um lado, a ciência é tradicionalmente vista como um empreendimento universal que transcende fronteiras nacionais. Por outro lado, dados científicos podem ter implicações estratégicas, econômicas, e de segurança que justificam preocupações sobre soberania nacional. Encontrar o equilíbrio entre abertura científica e soberania nacional é um desafio crescente para formuladores de políticas.

As tensões técnicas na implementação da Ciência Aberta revelam-se como um dos aspectos mais sutis, porém fundamentais, da contradição entre abertura e controle. A aparente neutralidade da infraestrutura tecnológica mascara frequentemente relações de poder que se cristalizam em arquiteturas, protocolos e padrões técnicos.

A adoção de padrões técnicos na Ciência Aberta, embora necessária para garantir interoperabilidade e acesso, cria inevitavelmente hierarquias de conhecimento e formas de exclusão. Os padrões de metadados, por exemplo, não são meramente técnicos, mas refletem ontologias específicas sobre como o conhecimento deve ser organizado, categorizado e descoberto. Quando organizações internacionais ou consórcios dominados por instituições do Norte Global estabelecem esses padrões, eles podem inadvertidamente marginalizar formas de conhecimento que não se adequam facilmente a essas categorias.

O Dublin Core Metadata Initiative, amplamente adotado em repositórios de Ciência Aberta, exemplifica essa tensão. Embora tenha facilitado a descoberta e o compartilhamento de recursos digitais, seus elementos básicos refletem uma visão ocidental de como a informação deve ser estruturada. Conceitos como “criador”, “assunto” e “tipo” podem não capturar adequadamente as nuances de sistemas de conhecimento indígenas ou tradicionais, onde a autoria pode ser coletiva, os assuntos podem estar interconectados de maneiras não hierárquicas, e os tipos de conhecimento podem não se adequar a categorias acadêmicas convencionais.

A questão dos identificadores únicos ilustra outro aspecto dessa tensão. Sistemas como DOI (Digital Object Identifier), ORCID (Open Researcher and Contributor ID) e ROR (Research Organization Registry) prometem resolver problemas de ambiguidade e facilitar a descoberta de pesquisa. No entanto, eles também criam pontos centralizados de controle sobre a identidade científica. O ORCID, por exemplo, é gerenciado por uma organização sem fins lucrativos, mas sua governança é dominada por editores comerciais e instituições do Norte Global. Isso levanta questões sobre quem tem o poder de definir e validar identidades científicas, e como essas definições podem excluir ou marginalizar certas comunidades de pesquisa.

A migração para plataformas digitais de publicação e compartilhamento de dados representa uma das transformações mais significativas na infraestrutura científica das últimas décadas. Plataformas como arXiv, bioRxiv, e Zenodo prometem democratizar o acesso ao conhecimento científico, eliminando barreiras tradicionais de publicação e permitindo o compartilhamento rápido de resultados de pesquisa. No entanto, essa migração também cria novas formas de dependência e vulnerabilidade.

A dependência de plataformas comerciais para serviços essenciais de pesquisa ilustra essa tensão. Muitas instituições acadêmicas dependem de serviços como Google Scholar para

descoberta de literatura, Mendeley para gestão de referências, e ResearchGate para networking acadêmico. Embora esses serviços sejam frequentemente gratuitos para usuários individuais, eles são controlados por empresas comerciais que podem alterar termos de serviço, descontinuar funcionalidades, ou monetizar dados de usuários sem aviso prévio.

O caso do Mendeley é particularmente ilustrativo. Originalmente uma startup independente focada em gestão de referências acadêmicas, foi adquirida pela Elsevier em 2013 por \$100 milhões. Após a aquisição, muitos usuários expressaram preocupações sobre como seus dados de pesquisa poderiam ser utilizados pela Elsevier para fins comerciais, incluindo análise de tendências de pesquisa e desenvolvimento de produtos de inteligência de mercado. Embora a Elsevier tenha mantido o serviço gratuito, a aquisição demonstra como plataformas aparentemente neutras podem se tornar ativos estratégicos para grandes corporações.

2.5 ALGORITMOS E CURADORIA: O PODER INVISÍVEL DA DESCOBERTA

Os algoritmos que governam a descoberta e recomendação de conteúdo científico exercem uma influência crescente sobre quais pesquisas são visíveis e influentes. Plataformas como Google Scholar, Semantic Scholar, e até mesmo repositórios institucionais utilizam algoritmos complexos para ranquear e recomendar conteúdo. Esses algoritmos, embora apresentados como neutros e objetivos, incorporam inevitavelmente vieses e pressupostos sobre o que constitui pesquisa de qualidade ou relevante.

A opacidade desses algoritmos cria uma forma de controle invisível sobre o fluxo de conhecimento científico. Pesquisadores

podem não estar cientes de como suas escolhas de palavras-chave, afiliações institucionais, ou padrões de citação afetam a visibilidade de seu trabalho. Estudos têm demonstrado que algoritmos de descoberta podem perpetuar vieses existentes, favorecendo pesquisas de instituições prestigiosas, autores com histórico de citações elevadas, ou tópicos que já recebem atenção significativa.

O fenômeno do “efeito Mateus” na ciência, onde pesquisadores já estabelecidos recebem crédito desproporcional por descobertas, pode ser amplificado por algoritmos de recomendação que favorecem conteúdo já popular ou altamente citado. Isso pode criar ciclos de feedback que marginalizam pesquisas emergentes, perspectivas minoritárias, ou trabalhos de instituições menos prestigiosas.

A concentração do mercado de publicação científica representa uma das manifestações mais tangíveis da tensão entre abertura e controle na Ciência Aberta. Dados recentes revelam a extensão alarmante dessa concentração e seus impactos mensuráveis na disseminação do conhecimento científico.

Pesquisa empírica conduzida por An, Williams e Xiao (2024) fornece evidências quantitativas precisas sobre os impactos da concentração corporativa na publicação científica. O estudo, que analisou mais de um milhão de artigos em economia, física e engenharia eletrônica entre 2009 e 2018, demonstra que a indústria de publicação acadêmica, dominada por empresas como Elsevier, Wiley-Blackwell, Taylor & Francis, Springer Nature e SAGE, gerou \$19 bilhões em receita em 2020. Particularmente notável é a margem de lucro da Elsevier, de aproximadamente 40%, comparável a empresas altamente lucrativas como Microsoft, Google e Coca-Cola (AN, 2024).

Os dados revelam que cinco empresas controlam mais de 50% da publicação acadêmica desde 2006, e mais de 75% da produção de artigos de pesquisa de qualquer país é publicada em revistas de aproximadamente 20 editoras (ESAC, 2025). Esta concentração

não é meramente estatística - ela tem impactos mensuráveis na disseminação do conhecimento. O estudo demonstra que um aumento de 1% nas taxas de assinatura de revistas leva a uma diminuição de 0,83% nas citações em economia, 1,46% em física, e efeitos ainda mais pronunciados em engenharia eletrônica, onde um aumento de 1% na participação de mercado de uma editora resulta em uma diminuição de 4,67% nas citações (AN, 2024).

A integração vertical das grandes editoras científicas representa uma estratégia sofisticada de captura de valor que vai muito além da publicação tradicional. A Elsevier, por exemplo, expandiu suas operações para incluir análise de dados (através de plataformas como SciVal e Scopus), avaliação de pesquisa (Mendeley e Pure), e até mesmo inteligência artificial para descoberta de medicamentos (através de aquisições como Reaxys). Esta diversificação permite que a empresa capture valor em múltiplos pontos da cadeia de produção científica, desde a geração inicial de conhecimento até sua aplicação comercial.

O modelo de negócios dessas empresas revela uma contradição fundamental com os princípios da Ciência Aberta. Pesquisadores produzem conhecimento frequentemente com financiamento público, outros pesquisadores fornecem revisão por pares sem remuneração, pesquisadores pagam taxas de submissão para publicar, e universidades pagam taxas de assinatura para acessar o conteúdo. Essencialmente, o público paga duas vezes: primeiro para financiar a pesquisa e depois para financiar o acesso através dos orçamentos de bibliotecas universitárias (AN, 2024).

A pesquisa de An, Williams e Xiao (2024) fornece evidências empíricas convincentes de que o acesso é o gargalo fundamental para a disseminação do conhecimento [1]. Utilizando a disponibilização de artigos no JSTOR como experimento natural, os pesquisadores demonstraram que quando artigos se tornam disponíveis através desta plataforma de baixo custo, há um aumento imediato no

crescimento anual de citações. Dois anos após a disponibilização no JSTOR, o crescimento nas citações aumenta aproximadamente 5%, chegando a 15% cinco anos depois. O efeito no número de autores que citam esses artigos é ainda maior (AN, 2024).

Estes dados contradizem argumentos de que a concentração corporativa é benéfica para a ciência através de economias de escala ou melhor curadoria. Pelo contrário, eles demonstram que preços elevados e poder de mercado concentrado reduzem sistematicamente a disseminação do conhecimento, com efeitos particularmente pronunciados em instituições de menor prestígio e países em desenvolvimento.

A análise de contratos entre a Elsevier e universidades públicas revela a mecânica concreta através da qual o poder de mercado se traduz em vantagem negocial (AN, 2024). Universidades classificadas pelo Quacquarelli Symonds assinam mais que o dobro de revistas comparado às não classificadas, mas universidades não classificadas pagam ligeiramente mais por revista em média. Esta dinâmica ilustra como o poder de negociação varia com o prestígio institucional, criando um sistema estratificado onde instituições menos prestigiosas enfrentam custos proporcionalmente maiores.

Os contratos revelam também a estratégia de “grandes pacotes” (big deals) utilizada pelas editoras. Em vez de permitir que bibliotecas selecionem revistas individuais, as editoras pressionam por pacotes abrangentes que consomem uma proporção significativa dos orçamentos de bibliotecas, deixando menos recursos para aquisições de editoras menores ou independentes. Esta estratégia efetivamente cria barreiras de entrada para novos competidores e consolida o domínio das grandes editoras.

Os efeitos da concentração corporativa não são uniformemente distribuídos. A pesquisa demonstra que os impactos negativos de preços elevados e poder de mercado concentrado são significativamente

maiores para instituições de menor prestígio e países em desenvolvimento (AN, 2024). Esta distribuição desigual de impactos exacerba desigualdades existentes no sistema científico global.

Para instituições em países em desenvolvimento, os custos de acesso representam uma proporção muito maior de seus orçamentos de pesquisa. Um estudo sobre pesquisadores colombianos revelou que mais de 90% enfrentam custos adicionais significativos devido à necessidade de se comunicar em inglês e acessar literatura científica através de plataformas comerciais (Ramírez-Castañeda, 2020). Esses custos envolvem assinaturas de revistas, serviços de tradução, revisão linguística e taxas de processamento de artigos.

A concentração corporativa também afeta diferentemente campos disciplinares. Áreas como ciências biomédicas, onde a Elsevier tem uma presença particularmente forte através de revistas como *The Lancet* e *Cell*, podem enfrentar custos de acesso proporcionalmente maiores. Campos emergentes ou interdisciplinares podem ter dificuldade em encontrar versões de publicação adequados se não se adequarem às categorias estabelecidas pelas grandes editoras.

2.6 CIÊNCIA CIDADÃ E PARTICIPAÇÃO DEMOCRÁTICA

A ciência cidadã representa uma das dimensões mais promissoras e transformadoras da Ciência Aberta, oferecendo possibilidades para democratizar tanto o acesso ao conhecimento científico quanto a própria produção desse conhecimento. Através da ciência cidadã, pessoas sem treinamento científico formal podem contribuir para pesquisas científicas, desde a coleta de dados até a análise e interpretação de resultados (Bonney *et al.* 2009; Finger *et al.* 2023).

Projetos de ciência cidadã variam enormemente em escopo, metodologia e grau de participação. Em um extremo, encontramos projetos de “crowdsourcing” onde cidadãos contribuem principalmente através da provisão de poder computacional ou realização de tarefas simples de classificação, como Galaxy Zoo, onde voluntários classificam galáxias em imagens astronômicas, ou eBird, onde observadores de aves contribuem dados sobre distribuição e abundância de espécies (Raddick *et al.* 2013; eBird 2025).

Em outro extremo, encontramos projetos de “ciência participativa” onde comunidades estão envolvidas em todas as fases do processo de pesquisa, desde a definição de questões de pesquisa até a interpretação e aplicação de resultados. Esses projetos frequentemente emergem de necessidades e preocupações comunitárias específicas, como monitoramento de qualidade ambiental, avaliação de riscos de saúde ou documentação de conhecimentos tradicionais (Ramirez-Andreotta *et al.* 2015).

A ciência cidadã oferece benefícios mútuos tanto para a ciência quanto para os participantes. Para a ciência, ela pode fornecer acesso a escalas de coleta de dados que seriam impossíveis através de métodos tradicionais, diversificar perspectivas e abordagens de pesquisa, e aumentar a relevância social da pesquisa científica. Para os participantes, a ciência cidadã pode oferecer oportunidades de aprendizado, empoderamento cívico e contribuição para questões de interesse pessoal ou comunitário (Den Broeder *et al.* 2017; Danielsen *et al.* 2021).

No entanto, a ciência cidadã também enfrenta desafios significativos relacionados à qualidade de dados, sustentabilidade de projetos e equidade de participação. Preocupações sobre qualidade de dados levaram ao desenvolvimento de protocolos rigorosos de treinamento, validação e controle de qualidade, mas esses protocolos podem criar barreiras à participação e reproduzir hierarquias entre especialistas e amadores (Losey 2014; Danielsen *et al.* 2009).

A sustentabilidade de projetos de ciência cidadã é frequentemente problemática, pois muitos projetos dependem de financiamento de curto prazo e trabalho voluntário que pode ser difícil de manter ao longo do tempo. A falta de reconhecimento institucional para ciência cidadã pode desencorajar pesquisadores de se envolver em tais projetos, limitando seu desenvolvimento e impacto (FRONTIERS IN ENVIRONMENTAL SCIENCE, 2022; Bonney *et al.* 2009).

Questões de equidade são particularmente importantes na ciência cidadã. Embora a ciência cidadã prometa democratizar a participação na ciência, na prática, a participação é frequentemente enviesada em direção a populações com maior educação, renda e tempo livre. Isso pode resultar em uma ciência cidadã que na verdade reproduz exclusões existentes em vez de superá-las (Dawson 2019; Pandya *et al.* 2021).

Além disso, existe o risco de que a ciência cidadã seja utilizada como forma de trabalho gratuito que beneficia principalmente pesquisadores e instituições profissionais. Quando cidadãos contribuem dados e análises, mas não recebem reconhecimento adequado ou benefícios dos resultados, isso pode constituir exploração mascarada pela retórica da participação democrática (Reiheld; Gay, 2019).

Projetos de ciência cidadã mais transformadores buscam abordar essas limitações através de abordagens que enfatizam reciprocidade, empoderamento comunitário e justiça social. Isso inclui garantir que comunidades participantes tenham controle sobre como seus dados são utilizados, que recebam benefícios tangíveis dos resultados de pesquisa e tenham oportunidades de desenvolver suas próprias capacidades de pesquisa (Ramirez-Andreotta *et al.*, 2015).

2.7 PARADOXOS DA DEMOCRATIZAÇÃO CIENTÍFICA

A Ciência Aberta está repleta de paradoxos que revelam as complexidades e contradições inerentes aos esforços de democratizar o conhecimento científico. Esses paradoxos não são meras curiosidades intelectuais, mas tensões fundamentais que moldam as possibilidades e limitações da Ciência Aberta contemporânea.

O primeiro paradoxo é o da “abertura excludente”. Embora a Ciência Aberta prometa tornar o conhecimento científico acessível a todos, na prática, a capacidade de se beneficiar dela depende de recursos e capacidades que não estão igualmente distribuídos; acesso à internet de alta velocidade, dispositivos adequados, literacia digital, e tempo livre são pré-requisitos que muitas pessoas não possuem (Piwowar *et al.*, 2018).

O segundo paradoxo é o da “participação hierárquica”. Embora a Ciência Aberta prometa democratizar a produção de conhecimento, na prática diferentes formas de participação são estruturadas hierarquicamente. Muitos cidadãos são convidados somente a contribuir com dados ou tarefas simples, mas raramente participam na definição de questões de pesquisa, interpretação de resultados, ou decisões sobre o uso do conhecimento (Haklay *et al.*, níveis de participação; mais recente revisão).

O terceiro paradoxo é o da “transparência opaca”. A Ciência Aberta enfatiza transparência e reprodutibilidade, mas muitas das infraestruturas e algoritmos que a sustentam são propriedade de empresas privadas e não estão sujeitos a escrutínio público. Isso cria situações onde os dados e publicações podem estar “abertos”, mas os sistemas que os organizam permanecem opacos (Maedche *et al.*, 2024).

O quarto paradoxo é o da “colaboração competitiva”. Embora a Ciência Aberta prometa facilitar colaboração e compartilhamento, ela opera dentro de um sistema científico fundamentalmente competitivo. Pesquisadores são avaliados por publicações individuais, citações, e métricas que incentivam competição; isso pode resultar em formas superficiais de abertura que não transformam as dinâmicas competitivas subjacentes (Royal Society of Open Science, 2023).

O quinto paradoxo é o da “universalidade particular”. A Ciência Aberta é frequentemente apresentada como um movimento universal que beneficia toda a humanidade, mas na prática ela é moldada por valores, prioridades e interesses específicos de grupos dominantes. Padrões considerados “universais” podem privilegiar formas de organização do conhecimento familiares para comunidades científicas dominantes, mas estranhas para outras.

2.8 ALTERNATIVAS E RESISTÊNCIAS À CIÊNCIA ABERTA HEGEMÔNICA

A Ciência Aberta, em sua concepção mais difundida, propõe uma transformação radical na maneira como o conhecimento é produzido e compartilhado, mas sob essa aparente unidade esconde-se um modelo hegemônico que enfrenta crescentes resistências. Esse modelo dominante, frequentemente centrado nas agendas e tecnologias do Norte Global e por vezes cooptado por grandes interesses comerciais, é desafiado por movimentos que buscam não apenas o acesso livre, mas uma reestruturação fundamental das práticas científicas em nome da justiça e da equidade. Essas alternativas e resistências não rejeitam a partilha do saber, mas questionam quem define as regras do jogo e quem realmente se beneficia dele.

Nesse cenário, emergem iniciativas de “Ciência Aberta decolonial” que desenvolvem práticas para reconhecer epistemologias não-ocidentais e promover justiça epistêmica, incluindo projetos que documentam conhecimentos tradicionais e promovem colaboração equitativa entre pesquisadores do Norte e Sul Global (Number Analytics, 2025; Southern interventions, 2021). De forma complementar, movimentos de “Ciência Aberta feminista” aplicam perspectivas críticas para transformar as práticas estabelecidas, prestando atenção aos modos como gênero, raça e classe moldam as experiências de participação e desenvolvendo metodologias que incorporam valores de cuidado e reciprocidade (Samper, 2023; Data Feminism for AI, 2024).

Outra frente de resistência se organiza através de iniciativas de “commons científicos”, que buscam criar formas de governança coletiva de recursos que vão além de modelos privados ou estatais, oferecendo infraestrutura para *software* livre e publicação cooperativa em espaços de ação colaborativa e autogerida (Comunes Collective, 2009). Essa busca por autonomia reflete-se também em projetos de “tecnologia apropriada” para a Ciência Aberta, que desenvolvem ferramentas e infraestruturas adequadas a contextos específicos, como interfaces em línguas locais e tecnologias que funcionam com recursos limitados, em vez de soluções universais (Open Science for Decolonial Futures, 2025).

Embora esses movimentos alternativos enfrentem enormes desafios, como a falta de recursos e reconhecimento institucional, eles representam muito mais do que simples críticas. São esforços, muitas vezes marginalizados, que indicam caminhos potentes para que a promessa de uma Ciência Aberta verdadeiramente justa, democrática e transformadora possa um dia ser plenamente realizada.

2.9 PERSPECTIVAS FUTURAS

O futuro da Ciência Aberta será moldado pelas tensões e contradições discutidas ao longo deste capítulo, bem como por desenvolvimentos tecnológicos, políticos e sociais mais amplos. Várias tendências emergentes prometem intensificar tanto as oportunidades quanto os riscos associados à Ciência Aberta contemporânea.

O desenvolvimento de tecnologias de inteligência artificial está criando novas possibilidades para automatizar aspectos da produção científica, desde a coleta e análise de dados até a escrita e revisão de publicações. Embora isso possa acelerar o progresso científico e reduzir custos, também levanta questões sobre qualidade, transparência e controle humano sobre o processo científico (OpenAIRE Blog, 2023). Estudos recentes alertam que modelos como LLMs tendem a simplificar ou distorcer os resultados científicos, o que pode comprometer a confiança na pesquisa automatizada (The Guardian, 2025).

A emergência de tecnologias de *blockchain* e outras plataformas descentralizadas oferece possibilidades para criar sistemas de Ciência Aberta menos dependentes de controle corporativo centralizado, com infraestrutura de timestamping imutável, identidade digital auto-soberana e revisões científicas mais transparentes (Wired, 2024; MDPI Reviews, 2025). Por exemplo, pesquisas têm explorado o uso de *blockchain* para tornar o peer review mais íntegro e verificável (MDPI, 2025) e para facilitar computações científicas complexas fora de infraestrutura tradicional, como em simulações na origem da vida.

Mudanças nas políticas governamentais e institucionais estão moldando a Ciência Aberta. Mandatos de acesso aberto estão se tornando mais comuns, mas a implementação varia amplamente (Plan S, cOAlition S, 2021). Regulamentações sobre proteção de dados, como o GDPR, criam novas restrições sobre compartilhamento, mas também oferecem proteção à privacidade individual

A crescente conscientização sobre mudanças climáticas e outros desafios globais está pressionando por uma ciência mais relevante e responsiva às necessidades sociais urgentes. Isso favorece formas de Ciência Aberta que enfatizam participação comunitária e aplicação prática, embora exista o risco de instrumentalização da ciência para agendas políticas específicas (UNESCO, 2024).

O futuro da Ciência Aberta não está predeterminado. Ele será moldado pelas escolhas que fazemos hoje sobre como desenvolver, implementar, e governar sistemas de Ciência Aberta. Compreender as tensões e contradições atuais é fundamental para tomar decisões informadas que realizem os potenciais transformadores da Ciência Aberta, evitando suas armadilhas e limitações.

Parte



ANÁLISE CRÍTICA E IMPACTOS

3

**O IMPACTO DO
COLONIALISMO DE DADOS
NA PRODUÇÃO CIENTÍFICA
E NO CONHECIMENTO**

O colonialismo de dados não opera apenas no nível individual ou corporativo, mas permeia profundamente as estruturas de produção e circulação do conhecimento científico global. Este capítulo examina como as dinâmicas coloniais de extração e apropriação de dados afetam quem produz conhecimento, quem tem acesso a ele, quais narrativas predominam, e como essas assimetrias reproduzem e amplificam desigualdades históricas entre o Norte e Sul Global.

A produção científica contemporânea está intrinsecamente ligada à capacidade de coletar, processar e analisar grandes volumes de dados. Países e instituições com maior capacidade computacional, infraestrutura digital avançada e recursos financeiros substanciais possuem vantagens estruturais significativas na produção de conhecimento científico (UNESCO, 2021). Essa vantagem não é meramente técnica, mas reflete relações de poder com raízes históricas profundas no colonialismo tradicional.

3.1. DESIGUALDADES NA PRODUÇÃO CIENTÍFICA GLOBAL (NORTE GLOBAL VS. SUL GLOBAL)

As desigualdades na produção científica global são um reflexo direto das assimetrias de poder que caracterizam o sistema mundial contemporâneo. Dados da UNESCO apontam que os países do Norte Global, que representam aproximadamente 20% da população mundial, são responsáveis por mais de 80% da produção científica global. Essa concentração não é apenas quantitativa, mas também qualitativa, pois os critérios de qualidade, relevância e impacto científico são definidos por instituições do Norte Global, frequentemente marginalizando formas de conhecimento prioritárias no Sul Global (UNESCO, 2021).

Em termos práticos, isso se traduz em dependência de agendas, padrões e métricas que não necessariamente respondem às necessidades locais—criando um ciclo de path dependence: países com maior acúmulo histórico de capital científico seguem ditando o que “conta” como ciência de excelência, enquanto grupos periféricos lutam para fazer reconhecer problemas e metodologias contextualizados.

A concentração geográfica da infraestrutura digital científica também amplifica essas desigualdades. A maioria dos servidores que hospedam bases de dados, repositórios e plataformas de colaboração científica está localizada nos Estados Unidos e Europa, criando dependência estrutural e vulnerabilidade para países periféricos (UNESCO, 2021).

Consequências típicas incluem: exposição a mudanças unilaterais de termos de uso e preços; risco de interrupções e bloqueios que afetam a continuidade de pesquisa; latências regulatórias (conflitos entre regras de proteção de dados e políticas de armazenamento transfronteiriço); e assimetrias na própria governança das infraestruturas (comitês e consórcios sediados no Norte Global).

A concentração de recursos para pesquisa científica é impressionante. Os Estados Unidos sozinhos investem cerca de 28% do gasto global em Pesquisa e Desenvolvimento (P&D), enquanto América Latina, Caribe e África combinadas representam aproximadamente 8% (NSB 2023; WIPO 2023). Essa concentração de recursos traduz-se em capacidades diferenciadas para conduzir pesquisas que requerem infraestruturas caras, como sequenciamento genômico, simulações computacionais complexas ou experimentos com equipamentos sofisticados. Na prática, isso significa assimetria de entrada em áreas intensivas em capital (HPC, grandes coleções de dados, laboratórios de grande porte), custos de manutenção difíceis de sustentar (contratos, licenças, atualizações) e desigualdade na formação avançada (programas, estágios e redes de elite frequentemente concentrados no Norte).

A infraestrutura digital científica também está altamente concentrada. Empresas como Elsevier, Springer Nature e Wiley controlam grande parte das publicações científicas globais, enquanto plataformas como Web of Science e Scopus dominam a indexação e avaliação da produção científica (Mongeon; Paul-Hus, 2015). Essa concentração permite que poucas empresas influenciem significativamente quais pesquisas são visíveis, citadas e consideradas impactantes. Esse arranjo reforça mecanismos de gatekeeping: critérios editoriais e de indexação favorecem certos campos, países e idiomas; taxas de processamento de artigos (APCs) tornam-se barreiras de participação; e estruturas de avaliação (que dependem dessas bases) retroalimentam a centralidade de centros já dominantes.

A dependência de infraestruturas digitais proprietárias cria vulnerabilidades específicas para países do Sul Global. Interrupções de serviço, mudanças em políticas de acesso ou aumentos de preços podem afetar drasticamente a capacidade de pesquisadores nesses países de acessar literatura científica ou disseminar seus próprios resultados. Durante a pandemia de COVID-19, várias editoras ofereceram acesso gratuito temporário a suas publicações, mas essa "generosidade" destacou precisamente a dependência estrutural que caracteriza o sistema atual (NSB 2023). Ou seja, em contextos críticos, o acesso pode ser revogável e contingente, o que é incompatível com uma visão de ciência como bem comum e com demandas de continuidade para pesquisa e ensino.

O domínio do inglês como língua franca da ciência representa outra dimensão importante das desigualdades na produção científica global. Estudos indicam que cerca de 90–98% das publicações indexadas nas principais bases como Web of Science e Scopus são publicadas em inglês, marginalizando trabalhos em outras línguas (PLOS ONE 2020; Huttner-Koros 2015). Efeitos correlatos incluem: custos adicionais de tradução e revisão; pressão por conformidade retórica (formatos, estilos argumentativos e referenciais teóricos

dominantes); e redução da diversidade epistemológica, com perda de nuances locais, terminologias e tradições acadêmicas regionais.

Pesquisadores cujo primeiro idioma não é o inglês enfrentam barreiras adicionais para publicar em revistas internacionais de prestígio. Essas barreiras envolvem tanto a competência linguística quanto a familiaridade com convenções retóricas e epistemológicas da ciência anglófona. Isso pode resultar na exclusão de perspectivas importantes e na homogeneização do discurso científico global (Huttner-Koros, 2015).

Além da dimensão linguística, há assimetria de redes: editoriais e conselhos científicos concentram-se em instituições do Norte, o que afeta decisões de pauta, indicação de pareceristas e circulação de reconhecimento (convites, prêmios, colaborações).

Além disso, muitas questões de pesquisa prioritárias para o Sul Global podem não ser consideradas relevantes por editores e revisores baseados no Norte Global. Isso cria um viés sistemático que favorece pesquisas alinhadas às prioridades do Norte e marginaliza problemas locais urgentes (Nordling, 2024).

Esse desalinhamento aparece no recorte temático (o que merece atenção), nos métodos legitimados (o que é considerado robusto) e nos indicadores de impacto (o que "conta" para carreira e financiamento), perpetuando sub-representação de agendas como doenças negligenciadas, impactos socioambientais regionais, multilinguismo científico e tecnologias apropriadas.

Em síntese, os elementos destacados por UNESCO (2021), NSB (2023), WIPO (2023), Mongeon; Paul-Hus (2015), PLOS ONE (2020), Huttner-Koros (2015) e Nordling (2024) compõem um circuito de desigualdades que atua em três níveis interligados:

- **Recursos e infraestrutura** (financiamento, equipamentos, plataformas, governança);
- **Regras do jogo** (métricas, indexação, padrões editoriais, idioma);

- **Agendas e reconhecimento** (o que é pertinente, como se avalia, quem é visto).

Para mitigar essas assimetrias, sem contrariar os dados e diagnósticos já citados, caminhos consensuais incluem: fortalecimento de infraestruturas abertas e distribuídas (repositórios, indexadores e métricas responsáveis), políticas multilíngues de publicação e inclusão de editoras e conselhos locais, modelos de financiamento que reduzam barreiras de APC, e avaliação que reconheça relevância social e reprodutibilidade, permitindo que o Sul Global não apenas participe, mas coproduza critérios e direções do sistema científico.

3.2 APROPRIAÇÃO DE DADOS CIENTÍFICOS E PARACHUTE SCIENCE

A apropriação de dados científicos representa uma das manifestações mais diretas do colonialismo de dados na ciência contemporânea. Essa apropriação ocorre quando dados coletados em determinados contextos são utilizados para fins diferentes daqueles para os quais foram originalmente coletados, frequentemente sem o consentimento ou benefício das comunidades que forneceram os dados.

O fenômeno da “parachute science” (também chamado de pesquisa tipo helicóptero) descreve práticas em que cientistas de países desenvolvidos realizam estudos em países do Sul Global com pouca ou nenhuma colaboração substantiva de pesquisadores locais. Pesquisadores “pousam” para coletar dados e amostras, “decolam” para analisar e publicar em suas instituições de origem e, com frequência, não retornam benefícios tangíveis às comunidades envolvidas. Essa dinâmica, vinculada a assimetrias históricas e coloniais, pode incluir apropriação de conhecimentos tradicionais e

recursos biológicos sem reconhecimento ou repartição de benefícios, além da imposição de metodologias descontextualizadas que geram riscos sociais e ambientais (Dahdouh-Guebas *et al.*, 2003; *Havasupai Tribe v. Arizona Board of Regents*, 2010; “Neo-colonial science, 2003–2020”).

Evidências empíricas indicam a dimensão do problema: meta-análises mostram que, em estudos ecológicos, a proporção de autores locais incluídos em pesquisas sobre seus próprios países é de cerca de 0,43 no Sul Global, contra 0,73 no Norte Global — um indicador das barreiras de participação e reconhecimento enfrentadas por pesquisadoras e pesquisadores locais (Economou-García, 2022). Em síntese, a *parachute science* materializa desigualdades de poder no fazer científico e exige respostas que priorizem parcerias equitativas, repartição justa de benefícios, autoria responsável e fortalecimento de capacidades locais.

Essa prática é particularmente comum em áreas como biodiversidade, saúde pública, antropologia e estudos ambientais. Pesquisadores de universidades prestigiosas frequentemente têm acesso a financiamento para conduzir pesquisas em países tropicais, onde a biodiversidade é maior, ou em comunidades com características específicas que são de interesse científico. No entanto, os benefícios dessa pesquisa, incluindo publicações, reconhecimento acadêmico, e potenciais aplicações comerciais, frequentemente permanecem concentrados no Norte Global (Harmon 2009; *Havasupai Tribe v. Arizona Board of Regents* 2010).

Além disso, o caso dos dados genéticos da tribo Havasupai ilustra como a apropriação de dados pode violar direitos fundamentais e causar danos significativos a comunidades vulneráveis. Pesquisadores da Universidade do Arizona coletaram amostras de sangue de membros da tribo para estudos sobre diabetes, mas posteriormente utilizaram essas amostras para pesquisas sobre migração, consanguinidade e esquizofrenia sem o consentimento da tribo (Harmon 2009; *Havasupai Tribe v. Arizona Board of Regents* 2010).

Casos similares ocorreram em várias partes do mundo. Na Islândia, a empresa deCODE Genetics obteve acesso a dados genéticos de uma parcela significativa da população islandesa para pesquisas comerciais, levantando questões sobre consentimento coletivo e propriedade de dados genéticos populacionais (Stefansson *et al.* 2000; Act on Health Sector Database 1998). Na África, várias empresas farmacêuticas conduziram ensaios clínicos utilizando populações locais como sujeitos de pesquisa, mas desenvolveram medicamentos que são posteriormente vendidos a preços inacessíveis para essas mesmas populações (ex.: *generic patterns*, múltiplas reportagens).

Esses casos revelam como a apropriação de dados não é apenas uma questão técnica ou legal, mas uma questão de justiça social e direitos humanos. A extração de dados de comunidades vulneráveis para benefício de atores mais poderosos reproduz padrões históricos de exploração e pode causar danos culturais, sociais e econômicos significativos (Odeny; Bosurgi, 2022).

O caso dos dados genéticos da tribo Havasupai, que será examinado em detalhes no Capítulo 5, ilustra como a apropriação de dados pode violar direitos fundamentais e causar danos significativos a comunidades vulneráveis. Pesquisadores da Universidade do Arizona coletaram amostras de sangue de membros da tribo para estudos sobre diabetes, mas posteriormente utilizaram essas amostras para pesquisas sobre migração, consanguinidade e esquizofrenia sem o consentimento da tribo (Community Research Ethics, 2022).

Casos similares ocorreram em várias partes do mundo. Na Islândia, a empresa deCODE Genetics obteve acesso a dados genéticos de uma parcela significativa da população islandesa para pesquisas comerciais, levantando questões sobre consentimento coletivo e propriedade de dados genéticos populacionais (The Instructive Icelandic Saga, 2024). Na África, várias empresas farmacêuticas conduziram ensaios clínicos utilizando populações

locais como sujeitos de pesquisa, mas desenvolveram medicamentos que são posteriormente vendidos a preços inacessíveis para essas mesmas populações.

Esses casos revelam como a apropriação de dados não é apenas uma questão técnica ou legal, mas uma questão de justiça social e direitos humanos. A extração de dados de comunidades vulneráveis para benefício de atores mais poderosos reproduz padrões históricos de exploração e pode causar danos culturais, sociais e econômicos significativos (Garrison; Cho 2013; Community Research Ethics, 2022).

A apropriação de dados científicos opera através de vários mecanismos que frequentemente são legalmente válidos, porém são profundamente eticamente questionáveis.

Termos de consentimento extensos e complexos podem permitir usos de dados além da compreensão dos participantes, inclusive usos não previstos ou jamais autorizados por eles (Sedenberg; Hoffmann, 2016). Esses termos geralmente recorrem a linguagem jurídica, obscura ou técnica, que impede a compreensão sobre os riscos reais e a destinação futura dos dados (Hutton; Henderson, 2017).

Cláusulas de transferência de propriedade intelectual podem legalmente conceder às instituições controle sobre dados ou saberes tradicionais, sem que os fornecedores originais, frequentemente comunidades indígenas ou locais, tenham controle, direitos ou compensação (Duke Project 2004; WIPO 2014). Essas cláusulas vêm aparecendo em acordos de pesquisa que exigem cessão de direitos sobre amostras biológicas ou dados comunitários.

A digitalização e agregação de dados intensificam esse processo. Dados coletados para objetivos específicos podem ser combinados com outros conjuntos (incluindo dados sensíveis) para gerar novos conhecimentos, produtos ou patentes, sem que os participantes originais estejam cientes dessa reutilização (Roberts; Montoya, 2023).

Algoritmos de aprendizado de máquina potencializam isso ao extrair padrões latentes que jamais foram descritos durante a coleta, ampliando usos previstos e autorizados (Roberts; Montoya, 2023).

A globalização das infraestruturas de pesquisa facilita ainda mais esses processos. Dados coletados em um país podem ser armazenados ou processados em servidores em outras jurisdições com regulamentação mais brandas. Isso permite que pesquisadores contornem normas locais, como proteção de dados ou direitos das comunidades, e utilizem os dados de forma que nunca teria sido possível no contexto original (Roberts; Montoya 2023; Menlo Report, 2012).

Esse cenário torna evidente que muitos mecanismos de apropriação são formalmente legais, mas moralmente críticos.

3.3. EPISTEMOLOGIAS MARGINALIZADAS E COLONIALIDADE DO SABER NA CIÊNCIA

A discussão sobre a colonialidade do saber na ciência nos convida a desvendar as estruturas de poder invisíveis que governam o que consideramos conhecimento válido. Este conceito, fundamental para o pensamento decolonial, transcende a noção de colonialismo como um período histórico encerrado. Em vez disso, refere-se a uma lógica de poder duradoura, uma “matriz colonial” que sobreviveu às independências formais e continua a operar na era digital, hierarquizando e classificando o mundo a partir de uma ótica eurocêntrica (Quijano 1992; Maldonado-Torres 2007). Trata-se de um sistema que não apenas privilegiou certas formas de saber, mas que ativamente marginalizou e deslegitimou outras, num processo contínuo de dominação epistemológica.

Historicamente, a ciência moderna ocidental consolidou sua hegemonia através da desqualificação sistemática de outras epistemologias. Saberes indígenas, conhecimentos tradicionais sobre a terra, práticas locais de cura e cosmologias não-ocidentais foram categorizados como “não-científicos”, “primitivos” ou “folclore”. Essa classificação não foi um mero debate acadêmico, mas uma ferramenta política que justificou a exploração e a exclusão, negando a esses saberes o acesso aos circuitos formais de validação (Santos 2007; Quijano 1992). Esse processo pode ser entendido como um “epistemicídio”, a destruição sistemática de saberes não-hegemônicos, onde a afirmação universal de um único modelo de ciência levou à invalidação e ao desaparecimento de incontáveis formas de conhecer e de se relacionar com o mundo (Santos, 2007).

Na era contemporânea, essa colonialidade do saber se reencarna e se perpetua em nossas tecnologias digitais. Algoritmos de classificação, sistemas de inteligência artificial e grandes bases de dados, desenvolvidos predominantemente no Norte Global, inevitavelmente refletem os vieses epistemológicos de seus criadores. Eles impõem categorias e ontologias que muitas vezes não encontram correspondência nas realidades locais, achatando a complexidade do mundo para que ela caiba em modelos de dados pré-definidos (Roberts; Montoya 2023; Nyaaba *et al.* 2024). Conceitos, relações e saberes que não possuem equivalentes diretos em línguas dominantes ou que não se alinham à lógica cartesiana são frequentemente distorcidos, ignorados ou simplesmente tornam-se invisíveis, realizando uma forma digital de epistemicídio.

É precisamente essa pretensão de validade universal que a crítica decolonial problematiza. Pensadores como Aníbal Quijano (1992) argumentam que a associação da ciência exclusivamente à racionalidade ocidental é uma construção de poder que subordina todas as outras epistemologias. Essa perspectiva foi aprofundada por autores como Enrique Dussel, Walter Dignolo e Santiago Castro-Gómez, que demonstraram como a hegemonia científica

ocidental não é apenas um modelo entre outros, mas uma lógica global de dominação que ativamente desqualifica e silencia narrativas, histórias e saberes não-europeus, mantendo-os em uma posição de subalternidade (Dussel 2009; Castro-Gómez 2005; Tucker, 2018).

Essa dinâmica se manifesta de forma concreta nas práticas de pesquisa, que muitas vezes se tornam extrativas. A aplicação de metodologias eurocêntricas em contextos não-ocidentais pode perpetuar uma relação colonial, na qual pesquisadores do Norte Global se posicionam como a autoridade final sobre o conhecimento local. Eles “extraem” dados e saberes das comunidades, os processam em seus centros acadêmicos e os publicam em seus próprios termos, muitas vezes sem um retorno significativo ou uma coautoria real com os detentores originais daquele conhecimento. Essa prática não apenas se apropria de significados culturais, mas também reforça as hierarquias de poder que dizem estudar (Shepherd 2019; Nuñez Pardo, 2019).

Como um ato de resistência e reparação, surge o movimento por metodologias de pesquisa descolonizadas. Essa abordagem representa uma mudança radical de poder, que busca substituir a lógica extrativa pela colaboração genuína. Práticas como a participação direta das comunidades no desenho da pesquisa, o reconhecimento dos saberes locais e de seus portadores como coautores intelectuais, e a busca por uma “reciprocidade epistemológica” são fundamentais. O objetivo é estabelecer um diálogo horizontal, onde o conhecimento da comunidade não é tratado como matéria-prima, mas como um parceiro igualitário no processo de produção de saber (Keikelame; Swartz, 2019).

Em última análise, o projeto decolonial visa à construção de um mundo onde muitos mundos possam coexistir. A proposta é avançar para sistemas de conhecimento pluralistas, que legitimem os saberes locais, valorizem a produção científica em línguas maternas e construam o que Boaventura de Sousa Santos (2007) chama

de “ecologias do conhecimento”. Nessa visão, a ciência ocidental é “provincializada” — reconhecida como uma forma de conhecimento poderosa e útil, mas não a única. O objetivo não é negá-la, mas sim colocá-la em um diálogo respeitoso e produtivo com a vasta diversidade de saberes humanos, permitindo que colaborem para enfrentar os complexos desafios globais sem que uma precise aniquilar a outra (Keikelame, 2019).

Um dos efeitos mais perniciosos e persistentes da colonialidade do saber é a sistemática invisibilização de conhecimentos locais, um processo que se intensifica na era digital. Este fenômeno ocorre quando os sistemas de organização do mundo — sejam eles bases de dados, algoritmos de inteligência artificial ou plataformas digitais — são construídos sobre uma base epistemológica restrita e, por consequência, falham em capturar, representar ou sequer reconhecer a validade de formas de conhecimento que não se encaixam nos padrões dominantes do Norte Global (Roberts; Montoya 2023; Nyaaba *et al.*, 2024). Essa invisibilidade não é um mero defeito técnico; é a manifestação de uma hierarquia de poder que decreta quais realidades merecem ser vistas e quais podem ser ignoradas.

Um exemplo contundente desse processo se encontra nos sistemas de classificação taxonômica que alimentam os grandes bancos de dados sobre biodiversidade. A taxonomia ocidental, com sua estrutura lineana, é imposta como um padrão universal, desconsiderando ou tratando como secundários os sofisticados sistemas de classificação indígenas. Estes sistemas locais, muitas vezes, não se baseiam apenas na morfologia das espécies, mas em categorias ecológicas complexas, relações espirituais e usos práticos que são vitais para as comunidades. Ao impor um modelo único, os sistemas digitais acabam por apagar nomenclaturas e classificações construídas ao longo de séculos de diálogo profundo com o ambiente, resultando na perda de um conhecimento contextual, relacional e de imenso valor para a gestão sustentável dos ecossistemas (Vargas-Solar, 2022; Nyaaba *et al.*, 2024).

De forma análoga, os algoritmos de reconhecimento de imagem, que parecem operar com uma objetividade matemática, na verdade reproduzem uma visão de mundo extremamente limitada. Treinados predominantemente com vastos conjuntos de dados provenientes de países desenvolvidos, esses sistemas aprendem a “ver” um mundo que reflete a realidade de seus criadores. Consequentemente, eles demonstram uma acentuada inaptidão para identificar objetos, práticas ou contextos culturais que fogem desse padrão. Rostos de etnias sub-representadas, vestimentas tradicionais, gestos rituais ou práticas agrícolas não-ocidentais são frequentemente ignorados ou classificados de forma errônea. Isso não só produz altos índices de erro, mas resulta em uma exclusão digital ativa de grupos já marginalizados, reforçando sua invisibilidade no cenário global (Roberts; Montoya, 2023; Buolamwini; Gebru, 2018).

A própria iniciativa de digitalizar saberes tradicionais, embora muitas vezes bem-intencionada, carrega o risco da descontextualização e da apropriação indevida. Conhecimentos que são transmitidos oralmente, que vivem em práticas comunitárias e que estão intrinsecamente ligados a uma dimensão espiritual e relacional, sofrem uma profunda transformação ao serem convertidos em formatos digitais. Essa “tradução” para o meio digital inevitavelmente aplaina suas nuances, remove seu contexto sagrado e os despoja das relações de responsabilidade que os governam. Uma vez transformados em “dados”, esses conhecimentos podem ser utilizados de maneiras completamente desalinhadas com seus valores originais, configurando uma forma sutil, mas profunda, de violação cultural e epistemicídio (Santos, 2007; Nyaaba *et al.*, 2024).

Por fim, essa invisibilização é cimentada pela tirania da tradução e da categorização universal. Sistemas automatizados de tradução e taxonomias globais, ao tentarem criar uma equivalência entre todas as línguas e culturas, acabam por impor as categorias das línguas dominantes como a norma universal. Expressões, conceitos e visões de mundo que não possuem uma tradução direta são distorcidos, simplificados ou simplesmente perdidos no processo.

Ao forçar a diversidade epistemológica do mundo a caber em um vocabulário restrito, esses sistemas reduzem não só a visibilidade de saberes locais, como também corroem ativamente sua validade, perpetuando a ideia de que o que não pode ser nomeado ou categorizado pelo centro não possui legitimidade (Roberts; Montoya, 2023; Nyaaba *et al.*, 2024).

A busca incessante por interoperabilidade e padronização no universo digital, embora motivada pela nobre intenção de facilitar a comunicação e o intercâmbio de informações, carrega consigo um profundo e muitas vezes invisível risco: a homogeneização das formas de conhecimento. Padrões globais de metadados, ontologias e sistemas de classificação, essenciais para que sistemas diferentes possam “conversar” entre si, raramente são universais em sua concepção. Na prática, eles refletem as perspectivas epistemológicas específicas de onde foram criados, falhando em acomodar a imensa diversidade de saberes que existem no mundo (Montenegro, 2019; Carroll *et al.*, 2020). Ferramentas como o Dublin Core, por exemplo, são frequentemente implementadas sob um véu de neutralidade técnica, mas na verdade carregam vieses eurocêntricos que ignoram ou se mostram incompatíveis com os protocolos culturais e as visões de mundo de comunidades indígenas e outras populações marginalizadas (Montenegro, 2019).

Essa imposição de uma norma única é particularmente prejudicial para comunidades cujo modo de vida, saúde e identidade cultural dependem intrinsecamente de saberes locais. Ao serem forçadas a adaptar seus conhecimentos a padrões externos para que sejam “compatíveis” ou “legíveis” por sistemas globais, essas comunidades veem seus saberes serem despojados de aspectos críticos de significado e utilidade contextual (Carroll *et al.*, 2020; Riley *et al.*, 2023). Um sistema de classificação indígena, por exemplo, pode organizar plantas não apenas por sua biologia, mas por suas dimensões espirituais, seus usos rituais, suas relações ecológicas com outros seres e suas histórias orais. Nos sistemas padronizados,

essas dimensões relacionais e sagradas, que são o cerne do conhecimento, são frequentemente reduzidas a categorias irrelevantes, achatadas em campos de metadados inadequados ou simplesmente omitidas por completo, pois não há um “lugar” para elas na estrutura pré-definida (Indigenous Metadata Bundle, 2023).

Dessa forma, a padronização instituída deixa de ser uma ferramenta técnica para se tornar um mecanismo que estabelece e reforça hierarquias epistemológicas implícitas. Os saberes que se ajustam facilmente aos padrões dominantes — aqueles que podem ser quantificados, categorizados e descontextualizados sem grande perda — são implicitamente validados como mais “confiáveis”, “organizados” ou “científicos”. Em contrapartida, as formas de conhecimento que resistem a essa padronização, por sua natureza holística, relacional ou oral, correm o risco concreto de serem marginalizadas, desvalorizadas e, por fim, tornarem-se institucionalmente invisíveis, reforçando a ideia de que apenas um modo de organizar a realidade é legítimo (Riley *et al.*, 2023).

Como um movimento de resistência a essa colonização digital, surgiram iniciativas que buscam subverter essa lógica. As *Traditional Knowledge Labels* (TK Labels), ou Etiquetas de Conhecimento Tradicional, são um exemplo poderoso. Elas funcionam como camadas de metadados adicionais, criadas e controladas pelas próprias comunidades, que podem ser aplicadas a itens em arquivos e repositórios digitais. Essas etiquetas informam aos usuários sobre a proveniência cultural do conhecimento, indicam protocolos comunitários específicos, impõem restrições de uso ou sinalizam o caráter sagrado de um determinado saber. Mais do que um complemento técnico, as TK Labels são ferramentas anticoloniais que promovem a soberania de dados, permitindo que as comunidades indígenas e locais reafirmem o controle, o reconhecimento e a autoridade sobre seu próprio patrimônio cultural no ambiente digital (Montenegro, 2019; Anderson; Christen, 2017). Tais alternativas representam um esforço concreto de descolonização metodológica, abrindo caminhos para que os padrões tecnológicos não sufoquem, mas sim dialoguem com a diversidade de epistemologias locais.

3.4. RELAÇÃO ENTRE DADOS ABERTOS E ASSIMETRIA DE PODER

A relação entre a política de dados abertos e as assimetrias de poder é profundamente complexa e marcada por um paradoxo central. Por um lado, a abertura de dados carrega a promessa de democratizar o acesso ao conhecimento, aumentar a transparência governamental e reduzir desigualdades ao nivelar o campo da informação. Por outro, em cenários sociais e econômicos já desiguais, a simples disponibilização de dados pode ser cooptada por atores mais poderosos, que os utilizam de formas que não só refletem, mas também amplificam as disparidades existentes (Abebe *et al.*, 2021).

O conceito de “abertura”, portanto, não pode ser compreendido como um ato neutro. Ele está sempre inserido em contextos específicos de poder, e seus resultados dependem diretamente da distribuição de recursos na sociedade. Quando dados são abertos em um ambiente com profundas disparidades técnicas, financeiras e políticas, os benefícios tendem a ser desproporcionalmente capturados por aqueles que já são privilegiados. São as grandes corporações, as instituições de pesquisa de elite e os atores com capital e infraestrutura que possuem a capacidade de processar, analisar e transformar esses dados brutos em valor estratégico (Abebe *et al.*, 2021).

Essa dinâmica pode facilitar novas e sutis formas de exploração. Políticas que obrigam instituições públicas a liberar seus dados podem, na prática, fornecer matéria-prima gratuita para o setor privado. Empresas com recursos computacionais e analíticos superiores podem usar dados públicos sobre transporte, saúde ou educação para desenvolver produtos comerciais lucrativos, sem oferecer um retorno justo ou benefício direto às comunidades que, em primeiro lugar, geraram esses dados, muitas vezes através de serviços financiados com dinheiro público. Essa extração de valor é particularmente problemática quando os dados envolvem populações vulneráveis, transformando um bem público em um ativo privado.

Mesmo quando o acesso aos dados é genuinamente universal, as capacidades desiguais para seu uso efetivo criam uma lacuna de benefícios. Este cenário configura o chamado “paradoxo da abertura”: quanto mais dados são disponibilizados, maior pode se tornar a vantagem competitiva daqueles que já possuem os recursos para aproveitá-los. A abertura, em vez de nivelar, pode acelerar a concentração de poder em quem já está na frente. Essa lacuna, conhecida como “data divide” (o abismo de dados), não se refere apenas à diferença no acesso, mas, mais importante, na capacidade de uso, interpretação e controle dos dados, refletindo e reforçando as assimetrias socioeconômicas e geopolíticas já existentes. Afinal, os dados hoje configuram poder: ter acesso e capacidade de análise posiciona atores estratégicos para dominar decisões econômicas, científicas e de governança (KITCHIN, 2014).

Fica claro, portanto, que a simples abertura de dados, sem um arcabouço regulatório e infraestrutural adequado, não garante o empoderamento coletivo. Para que a promessa democrática dos dados abertos se concretize, são necessárias ações complementares e intencionais. Isso inclui a criação de portais de dados verdadeiramente acessíveis, investimentos maciços em literacia digital para capacitar cidadãos e organizações locais, e a realização de consultas públicas para alinhar as políticas de dados com as necessidades comunitárias. Além disso, é essencial desenvolver modelos de governança que combatam a captura privada de valor gerado por bens públicos. Isso pode envolver regulações que incentivem ecossistemas de dados mais justos, que recompensem a criação de conhecimento compartilhado e que garantam que os benefícios gerados a partir de dados públicos sejam reinvestidos na sociedade que os produziu (MDPI Sustainability 2021).

3.5 ALGORITMOS, VIESES E REPRODUÇÃO DE DESIGUALDADES

A infraestrutura digital que hoje sustenta a ciência ergue-se sobre uma premissa fundamentalmente falha: a de que seus algoritmos são ferramentas neutras. Longe disso, eles operam como verdadeiros artefatos culturais, codificando e refletindo as práticas históricas e as relações de poder da sociedade que os criou (Pasquale 2016). O que antes era visto como uma promessa de democratização do conhecimento tornou-se, em muitos aspectos, um sofisticado mecanismo de reprodução de desigualdades. Essa dinâmica invisível, mas poderosa, determina quais vozes são amplificadas e quais são silenciadas, influenciando desde a visibilidade de uma pesquisa até o sucesso de uma carreira acadêmica inteira.

No coração desse sistema está a exacerbação algorítmica do “efeito Matthew”, um fenômeno que transforma uma observação sociológica numa regra de engenharia de *software* (Kacperski *et al.* 2023). Plataformas como Google Scholar ou Web of Science não apenas recomendam o que é relevante; elas ativamente constroem a relevância ao criar um ciclo de retroalimentação contínua. Artigos de pesquisadores em instituições de elite, já beneficiados por redes e recursos, recebem prioridade, acumulam citações e, por consequência, são ainda mais destacados. O algoritmo, em sua busca por eficiência, aprende a associar prestígio com qualidade, gerando uma espiral que concentra o capital científico e marginaliza contribuições inovadoras da periferia, que permanecem afogadas num oceano de informação.

Esta arquitetura de privilégios é reforçada por barreiras linguísticas e geográficas que funcionam como uma espécie de colonialismo digital. A predominância de dados em inglês no treinamento desses sistemas faz com que os algoritmos não se limitem a refletir, mas reforcem ativamente a hegemonia de uma única língua,

tratando outras formas de produção de conhecimento como secundárias (Noble 2018). Essa uniformização é agravada pelas “bolhas de filtro”, que nos isolam em ecossistemas intelectuais cada vez mais homogêneos (Pariser 2010). Para um cientista, isso representa uma ameaça direta ao próprio método científico, que depende da exposição a ideias contraditórias, da falsificabilidade e do diálogo entre diferentes escolas de pensamento. Ao nos alimentar com o que já conhecemos e com quem já concordamos, os algoritmos arriscam fossilizar o conhecimento, em vez de catalisar sua evolução.

O impacto torna-se ainda mais direto e perigoso quando esses sistemas são usados para tomar decisões de alto impacto, como a seleção de currículos ou a avaliação de projetos. Um algoritmo treinado com décadas de dados históricos sobre contratações em uma instituição aprenderá a replicar os vieses do passado, efetivamente “lavando” a discriminação de gênero, raça ou afiliação institucional sob um verniz de objetividade tecnológica (Pasquale 2016; Greenlining Institute 2021). O estudo “Gender Shades” é a prova cabal dessa falha, ao demonstrar com números incontestáveis como sistemas de reconhecimento facial de gigantes da tecnologia eram praticamente cegos para mulheres negras (Buolamwini; Gebru, 2018). Este trabalho não foi apenas um estudo técnico; foi uma denúncia que expôs como a negligência na diversidade dos dados de treinamento gera tecnologias que ativamente excluem e prejudicam.

Essa realidade nos confronta com o paradoxo da “caixa-preta”. Por um lado, a opacidade de muitos algoritmos complexos cria uma crise de responsabilidade: se nem mesmo seus criadores conseguem explicar por que uma decisão foi tomada, como pode um indivíduo contestá-la? (Pasquale 2016). Por outro lado, a análise do comportamento agregado desses mesmos sistemas oferece um espelho sem precedentes para a sociedade. Ao revelar padrões de discriminação em larga escala, os algoritmos podem expor vieses sistêmicos que antes eram facilmente negados ou permaneciam ocultos na subjetividade de decisões individuais (Morewedge, 2024).

O código falho, nesse sentido, torna-se um inesperado delator das nossas próprias falhas sociais.

Portanto, a solução não reside em simples ajustes técnicos, pois até as mais sofisticadas tentativas de mitigar vieses podem fracassar se ignorarem os contextos culturais e as dinâmicas de poder em que operam (Fiorotto, 2024; Friedler *et al.*, 2021). O caminho a seguir exige uma mudança de paradigma em direção a uma “justiça algorítmica”. Isso implica incorporar valores éticos desde a concepção dos sistemas, através de auditorias rigorosas, simulações de impacto e, essencialmente, o envolvimento direto das comunidades afetadas nos processos de design e regulação (Caruana; Microsoft; Brookings, 2018). Trata-se de abandonar a fé ingênua na neutralidade da tecnologia e assumir a responsabilidade madura de governar as ferramentas que hoje moldam o futuro do conhecimento humano.

4

**FRAUDES CIENTÍFICAS,
MANIPULAÇÃO DE DADOS
E PRÁTICAS ANTIÉTICAS**

A ciência do século XXI encontra-se em uma encruzilhada crítica que define tanto o futuro da produção de conhecimento quanto a própria natureza da confiança epistêmica nas sociedades contemporâneas. A integridade científica enfrenta uma crise sem precedentes na era digital, caracterizada por um aumento exponencial de fraudes sofisticadas, manipulação de dados e práticas antiéticas facilitadas por tecnologias emergentes. Esta crise não se limita a um crescimento quantitativo das irregularidades, mas representa uma transformação qualitativa dos mecanismos de fraude, exigindo respostas igualmente transformadoras da comunidade científica global.

Por um lado, as tecnologias digitais oferecem possibilidades sem precedentes para acelerar descobertas, democratizar o acesso ao conhecimento e facilitar colaborações globais que transcendem fronteiras geográficas, linguísticas e institucionais. Por outro, essas mesmas tecnologias criaram novas vulnerabilidades e oportunidades para fraudes sofisticadas que ameaçam os fundamentos da integridade científica e comprometem, ao mesmo tempo, a legitimidade social da ciência como fonte confiável de conhecimento sobre o mundo natural e social.

Como observa Gerhard (2024), “o número de retrações científicas aumentou sete vezes desde 2000, e a inteligência artificial está sendo cada vez mais utilizada tanto para cometer quanto para detectar má conduta científica”. Esta observação captura a essência paradoxal do momento atual: as mesmas tecnologias que prometem revolucionar a pesquisa científica também facilitam formas de fraude de sofisticação sem precedentes, criando uma corrida armamentista entre fraudadores e detectores que redefine continuamente os parâmetros da integridade científica.

Os dados estatísticos revelam a magnitude alarmante desta crise contemporânea. Em 2023, foram retratados mais de 9.723 artigos científicos, equivalente ao total de retrações dos três anos anteriores, indicando uma aceleração exponencial que desafia

explicações baseadas meramente em melhorias na detecção (The Conversation, 2025). Estimativas conservadoras sugerem que até 400.000 artigos fraudulentos infiltraram-se na literatura científica nas últimas duas décadas, enquanto apenas 7.275 retrações foram oficialmente registradas como relacionadas a essas organizações, revelando a magnitude oculta do problema (Zein, 2024).

Petrou (2024) demonstra que a taxa global de retrações saltou de 3,5 para 11,2 por 10 mil artigos entre 2014 e 2022, praticamente triplicando em menos de uma década. Esta tendência não pode ser explicada apenas por melhorias na detecção, sugerindo uma transformação qualitativa fundamental na natureza e escala da má conduta científica. Mais preocupante ainda, cerca de metade de todas as retrações recentes ocorreu em poucos periódicos afetados por operações de artigos falsos, indicando que grande parte da elevação nas retrações deve-se à detecção aprimorada de fraudes em massa, mais do que a uma mudança súbita na ética individual dos cientistas.

A análise da industrialização da fraude através de artigos fraudulentos, combinada com a sofisticação crescente de manipulações assistidas por inteligência artificial, constitui uma ameaça existencial aos fundamentos epistemológicos da validação científica contemporânea. Estimativas recentes indicam que aproximadamente 50.000 artigos foram produzidos com auxílio de inteligência artificial em 2023, muitos dos quais podem conter elementos fraudulentos ou enganosos (Folha de S.Paulo, 2024). O número de artigos científicos considerados falhos ou fraudulentos subiu 500% nos últimos dez anos, demonstrando uma escalada sem precedentes na crise de integridade (The Conversation, 2024).

Esta transformação não ocorre em um vácuo histórico ou cultural. A ciência moderna, com suas raízes no Iluminismo europeu dos séculos XVII e XVIII, desenvolveu-se paralelamente ao colonialismo, criando estruturas de poder e conhecimento que persistem até hoje e moldam profundamente os padrões contemporâneos de produção,

validação e disseminação do conhecimento científico. A atual crise de integridade científica deve ser compreendida não apenas como um problema técnico ou ético isolado, mas como parte de um sistema mais amplo de desigualdades globais no acesso, produção e validação do conhecimento científico que reproduz e amplifica assimetrias de poder estabelecidas durante períodos coloniais.

A crise transcende questões técnicas ou acadêmicas isoladas para representar uma ameaça fundamental à confiança pública na ciência como instituição social, com implicações profundas para a capacidade das sociedades democráticas de tomar decisões baseadas em evidências sobre questões críticas como saúde pública, mudanças climáticas e desenvolvimento tecnológico. A erosão da confiança científica, documentada através de pesquisas de opinião que mostram declínios significativos na credibilidade da ciência, não é meramente um problema de relações públicas, mas uma crise de legitimidade epistêmica que questiona os próprios fundamentos sobre os quais sociedades modernas baseiam decisões coletivas.

Pesquisas de opinião pública revelam um declínio preocupante na confiança científica: nos Estados Unidos, a confiança na ciência caiu dramaticamente de 87% em 1973 para 73% em 2018, com declínios particularmente acentuados entre grupos politicamente conservadores (Pew Research Center, 2024). Esta erosão da confiança não é limitada a grupos específicos, mas representa uma tendência mais ampla e sistêmica que transcende divisões políticas, educacionais e demográficas tradicionais.

A dimensão global desta crise exige uma análise que transcenda perspectivas nacionais ou disciplinares isoladas. A má conduta científica contemporânea opera através de redes transnacionais que exploram diferenças regulatórias, culturais e econômicas entre países, criando um mercado global de fraudes científicas que desafia abordagens de governança baseadas em jurisdições nacionais.

Esta globalização da má conduta científica está intrinsecamente conectada a processos mais amplos de colonialismo de dados e extração de conhecimento que reproduzem assimetrias de poder globais no domínio científico.

O presente capítulo oferece uma análise abrangente e crítica desta crise multifacetada, examinando suas raízes históricas e teóricas, suas manifestações contemporâneas, e suas implicações para o futuro da ciência como instituição social. Através de uma abordagem interdisciplinar que combina análise empírica, reflexão teórica e perspectiva crítica sobre estruturas de poder global, buscamos contribuir para uma compreensão mais nuançada dos desafios enfrentados pela integridade científica na era digital, bem como para o desenvolvimento de estratégias eficazes para enfrentar esses desafios de forma sistemática e equitativa.

A análise incorpora dimensões neurobiológicas e psicológicas da tomada de decisão ética, reconhecendo que pesquisadores enfrentam pressões sistêmicas que podem afetar sua capacidade de julgamento moral. Estudos recentes indicam que pesquisadores têm níveis de ansiedade e depressão seis vezes maiores que a população geral, criando condições que podem facilitar comportamentos antiéticos sob pressão (ESTADÃO, 2025). A neurobiologia da desonestidade revela como o estresse pode comprometer os mecanismos neurais responsáveis pelo julgamento ético, particularmente no córtex frontal, que é fundamental para a integração mental ética (Revista Interações, 2024).

Além disso, o capítulo examina os custos econômicos substanciais da má conduta científica, incluindo tanto os custos diretos de investigação e retratação quanto os custos indiretos decorrentes do desperdício de recursos, da perda de confiança pública e do impacto nos sistemas de inovação. No Brasil, por exemplo, pesquisadores perdem 35% de seu tempo com atividades burocráticas, representando um desperdício estimado de R\$ 9 bilhões anuais em recursos científicos (CONFIES, 2019).

O capítulo também aborda tecnologias emergentes tanto como facilitadoras quanto como soluções para a crise de integridade. Enquanto a inteligência artificial generativa permite criar falsificações quase indistinguíveis do material real, ameaçando a ciência (Olhar Digital, 2024), também oferece ferramentas sofisticadas para detecção de fraudes através de análise de padrões linguísticos, visuais e estatísticos que seriam impossíveis de identificar manualmente.

Finalmente, a análise considera as dimensões de governança global necessárias para enfrentar eficazmente esta crise, incluindo propostas de tratados internacionais, frameworks de cooperação científica equitativa, e modelos de soberania científica que possam proteger a integridade da pesquisa enquanto promovem colaboração internacional responsável. O objetivo não é apenas prevenir má conduta, mas criar um sistema científico mais transparente, equitativo e sustentável que possa servir adequadamente às necessidades da humanidade no século XXI.

4.1. FUNDAMENTOS HISTÓRICOS E TEÓRICOS DA INTEGRIDADE CIENTÍFICA

4.1.1 EVOLUÇÃO HISTÓRICA DOS CONCEITOS DE INTEGRIDADE CIENTÍFICA

A noção contemporânea de integridade científica, longe de ser um conceito estático, é o resultado de uma longa e conturbada evolução histórica, moldada tanto por ideais filosóficos quanto por crises e escândalos que abalaram a confiança na ciência. Suas raízes podem ser traçadas até o século XIX, com a figura surpreendentemente presciente de Charles Babbage. Em sua obra de 1830, *Reflections on the Decline of Science in England*, Babbage não se

limitou a antecipar o debate moderno — ele estabeleceu uma tipologia de má conduta que ressoa até hoje. Identificou o *forging* (a fabricação completa de dados), o *trimming* (a falsificação seletiva pela eliminação de dados inconvenientes ou extremos) e o *cooking* (a manipulação de resultados para “embelezar” e forçar o encaixe em uma conclusão desejada) (Babbage, 1830).

A genialidade da análise de Babbage, no entanto, não residiu apenas em classificar fraudes, mas em diagnosticar sua origem. Ele argumentou que a má conduta não emanava primariamente de falhas morais individuais, mas era uma consequência direta das pressões institucionais e culturais da época, que já valorizavam a publicação rápida e o prestígio em detrimento da busca paciente e honesta pela verdade. Sua crítica representou um marco fundamental, sugerindo que a ética da pesquisa não poderia ser dissociada da estrutura de incentivos na qual os cientistas operam, uma reflexão que se tornaria central mais de um século depois.

Avançando para o século XX, a discussão sobre a ética na ciência foi profundamente marcada por dois movimentos paralelos. De um lado, o sociólogo Robert Merton, em 1973, formulou o que seria o *ethos* ideal da ciência, um conjunto de quatro normas que deveriam guiar a conduta científica: o comunalismo, que prega o compartilhamento aberto do conhecimento como um bem comum; o universalismo, que exige que as alegações científicas sejam julgadas por critérios impessoais, independentemente da origem do pesquisador; o desinteresse, que postula que a busca pelo conhecimento deve se sobrepor a interesses pessoais ou financeiros; e o ceticismo organizado, que impõe a submissão de todas as ideias a um escrutínio crítico e rigoroso pela comunidade (Merton, 1973). As normas de Merton se tornaram a base filosófica para inúmeros códigos de conduta.

Do outro lado, a dura realidade expunha a fragilidade dessa autorregulação idealizada. Escândalos de grande repercussão,

como o infame estudo de Tuskegee (1932-1972), no qual médicos do governo americano deliberadamente negaram tratamento para sífilis a centenas de homens negros para estudar a progressão natural da doença, chocaram a opinião pública. Tais violações éticas brutais demonstraram que a comunidade científica não era imune a preconceitos e a abusos de poder, forçando a criação de estruturas formais de governança, como os comitês de ética em pesquisa e regulamentações federais, para proteger os participantes de estudos.

Nas últimas décadas, as tensões sobre o *ethos* científico se intensificaram com a ascensão do que se chama de ciência “pós-acadêmica”. A crescente intersecção com interesses comerciais, através de parcerias público-privadas, o foco em patentes e a proteção da propriedade intelectual entraram em choque direto com a norma mertoniana do comunismo. O conhecimento, antes visto como um bem a ser compartilhado, passou a ser cada vez mais tratado como um ativo a ser protegido e explorado comercialmente. Simultaneamente, a pressão por produtividade acadêmica, imortalizada no lema “publique ou pereça” (*publish or perish*), popularizado após a Segunda Guerra, criou um sistema global de incentivos que recompensa a quantidade de publicações em detrimento da qualidade e do rigor, fomentando um ambiente propício para práticas questionáveis.

É nesse contexto que surge a mais recente e sofisticada compreensão do problema, encapsulada na teoria da «seleção natural da má ciência» de Smaldino e McElreath (2016). Eles argumentam que, dadas as atuais pressões institucionais, os laboratórios e pesquisadores que utilizam métodos menos rigorosos, mas que produzem resultados mais “publicáveis” (como resultados positivos ou estatisticamente significativos, mesmo que frágeis), tendem a ser mais bem-sucedidos em obter financiamento e posições. Com o tempo, essas práticas metodológicas mais frágeis se proliferam, não necessariamente por uma intenção consciente de enganar, mas porque o próprio sistema as seleciona e recompensa. A má ciência, portanto, evolui e se espalha como um traço adaptativo em um

ambiente acadêmico hostil à pesquisa lenta, cuidadosa e, muitas vezes, inconclusiva. A análise completa o ciclo iniciado por Babbage, confirmando que a integridade científica é, em grande medida, uma propriedade emergente do sistema em que a ciência é praticada.

4.1.2 MARCOS NORMATIVOS CONTEMPORÂNEOS E CODIFICAÇÃO INTERNACIONAL

Em um esforço consciente para combater as pressões sistêmicas e os escândalos de má conduta, a comunidade científica global mobilizou-se para criar e consolidar um conjunto de princípios e códigos de conduta. Esses documentos não são meras listas de regras, mas sim marcos normativos que refletem uma compreensão cada vez mais sofisticada dos desafios éticos contemporâneos, representando um movimento deliberado da diagnose do problema para a construção ativa de uma cultura de pesquisa mais saudável e confiável.

A Declaração de Singapura sobre Integridade em Pesquisa (Singapore Statement, 2010) representa o primeiro grande marco desse esforço globalizante. Nascida da 2ª Conferência Mundial sobre Integridade em Pesquisa, com a colaboração de profissionais de 51 países, sua ambição foi estabelecer uma base ética universal. A declaração se assenta em quatro princípios fundamentais: honestidade em todos os aspectos da pesquisa; responsabilidade pela própria conduta; cortesia profissional e justiça nas interações; e boa gestão (*good stewardship*) dos recursos e do conhecimento. O documento detalha 14 responsabilidades específicas, que vão desde a condução de pesquisa com métodos rigorosos e análise honesta, até a comunicação pública responsável e a essencial divulgação de conflitos de interesse, servindo como uma inspiração fundamental para a criação de políticas institucionais e nacionais (World Conference on Research Integrity, 2010).

Na Europa, a Carta Europeia para Pesquisadores e, especialmente, os Códigos Europeus de Conduta para Integridade em Pesquisa (All European Academies - ALLEA), em suas sucessivas revisões, avançaram para além dos princípios gerais, integrando de forma explícita as práticas da Ciência Aberta como um pilar da própria integridade. O Código Europeu é considerado um dos mais avançados nesse aspecto, pois não trata a gestão de dados (seguindo os princípios FAIR - Findable, Accessible, Interoperable, Reusable) e a publicação em acesso aberto como meras recomendações, mas como componentes essenciais da conduta responsável (Vicente-Saez; Martínez-Fuentes, 2018; European Code of Conduct 2017/2023). Essa abordagem representa uma evolução significativa: a integridade deixa de ser apenas sobre “não fazer o mal” (evitar fraude) para se tornar sobre “fazer o bem ativamente” (promover a transparência, a colaboração e a reprodutibilidade).

Contudo, enquanto as declarações de Singapura e da Europa focavam na conduta do pesquisador, os Princípios de Hong Kong (Hong Kong Principles, 2020) deram um passo adiante, direcionando sua crítica e propostas para a raiz do problema: os sistemas de avaliação acadêmica. Lançados na 6ª Conferência Mundial sobre Integridade em Pesquisa, esses princípios argumentam que a cultura do “publique ou pereça” só será superada se as instituições mudarem o que elas medem e recompensam. Seus cinco pilares são um chamado à ação para os avaliadores: valorizar práticas de pesquisa responsáveis, não apenas os resultados; reconhecer a importância do relato completo e da replicação de estudos; promover ativamente a ciência aberta; valorizar uma gama diversa de produtos de pesquisa (como conjuntos de dados, softwares, patentes, políticas públicas); e, crucialmente, recompensar tarefas essenciais para a saúde do ecossistema científico, como a revisão por pares de qualidade e a mentoria de jovens pesquisadores (Moher *et al.*, 2020).

Em conjunto, esses marcos normativos revelam uma estratégia multifacetada e madura para fortalecer a integridade científica.

Eles estabelecem um padrão global de responsabilidade individual (Singapura), promovem uma cultura de transparência e colaboração proativa (Europa) e, de forma mais transformadora, buscam reformar os sistemas de incentivos institucionais que fomentam a má conduta (Hong Kong). Eles representam o reconhecimento coletivo de que a integridade não pode ser delegada apenas à moralidade do cientista individual, mas deve ser ativamente cultivada por meio de uma governança ética que abranja desde o laboratório até as mais altas esferas da avaliação acadêmica global.

4.1.3 DIMENSÕES NEUROBIOLÓGICAS E PSICOLÓGICAS DA ÉTICA CIENTÍFICA

Uma compreensão verdadeiramente contemporânea da integridade científica não pode mais se limitar a modelos puramente filosóficos ou regulatórios; ela deve, necessariamente, incorporar as descobertas, por vezes fascinantes e desconcertantes, da neurociência. Pesquisas emergentes no campo da neuroética revelam que a tomada de decisão moral está longe de ser um processo puramente racional e deliberado. Pelo contrário, ela emerge de uma complexa interação de circuitos neurais que são profundamente vulneráveis a fatores ambientais como estresse crônico, pressão por resultados e o próprio clima institucional em que um pesquisador está inserido (PubMed, 2024).

O epicentro neurobiológico do julgamento ético reside, em grande parte, no córtex pré-frontal. Esta região do cérebro, especialmente suas áreas ventromedial (CPFvm) e dorsolateral (CPFdl), funciona como um centro de integração essencial. O córtex pré-frontal ventromedial está mais associado ao processamento do valor emocional e social de uma decisão, conectando a escolha a sentimentos de empatia, culpa ou justiça. Já o córtex pré-frontal dorsolateral está mais envolvido no controle cognitivo, no planejamento de longo prazo e na aplicação de regras abstratas (Revista Bioética, 2024).

Uma decisão ética robusta depende, portanto, do diálogo equilibrado entre essas sub-regiões. No entanto, estudos de neuroimagem demonstram que, sob condições de estresse agudo ou exaustão cognitiva, a atividade nesses circuitos pode ser significativamente comprometida, levando o cérebro a adotar “atalhos” que priorizam a sobrevivência ou o ganho imediato em detrimento de considerações éticas mais complexas.

Esta compreensão neurobiológica oferece uma nova e poderosa lente para analisar a má conduta científica. O ambiente acadêmico moderno, com sua intensa pressão por publicações (“publique ou pereça”), competição acirrada por financiamento escasso e insegurança profissional, cria um cenário ideal para o estresse crônico. A alarmante constatação de que pesquisadores podem apresentar níveis de ansiedade e depressão até seis vezes maiores que os da população geral (UOL, 2023) não é apenas uma crise de saúde mental, mas também uma crise de integridade em potencial. O estresse crônico inunda o cérebro com hormônios como o cortisol, que podem prejudicar a função do córtex pré-frontal, diminuindo a capacidade de autorregulação e de pensamento a longo prazo, e criando assim um terreno neurobiológico fértil para lapsos éticos.

A neurobiologia da desonestidade sugere que o comportamento antiético raramente começa com uma grande fraude. Muitas vezes, ele se desenvolve através de uma “ladeira escorregadia” (*slippery slope*), onde pequenas transgressões, inicialmente justificadas para lidar com a pressão, vão gradualmente dessensibilizando a resposta emocional do cérebro à desonestidade, tornando transgressões maiores mais fáceis de cometer no futuro (UCDB, 2019). O ato de má conduta, visto por essa ótica, não é apenas uma falha de caráter, mas pode ser o ponto final de um processo de adaptação neural a um ambiente insustentável. Isso implica que estratégias de prevenção focadas exclusivamente em educação e punição são insuficientes, pois falham em abordar as causas estruturais que empurram os cérebros dos pesquisadores para um estado de

vulnerabilidade ética. A verdadeira promoção da integridade científica, portanto, deve incluir a criação de ambientes de pesquisa mais saudáveis, a reforma dos sistemas de avaliação e o fornecimento de suporte robusto à saúde mental como pilares centrais e não periféricos da governança científica.

4.1.4 TEORIAS CONTEMPORÂNEAS DE INTEGRIDADE CIENTÍFICA

Com certeza. Apresento aqui uma explanação mais aprofundada e com maior contribuição discursiva sobre as abordagens teóricas contemporâneas da integridade científica, mantendo as citações e aprofundando os conceitos.

A busca por uma compreensão robusta da integridade científica na contemporaneidade levou ao desenvolvimento de um arcabouço teórico multifacetado, que reconhece a insuficiência de uma única perspectiva. As teorias atuais evoluíram para integrar três abordagens principais, baseadas em regras, em virtudes e em valores que, juntas, oferecem um modelo mais completo e resiliente para guiar a conduta científica em um mundo cada vez mais complexo (Mastroleo; Bianchini, 2024).

A abordagem baseada em regras é a mais tradicional e institucionalizada. Ela constitui o alicerce da governança científica, focando na codificação de normas, proibições e procedimentos de *compliance* para prevenir as formas mais graves de má conduta, como fabricação, falsificação e plágio. Sua grande força reside na clareza e na aplicabilidade, fornecendo um guia explícito sobre o que é inaceitável e estabelecendo as consequências para as transgressões. No entanto, sua principal limitação emerge justamente de sua rigidez. As regras são eficazes para lidar com problemas conhecidos e bem definidos, mas frequentemente se mostram inadequadas para navegar em “zonas cinzentas”, dilemas éticos ambíguos ou desafios inéditos trazidos por novas tecnologias, onde não existe uma norma pré-escrita para consultar (Mastroleo; Bianchini, 2024).

Em contraste, a abordagem baseada em virtudes desloca o foco do “o que não se deve fazer” para “que tipo de pessoa se deve ser”. Com raízes profundas na tradição ética de Aristóteles, este modelo enfatiza o desenvolvimento do caráter moral e das competências éticas do próprio pesquisador. A integridade, aqui, não é apenas a adesão a um código, mas o cultivo de virtudes como a honestidade intelectual, a coragem moral para desafiar o status quo ou denunciar uma má prática, e a modéstia epistemológica — a humildade de reconhecer os limites do próprio conhecimento. Estudos recentes destacam um conjunto de virtudes centrais para a formação do cientista, incluindo a confiança, a prudência no julgamento e o cuidado com os participantes da pesquisa, com os dados e com o impacto social da ciência (Mastroleo; Bianchini, 2024). Essa abordagem aposta que um pesquisador virtuoso saberá como agir eticamente mesmo na ausência de uma regra específica.

A terceira via, a abordagem baseada em valores, reconhece que a ciência nunca é uma atividade neutra ou isolada da sociedade. Ela é inevitavelmente guiada por valores, e esta perspectiva defende que eles devem ser articulados e debatidos de forma aberta e transparente. Esses valores são tanto epistêmicos (internos à prática científica, como a busca pela verdade, objetividade, rigor e reprodutibilidade) quanto sociais (relacionados ao papel da ciência no mundo, como a busca por justiça social, equidade, sustentabilidade e responsabilidade para com as comunidades). Ao trazer essa discussão para o centro, a abordagem de valores força a comunidade científica a refletir sobre os propósitos e as consequências de seu trabalho, garantindo que a pesquisa esteja alinhada não apenas com o rigor metodológico, mas também com o bem comum (Mastroleo; Bianchini, 2024).

4.1.5 COLONIALISMO CIENTÍFICO E ESTRUTURAS DE PODER GLOBAL

Uma análise verdadeiramente crítica da integridade científica contemporânea é incompleta se não confrontar as profundas dimensões de poder e desigualdade que determinam quem produz, quem valida e quem dissemina o conhecimento em escala global. O conceito de colonialismo científico, como articulado por Couldry e Mejias (2019), é uma lente poderosa para essa análise. Ele descreve o processo contínuo de extração de conhecimentos, dados e recursos de países do Sul Global, que são canalizados para beneficiar instituições, corporações e pesquisadores do Norte Global. Longe de ser uma metáfora, trata-se de uma continuidade direta das assimetrias de poder estabelecidas durante o colonialismo histórico, adaptada para a era da informação.

Essa dinâmica colonial se manifesta de formas variadas, desde as mais explícitas até as mais sutis. A chamada “ciência para-quedas” (*parachute science*) é uma de suas faces mais conhecidas: pesquisadores de países desenvolvidos “pousam” em comunidades do Sul Global para coletar dados e amostras, e depois “decolam” para analisá-los e publicá-los em seus países de origem, com pouca ou nenhuma colaboração significativa com os cientistas locais. Essa prática não apenas se apropria de conhecimentos tradicionais e recursos biológicos sem o devido reconhecimento ou partilha de benefícios, mas muitas vezes impõe metodologias inadequadas que podem gerar danos sociais ou ambientais, deixando as comunidades com os riscos da pesquisa, mas sem participação nos seus ganhos (Neo-colonial science 2003–2020).

Uma evolução dessa lógica é o “colonialismo de dados” (*data colonialism*), um processo mais profundo e insidioso. Nesta nova fase, o recurso extraído é o próprio dado, seja ele genético, ambiental, social ou comportamental. Populações do Sul Global tornam-se

fontes de matéria-prima para uma economia digital cujas tecnologias de processamento, algoritmos e plataformas são controladas quase exclusivamente por corporações do Norte. Essa extração massiva ocorre frequentemente sem um consentimento verdadeiramente informado, sem mecanismos de partilha de benefícios e, crucialmente, sem qualquer controle local sobre como os dados serão usados, perpetuando um ciclo de dependência tecnológica e exploração econômica (Couldry; Mejias, 2019).

Essas práticas extrativas são sustentadas por estruturas de poder globais que dominam a arquitetura da ciência. O controle sobre a infraestrutura editorial é um pilar central: um pequeno número de editoras comerciais do Norte detém o monopólio das revistas de maior prestígio, impondo *paywalls* que criam barreiras financeiras intransponíveis para muitos pesquisadores e instituições com menos recursos. Somado a isso, os sistemas de avaliação acadêmica (rankings, métricas de impacto) privilegiam esmagadoramente publicações e agendas de pesquisa centradas no Norte, estabelecendo uma hierarquia de prestígio que sistematicamente marginaliza vozes, temas e epistemologias periféricas. Isso é uma manifestação direta da “colonialidade do saber” (Quijano, 1992), onde um padrão específico de fazer ciência é imposto como universal, desvalorizando todo o resto.

Essas desigualdades estruturais geram vulnerabilidades específicas e perversas no Sul Global. A imensa pressão para publicar em periódicos “internacionais” de alto impacto, imposta por esses mesmos sistemas de avaliação, torna os pesquisadores locais alvos preferenciais para operações fraudulentas como as fábricas de artigos. Essas “empresas” oferecem atalhos para a publicação em revistas prestigiosas, uma tentação que se torna muito mais forte para quem carece de recursos, mentoria e reconhecimento institucional (NCBI-paper mills, 2023). Ao mesmo tempo, a dependência da literatura científica produzida externamente expõe os países em desenvolvimento aos impactos de fraudes científicas globais, que

podem influenciar diretamente políticas públicas em áreas críticas como saúde e meio ambiente, muitas vezes sem que haja capacidade local para verificar ou contestar esses resultados.

Portanto, para enfrentar essas questões de forma séria, é imperativo compreender que a integridade científica não pode ser dissociada da justiça epistêmica. Não se trata apenas de uma questão de conduta individual, mas de uma luta por soberania científica e equidade global no acesso e na produção do conhecimento. Estratégias efetivas devem ir além de códigos de conduta e focar na redistribuição do poder científico, na inclusão significativa e na liderança de pesquisadores locais, e no reconhecimento ativo da pluralidade de epistemologias do Sul Global, desafiando a própria noção de que a ciência de qualidade só pode emanar do centro (Coloniality of knowledge, 1992; Data colonialismo, 2019).

4.2. A CRISE CONTEMPORÂNEA: DIMENSÕES E MAGNITUDE DO PROBLEMA

4.2.1. ANÁLISE ESTATÍSTICA DA ESCALADA DE RETRAÇÕES CIENTÍFICAS

A magnitude da crise contemporânea de integridade científica torna-se evidente através de uma análise dos dados estatísticos sobre retrações científicas, que revelam tanto um aumento quantitativo dramático quanto transformações qualitativas fundamentais na natureza e escala da má conduta científica. Os dados de 2023 estabeleceram um recorde histórico alarmante: mais de 10.000 artigos científicos foram retratados globalmente, superando o total de retrações dos três anos anteriores combinados (The Conversation, 2025).

Este marco não é uma simples anomalia estatística, mas um indicador de transformações estruturais profundas no ecossistema científico global.

A análise longitudinal revela uma aceleração exponencial preocupante. Enquanto em 2000 foram registradas aproximadamente 40 retrações globalmente, este número saltou para mais de 4.000 em 2020, e posteriormente explodiu para mais de 10.000 em 2023 (Jornal da Unesp, 2024). Esta progressão não linear sugere que estamos testemunhando não só melhorias na detecção de má conduta, como também uma transformação qualitativa fundamental nos mecanismos e na escala da fraude científica.

Petrou (2024) fornece uma análise particularmente reveladora ao demonstrar que a taxa global de retrações saltou de 3,5 para 11,2 por 10.000 artigos entre 2014 e 2022, representando um aumento de mais de 300% em menos de uma década. Esta escalada não pode ser explicada exclusivamente por melhorias na detecção ou mudanças nos padrões editoriais, sugerindo um aumento real e substancial na prevalência de má conduta científica.

Tabela 3 - Evolução das Retrações Científicas Globais (2000-2023)

Ano	Número de Retrações	Taxa por 10.000 Artigos	Principais Causas
2000	~40	0.8	Plágio, fabricação isolada
2005	~150	1.2	Falsificação de dados, duplicação
2010	~400	2.1	Artigos fraudulentos emergentes, manipulação de imagens
2015	~800	3.8	Operações industriais de fraude
2020	~4.000	8.5	IA generativa, artigos fraudulentos sofisticados
2023	~10.000+	15.2	Fraudes assistidas por IA, operações transnacionais

Fonte: Compilação baseada em Petrou (2024), The Conversation (2025), Retraction Watch (2024).

A distribuição geográfica das retrações revela padrões complexos que refletem tanto vulnerabilidades sistêmicas quanto capacidades de detecção diferenciadas. Países com sistemas científicos em rápida expansão, como China, Índia, Irã e Brasil, aparecem desproporcionalmente representados nas estatísticas de retração, mas esta sobre-representação deve ser interpretada cuidadosamente, considerando fatores como pressões por produtividade, recursos limitados para treinamento ético, e possível viés de detecção que pode focar mais intensamente em periódicos e autores de certas regiões (Portal Tela, 2025).

4.2.2 PAPER MILLS: A INDUSTRIALIZAÇÃO SISTEMÁTICA DA FRAUDE

A emergência das “*paper mills*” (fábricas de artigos) representa uma transformação paradigmática e profundamente alarmante na ecologia da má conduta científica. O problema evoluiu de atos isolados de fraude, cometidos por pesquisadores individuais, para operações comerciais sistemáticas e transnacionais que produzem artigos científicos fraudulentos em uma escala industrial (Else & Van Noorden, 2021; Abalkina & Bishop, 2023). Essas entidades não são meros serviços de “ajuda” acadêmica; são empreendimentos criminosos sofisticados que exploram as vulnerabilidades estruturais do sistema de publicação científica para gerar lucro, transformando a natureza da fraude de uma atividade individual para um empreendimento sistemático (Zein, 2024). A magnitude do problema é assustadora: estimativas sugerem que até 400.000 artigos publicados entre 2000 e 2022 podem ter se originado dessas operações, com apenas uma fração identificada, revelando um “iceberg” de poluição científica (Van Noorden, 2023).

O *modus operandi* de uma paper mill assemelha-se a uma linha de montagem. Elas empregam equipes com uma divisão de trabalho especializada — redatores para o texto, estatísticos para forjar dados,

designers para manipular imagens e “especialistas” em submissão que conhecem as brechas dos sistemas editoriais. Operando em um mercado globalizado, oferecem um cardápio de serviços que vai da venda de posições de autoria em artigos pré-fabricados à produção de manuscritos sob encomenda, com preços que variam de centenas a milhares de dólares (Abalkina & Bishop, 2025). Embora as operações se concentrem em países com custos mais baixos, sua clientela é global e inclui pesquisadores de nações desenvolvidas, indicando que a pressão por publicação é um fenômeno mundial (Zein, 2024). Os modelos mais ousados envolvem a criação de identidades acadêmicas fictícias, a corrupção de editores e a infiltração em conselhos editoriais para garantir a aprovação dos artigos.

Existe uma corrida armamentista contínua entre os fraudadores e os sistemas de detecção. As *paper mills* demonstram notável capacidade de adaptação, incorporando tecnologias avançadas para evadir o escrutínio. Se as primeiras fraudes eram marcadas por “frases torturadas” (traduções automáticas bizarras), hoje, modelos de linguagem avançados (IA) geram textos fluentes e tecnicamente corretos. A fraude visual evoluiu da manipulação de imagens para a criação de evidências totalmente sintéticas (micrografias, western blots) por meio de Inteligência Artificial, como as GANs. Até mesmo os padrões de coautoria, antes um sinal de alerta, são agora fabricados para parecerem verossímeis (Zein, 2024; Cabanac & Labbé, 2021).

As consequências dessa indústria são devastadoras. A crise da editora Hindawi, adquirida pela Wiley, serve como um estudo de caso. A infiltração sistemática de artigos de *paper mills* em suas edições especiais levou a um escândalo massivo, com mais de 11.300 retratações, o encerramento de 19 periódicos e perdas estimadas em até 40 milhões de dólares (Wiley crisis 2024). Este episódio expôs como o modelo de Acesso Aberto baseado em taxas de processamento (APCs) cria um incentivo perverso para editores aceitarem um alto volume de artigos, relaxando padrões de revisão para maximizar o lucro — uma falha explorada com maestria pelas *paper mills* (Hanson *et al.*, 2023).

O impacto sistêmico, com um custo global estimado em centenas de milhões de dólares, vai além das perdas financeiras (The Guardian, 2025). Inclui o desperdício de milhões de horas de trabalho voluntário de revisores, a corrupção de meta-análises com dados falsos — criando uma “poluição epistêmica” — e a alocação indevida de fundos de pesquisa. Cada artigo fraudulento gera custos sistêmicos médios de \$15.000 a \$25.000 (Economic Analysis of Research Integrity, 2024). Acima de tudo, essa indústria corrói profundamente a confiança pública na ciência.

Em resposta, uma contraofensiva começa a se formar. Ferramentas de análise forense de imagens e texto ajudam a identificar e agrupar manuscritos suspeitos (Cardenuto *et al.*, 2024). No entanto, a tecnologia por si só é insuficiente. Organizações como o COPE e a STM enfatizam que o combate eficaz exige uma resposta coletiva e coordenada, envolvendo auditorias rigorosas, formação de editores e, fundamentalmente, a colaboração entre editoras, instituições, agências de fomento e pesquisadores para reformar os incentivos sistêmicos que permitiram que esta indústria da desonestidade florescesse (COPE & STM, 2022; Open Science Unimi, 2024).

4.2.3 IMPACTOS DA PANDEMIA COVID-19 NA INTEGRIDADE CIENTÍFICA

A pandemia de COVID-19 funcionou como um “teste de estresse” em escala global para todo o ecossistema científico, submetendo suas estruturas, processos e atores a uma pressão sem precedentes. A urgência por respostas imediatas, combinada com uma intensa politização da ciência e enormes interesses econômicos, criou um ambiente volátil onde as boas práticas de pesquisa foram frequentemente tensionadas, e por vezes rompidas, em nome da velocidade e da visibilidade. Este período não apenas exacerbou vulnerabilidades já conhecidas no sistema, mas também revelou novas e complexas dimensões da crise de integridade contemporânea.

Um dos sintomas mais evidentes dessa crise foi o aumento dramático no número de retratações de artigos científicos. Dados do portal *Retraction Watch* indicam um aumento significativo, chegando a 40%, nas retratações de pesquisas relacionadas à COVID-19 em comparação com as taxas históricas para outras áreas da medicina (*Retraction Watch*, 2024). Este número não reflete apenas um maior escrutínio público sobre um tema de altíssimo interesse, mas é também um indicativo de problemas sistêmicos mais profundos. A necessidade de acelerar a disseminação do conhecimento levou a um relaxamento generalizado dos processos de revisão por pares, que se tornaram mais rápidos, mas frequentemente mais superficiais. A corrida por financiamento, prestígio e influência política intensificou os incentivos para a publicação de resultados, mesmo que preliminares ou frágeis, criando um terreno fértil para a má conduta.

Os casos de má conduta durante a pandemia foram variados, incluindo a fabricação de dados sobre a eficácia de tratamentos, a manipulação de estatísticas para fins políticos, o plágio em meio a uma avalanche de publicações e o uso inadequado de *preprints* (artigos ainda não revisados por pares) para influenciar políticas públicas e a opinião popular. O caso do estudo sobre a hidroxiclороquina, publicado em uma das revistas mais prestigiosas do mundo, a *The Lancet*, e posteriormente retratado, tornou-se o exemplo emblemático dessa crise. A publicação, baseada em dados de uma empresa obscura (*Surgisphere*), foi rapidamente contestada pela comunidade científica global devido a inconsistências gritantes. Sua retratação humilhante expôs como a busca desesperada por uma “solução mágica”, somada à pressão por uma publicação de alto impacto, pôde levar até mesmo os mais rigorosos “guardiões” da ciência a falharem em seus processos de validação (*The Lancet*, 2020).

A pandemia também funcionou como um catalisador para a adoção da inteligência artificial (IA) na pesquisa, revelando seu potencial como uma faca de dois gumes. Por um lado, a IA foi uma ferramenta poderosa e legítima, utilizada de forma produtiva para

acelerar a descoberta de medicamentos, analisar complexos dados epidemiológicos e modelar a estrutura de proteínas virais, como no caso do projeto AlphaFold da DeepMind. Por outro lado, a mesma tecnologia foi cooptada para fins fraudulentos. Ferramentas de IA generativa provavelmente facilitaram a produção em massa de artigos científicos de baixa qualidade ou totalmente fabricados sobre a COVID-19, explorando o imenso interesse público e a urgência por informações. Essa prática contribuiu para a chamada “infodemia” — uma inundação de informações, muitas delas falsas ou enganosas, que poluiu o ambiente informacional e dificultou a distinção entre ciência robusta e pseudociência oportunista.

Em retrospecto, a experiência da pandemia oferece lições cruciais. Ela demonstrou a incrível capacidade de mobilização e colaboração da ciência global, mas também expôs suas fragilidades de forma inequívoca. A crise sanitária revelou a necessidade urgente de fortalecer os mecanismos de controle de qualidade, de desenvolver modelos de revisão por pares mais transparentes e resilientes, e de aprimorar a comunicação da ciência com o público, especialmente no que diz respeito às incertezas inerentes ao processo científico. Garantir que a ciência possa responder a futuras crises com velocidade e, ao mesmo tempo, com integridade, é um dos maiores desafios legados por este período histórico.

4.2.4 DIMENSÕES ECONÔMICAS E DESPERDÍCIO DE RECURSOS SISTÊMICOS

A crise de integridade científica, quando analisada sob uma ótica econômica, revela um cenário de desperdício sistêmico em uma escala massiva. Os custos transcendem em muito os lucros obtidos por operações fraudulentas, como as *paper mills*, e se manifestam como uma profunda hemorragia de recursos públicos e privados, uma distorção de incentivos que mina a produtividade e

uma erosão de confiança com consequências econômicas de longo prazo. Trata-se de uma questão não apenas ética, mas de eficiência e de responsabilidade na gestão de investimentos cruciais para o desenvolvimento social e tecnológico.

O desperdício de recursos financeiros é evidente e multifacetado. Um exemplo contundente vem do Brasil, onde a burocracia excessiva consome uma porção alarmante do tempo dos pesquisadores. Estima-se que cientistas brasileiros percam cerca de 35% de seu tempo de trabalho com atividades burocráticas que não agregam valor à pesquisa, o que representa um desperdício anual projetado em R\$ 9 bilhões em salários e bolsas pagos por tempo que poderia ser dedicado à produção de conhecimento (CONFIES, 2019). A este fardo somam-se os custos diretos da má conduta: universidades públicas e agências de fomento acabam por direcionar recursos substanciais para recompensar a fraude, seja através de bônus por produtividade baseados em métricas infladas artificialmente por artigos falsos, seja pelos altos custos administrativos para processar alegações e conduzir investigações de má conduta. Cada real gasto nesses processos é um recurso que deixa de ser investido em pesquisa legítima, na melhoria da infraestrutura ou no apoio a estudantes e cientistas que trabalham com integridade.

Igualmente significativo é o desperdício de capital humano, o ativo mais valioso da ciência. Revisores por pares, que em sua maioria trabalham de forma voluntária, dedicam horas preciosas de sua expertise para avaliar manuscritos fraudulentos que nunca deveriam ter chegado a seus e-mails, um tempo que poderia ser usado para suas próprias pesquisas ou para a revisão de trabalhos legítimos. Editores, por sua vez, gastam seus recursos limitados gerenciando o fluxo de submissões falsas, investigando alegações complexas e lidando com o processo custoso e demorado das retrações. Talvez o custo mais devastador, no entanto, seja o efeito de “cascata de erro”: quando pesquisadores honestos, de boa-fé, baseiam seus próprios estudos em resultados fabricados, eles iniciam uma linha

de pesquisa fadada ao fracasso. Isso pode levar ao desperdício de anos de trabalho e de milhões em financiamento, não em um único laboratório, mas em dezenas de outros ao redor do mundo, antes que o erro original seja finalmente identificado e corrigido.

O impacto econômico dessas práticas é desproporcionalmente devastador em países em desenvolvimento, onde os recursos para a ciência já são extremamente escassos. Nesses contextos, cada dólar ou real desperdiçado tem um custo de oportunidade muito maior. Fundos limitados que poderiam financiar a pesquisa local e fortalecer a capacidade nacional são, em vez disso, drenados para o pagamento de taxas de publicação (APCs) a editoras internacionais por artigos fraudulentos que não geram qualquer avanço real. Ao mesmo tempo, os pesquisadores legítimos, que se recusam a participar desses esquemas, encontram-se em desvantagem competitiva, lutando ainda mais para obter financiamento, reconhecimento e visibilidade internacional. A fraude, portanto, não se limita a desviar recursos, mas também aprofunda o abismo científico entre o Norte e o Sul Global.

4.3 INTELIGÊNCIA ARTIFICIAL: A TRANSFORMAÇÃO FUNDAMENTAL DA MÁ CONDUTA CIENTÍFICA

A ascensão da inteligência artificial (IA) no ecossistema científico marca o início de uma nova era, tanto para a descoberta quanto para a própria natureza da integridade e da má conduta. Diferente dos desafios anteriores, a IA não é simplesmente mais uma ferramenta passível de uso indevido; trata-se de uma força transformadora que redefine as capacidades tanto da pesquisa honesta quanto da fraude sistemática, alterando fundamentalmente o cenário ético da ciência.

Esta seção explora como a integração da IA na pesquisa introduz um paradoxo fundamental: ao mesmo tempo que oferece promessas de avanços revolucionários, ela também fornece um arsenal de ferramentas para a fabricação de dados e textos fraudulentos com um nível de sofisticação e escala nunca antes visto. Analisar este cenário exige ir além de uma visão puramente instrumental da tecnologia para compreender como ela desafia os próprios pilares da validação científica, da confiança e da epistemologia em um mundo digitalmente mediado.

4.3.1 A DUPLA FACE DA REVOLUÇÃO TECNOLÓGICA: PROMESSA E AMEAÇA SIMULTÂNEAS

De um lado, a face luminosa da IA revela um potencial transformador que já é indispensável em muitas áreas. No processamento de *big data*, algoritmos de aprendizado de máquina conseguem identificar padrões sutis em conjuntos de dados genômicos, climatológicos ou astronômicos que seriam humanamente impossíveis de detectar. O desenvolvimento de sistemas como AlphaFold 3, da DeepMind e Isomorphic Labs, representa um avanço monumental. Lançado em maio de 2024, esse modelo de IA pode prever estruturas de proteínas, DNA, RNA, ligantes e íons com precisão significativamente maior do que seus predecessores — incluindo cerca de 50 % mais acerto em interações moleculares complexas (Jumper et al., 2024). Essas previsões têm acelerado o design de medicamentos, reduzindo etapas que antes demoravam anos para apenas meses.

No campo da automatização científica, o Robot Scientist “Adam”, desenvolvido por Ross D. King e equipe na década de 2000, foi o primeiro sistema capaz de conduzir todo o método científico — formulando hipóteses, desenhando experimentos, executando análises em laboratório robótico e interpretando resultados — sem intervenção humana (King et al., 2009). Estudos posteriores como

o “Gower *et al.* (2024)” refletem a evolução para sistemas de nova geração como o “Genesis”, capazes de realizar milhares de ciclos de experimentação diários.

Essas tecnologias não apenas representam progresso incremental; elas expandem os limites do que é possível em ciência: do domínio da estrutura molecular à automação completa do processo experimental, proporcionando precisão maior, escalabilidade contínua e reproducibilidade aprimorada.

No campo da automação de laboratório, plataformas como labs autônomos guiados por IA transformam radicalmente o ritmo e a qualidade da experiência científica. Estudos recentes mostram que laboratórios robóticos podem funcionar de forma contínua, eliminando fadiga humana e oferecendo maior precisão e segurança em testes repetitivos (UNC Chapel Hill, Science Robotics 2024). Plataformas sofisticadas de orquestração como Artificial, integrando IA, sensores e pneumáticas de robôs, demonstram replicabilidade superior e agilidade em descobertas farmacológicas (Fehlis *et al.*, 2025).

Projetos como o Robot Scientist Adam/Eve, concebidos por Ross King e colaboradores, representam a automação completa do método científico: formula hipóteses, desenha experimentos, executa testes e interpreta alegações com mínima intervenção humana, tornando possível registrar todo o processo de forma auditável para replicação futura (King 2009; Gower *et al.*, 2024)

No entanto, o reverso dessa capacidade é sombrio e representa uma ameaça existencial à integridade da pesquisa. A mesma IA que analisa dados reais pode ser usada para gerar dados sintéticos que são estatisticamente indistinguíveis de observações empíricas. Isso transcende a fraude tradicional. Um cientista mal-intencionado não precisa mais “inventar” números em uma planilha; ele pode instruir um algoritmo a gerar um conjunto de dados completo que não

apenas pareça realista, mas que seja otimizado para apresentar os resultados desejados — por exemplo, com um valor específico que garanta a “significância” estatística ou com correlações que confirmem perfeitamente uma hipótese (Science, 2024). A fraude deixa de ser um ato de invenção para se tornar um ato de engenharia reversa da prova científica, criando uma falsificação quase perfeita.

A geração de texto por IA representa outro campo de batalha onde as fronteiras entre o uso legítimo e a má conduta se tornam perigosamente difusas. Grandes Modelos de Linguagem (LLMs) são capazes de redigir artigos acadêmicos inteiros com uma fluidez e uma coerência impressionantes. Eles podem gerar introduções, descrever metodologias detalhadas, apresentar análises sofisticadas e formular conclusões plausíveis, tudo isso sem qualquer conexão com um trabalho empírico real. Como observa o especialista em detecção de fraudes Guillaume Cabanac, esses sistemas são “muito mais inteligentes e geram textos realmente bons”, tornando obsoletos os detectores de plágio tradicionais, que foram projetados para identificar cópias diretas ou paráfrases simples (The Scientist, 2024). O desafio agora é muito maior: não se trata de detectar cópias, mas de discernir entre um texto que descreve uma realidade e um que apenas simula a aparência de uma descrição da realidade.

Em última análise, essa dupla face da IA nos coloca diante de uma crise epistemológica. Se uma máquina pode gerar dados e textos que são indistinguíveis daqueles produzidos por uma pesquisa honesta e baseada em evidências, como podemos continuar a confiar no que lemos? A confiança no *produto* final — o artigo publicado — torna-se cada vez mais precária. A resposta a esse desafio existencial exigirá uma mudança de foco: da validação do produto para a verificação do *processo*. A procedência dos dados, a transparência dos métodos e a capacidade de auditoria dos fluxos de trabalho se tornarão ainda mais cruciais. A revolução da IA nos força a reconstruir nossas defesas e a questionar os próprios fundamentos de como validamos o conhecimento em um mundo onde a simulação da verdade pode se tornar indistinguível da própria verdade.

4.3.2 TIPOLOGIA EXPANDIDA DA MÁ CONDUTA CIENTÍFICA DIGITALMENTE AMPLIFICADA

A era digital não apenas amplificou formas tradicionais de má conduta científica, mas também criou categorias inteiramente novas de fraude que transcendem as classificações convencionais de fabricação, falsificação e plágio. Esta transformação qualitativa exige uma reconceituação fundamental da tipologia da má conduta científica que reconheça tanto as continuidades com práticas históricas quanto as inovações radicais facilitadas por tecnologias emergentes (Research Integrity, 2024).

Tabela 4 - Tipologia expandida da má conduta científica na era digital

Categoria	Manifestação Tradicional	Manifestação Digitalmente Amplificada	Tecnologias Envolvidas
Fabricação	Inventar dados manualmente	Gerar conjuntos de dados sintéticos com IA que mimetizam padrões estatísticos específicos	GANs, MLMs, Simulação Computacional
Falsificação	Omitir dados anômalos manualmente	Usar <i>software</i> para remover/duplicar elementos; aplicar <i>p-hacking</i> algorítmico; gerar imagens fotorrealistas	Editores de imagem IA, Algoritmos de otimização
Plágio	Copiar texto diretamente	Usar LLMs para parafrasear extensivamente; criar "plágio disfarçado" com frases torturadas	GPT, Claude, Sistemas de parafrase
Manipulação de Autoria	Adicionar autores fictícios	Criar identidades acadêmicas completas com históricos fabricados	Geradores de identidade, Redes sociais falsas
Fraude de Revisão	Sugerir revisores cúmplices	Criar perfis falsos de revisores com históricos convincentes	Bots conversacionais, Perfis sintéticos

Fonte: Elaboração própria baseada em Research Integrity (2024), Nature Machine Intelligence (2024).

A fabricação digital permite a criação de dados sintéticos que podem ser estatisticamente indistinguíveis de dados empíricos reais.

Enquanto a fabricação histórica deixava frequentemente pistas detectáveis, a fabricação assistida por IA permite que fraudadores atuem como “curadores algorítmicos” que podem gerar múltiplos resultados plausíveis e selecionar aqueles que melhor se alinham a narrativas desejadas. Esta capacidade transforma fundamentalmente a natureza da fabricação, tornando a detecção exponencialmente mais difícil (Nature Biotechnology, 2024).

A falsificação digital expandiu dramaticamente através de ferramentas sofisticadas de edição de imagem e análise estatística automatizada. Redes Adversariais Generativas podem criar imagens científicas convincentes, incluindo micrografias e western blots indistinguíveis de fotografias reais mesmo para especialistas experientes. O *p-hacking* algorítmico permite que fraudadores automatizem a busca por significância estatística através de múltiplas análises até encontrar resultados publicáveis (PLOS ONE, 2024).

O plágio digital evoluiu para formas sofisticadas de paráfrase automatizada que evitam detecção por *software* convencional. LLMs podem reescrever textos mantendo conteúdo essencial enquanto alteram suficientemente a linguagem para evitar detecção algorítmica. Esta forma de “plágio disfarçado” é particularmente insidiosa porque preserva ideias originais enquanto obscurece sua origem (Computers; Education, 2024).

4.3.3 GERAÇÃO DE CONTEÚDO FRAUDULENTO: ANÁLISE MULTIMODAL

A capacidade da inteligência artificial de gerar conteúdo científico convincente em múltiplas modalidades — texto, imagem e dados — representa uma virada de paradigma que desafia as premissas mais fundamentais da validação científica. A verdadeira ameaça não reside na habilidade de falsificar um único elemento,

mas na capacidade de orquestrar uma fraude holística: a criação de artigos científicos inteiramente fabricados que são internamente consistentes, com metodologias detalhadas, resultados plausíveis, figuras convincentes e discussões sofisticadas, tudo gerado algoritmicamente sem qualquer fundamento na realidade empírica (Nature Communications, 2024).

No domínio do texto, os modelos de linguagem avançados (LLMs) são capazes de produzir uma prosa acadêmica que é, em muitos casos, indistinguível daquela escrita por humanos. Treinados em vastos corpora textuais que incluem milhões de artigos científicos, esses sistemas aprendem a mimetizar o estilo, a terminologia técnica e as estruturas argumentativas da escrita científica. Uma assinatura involuntária deixada por modelos mais antigos ou menos sofisticados é o fenômeno das “frases torturadas” (*tortured phrases*), identificado por Guillaume Cabanac e Cyril Labbé (2021). Expressões bizarras como “consciência falsa” em vez de “inteligência artificial” ou “doença do seio” em vez de “câncer de mama” surgem de tentativas rudimentares do algoritmo para parafrasear e evitar sistemas anti-plágio. Embora essas frases sirvam atualmente como um “sinal vermelho” para detectores de fraude, os modelos mais recentes estão se tornando cada vez mais hábeis em evitar esses erros, tornando a detecção baseada apenas no texto uma tarefa cada vez mais difícil.

A geração de imagens científicas pela IA introduz outra camada de complexidade e risco. O caso notório da publicação de um artigo contendo uma imagem absurdamente manipulada de um rato, com órgãos genitais fictícios e anotações sem sentido, alertou a comunidade científica de forma chocante. O fato de que uma fabricação tão grosseira, claramente gerada por IA, conseguiu passar pela revisão por pares e ser publicada antes de ser retratada, expôs a fragilidade dos processos de verificação visual (The Scientist, 2024). Para além de diagramas, a ameaça mais sutil vem de Redes Adversariais Generativas (GANs), que podem criar imagens fotorealistas de micrografias, géis de eletroforese ou *blots* que enganam até

mesmo a inspeção de especialistas. A distinção entre uma imagem real e uma falsificação algorítmica perfeita torna-se quase impossível a olho nu, exigindo o desenvolvimento de uma nova geração de ferramentas de análise forense digital para garantir a autenticidade da evidência visual (Computer Vision and Image Understanding, 2024).

Contudo, a capacidade de gerar dados sintéticos que mimetizam padrões estatísticos reais representa, talvez, a ameaça mais fundamental e existencial à integridade científica. Isso vai muito além de simplesmente “inventar números.” Algoritmos podem ser instruídos a produzir conjuntos de dados completos que não apenas pareçam plausíveis, mas que sejam customizados para atender a critérios estatísticos específicos. Um fraudador pode, por exemplo, “encomendar” um conjunto de dados que, quando analisado, resulte em um valor-p de 0.049, garantindo assim a tão cobiçada “significância estatística”. Essa capacidade de engenharia reversa da prova estatística permite a criação de fraudes que são projetadas para passar nos testes convencionais de validação, apresentando distribuições de dados apropriadas, correlações esperadas e outras características de um estudo legítimo (Journal of Computational Science, 2024).

Em última análise, essa capacidade multimodal de fabricação cria uma crise epistemológica. Se um artigo pode apresentar um texto eloquente, imagens convincentes e dados estatisticamente “perfeitos”, todos gerados artificialmente, como podemos continuar a confiar no registro publicado? A presunção de boa-fé, um pilar da ciência, fica abalada. A validação científica precisará se adaptar a essa nova realidade, possivelmente movendo-se para um modelo onde a disponibilização obrigatória e a auditoria de dados brutos, códigos de análise e processos de geração de imagens não sejam mais uma opção de “ciência aberta”, mas um requisito fundamental para a confiança e a publicação.

4.3.4 IMPLICAÇÕES EPISTEMOLÓGICAS: REDEFININDO A NATUREZA DA EVIDÊNCIA

A capacidade da inteligência artificial de gerar conteúdo científico fraudulento não representa apenas um avanço técnico na arte da enganação; ela provoca uma ruptura paradigmática que atinge o coração da epistemologia científica. O que está em jogo não é apenas a integridade de artigos individuais, mas a própria natureza do que consideramos “evidência” e os fundamentos sobre os quais construímos o conhecimento. Tradicionalmente, a fraude científica, por mais grave que fosse, era um ato artesanal. A fabricação manual de dados era um processo trabalhoso e propenso a erros, que frequentemente deixava “pegadas” — inconsistências estatísticas, padrões anômalos ou implausibilidades metodológicas que um revisor atento poderia detectar (Science and Engineering Ethics, 2024).

Hoje, a IA transforma o fraudador de um falsificador manual em um “otimizador algorítmico”. A fraude deixa de ser um ato de invenção para se tornar um ato de simulação e otimização. O ator mal-intencionado não precisa mais criar dados falsos do zero; ele pode simplesmente especificar os parâmetros de um resultado desejado — como um nível de significância estatística, uma correlação específica ou uma tendência particular — e instruir um sistema de IA a gerar o conjunto de dados, as imagens e o texto mais plausíveis que se alinhem a essa narrativa. O engano torna-se um produto de engenharia, customizado para ser o mais convincente possível.

Esta transformação radical levanta questões filosóficas profundas e desconcertantes sobre autoria, responsabilidade e a natureza da evidência. Quando um artigo fraudulento é gerado por uma IA sob a instrução de um humano, quem é o verdadeiro autor da fraude? Onde termina a responsabilidade do programador e começa a do usuário? Mais fundamentalmente, se a evidência sintética pode ser tornada estatisticamente equivalente à evidência empírica,

como podemos continuar a fundamentar a ciência na observação do mundo real? A relação ontológica entre a evidência (como um registro do mundo) e o conhecimento (como uma interpretação desse registro) é profundamente abalada (Philosophy of Science, 2024).

Consequentemente, os nossos sistemas tradicionais de validação científica são postos em xeque. A revisão por pares (*peer review*), o pilar central do controle de qualidade da ciência, baseia-se em grande parte na avaliação da plausibilidade, da consistência interna e da solidez metodológica de um trabalho. No entanto, este sistema é inerentemente vulnerável a fraudes que são projetadas precisamente para serem plausíveis e consistentes. Se um algoritmo pode gerar um artigo que “passa” em todos esses testes convencionais, a capacidade dos revisores de funcionar como guardiões da integridade fica severamente comprometida. Esta não é uma falha meramente técnica; é uma crise existencial para a confiança. Se a distinção entre o real e o simulado se torna difusa, a confiança pública na ciência — e, por extensão, a capacidade das sociedades de tomar decisões baseadas em evidências em áreas como saúde pública e política ambiental — fica sob grave ameaça (Social Studies of Science, 2024).

A sobrevivência da integridade científica na era da IA exige, portanto, mais do que apenas novas ferramentas de detecção; ela demanda o desenvolvimento de novos frameworks epistemológicos. Precisamos de uma nova filosofia da ciência que possa navegar neste novo terreno, estabelecendo critérios claros para distinguir entre simulação e realidade, e entre conhecimento gerado por algoritmos e descoberta científica genuína. Isso se traduz na necessidade prática de criar novos padrões para a transparência algorítmica, exigindo que os pesquisadores revelem quando e como a IA foi usada. Implica também desenvolver métodos para a verificação independente de resultados gerados ou analisados por IA, possivelmente através do compartilhamento obrigatório de dados brutos e códigos.

O objetivo final não é frear a inovação, mas garantir que possamos aproveitar o potencial transformador da IA sem sacrificar os valores fundamentais de verdade, objetividade e reprodutibilidade que definem a própria essência do empreendimento científico (IEEE Standards, 2024).

4.3.5 CARACTERÍSTICAS IDENTIFICÁVEIS, PADRÕES DE DETECÇÃO E EVOLUÇÃO ESTRATÉGICA: UMA ANÁLISE INTEGRADA DA SOFISTICAÇÃO OPERACIONAL

Os artigos produzidos por *paper mills* exibem um conjunto complexo e evolutivo de características distintivas que refletem tanto as limitações inerentes da produção industrial quanto as tentativas sofisticadas de evitar detecção, criando um panorama dinâmico onde padrões identificáveis emergem, são detectados e, subsequentemente, adaptados em uma corrida armamentista contínua entre fraudadores e sistemas de detecção (Zein, 2024). Esta dinâmica evolutiva revela não só as assinaturas técnicas dessas operações, como também sua notável capacidade de adaptação e inovação em resposta a medidas de contradetecção, demonstrando um nível de sofisticação organizacional que transcende operações fraudulentas tradicionais para constituir verdadeiros ecossistemas de inovação criminal.

A análise contemporânea de padrões de detecção deve ser compreendida dentro do contexto mais amplo da evolução tecnológica e estratégica dessas organizações. Enquanto as primeiras gerações de *paper mills* dependiam de métodos relativamente rudimentares que deixavam assinaturas detectáveis óbvias, as operações contemporâneas demonstram sofisticação crescente que incorpora inteligência artificial, análise de big data, e estratégias de evasão que rivalizam com as capacidades de detecção mais avançadas disponíveis para a comunidade científica legítima.

Uma das assinaturas mais reveladoras de operações de *paper mills* reside nos padrões implausíveis de colaboração que emergem quando a produção industrial de artigos científicos tenta simular redes orgânicas de pesquisa colaborativa. Artigos sobre tópicos altamente específicos escritos por autores sem conexão lógica com o tema representam um dos indicadores mais óbvios, mas a análise deve ir além de casos flagrantes para examinar padrões mais sutis que podem indicar operações sofisticadas (Zein, 2024).

A análise de redes de co-autoria revela anomalias estruturais características onde autores listados no mesmo artigo frequentemente não se conhecem, não têm histórico de colaboração prévia, não compartilham afiliações geográficas ou institucionais, e podem estar localizados em continentes diferentes sem justificativa metodológica plausível para colaboração internacional. Estas redes artificiais frequentemente exibem propriedades topológicas distintivas que diferem significativamente de redes de colaboração orgânicas, incluindo ausência de clustering natural, distribuições anômalas de graus de conectividade, e padrões temporais de formação que são inconsistentes com o desenvolvimento gradual de relacionamentos de pesquisa genuínos.

Operações mais sofisticadas têm respondido a essas vulnerabilidades através do desenvolvimento de redes de autoria mais plausíveis que incorporam colaborações graduais, históricos de publicação construídos ao longo do tempo, e afiliações geográficas que justificam colaborações internacionais. Algumas operações chegaram ao ponto de criar conferências acadêmicas fictícias onde “pesquisadores” fabricados podem estabelecer conexões aparentemente orgânicas, demonstrando um nível de planejamento estratégico que transcende fraudes oportunistas para constituir operações de longo prazo com investimento substancial em credibilidade artificial.

A análise textual e linguística de artigos de *paper mills* revela um espectro evolutivo de sofisticação que reflete tanto as capacidades

tecnológicas disponíveis quanto as estratégias adaptativas desenvolvidas em resposta a sistemas de detecção. As primeiras gerações de artigos fraudulentos eram caracterizadas por “frases torturadas” óbvias, onde sistemas de paráfrase automatizada produziam substituições bizarras como “bosom peril” substituindo “breast cancer” ou “counterfeit consciousness” no lugar de “artificial intelligence”, criando assinaturas linguísticas facilmente detectáveis por análise computacional (Cabanac e Labbé, 2021).

Contudo, a evolução tecnológica das operações de *paper mills* tem produzido gerações sucessivas de conteúdo textual cada vez mais sofisticado. A integração de modelos de linguagem avançados permite a produção de texto que não apenas evita substituições bizarras óbvias, mas também incorpora terminologia técnica apropriada, estruturas argumentativas convincentes, e fluxos narrativos que podem parecer naturais mesmo para revisores experientes. Esta evolução representa uma transformação qualitativa fundamental na natureza da fraude textual, onde a detecção baseada em anomalias linguísticas óbvias torna-se progressivamente menos eficaz.

A sofisticação contemporânea inclui não apenas melhorias na qualidade do texto individual, mas também na coerência interna dos artigos, onde metodologias, resultados, discussões e conclusões são integrados de forma logicamente consistente. Algumas operações desenvolveram bibliotecas de componentes metodológicos modulares que podem ser combinados de formas aparentemente originais, criando artigos que exibem diversidade superficial enquanto mantêm estruturas subjacentes padronizadas. Esta modularidade permite produção eficiente em escala enquanto reduz a detectabilidade através de análise de similaridade textual.

A evolução da manipulação de imagens em operações de *paper mills* representa uma das dimensões mais tecnologicamente sofisticadas da fraude científica contemporânea, progredindo de duplicação simples e manipulação básica para geração sintética

de evidência visual que pode ser indistinguível de dados empíricos reais. Esta progressão reflete não apenas avanços tecnológicos, mas também compreensão crescente da importância da evidência visual na validação científica e dos pontos vulneráveis nos sistemas de revisão por pares.

As primeiras gerações de manipulação de imagens eram caracterizadas por duplicação óbvia, rotação simples, e edição rudimentar que deixava artefatos detectáveis através de análise forense digital básica. Contudo, a integração de ferramentas de edição avançadas e, mais recentemente, de inteligência artificial generativa, tem produzido capacidades de manipulação que desafiam fundamentalmente os pressupostos sobre os quais a validação de evidência visual se baseia.

Redes Adversariais Generativas (GANs) e outros sistemas de IA generativa podem agora criar micrografias, western blots, gráficos, e outras formas de evidência visual científica que são estatisticamente indistinguíveis de imagens reais mesmo quando analisadas por especialistas experientes usando ferramentas forenses digitais avançadas. Esta capacidade representa uma transformação paradigmática que vai além da manipulação de evidência existente para a criação de evidência completamente sintética que pode sustentar narrativas científicas inteiramente fabricadas.

A sofisticação contemporânea inclui não só a qualidade individual das imagens geradas, como também sua integração coerente com narrativas textuais, nas quais figuras sintéticas são projetadas especificamente para sustentar alegações metodológicas e resultados descritos no texto. Algumas operações desenvolveram pipelines automatizados capazes de gerar conjuntos completos de evidência visual com base em especificações textuais, criando artigos em que texto e imagens são co-gerados de forma integrada para maximizar a plausibilidade e a coerência interna.

A análise de padrões temporais e estratégias de submissão revela dimensões operacionais sofisticadas que refletem planejamento estratégico e compreensão profunda dos sistemas editoriais e de revisão por pares. *Paper mills* frequentemente exibem padrões temporais característicos que incluem submissões em lotes coordenados, cronogramas de produção que são implausíveis para pesquisa empírica genuína, e estratégias de timing que exploram vulnerabilidades sazonais ou cíclicas nos sistemas de revisão.

Operações sofisticadas desenvolveram estratégias de infiltração editorial que vão além da submissão passiva para incluir influência ativa sobre processos de revisão e decisão editorial. Isto inclui criação de perfis falsos de revisores com históricos convincentes, infiltração de agentes em conselhos editoriais de periódicos legítimos, e até mesmo criação de periódicos inteiros controlados por operações fraudulentas. Estas estratégias representam uma evolução qualitativa fundamental que transforma *paper mills* de produtores passivos de conteúdo fraudulento para agentes ativos que moldam os próprios sistemas através dos quais a validação científica ocorre.

A sofisticação estratégica inclui também análise de vulnerabilidades específicas de periódicos individuais, onde operações desenvolvem perfis detalhados de preferências editoriais, padrões de revisão, e critérios de aceitação que permitem customização de submissões para maximizar probabilidade de aceitação. Algumas operações mantêm bancos de dados de revisores, editores, e suas preferências conhecidas, permitindo targeting estratégico que pode incluir submissão de artigos projetados especificamente para explorar vieses ou limitações de revisores específicos.

A evolução contemporânea de *paper mills* é caracterizada por integração crescente de tecnologias emergentes e notável capacidade de adaptação a medidas de contradetecção, criando uma dinâmica de corrida armamentista digital em que cada avanço em detecção é seguido por contramedidas de evasão que

frequentemente incorporam as mesmas tecnologias utilizadas para detectar fraudes. Essa dinâmica revela não só a sofisticação técnica dessas operações, como também sua capacidade organizacional de inovação contínua e adaptação estratégica.

A integração de inteligência artificial em operações de *paper mills* vai além da simples automação de tarefas existentes, incluindo inovações fundamentais nas capacidades fraudulentas. Sistemas de IA são utilizados tanto para gerar conteúdo quanto para analisar padrões de detecção, otimizar estratégias de evasão e desenvolver novas modalidades de fraude que antecipam futuras capacidades de detecção. Essa aplicação da IA à meta-análise de sistemas de detecção representa um nível de sofisticação estratégica que rivaliza com as capacidades da própria pesquisa legítima em inteligência artificial.

As operações contemporâneas demonstram capacidade de adaptação em tempo real às medidas de detecção, modificando estratégias, tecnologias e até mesmo modelos de negócio em resposta a mudanças no ambiente regulatório e tecnológico. Essa adaptabilidade envolve não só respostas reativas às detecções, como também a antecipação proativa de futuras medidas de contradetecção, revelando capacidades de planejamento estratégico e inteligência competitiva que transcendem operações criminosas tradicionais.

A análise de casos específicos revela que algumas operações mantêm equipes dedicadas de pesquisa e desenvolvimento que monitoram literatura científica sobre detecção de fraudes, participam de conferências acadêmicas sobre integridade científica, e até mesmo publicam pesquisa legítima sobre métodos de detecção como forma de inteligência competitiva. Esta integração entre atividades legítimas e fraudulentas cria desafios fundamentais para detecção e prevenção que vão além de questões técnicas para incluir dimensões de infiltração e subversão dos próprios sistemas de conhecimento usados para combater fraudes.

O impacto econômico de *paper mills* estende-se muito além dos lucros diretos dessas operações para incluir custos sistêmicos massivos que afetam toda a infraestrutura científica global, criando efeitos cascata que se propagam através de múltiplas dimensões do ecossistema científico e podem persistir por décadas após a detecção inicial de fraudes. Esta análise econômica revela que o verdadeiro custo de *paper mills* não reside apenas nos recursos desperdiçados diretamente, mas na erosão sistemática da eficiência, confiabilidade, e sustentabilidade do sistema científico como um todo.

Uma análise econômica abrangente de 2024 estimou que cada artigo fraudulento de *paper mill* gera custos sistêmicos médios entre US\$ 15.000 e US\$ 25.000, quando todos os efeitos diretos e indiretos são considerados (Economic Analysis of Research Integrity, 2024). Esses custos envolvem desperdício de tempo de revisão, uso inadequado de recursos editoriais, investigações subsequentes, retrações, impacto em pesquisas que citam resultados fraudulentos e a erosão da confiança pública, um fator que pode comprometer o financiamento científico e o apoio político à pesquisa.

O desperdício de recursos humanos especializados representa uma dimensão particularmente devastadora desses custos sistêmicos. Revisores voluntários experientes, cujo tempo é um recurso escasso e valioso para o sistema científico, dedicam coletivamente centenas de horas avaliando manuscritos fraudulentos que nunca deveriam ter entrado no sistema de revisão por pares. Esse tempo desperdiçado não representa apenas um custo de oportunidade direto, como também contribui para a sobrecarga do processo de avaliação, podendo comprometer a qualidade da revisão por pares de manuscritos legítimos.

Editores gastam recursos editoriais limitados processando submissões falsas, investigando alegações de má conduta, gerenciando retrações, e implementando medidas de detecção que, embora necessárias, desviam recursos de atividades que poderiam

melhorar a qualidade e eficiência da publicação científica. Estes custos administrativos são frequentemente subestimados, mas podem representar uma fração substancial dos orçamentos editoriais, particularmente para periódicos que são desproporcionalmente afetados por operações de *paper mills*.

Os efeitos cascata de *paper mills* na pesquisa legítima representam talvez a dimensão mais insidiosa e duradoura de seu impacto sistêmico. O desperdício de recursos humanos é igualmente significativo: revisores voluntários experientes dedicam tempo precioso avaliando manuscritos fraudulentos que nunca deveriam ter entrado no sistema, e editores gastam recursos processando submissões falsas, investigando alegações de má conduta e gerenciando retratações. Estes efeitos multiplicam exponencialmente os custos iniciais da fraude, criando redes de contaminação que podem afetar campos inteiros de pesquisa.

A análise de redes de citação revela que artigos fraudulentos podem ser citados centenas ou milhares de vezes antes de serem retratados, e que mesmo após retratação, continuam sendo citados por pesquisadores que podem não estar cientes da retratação ou que citam trabalhos que por sua vez citaram o artigo retratado. Esta persistência de informação fraudulenta na literatura científica cria um problema de “poluição epistêmica” que pode distorcer a compreensão científica por períodos prolongados.

O impacto é particularmente severo em campos emergentes ou interdisciplinares onde a base de literatura é limitada e cada contribuição tem peso desproporcional na formação de consenso científico. Quando artigos fraudulentos constituem uma fração significativa da literatura inicial em um campo, podem influenciar fundamentalmente a direção de pesquisa futura, criando paradigmas baseados em fundações falsas que podem ser extremamente difíceis de corrigir uma vez estabelecidos.

Em países em desenvolvimento, onde recursos para pesquisa são particularmente escassos e cada investimento em ciência representa uma oportunidade preciosa para desenvolvimento nacional, o impacto de *paper mills* é desproporcionalmente devastador e pode ter consequências duradouras para capacidade científica nacional. Fundos limitados para incentivos de publicação são desperdiçados em artigos fraudulentos que não contribuem para conhecimento científico real, capacidade nacional de pesquisa, ou desenvolvimento de recursos humanos especializados.

Esta dinâmica pode perpetuar e amplificar desigualdades globais na produção científica, criando círculos viciosos onde países com recursos limitados são simultaneamente mais vulneráveis à exploração por *paper mills* e menos capazes de desenvolver capacidades institucionais para detectar e prevenir má conduta científica. O resultado é um sistema que penaliza duplamente países em desenvolvimento: primeiro através da exploração direta por operações fraudulentas, e segundo através da erosão da credibilidade científica nacional que pode afetar colaborações internacionais e reconhecimento acadêmico por décadas.

A dependência de literatura científica internacional para informar políticas públicas, práticas médicas, e decisões de investimento torna países em desenvolvimento particularmente vulneráveis aos efeitos de fraudes científicas que podem influenciar decisões nacionais críticas baseadas em evidência contaminada. Quando políticas de saúde pública, estratégias de desenvolvimento econômico, ou investimentos em infraestrutura são baseados em pesquisa fraudulenta, as consequências podem incluir não apenas desperdício de recursos, mas também danos diretos ao bem-estar público e desenvolvimento nacional.

4.4 FERRAMENTAS DE DETECÇÃO E CONTRAMEDIDAS TECNOLÓGICAS: A EVOLUÇÃO DA RESPOSTA CIENTÍFICA À FRAUDE DIGITAL

4.4.1 EVOLUÇÃO SISTÊMICA DOS PARADIGMAS DE DETECÇÃO: DA ANÁLISE MANUAL À INTELIGÊNCIA ARTIFICIAL INTERPRETÁVEL

A resposta tecnológica à crise de integridade científica representa uma das mais fascinantes corridas armamentistas intelectuais da era digital, caracterizada por ciclos acelerados de inovação onde cada avanço em capacidades fraudulentas é seguido por desenvolvimentos correspondentes em tecnologias de detecção, criando uma dinâmica evolutiva que impulsiona inovação tanto em métodos de fraude quanto em contramedidas defensivas (The Scientist, 2024). Esta evolução transcende simples desenvolvimento tecnológico para constituir uma transformação fundamental na natureza da validação científica, onde sistemas automatizados assumem responsabilidades crescentes por tarefas que historicamente dependiam exclusivamente de expertise humana e julgamento contextual.

A primeira geração de ferramentas de detecção (1990s-2000s) operava através de princípios relativamente simples de correspondência textual e comparação com bases de dados publicadas, utilizando algoritmos básicos de string matching que podiam identificar casos óbvios de cópia direta mas eram facilmente evadidos através de paráfrase simples ou tradução. Esta abordagem, embora útil para casos flagrantes de plágio textual, mostrou-se fundamentalmente inadequada para enfrentar fraudes sofisticadas que envolviam manipulação de dados, geração de conteúdo original mas falso, ou operações coordenadas que transcendiam simples duplicação de texto.

A segunda geração (2000s-2010s) incorporou análise de imagens para detectar duplicação e manipulação visual, expandindo significativamente as capacidades de detecção mas ainda limitada a comparações baseadas em similaridade óbvia e análise forense digital básica. Esta expansão refletiu reconhecimento crescente de que fraude científica não se limitava a texto, mas incluía manipulação sofisticada de evidência visual que era fundamental para validação em muitas disciplinas científicas.

A dependência de literatura científica internacional para informar políticas públicas, práticas médicas e decisões de investimento torna países em desenvolvimento particularmente vulneráveis aos efeitos de fraudes científicas, que podem influenciar decisões nacionais críticas baseadas em evidência contaminada. Quando políticas de saúde pública, estratégias de desenvolvimento econômico ou investimentos em infraestrutura são baseados em pesquisa fraudulenta, as consequências podem incluir desperdício de recursos e, ainda mais gravemente, danos diretos ao bem-estar público e ao desenvolvimento nacional.

Os sistemas contemporâneos mais avançados operam através de abordagens multimodais integradas que analisam simultaneamente múltiplas dimensões de artigos científicos - texto, imagens, dados, metadados, padrões de citação, redes de colaboração, e cronogramas de submissão - para identificar inconsistências que podem ser invisíveis quando modalidades são analisadas isoladamente. Esta integração representa uma evolução fundamental de sistemas especializados que focavam em tipos específicos de fraude para plataformas holísticas que podem detectar operações coordenadas complexas.

A análise multimodal permite identificação de padrões de fraude que transcendem categorias tradicionais, como operações de *paper mills* que coordenam manipulação textual, visual, e de dados de forma integrada para criar artigos que podem passar por sistemas

de detecção especializados mas exibem assinaturas detectáveis quando analisados holisticamente. Esta capacidade é particularmente importante para detectar operações industriais sofisticadas que empregam divisão de trabalho especializada onde diferentes aspectos da fraude são realizados por equipes diferentes usando ferramentas diferentes.

Uma das inovações mais significativas na detecção contemporânea é o desenvolvimento de sistemas de inteligência artificial interpretável que vão além da simples identificação de possíveis fraudes, oferecendo também explicações compreensíveis sobre suas decisões. Isso permite que revisores humanos façam julgamentos informados com base em evidências específicas, em vez de simplesmente aceitar ou rejeitar recomendações algorítmicas opacas (Nature Machine Intelligence, 2024). Essa transparência algorítmica é essencial tanto para a eficácia da detecção quanto para a aceitação e a confiança da comunidade científica em sistemas automatizados.

Sistemas como o Pub-Guard-LLM utilizam Grandes Modelos de Linguagem não só para detectar fraudes, como também para gerar explicações detalhadas sobre características específicas que indicam possível má conduta, níveis de confiança calibrados para diferentes tipos de evidência e recomendações contextuais sobre investigações adicionais apropriadas. Essa capacidade explicativa permite uma colaboração efetiva entre sistemas algorítmicos e a expertise humana, em que os algoritmos processam volumes de informação impossíveis de serem analisados manualmente, enquanto os humanos mantêm o controle sobre a interpretação final e as decisões relativas às ações a serem tomadas.

A interpretabilidade também facilita melhoria contínua de sistemas de detecção através de feedback humano, onde especialistas podem identificar falsos positivos e negativos, compreender limitações algorítmicas, e contribuir para refinamento de modelos. Esta colaboração humano-algoritmo representa uma abordagem

mais sustentável e robusta do que sistemas completamente automatizados que podem ser vulneráveis a evasão adversarial ou sistemas puramente manuais que não podem escalar para volumes contemporâneos de literatura científica.

4.4.2 DETECÇÃO AVANÇADA DE MANIPULAÇÃO VISUAL: FORENSE DIGITAL E ANÁLISE DE AUTENTICIDADE

A detecção automatizada de manipulação de imagens científicas representa uma das fronteiras mais tecnologicamente sofisticadas e filosoficamente complexas na luta contra fraude científica, incorporando técnicas de visão computacional, análise forense digital, aprendizado profundo, e até mesmo princípios de física e óptica para identificar alterações que podem ser imperceptíveis ao olho humano mas deixam assinaturas detectáveis em propriedades estatísticas, geométricas, ou físicas de imagens (Revista Pesquisa FAPESP, 2024).

O sistema Sila (Scientific Image AnaLysis), desenvolvido no Brasil, exemplifica esta abordagem através da combinação de extração automática de imagens de artigos científicos com algoritmos especializados que podem identificar clonagens sutis, reuso não declarado entre artigos diferentes, manipulações que alteram aspectos específicos mantendo aparência geral convincente, e até mesmo inconsistências de iluminação ou perspectiva que podem indicar composição artificial. O sistema demonstrou capacidade de detectar manipulações que passaram despercebidas por revisores experientes e editores especializados, revelando a prevalência subestimada de manipulação visual na literatura científica.

Ferramentas comerciais como *Image Twin*, *Proofing AI* e sistemas desenvolvidos por empresas de segurança digital têm sido implementados por grandes periódicos científicos, demonstrando tanto a viabilidade técnica quanto a aceitação crescente, por parte

da indústria editorial, de que a detecção automatizada é necessária para manter a integridade diante do volume contemporâneo de submissões. Estudos empíricos revelam que a manipulação de imagens é alarmantemente comum: uma análise automatizada de 1.364 artigos biomédicos identificou manipulações suspeitas em 5,7% dos papers, com quase 22% dos artigos que continham géis apresentando imagens potencialmente adulteradas.

Sistemas baseados em visão computacional avançada utilizam algoritmos de aprendizado profundo treinados em grandes conjuntos de imagens científicas para identificar tanto duplicações evidentes quanto manipulações sofisticadas que alteram aspectos específicos mantendo uma aparência geral convincente. Essas ferramentas são capazes de detectar inconsistências de iluminação que indicam composição artificial, artefatos de compressão que revelam edição digital e transformações geométricas que sugerem manipulação mesmo quando visualmente imperceptíveis (Bik *et al.*, 2016; Coorey *et al.*, 2022).

Algoritmos de análise espectral podem identificar assinaturas de diferentes dispositivos de captura, revelando quando imagens supostamente do mesmo experimento foram na verdade capturadas usando equipamentos diferentes ou em momentos diferentes. Análise de ruído pode detectar padrões inconsistentes que indicam edição digital, enquanto análise de metadados pode revelar informações sobre software de edição, timestamps, e histórico de modificações que podem contradizer alegações sobre origem e autenticidade de imagens.

Técnicas de análise de redes adversariais podem detectar imagens geradas por IA através de identificação de assinaturas estatísticas características de diferentes modelos generativos. Esta capacidade é particularmente importante dado o uso crescente de GANs e outros sistemas de IA generativa para criar evidência visual completamente sintética que pode ser indistinguível de fotografias reais através de inspeção visual.

A implementação de detecção de manipulação de imagens em escala editorial enfrenta desafios significativos que vão além de questões técnicas para incluir considerações de custo, velocidade, integração com fluxos de trabalho existentes, e balanceamento entre sensibilidade e especificidade para minimizar falsos positivos que poderiam prejudicar pesquisadores legítimos. Sistemas devem processar milhares de imagens diariamente com velocidade suficiente para não atrasar processos editoriais, mantendo precisão suficiente para detectar manipulações sutis sem gerar alarmes falsos excessivos.

A integração com sistemas editoriais existentes requer desenvolvimento de interfaces que permitam que editores e revisores compreendam e atuem sobre resultados de detecção sem necessidade de expertise técnica especializada. Isto inclui visualizações intuitivas de possíveis manipulações, explicações compreensíveis de evidência de alteração, e recomendações contextuais sobre investigação adicional apropriada.

Considerações de privacidade e propriedade intelectual também complicam implementação, particularmente quando detecção requer comparação com bases de dados de imagens que podem conter material proprietário ou sensível. Sistemas devem equilibrar necessidade de detecção abrangente com proteção de direitos autorais e confidencialidade de pesquisa não publicada.

4.4.3 ANÁLISE LINGÜÍSTICA AVANÇADA E DETECÇÃO DE CONTEÚDO GERADO POR IA: ALÉM DA DETECÇÃO DE PLÁGIO TRADICIONAL

A detecção de conteúdo científico gerado por inteligência artificial representa um dos desafios mais complexos e filosoficamente profundos na validação científica contemporânea, exigindo desenvolvimento de técnicas que vão muito além de detectores tradicionais de plágio para incorporar análise linguística sofisticada,

modelagem estatística avançada, e compreensão profunda das assinaturas características deixadas por diferentes sistemas de IA generativa (The Scientist, 2024). Esta área de pesquisa evolui rapidamente em resposta à sofisticação crescente de modelos de linguagem e à necessidade urgente de distinguir entre uso legítimo e fraudulento de IA na produção científica.

O Problematic Paper Screener, desenvolvido por Guillaume Cabanac e Cyril Labbé, representa uma das ferramentas mais avançadas e amplamente utilizadas para detecção de texto científico problemático, incorporando módulos especializados na identificação de “frases torturadas” que se tornaram assinatura involuntária de texto gerado por IA de primeira geração. O sistema sinaliza automaticamente substituições bizarras como “bosom peril” substituindo “breast cancer”, “counterfeit consciousness” no lugar de “artificial intelligence”, ou “fake consciousness” ao invés de “artificial intelligence”, demonstrando como tentativas algorítmicas de evitar detecção de plágio podem criar assinaturas linguísticas ainda mais detectáveis.

Contudo, a evolução rápida de modelos de linguagem tem tornado detecção baseada em anomalias linguísticas óbvias progressivamente menos eficaz, exigindo desenvolvimento de técnicas mais sofisticadas que analisam propriedades estatísticas sutis, padrões sintáticos complexos, distribuições de vocabulário, e outras características que podem ser preservadas mesmo quando texto superficialmente parece natural e convincente.

Detectores de estilo mais sofisticados analisam distribuições estatísticas de palavras, padrões sintáticos complexos, uso de conectores lógicos, estruturas argumentativas, frequências de n-gramas, e outras assinaturas linguísticas que podem ser características de modelos generativos específicos. Diferentes sistemas de IA - GPT, Claude, Gemini, e outros - podem deixar “impressões digitais” linguísticas detectáveis através de análise estatística avançada que examina não apenas palavras individuais, mas padrões de combinação, estruturas de dependência, e propriedades emergentes de texto gerado.

Análise de entropia linguística pode identificar texto que é estatisticamente “muito perfeito” - exibindo distribuições de características que são implausíveis para produção humana natural mas consistentes com otimização algorítmica. Modelos de linguagem frequentemente produzem texto com propriedades estatísticas que diferem sutilmente de escrita humana, incluindo distribuições de comprimento de sentença, variabilidade lexical, e padrões de repetição que podem ser detectáveis através de análise quantitativa sofisticada.

Técnicas de análise de redes semânticas podem identificar padrões de associação conceitual que são característicos de diferentes modelos de treinamento, revelando quando texto foi gerado por sistemas treinados em corpora específicos ou usando arquiteturas particulares. Esta análise pode ser particularmente útil para identificar texto gerado por modelos especializados que foram *fine-tuned* para produção científica.

Um dos desafios mais complexos na detecção de conteúdo gerado por IA é a necessidade de distinguir entre uso legítimo de IA como ferramenta de assistência editorial e uso fraudulento para geração de conteúdo científico substantivo sem base empírica. Esta distinção não é meramente técnica, mas envolve julgamentos sobre intenção, transparência, e contribuição intelectual que podem ser difíceis de automatizar e podem variar entre disciplinas e contextos.

Uso legítimo de IA pode incluir assistência com gramática e estilo, tradução entre idiomas, geração de resumos de literatura existente, ou criação de primeiros rascunhos que são subsequentemente revisados e validados por pesquisadores humanos. Uso fraudulento inclui geração de metodologias fictícias, resultados fabricados, ou discussões substantivas sem base em pesquisa real. A fronteira entre estes usos pode ser difusa e contextual, exigindo desenvolvimento de critérios nuançados que considerem não apenas características técnicas de texto, mas também transparência sobre uso de IA e validação independente de alegações substantivas.

Detectores devem equilibrar sensibilidade suficiente para identificar uso fraudulento com especificidade suficiente para evitar falsos positivos que poderiam prejudicar pesquisadores que usam IA legitimamente. Este balanceamento é particularmente desafiador dado que a linha entre uso legítimo e fraudulento pode depender de fatores como divulgação adequada, validação independente, e contribuição intelectual humana que podem não ser detectáveis através de análise textual isolada.

A detecção de conteúdo gerado por IA enfrenta o desafio fundamental de que sistemas de geração evoluem continuamente para evitar detecção, criando uma corrida armamentista onde cada avanço em capacidades de detecção é seguido por desenvolvimentos correspondentes em técnicas de evasão. Esta dinâmica adversarial exige que sistemas de detecção sejam projetados para adaptação contínua e robustez contra tentativas deliberadas de evasão.

Técnicas de evasão incluem uso de múltiplos modelos de linguagem para obscurecer assinaturas específicas, edição humana posterior para introduzir variabilidade natural, e desenvolvimento de modelos especializados treinados especificamente para evitar detecção por sistemas conhecidos. Algumas operações fraudulentas podem até mesmo usar sistemas de detecção como ferramentas de controle de qualidade, refinando texto gerado até que passe por detectores conhecidos.

Resposta a esta dinâmica adversarial requer desenvolvimento de sistemas de detecção que são robustos contra evasão conhecida e capazes de adaptação rápida a novas técnicas de evasão. Isto pode incluir uso de *ensemble methods* que combinam múltiplas abordagens de detecção, técnicas de aprendizado adversarial que antecipam tentativas de evasão, e sistemas de atualização contínua que incorporam novos dados sobre técnicas de evasão emergentes.

4.4.4 TECNOLOGIAS EMERGENTES E DIREÇÕES FUTURAS: BLOCKCHAIN, COMPUTAÇÃO QUÂNTICA E REALIDADE AUMENTADA

O horizonte tecnológico para detecção de má conduta científica inclui tecnologias emergentes que prometem capacidades transformadoras mas também apresentam desafios únicos de implementação, custo, e integração com infraestruturas existentes. *Blockchain* pode fornecer verificação criptográfica de integridade de dados que é independente de autoridades centralizadas e resistente a manipulação posterior, criando registros imutáveis de versões de dados, *timestamps* de criação, e histórico de modificações que podem ser verificados independentemente por qualquer parte interessada (Quantum Information Processing, 2024).

Computação quântica pode acelerar dramaticamente análises complexas de padrões em grandes conjuntos de dados, permitindo detecção de anomalias sutis que seriam computacionalmente intratáveis usando sistemas clássicos. Algoritmos quânticos podem ser particularmente úteis para análise de redes complexas de colaboração, detecção de padrões em dados de alta dimensionalidade, e otimização de sistemas de detecção multimodal que devem equilibrar múltiplos critérios simultaneamente.

Realidade virtual e aumentada podem facilitar visualização de padrões complexos de fraude que seriam difíceis de identificar através de interfaces tradicionais, permitindo que investigadores "naveguem" através de redes de citação, visualizem relações temporais entre submissões, e identifiquem anomalias através de representações espaciais intuitivas. Estas tecnologias podem ser particularmente úteis para análise de operações coordenadas complexas que envolvem múltiplos artigos, autores e periódicos.

Sistemas futuros provavelmente integrarão múltiplas modalidades de detecção em plataformas unificadas que podem analisar simultaneamente texto, imagens, dados, padrões de colaboração, metadados temporais, e até mesmo comportamento de submissão para identificar padrões de fraude que seriam invisíveis quando modalidades são analisadas isoladamente. Esta integração pode revelar operações coordenadas complexas que empregam divisão de trabalho especializada onde diferentes aspectos da fraude são realizados por equipes diferentes usando ferramentas diferentes.

Machine learning federado pode permitir que instituições compartilhem informações sobre padrões de fraude sem compartilhar dados sensíveis, criando sistemas de detecção globais que preservam a privacidade enquanto maximizam a eficácia por meio do aprendizado coletivo. Esta abordagem pode ser particularmente valiosa para detectar operações transnacionais de *paper mills* que exploram diferenças regulatórias entre jurisdições.

Sistemas de monitoramento em tempo real podem detectar anomalias à medida que artigos são submetidos, permitindo intervenção preventiva ao invés de detecção post-hoc. Esta capacidade pode incluir análise de padrões de submissão, detecção de campanhas coordenadas, e identificação de comportamento suspeito que pode indicar operações fraudulentas em desenvolvimento.

A implementação de tecnologias emergentes para detecção de fraude enfrenta desafios significativos que vão além de questões técnicas para incluir considerações de custo, privacidade, equidade, e potencial para uso indevido. Sistemas de detecção avançados podem ser caros de desenvolver e manter, criando disparidades entre instituições com recursos diferentes e potencialmente exacerbando desigualdades existentes na capacidade de detectar e prevenir má conduta.

Considerações de privacidade são particularmente importantes quando sistemas de detecção requerem acesso a dados sensíveis, comunicações privadas, ou informações proprietárias.

Balanceamento entre necessidade de detecção eficaz e proteção de privacidade e direitos autorais requer desenvolvimento de frameworks legais e técnicos sofisticados que podem variar entre jurisdições.

Potencial para uso indevido de tecnologias de detecção - incluindo vigilância excessiva, discriminação algorítmica, ou supressão de pesquisa legítima - requer desenvolvimento de salvaguardas robustas e supervisão ética. Sistemas de detecção devem ser projetados com transparência, responsabilização, e mecanismos de recurso que protegem direitos de pesquisadores enquanto mantêm eficácia na detecção de má conduta genuína.

4.5 DIMENSÕES ÉTICAS E SOCIAIS: IMPACTOS SISTÊMICOS NA CONFIANÇA PÚBLICA E COESÃO EPISTÊMICA

Quando pensamos em fraude científica, nossa mente tende a visualizar um problema contido dentro dos laboratórios e das revistas acadêmicas. Imaginamos pesquisadores desonestos, dados manipulados e artigos retratados. No entanto, a abordagem que trazemos aqui é que essa visão é perigosamente limitada. A verdadeira crise de integridade na ciência hoje transcende, e muito, o mundo acadêmico. Ela se tornou uma fratura exposta no próprio tecido social, com impactos sistêmicos que ameaçam os alicerces de como as sociedades modernas funcionam.

Primeiro, falemos da confiança. A ciência opera com base em um contrato social implícito: a sociedade investe recursos e concede autoridade à comunidade científica em troca de conhecimento confiável que sirva ao bem comum. Cada artigo fraudulento, cada escândalo de *paper mills*, cada retratação em massa é uma martelada nesse contrato.

Não se trata apenas de uma percepção ou de um problema de comunicação. Dados de instituições como o Pew Research Center documentam uma queda real e preocupante na confiança que as pessoas depositam na ciência. Essa tendência não é um fenômeno isolado; é global, indicando que não estamos lidando com uma crise local, mas com uma patologia sistêmica na relação entre conhecimento e sociedade.

Essa perda de confiança nos leva ao segundo ponto, que é ainda mais profundo: a fragmentação da coesão epistêmica. “Coesão epistêmica” soa como um termo acadêmico, mas seu significado é simples e vital: é a nossa capacidade coletiva de concordar sobre uma base de fatos. É o chão comum sobre o qual podemos debater soluções para nossos maiores desafios. Quando a ciência, que deveria fornecer esse chão, se torna ela mesma uma fonte de incerteza e controvérsia, esse solo se desfaz. O debate público sobre mudanças climáticas, vacinas ou políticas econômicas deixa de ser uma discussão sobre como interpretar os fatos e se torna uma batalha campal onde os próprios fatos são questionados. Sem uma realidade compartilhada, a tomada de decisão democrática se torna disfuncional.

Mas como essa destruição acontece na prática? Ela é impulsionada por um motor poderoso e moderno. A fraude científica fornece o combustível, mas é o ambiente de informação digital que atua como um acelerador. Um estudo fraudulento, mas sensacionalista, pode alcançar milhões de pessoas através das redes sociais em questão de horas. A sua eventual correção ou retratação, por outro lado, move-se a passos de tartaruga, alcançando uma fração ínfima da audiência original. Cria-se, assim, uma impressão duradoura baseada em falsidade.

Pior ainda, esse veneno não se espalha de maneira uniforme. Ele encontra terreno fértil em comunidades que já possuem um histórico de desconfiança com as instituições, aprofundando feridas sociais e exacerbando desigualdades em saúde e bem-estar.

E, finalmente, essa atmosfera de dúvida é estrategicamente explorada. Atores políticos e corporativos com interesses particulares aprenderam que não precisam provar que o consenso científico está errado; eles só precisam gerar dúvida suficiente sobre a integridade do processo científico. A existência de fraudes reais se torna a desculpa perfeita para desacreditar pesquisas legítimas que ameaçam seus lucros ou sua agenda ideológica.

Portanto, a questão que se coloca é muito maior do que “como pegamos os fraudadores?”. A verdadeira pergunta é: “Como reconstruímos a confiança em um ecossistema de informação que recompensa a desinformação?”. Como restauramos um chão comum de fatos em uma sociedade polarizada? A crise de integridade científica nos força a encarar que a busca pela verdade não é um exercício acadêmico isolado, mas uma condição essencial para a sobrevivência e o progresso de uma sociedade livre e democrática.

4.5.1 EROSÃO DA CONFIANÇA PÚBLICA E FRAGMENTAÇÃO DO CONSENSO CIENTÍFICO: UMA ANÁLISE MULTIDIMENSIONAL

A crise de integridade científica transcende questões técnicas ou acadêmicas isoladas para representar uma ameaça existencial à confiança pública na ciência como instituição social fundamental, com implicações profundas que se estendem muito além da comunidade científica para afetar a capacidade das sociedades democráticas de tomar decisões coletivas baseadas em evidências sobre questões críticas que definem o futuro da humanidade (Science, 2024). Esta erosão da confiança não é meramente um problema de relações públicas ou comunicação científica inadequada, mas representa uma crise de legitimidade epistêmica que questiona os próprios fundamentos sobre os quais sociedades modernas baseiam decisões sobre saúde pública, mudanças climáticas, desenvolvimento tecnológico, e políticas econômicas.

Pesquisas de opinião pública revelam um declínio sistemático e preocupante na confiança científica que transcende divisões políticas, educacionais, e demográficas tradicionais, sugerindo que estamos testemunhando uma transformação cultural mais ampla na relação entre ciência e sociedade. Nos Estados Unidos, a confiança na ciência caiu dramaticamente de 87% em 1973 para 73% em 2018, com declínios particularmente acentuados entre grupos politicamente conservadores, mas evidências crescentes de erosão em outros segmentos demográficos que tradicionalmente demonstravam alta confiança na autoridade científica (Pew Research Center, 2024).

Esta tendência não é limitada aos Estados Unidos, mas reflete padrões globais de ceticismo crescente em relação à autoridade epistêmica da ciência. Pesquisas em países europeus, asiáticos, e latino-americanos revelam declínios similares, embora com variações regionais que refletem contextos políticos, culturais, e históricos específicos. A universalidade desta tendência sugere que estamos enfrentando não apenas problemas locais de comunicação científica ou polarização política, mas uma crise sistêmica mais profunda na relação entre conhecimento especializado e confiança pública.

A proliferação de artigos fraudulentos contribui diretamente para erosão de confiança através de múltiplos mecanismos que são amplificados por características específicas do ambiente de informação digital contemporâneo. Quando estudos posteriormente retratados são citados em debates públicos, políticas governamentais, discussões midiáticas, ou redes sociais, a subsequente correção frequentemente recebe menos atenção, alcance, e credibilidade que a publicação original, deixando impressões duradouras baseadas em informações falsas que podem persistir indefinidamente no discurso público (Policy Studies Journal, 2024).

Este fenômeno é particularmente problemático em áreas de alta relevância pública e controvérsia política como medicina, mudanças climáticas, segurança alimentar, e tecnologias emergentes, onde

estudos fraudulentos podem ser seletivamente amplificados por grupos com interesses específicos, criando narrativas alternativas que competem com consenso científico legítimo. A estrutura de algoritmos de redes sociais, que frequentemente priorizam engajamento sobre precisão, pode amplificar informações sensacionais ou controversas independentemente de sua validade científica.

A velocidade e escala da disseminação de informação digital significa que artigos fraudulentos podem alcançar audiências massivas e influenciar opinião pública antes que processos de validação científica possam identificar e corrigir problemas. Esta dinâmica temporal cria uma assimetria fundamental onde informação falsa pode se espalhar mais rapidamente que correções, estabelecendo narrativas que são resistentes a correção posterior.

Os efeitos da erosão de confiança científica não são distribuídos uniformemente através da sociedade, mas afetam desproporcionalmente comunidades que já enfrentam marginalizações históricas, acesso limitado à educação científica de qualidade, ou experiências negativas com instituições científicas e médicas. Comunidades que foram historicamente sujeitas a experimentação não ética, discriminação médica, ou negligência científica podem ser particularmente vulneráveis a narrativas que questionam a confiabilidade da ciência institucional.

Esta vulnerabilidade diferenciada cria dinâmicas complexas onde erosão de confiança pode exacerbar desigualdades existentes em saúde, educação, e participação cívica. Quando comunidades marginalizadas perdem confiança na ciência, podem ser menos propensas a buscar cuidados médicos baseados em evidências, participar de programas de saúde pública, ou apoiar políticas baseadas em consenso científico, criando disparidades que podem perpetuar e amplificar desigualdades sociais existentes.

Simultaneamente, a erosão de confiança pode ser explorada por atores políticos e econômicos que têm interesse em questionar consenso científico sobre questões específicas. Indústrias que enfrentam regulamentação baseada em evidência científica podem amplificar ceticismo sobre integridade científica como estratégia para questionar pesquisa que ameaça seus interesses, criando alianças oportunísticas entre ceticismo legítimo sobre má conduta e negação estratégica de consenso científico estabelecido.

4.5.2 COLONIALISMO CIENTÍFICO E VULNERABILIDADES SISTÊMICAS: REPRODUÇÃO DE DESIGUALDADES GLOBAIS

Os efeitos da crise de integridade científica são desproporcionalmente severos em países em desenvolvimento, onde recursos limitados, infraestrutura científica menos desenvolvida, e dependência de literatura internacional tornam cada investimento em ciência particularmente precioso e cada erro potencialmente devastador para desenvolvimento nacional e bem-estar público (UNESCO, 2024). Esta vulnerabilidade diferenciada não é acidental, mas reflete estruturas sistêmicas de desigualdade global que são reproduzidas e amplificadas através de dinâmicas de colonialismo científico que extraem conhecimento, dados, e recursos de países do Sul Global para beneficiar instituições e pesquisadores do Norte Global.

Países em desenvolvimento frequentemente dependem de literatura científica internacional para informar políticas públicas críticas, práticas médicas, estratégias de desenvolvimento econômico, e decisões sobre investimentos em infraestrutura. Quando essa literatura é contaminada por fraudes, as consequências podem incluir adoção de tratamentos médicos ineficazes ou prejudiciais, implementação de políticas públicas mal-informadas que desperdiçam recursos escassos, e decisões de investimento baseadas em promessas tecnológicas falsas que podem desviar recursos de alternativas mais eficazes.

A dinâmica do colonialismo científico exacerba esses problemas através de mecanismos que reproduzem desigualdades históricas e criam novas formas de dependência epistêmica. Pesquisadores de países desenvolvidos podem conduzir “ciência paraquedas” em nações em desenvolvimento, extraindo dados sem envolvimento adequado de cientistas locais, apropriando-se de conhecimentos tradicionais sem reconhecimento ou compensação, e publicando resultados que podem não refletir adequadamente contextos locais, necessidades comunitárias, ou prioridades nacionais de desenvolvimento.

O colonialismo de dados representa uma manifestação contemporânea particularmente insidiosa das dinâmicas coloniais. Dados pessoais, genéticos, ambientais e culturais são coletados de populações do Sul Global e processados por algoritmos e infraestruturas controladas por corporações e instituições do Norte Global. Essa extração ocorre sem consentimento adequado, benefício compartilhado ou controle local sobre o uso dos dados, criando novas formas de dependência e exploração que podem ser mais sutis, mas igualmente prejudiciais, que as formas históricas de colonialismo.

Quando essa extração de dados é combinada com má conduta científica - incluindo fabricação de dados, manipulação de resultados, ou publicação de pesquisa de baixa qualidade - as comunidades locais sofrem consequências duplas: primeiro através da exploração de seus recursos sem benefício adequado, e segundo através de danos potenciais causados por pesquisa fraudulenta que pode influenciar políticas que os afetam diretamente.

As estruturas de poder global na ciência também se manifestam através do controle sobre infraestruturas de publicação, sistemas de revisão por pares, e métricas de avaliação científica que podem marginalizar contribuições de pesquisadores do Sul Global. Grandes editoras comerciais, predominantemente baseadas no Norte Global, controlam acesso à literatura científica através de *paywalls* que

limitam o acesso de pesquisadores e instituições com recursos limitados, criando barreiras que podem impedir participação equitativa no discurso científico global.

Estas dinâmicas de poder criam vulnerabilidades específicas à má conduta científica que afetam desproporcionalmente países e comunidades do Sul Global. *Paper mills* e outras operações fraudulentas frequentemente exploram essas assimetrias, oferecendo “atalhos” para publicação em periódicos de prestígio que podem ser particularmente atraentes para pesquisadores que enfrentam barreiras estruturais para reconhecimento internacional, recursos limitados para pesquisa de alta qualidade, ou pressões institucionais por produtividade que não são acompanhadas por suporte adequado.

Simultaneamente, a dependência de literatura científica produzida externamente torna países em desenvolvimento mais vulneráveis aos efeitos de fraudes científicas que podem influenciar políticas públicas, práticas médicas, e decisões de investimento sem que haja capacidade local adequada para validação independente ou detecção de problemas. Esta dependência pode criar círculos viciosos onde países com recursos limitados são simultaneamente mais vulneráveis à exploração por operações fraudulentas e menos capazes de desenvolver capacidades institucionais para detectar e prevenir má conduta científica.

O resultado é um sistema que penaliza duplamente países em desenvolvimento: primeiro através da exploração direta por operações fraudulentas que extraem recursos sem fornecer valor real, e segundo através da erosão da credibilidade científica nacional que pode afetar colaborações internacionais, reconhecimento acadêmico, e capacidade de participar equitativamente no sistema científico global por décadas.

4.5.3 DIMENSÕES NEUROBIOLÓGICAS E PSICOLÓGICAS: ESTRESSE SISTÊMICO E DEGRADAÇÃO DO JULGAMENTO ÉTICO

Uma compreensão contemporânea dos impactos sociais da má conduta científica deve incorporar evidências emergentes sobre os mecanismos neurobiológicos e psicológicos que influenciam o comportamento ético, reconhecendo que a integridade científica não é meramente uma questão de escolha moral individual, mas está profundamente entrelaçada com condições estruturais que afetam a capacidade cognitiva, o bem-estar mental e o funcionamento neurológico de pesquisadores (UOL, 2023). Esta perspectiva neurobiológica revela que estratégias eficazes de prevenção devem abordar não apenas aspectos educacionais e regulatórios, mas também as condições sistêmicas que geram estresse crônico e comprometem a capacidade de julgamento ético.

Pesquisas revelam que pesquisadores têm níveis de ansiedade e depressão seis vezes maiores que a população geral, criando condições neurobiológicas que podem facilitar comportamentos antiéticos sob pressão (UOL, 2023). Esta prevalência elevada de problemas de saúde mental não é acidental, mas reflete características estruturais do ambiente acadêmico contemporâneo que incluem competição intensa por recursos limitados, insegurança profissional crônica, pressões por produtividade que frequentemente excedem capacidades humanas sustentáveis, e culturas institucionais que podem valorizar resultados sobre bem-estar pessoal.

A neurobiologia da desonestidade revela que comportamentos desonestos frequentemente ocorrem quando indivíduos enfrentam pressões que ativam sistemas de resposta ao estresse, reduzindo capacidade de autorregulação ética e comprometendo funcionamento do córtex pré-frontal, que é fundamental para julgamento moral, planejamento de longo prazo, e integração de considerações éticas com processos de tomada de decisão (Revista Interações, 2024).

Sob estresse crônico, indivíduos podem experimentar degradação de capacidades cognitivas que normalmente sustentam comportamento ético, incluindo redução de empatia, foco excessivo em benefícios imediatos, e diminuição de capacidade para considerar consequências de longo prazo de ações.

O ambiente acadêmico contemporâneo, caracterizado por competição intensa, financiamento limitado, pressões por produtividade quantitativa, e insegurança profissional crônica, cria condições de estresse que podem comprometer capacidade de julgamento ético de formas que transcendem escolhas individuais para refletir falhas sistêmicas em design institucional e cultura organizacional (Occupational Health Psychology, 2024). Estudos mostram que estresse pode afetar tomada de decisão moral de formas específicas, levando indivíduos a priorizar benefícios imediatos sobre considerações éticas de longo prazo, reduzir capacidade de empatia e consideração por outros, e aumentar propensão a racionalizar comportamentos que normalmente considerariam inaceitáveis.

Esta compreensão sugere que estratégias de prevenção de má conduta científica devem abordar não apenas aspectos educacionais sobre princípios éticos, mas também condições estruturais que criam estresse crônico no ambiente de pesquisa. Isso inclui reforma de sistemas de avaliação que priorizam quantidade sobre qualidade, melhoria de suporte institucional para bem-estar de pesquisadores, criação de culturas que valorizam sustentabilidade de carreira sobre produtividade de curto prazo, e desenvolvimento de sistemas de financiamento que reduzem insegurança crônica e permitem planejamento de longo prazo.

A degradação institucional pode criar círculos viciosos onde estresse individual contribui para deterioração de normas coletivas, que por sua vez aumenta estresse individual através de erosão de confiança, cooperação, e suporte mútuo. Quando má conduta se torna normalizada ou tolerada institucionalmente, pode criar

ambientes onde comportamento ético se torna mais difícil de manter e onde indivíduos podem sentir pressão para comprometer padrões éticos para permanecer competitivos.

Evidência emergente sobre neurobiologia do comportamento ético sugere que intervenções eficazes devem ser projetadas com compreensão de como estresse, fadiga, e pressão afetam capacidade de julgamento moral. Isto pode incluir desenvolvimento de ambientes de trabalho que minimizam estresse desnecessário, criação de sistemas de suporte que ajudam pesquisadores a gerenciar pressões de forma saudável, e implementação de políticas que reconhecem limitações humanas e promovem sustentabilidade de carreira.

Intervenções podem incluir treinamento em técnicas de gerenciamento de estresse, criação de culturas institucionais que valorizam bem-estar e equilíbrio, desenvolvimento de sistemas de mentoria que fornecem suporte emocional além de orientação técnica, e implementação de políticas que reduzem pressões desnecessárias e promovem colaboração sobre competição destrutiva.

A prevenção sistêmica também requer reconhecimento de que integridade científica não pode ser separada de questões mais amplas de justiça social, equidade, e sustentabilidade do sistema científico. Quando pesquisadores enfrentam discriminação, marginalização, ou acesso desigual a recursos e oportunidades, podem experimentar estresse adicional que compromete capacidade de manter padrões éticos elevados.

4.5.4 RESPONSABILIDADE INSTITUCIONAL E TRANSFORMAÇÃO CULTURAL: ALÉM DA RESPONSABILIZAÇÃO INDIVIDUAL

A magnitude e sistematicidade da crise contemporânea de integridade científica revelam falhas estruturais profundas que transcendem responsabilidades individuais e exigem reformas

institucionais abrangentes que abordem as raízes sistêmicas de incentivos perversos, culturas disfuncionais, e estruturas organizacionais que facilitam ou normalizam má conduta (Research Policy, 2024). Esta perspectiva sistêmica reconhece que, embora a responsabilidade individual permaneça importante, focar exclusivamente em educação ética individual ou punição de transgressores é inadequado para enfrentar problemas que emergem de características estruturais do sistema científico contemporâneo.

Universidades, agências de financiamento, editoras científicas, organizações profissionais, e governos devem reconhecer responsabilidade coletiva na criação de incentivos, estruturas, e culturas que facilitam má conduta, bem como na implementação de reformas sistêmicas que promovem integridade como componente integral de excelência científica. Esta responsabilização institucional requer não apenas mudanças em políticas formais, mas transformação cultural profunda que redefine valores, prioridades, e práticas em todos os níveis do sistema científico.

Sistemas de avaliação acadêmica que priorizam quantidade sobre qualidade, métricas quantitativas sobre impacto real, produtividade individual sobre colaboração, e resultados de curto prazo sobre sustentabilidade de longo prazo criam pressões estruturais que incentivam práticas antiéticas e desencorajam comportamentos que promovem integridade científica genuína. Reformas necessárias incluem desenvolvimento de métricas qualitativas que capturam contribuições reais para conhecimento científico, valorização explícita de reprodutibilidade e transparência como critérios de excelência, reconhecimento de mentoria e colaboração como contribuições e criação de incentivos para Ciência Aberta e práticas que promovem verificação independente.

Agências de financiamento têm responsabilidade particular na promoção da integridade através de políticas que exigem compartilhamento de dados, pré-registro de estudos, treinamento em

ética, colaborações equitativas, e práticas que promovem transparência e reprodutibilidade. Políticas de financiamento podem criar incentivos sistêmicos positivos que vão além de prevenção de má conduta para promover ativamente excelência científica através de práticas íntegras.

Isto pode incluir financiamento específico para atividades que promovem integridade científica, como desenvolvimento de ferramentas de detecção, treinamento em ética, criação de infraestruturas para compartilhamento de dados, e apoio para pesquisa sobre métodos que promovem reprodutibilidade. Agências podem também implementar políticas que valorizam qualidade sobre quantidade, apoiam colaboração internacional equitativa, e reconhecem contribuições diversificadas para conhecimento científico.

Responsabilização institucional deve abordar como estruturas de poder facilitam ou desencorajam má conduta. Hierarquias rígidas, culturas de competição destrutiva, falta de transparência em processos de tomada de decisão, ausência de mecanismos seguros para reportar preocupações éticas, e tolerância de comportamentos abusivos ou discriminatórios podem criar ambientes onde má conduta floresce sem detecção ou onde indivíduos que tentam manter padrões éticos enfrentam retaliação ou marginalização.

Transformação cultural eficaz requer mudanças coordenadas em múltiplos níveis que incluem liderança institucional comprometida com integridade, políticas que alinham incentivos com valores éticos, criação de ambientes onde questões éticas podem ser discutidas abertamente sem medo de retaliação, e desenvolvimento de culturas que celebram e recompensam práticas íntegras como componentes essenciais de excelência científica.

Esta transformação deve reconhecer que cultura institucional é moldada não apenas por políticas formais, mas também por práticas informais, narrativas compartilhadas, e modelos de comportamento

demonstrados por líderes e mentores. Mudança cultural sustentável requer não apenas mudanças em regras e procedimentos, mas também transformação em valores, expectativas, e práticas cotidianas que definem como ciência é conduzida e avaliada.

A transformação deve também abordar dimensões de diversidade, equidade, e inclusão, reconhecendo que culturas que marginalizam ou discriminam contra grupos específicos podem criar condições que facilitam má conduta e desencorajam reportagem de problemas. Criação de ambientes verdadeiramente inclusivos e equitativos não é apenas uma questão de justiça social, mas também uma estratégia importante para promover integridade científica através de diversidade de perspectivas, redução de pressões que podem levar a comportamento antiético, e criação de culturas onde todos os membros se sentem valorizados e responsáveis por manter padrões éticos elevados.

4.6 PERSPECTIVAS FUTURAS E RECOMENDAÇÕES ESTRATÉGICAS: CONSTRUINDO UM SISTEMA CIENTÍFICO ÍNTEGRO E EQUITATIVO

Até este ponto de nossa análise, nossa jornada foi, em grande parte, um diagnóstico. Mapeamos a ascensão industrial da fraude científica, dissecamos as táticas sofisticadas das *paper mills* e, na seção anterior, confrontamos as consequências sociais devastadoras dessa crise: a profunda erosão da confiança pública e a perigosa fragmentação do nosso consenso sobre a realidade.

Diante de um quadro tão grave, a pergunta inevitável e urgente é: E agora? O que fazemos?

É para responder a essa pergunta que adentramos nas «Perspectivas futuras e recomendações estratégicas». Esta seção marca uma virada fundamental em nossa abordagem, saindo do campo do problema para o campo da solução. No entanto, a premissa central que guia esta parte final do nosso trabalho é uma rejeição clara a soluções simplistas.

Conforme visto anteriormente, a crise de integridade é um problema sistêmico, multifacetado e global. Portanto, seria ingênuo e ineficaz propor “remendos” isolados, como uma nova ferramenta de *software* ou uma única mudança regulatória.

A abordagem que será apresentada aqui é radicalmente diferente. Propomos uma visão integrada para a governança da integridade científica. O que isso significa?

Significa que a solução deve ser tão interconectada quanto o problema. Ela deve operar em múltiplas escalas simultaneamente:

- No nível individual, focando na ética e formação do pesquisador.
- No nível institucional, reformando os sistemas de incentivo dentro das universidades e centros de pesquisa.
- No nível nacional, através de políticas de fomento e regulação.
- E, crucialmente, no nível internacional, com uma cooperação transnacional capaz de enfrentar redes de fraude que não respeitam fronteiras.

Mais importante ainda, esta seção argumenta que a integridade não pode ser desassociada da equidade. Um sistema científico verdadeiramente íntegro deve ser também um sistema justo. Isso nos obriga a ir além da prevenção da fraude e a questionar as estruturas de poder, as desigualdades globais e as dinâmicas de “colonialismo científico” que, muitas vezes, criam a pressão e as vulnerabilidades que permitem que a má conduta floresça.

Portanto, o que se segue não é uma lista de desejos, mas um roteiro estratégico. O objetivo não é apenas restaurar o sistema científico ao que ele era, mas sim construir um ecossistema científico fundamentalmente novo: mais robusto, mais transparente e, acima de tudo, mais íntegro e equitativo, capaz de verdadeiramente servir ao bem público em escala global.

4.6.1 VISÃO INTEGRADA PARA GOVERNANÇA DA INTEGRIDADE CIENTÍFICA: COORDENAÇÃO MULTI-ESCALAR E ABORDAGEM SISTÊMICA

O enfrentamento eficaz da crise de integridade científica requer uma abordagem fundamentalmente integrada que reconheça a natureza complexa, multifacetada, e profundamente interconectada dos desafios contemporâneos, rejeitando soluções isoladas ou incrementais em favor de estratégias coordenadas que abordem simultaneamente dimensões tecnológicas, institucionais, culturais, econômicas, e geopolíticas através de uma arquitetura de governança que opera eficazmente em múltiplas escalas temporais e espaciais (Science Policy, 2024). Esta visão integrada reconhece que a integridade científica não pode ser tratada como um problema técnico isolado, mas deve ser compreendida como parte de um sistema mais amplo de produção, validação, e disseminação de conhecimento que está profundamente entrelaçado com estruturas de poder, desigualdades globais, e dinâmicas culturais que moldam como ciência é conduzida, avaliada e utilizada.

Uma governança eficaz deve operar através de múltiplas escalas coordenadas que incluem o nível individual (formação ética, desenvolvimento de competências, e responsabilização pessoal), institucional (políticas organizacionais, culturas que promovem integridade, e sistemas de incentivos alinhados), nacional (regulamentações governamentais, sistemas de financiamento,

e frameworks legais), e internacional (cooperação transnacional, padronização global, e mecanismos de coordenação que transcendem fronteiras jurisdicionais). Esta abordagem multi-escalar reconhece que fraudes contemporâneas, particularmente operações industriais como *paper mills*, operam através de redes transnacionais que exploram diferenças regulatórias, culturais, e econômicas entre países, exigindo respostas coordenadas que podem abordar essas operações em todas as escalas relevantes.

A integração eficaz requer desenvolvimento de mecanismos de coordenação sofisticados que possam alinhar incentivos através de diferentes níveis e setores, facilitar compartilhamento de informações e melhores práticas, e promover ação coletiva sem comprometer autonomia institucional, diversidade cultural, ou inovação científica que são essenciais para vitalidade e criatividade do empreendimento científico. Isto inclui criação de padrões internacionais flexíveis que podem ser adaptados a contextos locais, sistemas de compartilhamento de informações que protegem privacidade e propriedade intelectual, e mecanismos de financiamento que apoiem desenvolvimento de capacidades em países e instituições com recursos limitados.

A governança integrada deve reconhecer explicitamente que integridade científica está intrinsecamente conectada a questões mais amplas de justiça social, equidade global, e sustentabilidade do sistema científico como bem público que deve servir às necessidades da humanidade como um todo, não apenas interesses de elites científicas ou econômicas. Estratégias devem abordar não apenas comportamentos individuais problemáticos, mas estruturas sistêmicas de poder e desigualdade que moldam condições sob as quais ciência é produzida, validada, e disseminada, incluindo dinâmicas de colonialismo científico, extração de dados, e marginalização de conhecimentos tradicionais e perspectivas não-ocidentais.

Isso implica desenvolvimento de frameworks de governança que possam equilibrar autonomia científica com responsabilidade social, promover colaboração internacional genuinamente equitativa que beneficia todas as partes, proteger conhecimentos tradicionais e soberania científica de países em desenvolvimento, e criar mecanismos para participação significativa de comunidades afetadas por pesquisa científica em processos de tomada de decisão sobre prioridades, métodos, e aplicações de pesquisa.

O objetivo deve ser criar um sistema científico global que seja simultaneamente mais íntegro em seus métodos e práticas, mais justo em sua distribuição de benefícios e oportunidades, mais sustentável em seu uso de recursos e impacto ambiental, e mais responsivo às necessidades e valores de comunidades diversas ao redor do mundo. Esta visão requer reconhecimento de que integridade científica não é apenas uma questão técnica de prevenção de fraudes, mas um componente essencial de um sistema científico que pode contribuir efetivamente para bem-estar humano e justiça global.

4.6.2 RECOMENDAÇÕES TECNOLÓGICAS: INOVAÇÃO RESPONSÁVEL, INTERPRETABILIDADE E DEMOCRATIZAÇÃO

O desenvolvimento de ferramentas de detecção de próxima geração deve priorizar interpretabilidade, transparência algorítmica, e responsabilização humana como princípios fundamentais que guiam não apenas design técnico, mas também implementação, governança, e evolução contínua desses sistemas (Nature Machine Intelligence, 2024). Sistemas que podem explicar suas decisões de forma compreensível para usuários humanos, fornecer evidência específica para suas recomendações, e permitir supervisão e correção humana eficazes representam não apenas avanços técnicos, mas também requisitos éticos essenciais para manter confiança pública e responsabilização democrática em sistemas automatizados que afetam carreiras científicas e validação de conhecimento.

Esta ênfase em interpretabilidade vai além de questões técnicas para abordar preocupações fundamentais sobre poder algorítmico, vieses sistêmico, e potencial para uso indevido de tecnologias de detecção. Sistemas opacos que produzem recomendações sem explicação podem perpetuar ou amplificar vieses existentes, criar vulnerabilidades para manipulação ou sabotagem, e minar confiança pública em processos de validação científica. Transparência algorítmica permite não apenas detecção e correção de problemas, mas também melhoria contínua através de feedback da comunidade científica e adaptação a contextos e necessidades diversificadas.

Investimento em infraestrutura de dados científicos representa uma prioridade para criar fundações técnicas sólidas para verificação independente, reprodutibilidade, e colaboração científica responsável. Repositórios seguros, interoperáveis, e sustentáveis podem facilitar verificação independente de resultados, permitir reanálise de dados por pesquisadores independentes, e promover transparência científica enquanto protegem propriedade intelectual legítima, privacidade de participantes de pesquisa, e segurança nacional quando apropriado.

Desenvolvimento de padrões internacionais para uso ético de IA em pesquisa deve ser uma prioridade urgente que envolve não apenas comunidades técnicas e especialistas em ética, mas também representantes de comunidades afetadas, formuladores de políticas, e organizações da sociedade civil que podem fornecer perspectivas essenciais sobre implicações sociais e éticas de tecnologias emergentes (IEEE Standards, 2024). Padrões devem abordar tanto aplicações legítimas quanto prevenção de uso fraudulento, estabelecendo diretrizes claras para transparência sobre uso de IA, responsabilização por resultados gerados ou analisados por sistemas automatizados, e auditabilidade de processos que incorporam IA.

Isso inclui desenvolvimento de protocolos específicos para divulgação de uso de IA em pesquisa científica, padrões para validação

de resultados gerados por IA, mecanismos para garantir que sistemas não reproduzam ou ampliem vieses existentes, e frameworks para avaliação contínua de impactos sociais e éticos de tecnologias emergentes. Padrões devem ser desenvolvidos através de processos participativos que incluem vozes diversas e devem ser regularmente atualizados para refletir avanços tecnológicos e compreensão de implicações éticas.

Democratização de ferramentas de detecção é essencial para garantir que capacidades avançadas não sejam monopolizadas por instituições com recursos substanciais, mas estejam disponíveis para pesquisadores, editores, e instituições em países em desenvolvimento e contextos com recursos limitados. Isto pode incluir desenvolvimento de ferramentas open-source, criação de serviços baseados em nuvem que reduzem barreiras de entrada, e programas de transferência de tecnologia que apoiam desenvolvimento de capacidades locais.

4.6.3 REFORMAS INSTITUCIONAIS: TRANSFORMAÇÃO DE INCENTIVOS, CULTURAS, E ESTRUTURAS DE PODER

Sistemas de avaliação acadêmica requerem reforma fundamental que vá muito além de ajustes incrementais para abordar raízes sistêmicas de incentivos perversos que priorizam quantidade sobre qualidade, métricas simples sobre impacto real, e produtividade individual sobre colaboração e contribuições para bem comum científico (DORA, 2024). Esta reforma deve incluir desenvolvimento de métricas qualitativas sofisticadas que capturam contribuições reais para conhecimento científico, valorização explícita de reprodutibilidade e transparência como critérios centrais de excelência, reconhecimento de mentoria e desenvolvimento de recursos humanos como contribuições valiosas, e criação de incentivos robustos para Ciência Aberta, compartilhamento de dados, e práticas que facilitam verificação independente.

Reforma deve incluir mudanças estruturais profundas nos processos de revisão por pares para torná-los mais robustos, transparentes, equitativos, e capazes de detectar problemas sofisticados que caracterizam má conduta contemporânea. Isso pode incluir implementação de revisão por pares aberta que aumenta responsabilização e transparência, sistemas de revisão pós-publicação que permitem avaliação contínua de qualidade científica, mecanismos para reconhecer e recompensar revisores por contribuições de alta qualidade, e desenvolvimento de processos que podem lidar eficazmente com alegações de má conduta e investigações subsequentes.

Transformação de culturas institucionais requer mudanças coordenadas que vão além de políticas formais para abordar normas informais, expectativas implícitas, e práticas cotidianas que definem como ciência é realmente conduzida e avaliada. Isto inclui criação de ambientes onde questões éticas podem ser discutidas abertamente sem medo de retaliação, desenvolvimento de liderança que modela e promove comportamento ético, e estabelecimento de sistemas de suporte que ajudam pesquisadores a navegar dilemas éticos e pressões profissionais de forma saudável e sustentável.

Políticas institucionais abrangentes e responsabilização sistêmica

Universidades devem implementar políticas abrangentes que vão muito além de prevenção de má conduta para promover ativamente práticas íntegras como componentes essenciais de excelência científica e responsabilidade social. Isso inclui criação de escritórios de integridade com recursos adequados e autoridade real, implementação de treinamento obrigatório em ética que é relevante, atualizado, e contextualizado para desafios contemporâneos, estabelecimento de mecanismos seguros e eficazes para reportar preocupações éticas, e desenvolvimento de culturas institucionais que valorizam transparência, colaboração, e responsabilidade social como valores fundamentais.

Agências de financiamento têm papel fundamental na transformação de incentivos sistêmicos através de políticas que exigem e recompensam práticas que promovem integridade científica, incluindo compartilhamento obrigatório de dados e código, pré-registro de estudos para reduzir viés de publicação, colaborações equitativas que beneficiam todas as partes, e avaliação baseada em qualidade real e impacto social ao invés de métricas quantitativas simples. Isso pode incluir financiamento específico para atividades que promovem integridade científica, como desenvolvimento de ferramentas de detecção, treinamento em ética, criação de infraestruturas para Ciência Aberta, e apoio para pesquisa sobre métodos que promovem reprodutibilidade e transparência.

Responsabilização sistêmica deve abordar como estruturas de poder institucional facilitam ou desencorajam má conduta, incluindo hierarquias que podem silenciar vozes críticas, culturas de competição destrutiva que incentivam comportamento antiético, falta de transparência em processos de tomada de decisão que afetam carreiras científicas, e tolerância de comportamentos abusivos ou discriminatórios que podem criar ambientes onde má conduta floresce sem detecção ou onde indivíduos que tentam manter padrões éticos enfrentam marginalização ou retaliação.

4.6.4 COOPERAÇÃO INTERNACIONAL E SOBERANIA CIENTÍFICA: EQUIDADE GLOBAL E JUSTIÇA EPISTÊMICA

A natureza fundamentalmente transnacional da má conduta científica contemporânea, particularmente operações industriais como *paper mills* que exploram diferenças regulatórias e econômicas entre países, exige cooperação internacional robusta e coordenada que respeite soberania científica e diversidade cultural enquanto estabelece padrões mínimos comuns e mecanismos eficazes para ação coletiva (Research Integrity, 2024). Esta cooperação deve ir

além de acordos técnicos para abordar desigualdades estruturais que tornam alguns países e comunidades mais vulneráveis à exploração, incluindo desenvolvimento de capacidades, transferência de tecnologia, e apoio para fortalecimento de sistemas nacionais de integridade científica.

Cooperação deve ser baseada em princípios de equidade e benefício mútuo que reconhecem que países em desenvolvimento não são apenas recipientes passivos de assistência, mas parceiros ativos que podem contribuir conhecimentos, perspectivas, e inovações para esforços globais de promoção de integridade científica. Isto inclui reconhecimento de conhecimentos tradicionais e sistemas de conhecimento não-ocidentais como contribuições legítimas para compreensão científica, desenvolvimento de mecanismos para participação equitativa em governança científica global, e criação de oportunidades para liderança científica de países em desenvolvimento em áreas de sua expertise e prioridade.

Desenvolvimento de tratados internacionais e frameworks legais para combate à má conduta científica deve equilibrar necessidade de ação coordenada com respeito por soberania nacional e diversidade de sistemas legais e culturais. Isto pode incluir acordos para compartilhamento de informações sobre operações fraudulentas, protocolos para investigação coordenada de casos transnacionais, e mecanismos para responsabilização de operações que exploram diferenças jurisdicionais para evitar detecção ou punição.

Soberania Científica e Desenvolvimento de Capacidades

O conceito de soberania científica deve garantir que países tenham controle adequado sobre pesquisas conduzidas em seus territórios, dados coletados de suas populações, conhecimentos tradicionais de suas comunidades, e prioridades de pesquisa que refletem suas necessidades e valores nacionais. Isso inclui desenvolvimento de protocolos para colaboração científica equitativa que

garante benefício compartilhado, mecanismos para proteção de conhecimentos tradicionais contra apropriação indevida, e sistemas para garantir que pesquisa internacional contribua para desenvolvimento nacional e bem-estar comunitário.

Cooperação deve focar no desenvolvimento de capacidades científicas locais que possam reduzir dependência de literatura externa, criar sistemas mais resilientes de produção e validação de conhecimento, e promover inovação científica que é relevante para contextos e necessidades locais. Isso inclui apoio para educação científica de alta qualidade, desenvolvimento de infraestrutura de pesquisa, criação de redes de colaboração Sul-Sul que reduzem dependência de instituições do Norte Global, e transferência de tecnologias de detecção e prevenção de má conduta que podem ser adaptadas para contextos locais.

Programas de desenvolvimento de capacidades devem ser projetados com participação significativa de países recipientes e devem abordar não apenas aspectos técnicos, mas também desenvolvimento de liderança científica, criação de culturas institucionais que promovem integridade, e estabelecimento de sistemas de governança que são apropriados para contextos locais. O objetivo deve ser criar sistemas científicos nacionais que são simultaneamente conectados globalmente e enraizados localmente, capazes de participar eficazmente em colaboração internacional enquanto servem às necessidades e prioridades nacionais.

4.6.5 *ROADMAP* DE IMPLEMENTAÇÃO: ESTRATÉGIAS INTEGRADAS DE CURTO, MÉDIO E LONGO PRAZO PARA TRANSFORMAÇÃO SISTÊMICA DA INTEGRIDADE CIENTÍFICA

A superação da crise de integridade científica, em sua forma atual, exige uma resposta que espelhe a profundidade e a complexidade do problema. Não se trata de uma questão que possa ser

resolvida com soluções pontuais, mas sim de uma falha sistêmica que demanda um *roadmap* de implementação abrangente e cuidadosamente sequenciado. A premissa fundamental deste plano é o reconhecimento de que qualquer esforço de reforma deve ser robusto o suficiente para resistir às pressões políticas, econômicas e culturais que historicamente minaram tentativas anteriores. Para isso, a estratégia se desdobra em três horizontes temporais integrados — curto, médio e longo prazo — permitindo uma abordagem que equilibra a urgência da ação imediata com a necessidade de construir uma mudança sustentável e profunda (Science Policy Research, 2024). Esta abordagem integrada atua em múltiplos pontos de alavancagem do sistema simultaneamente, evitando a armadilha de soluções unidimensionais que, embora possam trazer alívio temporário, falham em corrigir as causas estruturais subjacentes.

Tabela 4 - Roadmap Visual para a Transformação Sistêmica da Integridade Científica

Eixo de Ação	Curto Prazo (Até 2 anos) - Contenção e Estabilização	Médio Prazo (2 a 5 anos) - Reconstrução e Fortalecimento	Longo Prazo (Acima de 5 anos) - Transformação Paradigmática
Tecnológico	• Implementar em larga escala ferramentas de detecção automatizada de fraudes (ex: <i>Problematic Paper Screener</i>) para conter a produção das <i>paper mills</i>.	• Desenvolver sistemas de detecção de próxima geração baseados em IA interpretável. • Criar infraestruturas de dados para transparência e reprodutibilidade (repositórios seguros, uso de <i>blockchain</i> para registros imutáveis).	• Consolidar novos modelos de produção científica onde a Ciência Aberta e a colaboração equitativa são o padrão. • Desenvolver novos <i>frameworks</i> epistemológicos para distinguir entre evidência empírica e simulação sintética na era da IA.
Institucional/ Político	• Criar redes de resposta rápida entre instituições para compartilhar inteligência sobre fraudes e interromper operações transnacionais. • Implementar moratórias em bônus financeiros baseados exclusivamente na quantidade de publicações.	• Reformar o sistema de revisão por pares, tornando-o mais aberto, transparente e robusto. • Estabelecer escritórios de integridade científica com recursos e autoridade para investigar e agir. • Promover programas de capacitação para o Sul Global e incentivar redes de colaboração Sul-Sul.	• Criar sistemas de validação de pesquisa mais inclusivos, que reconheçam saberes tradicionais e comunitários. • Integrar a governança da ciência a desafios globais, como direitos humanos e justiça climática.
Cultural	• Realizar treinamento intensivo de editores e revisores sobre novas táticas de fraude. • Valorizar explicitamente nos sistemas de avaliação atividades como mentoria ética e revisão por pares de qualidade.	• Implementar programas de transformação cultural nas instituições para questionar normas e estruturas de poder que incentivam a má conduta.	• Fomentar culturas científicas autorregulatórias, onde a integridade seja uma norma internalizada. • Adotar a relevância social como critério central de avaliação da pesquisa, ao lado do rigor técnico

Fonte: Elaborado pelo autor (2025).

No horizonte de curto prazo, o foco é na estabilização e na contenção de danos, uma espécie de “triagem de emergência” para o ecossistema científico. A prioridade mais imediata é a implementação ampla de ferramentas de detecção automatizada, como o *Problematic Paper Screener*. Esta é uma medida de defesa direta, projetada para “estancar a hemorragia” de artigos fraudulentos que poluem a literatura, atacando o modelo de negócio das *paper mills* em seu ponto mais vulnerável: a dependência de baixas taxas de detecção para sua viabilidade econômica. Contudo, a tecnologia por si só é insuficiente; sua implementação deve ser acompanhada por um treinamento intensivo de editores e revisores e pela criação de redes de resposta rápida, que funcionem como um sistema imunológico coordenado, capaz de compartilhar inteligência sobre novas táticas de fraude e de interromper operações criminosas que exploram fronteiras jurisdicionais. Simultaneamente, é preciso atacar a “demanda” por fraude, implementando reformas urgentes nos sistemas de avaliação acadêmica. Isso inclui ações como moratórias em bônus financeiros baseados apenas na quantidade de publicações, valorizando explicitamente atividades como a mentoria ética e a revisão por pares de qualidade como contribuições essenciais para a excelência acadêmica e alterando o cálculo de risco-recompensa para práticas antiéticas.

Superada essa fase de contenção, o *roadmap* avança para o horizonte de médio prazo, focado em reconstruir e fortalecer as fundações do sistema. Esta é a fase de investimentos substanciais em infraestrutura, cultura e capacidades institucionais. No pilar tecnológico, o plano prevê a criação de sistemas de detecção de próxima geração, baseados em IA interpretável, e o desenvolvimento de infraestruturas de dados que promovam a transparência e a reprodutibilidade, como repositórios seguros e o uso de tecnologias como blockchain para criar registros criptográficos imutáveis, fornecendo uma nova camada de confiança matemática. No pilar humano e organizacional, a ênfase está em programas de transformação cultural,

que vão além de simples treinamentos para abordar as normas e as estruturas de poder dentro das instituições. Isso se materializa na reforma da revisão por pares, tornando-a mais aberta e robusta, e no estabelecimento de escritórios de integridade científica com recursos e autoridade real para agir. De forma essencial, esta fase aborda a desigualdade global, promovendo programas de capacitação para o Sul Global e incentivando redes de colaboração Sul-Sul, como uma estratégia prática para reduzir a dependência epistêmica e construir uma ciência mais equitativa.

Finalmente, com as fundações fortalecidas, o *roadmap* projeta sua visão mais ambiciosa no horizonte de longo prazo: a transformação paradigmática da ciência. O objetivo aqui não é apenas consertar um sistema quebrado, mas criar um novo, fundamentalmente mais íntegro, justo e responsivo. Isso implica a implementação de novos modelos de produção científica onde a Ciência Aberta e a colaboração equitativa sejam o padrão. Requer a criação de sistemas de validação mais inclusivos, que reconheçam múltiplas formas de expertise, incluindo conhecimentos tradicionais e comunitários, avaliando não apenas o rigor técnico, mas a relevância social. Exige o desenvolvimento de novos frameworks epistemológicos para navegar na era da IA, distinguindo com clareza entre evidência empírica e simulação sintética. A meta final é fomentar culturas científicas autorregulatórias, onde a integridade seja uma norma internalizada, e integrar a governança da ciência a desafios globais mais amplos, como os direitos humanos e a justiça climática. A visão é a de uma ciência que, ao garantir sua própria integridade, se posiciona de forma mais eficaz para servir ao bem-estar da humanidade de maneira sustentável e justa.

Parte



ESTUDOS DE CASO E EXPERIÊNCIAS PRÁTICAS

5

**ESTUDOS DE CASO
INTERNACIONAIS**

Este capítulo examina casos concretos que ilustram as tensões entre Ciência Aberta e colonialismo de dados em diferentes contextos geográficos e disciplinares. Através da análise desses casos, podemos compreender melhor como essas dinâmicas operam na prática e quais lições podem ser extraídas para orientar políticas e práticas futuras. Após estabelecermos os fundamentos teóricos e analisarmos os impactos sistêmicos, como a crise de integridade, examinaremos agora como essas dinâmicas se manifestam em casos práticos ao redor do mundo, revelando a face concreta do colonialismo de dados.

Os casos selecionados representam diferentes dimensões do problema: desde violações diretas de direitos de comunidades vulneráveis até iniciativas bem-sucedidas de resistência e construção de alternativas. Cada caso oferece lições específicas sobre como navegar as complexidades entre abertura científica e justiça social, demonstrando que não existe uma solução única ou simples para esses desafios.

5.1. CASO 1: DADOS GENÉTICOS INDÍGENAS E DESRESPEITO AO CONSENTIMENTO (O CASO HAVASUPAI)

O caso da tribo Havasupai representa um dos exemplos mais emblemáticos de como a pesquisa científica pode violar direitos fundamentais de comunidades indígenas através da apropriação indevida de dados genéticos. Este caso, que se desenrolou ao longo de mais de uma década, oferece lições importantes sobre consentimento informado, soberania de dados, e a necessidade de frameworks éticos específicos para pesquisa com populações vulneráveis. Mais do que um simples caso de má conduta científica,

o caso Havasupai ilustra as dimensões estruturais do colonialismo de dados e como práticas aparentemente neutras de pesquisa podem perpetuar relações coloniais de poder.

A tribo Havasupai, cujo nome significa “povo da água azul-verde”, habita há séculos as profundezas do Grand Canyon no Arizona. Para compreender adequadamente as violações éticas que ocorreram, é fundamental entender o contexto histórico e cultural desta comunidade. Os Havasupai mantêm uma relação sagrada com o Grand Canyon, que consideram não apenas seu lar atual, mas o local de suas origens ancestrais e centro de sua cosmologia espiritual.

Como muitas comunidades indígenas americanas, os Havasupai enfrentaram séculos de colonização, deslocamento forçado e tentativas de assimilação cultural. Essas experiências históricas criaram uma desconfiança justificada em relação a instituições externas, incluindo universidades e pesquisadores. Quando a tribo decidiu colaborar com pesquisadores da Universidade do Arizona na década de 1990, foi um ato de confiança baseado na esperança de que a ciência poderia ajudar a resolver problemas de saúde que afetavam desproporcionalmente sua comunidade.

A tribo Havasupai, que vive no Grand Canyon no Arizona, enfrentava altas taxas de diabetes tipo 2 que afetavam significativamente a saúde e bem-estar da comunidade. Em 1990, líderes tribais procuraram pesquisadores da Universidade do Arizona para colaborar em estudos que pudessem ajudar a compreender e abordar esse problema de saúde pública. Como relatam Mello e Wolf, “em 1990, a tribo Havasupai do Arizona deu amostras de sangue para pesquisadores da Universidade do Arizona para estudos sobre diabetes, uma doença que afetava desproporcionalmente sua comunidade” (MELLO; WOLF, 2010).

O consentimento inicial foi obtido para pesquisas sobre “distúrbios comportamentais/médicos”, uma linguagem que a tribo interpretou como relacionada especificamente ao diabetes.

Esta interpretação não era irrazonável, dado o contexto das discussões iniciais e as preocupações específicas da comunidade. No entanto, os pesquisadores posteriormente utilizaram essas amostras para uma gama muito mais ampla de estudos, incluindo pesquisas sobre esquizofrenia, migração populacional, e consanguinidade.

A descoberta dessas violações ocorreu de forma quase acidental. Em 2003, Carletta Tilousi, membro da tribo, descobriu durante uma apresentação acadêmica que as amostras de sangue de sua comunidade estavam sendo utilizadas para pesquisas que iam muito além do escopo original. Esta descoberta desencadeou uma investigação mais ampla que revelou a extensão das violações éticas.

As violações no caso Havasupai foram múltiplas e interconectadas, demonstrando como o colonialismo de dados opera através de camadas sobrepostas de exploração. Primeiro, houve uma clara violação do consentimento informado, pois as amostras foram utilizadas para fins que não foram adequadamente explicados ou autorizados pela tribo. Esta violação não foi meramente técnica, mas refletiu uma compreensão fundamentalmente diferente sobre a natureza do consentimento e a propriedade dos dados biológicos.

Segundo, os estudos sobre migração contradiziam diretamente as crenças tradicionais da tribo sobre suas origens, causando angústia cultural significativa. Como observam Mello e Wolf, "os estudos sobre migração eram particularmente problemáticos porque contradiziam as crenças tradicionais da tribo sobre suas origens no Grand Canyon" (MELLO; WOLF, 2010). Para a tribo Havasupai, o Grand Canyon não é apenas seu lar atual, mas o local sagrado de suas origens ancestrais. Pesquisas que sugeriam migração de outras regiões não apenas contradiziam essas crenças, mas também poderiam potencialmente minar reivindicações territoriais legais e a identidade cultural fundamental da comunidade.

Terceiro, os estudos sobre consanguinidade e esquizofrenia levantaram preocupações sobre estigmatização e discriminação. Resultados que sugerissem predisposições genéticas para doenças mentais ou altos níveis de consanguinidade poderiam ser utilizados para justificar discriminação contra membros da tribo, perpetuando estereótipos racistas sobre povos indígenas.

O caso Havasupai ilustra várias dimensões do colonialismo de dados que vão além das violações éticas individuais. Primeiro, demonstra como o conhecimento científico pode ser extraído de comunidades marginalizadas sem benefício recíproco. Os dados genéticos dos Havasupai foram utilizados para produzir conhecimento que beneficiou principalmente pesquisadores e instituições acadêmicas, enquanto a comunidade que forneceu os dados não recebeu os benefícios prometidos.

Segundo, o caso revela como práticas aparentemente neutras de pesquisa podem perpetuar relações coloniais de poder. Os pesquisadores operaram dentro de um paradigma científico que privilegia a objetividade e a universalidade do conhecimento, ignorando as perspectivas e valores específicos da comunidade Havasupai. Esta abordagem reflete o que Santos (2014) chama de “epistemicídio” - a supressão de formas de conhecimento não-ocidentais em favor de paradigmas científicos dominantes.

Terceiro, o caso demonstra como dados sobre indivíduos podem afetar comunidades inteiras. Os dados genéticos não são meramente individuais, mas carregam informações sobre famílias, linhagens e populações. Quando esses dados são utilizados de forma inadequada, as consequências se estendem muito além dos indivíduos que forneceram as amostras originais.

O caso Havasupai resultou em uma batalha legal prolongada que se estendeu por mais de uma década. Em 2010, a Universidade do Arizona concordou em pagar \$700.000 em compensação para a tribo e devolver as amostras de sangue restantes (MELLO; WOLF, 2010).

Mais importante do que a compensação financeira, o caso estabeleceu precedentes importantes para pesquisa com populações indígenas e contribuiu para mudanças significativas nas práticas de pesquisa.

Como resultado do caso, muitas instituições desenvolveram protocolos mais rigorosos para pesquisa com comunidades indígenas, incluindo requisitos para consentimento comunitário além do consentimento individual, e maior envolvimento de comunidades na definição de prioridades de pesquisa. O caso também contribuiu para o desenvolvimento dos princípios CARE (Collective benefit, Authority to control, Responsibility, Ethics) para dados indígenas, que complementam os princípios FAIR com considerações específicas sobre soberania e autodeterminação indígena.

Os princípios CARE, desenvolvidos pela Global Indigenous Data Alliance, representam uma contribuição fundamental para repensar a governança de dados de forma mais inclusiva e equitativa (CARROLL *et al.*, 2020). O princípio do Benefício Coletivo estabelece que dados sobre povos indígenas devem ser utilizados de formas que beneficiem as comunidades que os produziram. A Autoridade para Controlar reconhece o direito de povos indígenas de controlar dados sobre suas comunidades, territórios, e recursos. O princípio da Responsabilidade estabelece que aqueles que trabalham com dados indígenas têm responsabilidades específicas de garantir que esses dados sejam utilizados de forma respeitosa e benéfica. A Ética reconhece que dados indígenas estão conectados a valores, conhecimentos, e práticas culturais específicas que devem ser respeitadas.

O caso Havasupai oferece lições importantes para a era digital, onde dados podem ser facilmente copiados, compartilhados e reutilizados. A digitalização de dados genéticos cria novas possibilidades para violações de consentimento, pois dados coletados para um propósito podem ser facilmente reutilizados para outros fins. A crescente integração de dados genéticos com outras formas de dados pessoais também amplifica os riscos de violações de privacidade e autonomia.

Quando dados genéticos são combinados com dados de localização, redes sociais, ou registros médicos, eles podem revelar informações sensíveis que não eram previsíveis no momento da coleta original. Esta capacidade de combinação e análise de dados cria novos riscos que vão muito além daqueles contemplados pelos frameworks éticos tradicionais.

Para comunidades indígenas, esses riscos são amplificados pela natureza coletiva de dados genéticos. Informações sobre um indivíduo podem revelar características de toda a comunidade, criando responsabilidades coletivas que vão além do consentimento individual. Isso destaca a necessidade de frameworks éticos que reconheçam tanto direitos individuais quanto coletivos, como proposto pelos princípios CARE.

O caso também ilustra a importância de processos de consentimento culturalmente apropriados. Como argumentam Castleden *et al.* (2007), o consentimento para pesquisa com comunidades indígenas deve ir além de formulários padronizados para incluir processos que respeitam protocolos culturais específicos e reconhecem a autoridade de lideranças comunitárias.

O legado do caso Havasupai continua a influenciar a pesquisa genética contemporânea. Garrison *et al.* (2013) documentam como o caso levou a mudanças significativas nas práticas de pesquisa genética com populações indígenas, incluindo maior ênfase no retorno de resultados para comunidades e no desenvolvimento de parcerias colaborativas genuínas.

No entanto, desafios persistem. A crescente comercialização de dados genéticos através de empresas como 23andMe e AncestryDNA cria novos riscos para comunidades indígenas. Esses serviços frequentemente operam com modelos de consentimento que não reconhecem adequadamente os direitos coletivos ou as implicações culturais específicas de dados genéticos indígenas.

Além disso, o desenvolvimento de tecnologias como CRISPR e outras ferramentas de edição genética levanta novas questões sobre propriedade intelectual e benefício compartilhado. Quando conhecimentos tradicionais sobre plantas medicinais ou características genéticas específicas são utilizados para desenvolver tratamentos comerciais, como garantir que as comunidades que contribuíram com esse conhecimento sejam adequadamente reconhecidas e beneficiadas?

O caso Havasupai também contribuiu para o desenvolvimento do conceito de “justiça genômica” - a ideia de que os benefícios e riscos da pesquisa genética devem ser distribuídos de forma equitativa, e que comunidades historicamente marginalizadas devem ter controle sobre como seus dados genéticos são utilizados (GARRISON *et al.*, 2013).

Este conceito reconhece que a genética não é neutra, mas está inserida em contextos sociais, políticos e econômicos específicos. Dados genéticos podem ser utilizados para reforçar hierarquias raciais existentes ou para desafiar essas hierarquias, dependendo de como são coletados, analisados e interpretados.

A justiça genômica também reconhece que diferentes comunidades podem ter diferentes perspectivas sobre o que constitui benefício ou dano. Para algumas comunidades, a preservação da privacidade pode ser mais importante do que o acesso a tratamentos médicos. Para outras, o reconhecimento cultural e a soberania podem ser prioridades maiores do que benefícios econômicos diretos.

O caso Havasupai, portanto, não é apenas uma história sobre violações éticas passadas, mas um caso paradigmático que continua a informar debates contemporâneos sobre como conduzir pesquisa genética de forma ética e equitativa. Suas lições são particularmente relevantes à medida que entramos em uma era de medicina personalizada e big data genômico, onde os riscos e oportunidades associados a dados genéticos são amplificados exponencialmente.

5.2. CASO 2: SCIELO E A DEMOCRATIZAÇÃO DA COMUNICAÇÃO CIENTÍFICA NA AMÉRICA LATINA

O SciELO (Scientific Electronic Library Online) representa um dos exemplos mais bem-sucedidos de iniciativa regional de Ciência Aberta, demonstrando como países do Sul Global podem desenvolver alternativas viáveis ao sistema dominante de comunicação científica. Lançado em 1997 no Brasil, o SciELO expandiu-se para incluir 16 países e tornou-se uma referência global para acesso aberto. Mais do que uma simples plataforma de publicação, o SciELO representa um modelo alternativo de produção e circulação de conhecimento científico que desafia as estruturas hegemônicas do sistema editorial acadêmico global.

O SciELO surgiu em um contexto específico de marginalização da produção científica latino-americana no cenário global. Durante décadas, pesquisadores da região enfrentaram barreiras significativas para publicar em revistas internacionais de prestígio, não necessariamente devido à qualidade de sua pesquisa, mas devido a vieses geográficos, linguísticos e temáticos que privilegiavam pesquisas produzidas no Norte Global.

Simultaneamente, a produção científica regional permanecia largamente invisível para a comunidade científica internacional, criando um círculo vicioso onde a baixa visibilidade levava a menor citação, que por sua vez reforçava a percepção de menor qualidade ou relevância. Este fenômeno, que podemos caracterizar como uma forma de colonialismo epistêmico, limitava não apenas o reconhecimento de pesquisadores latino-americanos, mas também a circulação de conhecimentos relevantes para os contextos regionais.

O SciELO surgiu de uma colaboração entre a FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) e o Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (BIREME). Como observa Fressoli (2018), o projeto foi motivado pela necessidade de “aumentar a visibilidade, acessibilidade e credibilidade da literatura científica produzida na região”.

A iniciativa foi pioneira em vários aspectos fundamentais. Primeiro, estabeleceu padrões técnicos rigorosos para publicação eletrônica que eram comparáveis aos utilizados por editoras comerciais internacionais. Esta decisão foi estratégica: em vez de aceitar padrões inferiores que poderiam reforçar percepções de menor qualidade, o SciELO investiu em infraestrutura técnica de primeira linha.

Segundo, desenvolveu critérios de qualidade específicos para revistas da região que reconheciam suas características e necessidades particulares. Estes critérios não eram simplesmente adaptações de padrões internacionais, mas refletiam uma compreensão das especificidades da produção científica regional, incluindo a importância de pesquisas sobre problemas locais e a relevância de publicação em idiomas regionais.

Terceiro, criou uma infraestrutura técnica sustentável que não dependia de empresas comerciais estrangeiras. Esta independência tecnológica foi fundamental para garantir que a plataforma pudesse evoluir de acordo com as necessidades da comunidade científica regional, sem estar sujeita a decisões comerciais de empresas multinacionais.

O modelo de financiamento do SciELO baseia-se em recursos públicos e cooperação internacional, evitando tanto taxas de assinatura quanto taxas de processamento de artigos (APCs). Esta escolha foi fundamental para garantir que revistas de países com recursos limitados pudessem participar da rede sem barreiras financeiras significativas. Como argumentam Babini e Machin-Mastromatteo (2017), este modelo reflete uma filosofia específica sobre ciência aberta que prioriza o acesso universal sobre a sustentabilidade comercial.

O SciELO teve impactos transformadores na comunicação científica latino-americana que vão muito além do simples aumento de visibilidade. Primeiro, contribuiu para uma mudança fundamental na percepção da qualidade da pesquisa regional. Ao estabelecer padrões técnicos rigorosos e processos de avaliação transparentes, o SciELO demonstrou que a pesquisa latino-americana podia competir em termos de qualidade com a produção internacional.

Estudos bibliométricos mostram que artigos publicados em revistas SciELO recebem mais citações do que artigos similares publicados em revistas não-indexadas, indicando que a plataforma efetivamente aumentou a visibilidade da pesquisa regional. Mais importante, análises de padrões de citação revelam que o SciELO facilitou a criação de redes de citação intra-regionais, fortalecendo o diálogo científico entre países latino-americanos.

Segundo, o SciELO contribuiu significativamente para a profissionalização da edição científica na região. Muitas revistas melhoraram seus padrões editoriais e processos de revisão por pares para atender aos critérios de inclusão na rede. Este processo de profissionalização não foi meramente técnico, mas envolveu a adoção de práticas éticas mais rigorosas, maior transparência nos processos editoriais, e melhor gestão de conflitos de interesse.

Terceiro, a plataforma facilitou a colaboração científica intra-regional de formas que não eram possíveis anteriormente. Pesquisadores latino-americanos passaram a citar mais frequentemente trabalhos de colegas da região, não apenas porque esses trabalhos se tornaram mais visíveis, mas porque o SciELO criou uma infraestrutura que facilitava a descoberta de pesquisas relevantes produzidas regionalmente.

Um aspecto particularmente importante do SciELO é seu modelo de governança, que difere significativamente das plataformas comerciais dominantes. A rede SciELO é governada por um consórcio de instituições públicas e organizações sem fins lucrativos,

com representação de diferentes países participantes. Esta estrutura de governança garante que decisões sobre o desenvolvimento da plataforma reflitam as necessidades da comunidade científica regional, não interesses comerciais.

Este modelo de governança também permite que o SciELO mantenha compromissos com princípios de acesso aberto mesmo quando estes conflitam com pressões comerciais. Por exemplo, a plataforma mantém todos os artigos em acesso aberto permanente, sem períodos de embargo ou restrições baseadas em modelos de negócio.

Além disso, o modelo de governança do SciELO incorpora princípios de transparência e prestação de contas que são frequentemente ausentes em plataformas comerciais. Dados sobre uso, citações, e desempenho das revistas são disponibilizados publicamente, permitindo que a comunidade científica monitore e avalie o desempenho da plataforma.

Apesar de seus sucessos notáveis, o SciELO também enfrenta desafios significativos que ilustram as dificuldades de construir alternativas sustentáveis ao sistema dominante. A dependência de financiamento público torna a plataforma vulnerável a mudanças políticas e crises econômicas. Vários países reduziram ou eliminaram seu apoio ao SciELO durante períodos de austeridade fiscal, demonstrando a fragilidade de modelos baseados em recursos públicos.

A competição com plataformas comerciais internacionais também representa um desafio contínuo. Pesquisadores da região frequentemente preferem publicar em revistas indexadas na Web of Science ou Scopus, que são percebidas como tendo maior prestígio internacional. Esta preferência reflete não apenas questões de prestígio individual, mas também pressões institucionais e sistemas de avaliação que privilegiam publicações em revistas "internacionais".

Questões de sustentabilidade técnica também são importantes. Manter uma infraestrutura digital moderna requer investimentos

contínuos em tecnologia e pessoal especializado, que podem ser difíceis de garantir com financiamento público limitado. À medida que as expectativas técnicas evoluem e novas funcionalidades se tornam necessárias, o SciELO precisa competir com recursos limitados contra plataformas comerciais com orçamentos muito maiores.

Um desafio particular que o SciELO enfrenta é a tensão entre seus objetivos de democratização da comunicação científica e a crescente importância de métricas bibliométricas globais. Sistemas de avaliação baseados em fatores de impacto e rankings internacionais frequentemente desfavorecem revistas regionais, mesmo quando estas publicam pesquisas de alta qualidade relevantes para contextos locais.

Esta tensão ilustra uma contradição fundamental no sistema científico global: enquanto há reconhecimento retórico da importância da diversidade epistêmica e da relevância local da pesquisa, os sistemas de recompensa e reconhecimento continuam privilegiando publicações em revistas “internacionais” que frequentemente ignoram ou marginalizam perspectivas do Sul Global.

O SciELO tem tentado navegar esta tensão através de várias estratégias, incluindo o desenvolvimento de métricas alternativas que capturam melhor o impacto regional da pesquisa e a promoção de critérios de avaliação mais plurais que reconhecem diferentes tipos de contribuição científica.

O caso SciELO oferece várias lições importantes para outras iniciativas de Ciência Aberta, particularmente aquelas que buscam desafiar estruturas hegemônicas do sistema científico global. Primeiro, demonstra a importância de liderança regional e apropriação local. O sucesso do SciELO deve-se em parte ao fato de ter sido desenvolvido por e para a comunidade científica latino-americana, refletindo suas necessidades e prioridades específicas.

Esta lição é particularmente relevante para iniciativas que buscam promover a ciência aberta no Sul Global. Modelos desenvolvidos no Norte Global podem não ser adequados para contextos diferentes, e iniciativas bem-sucedidas frequentemente requerem adaptação significativa às condições locais.

Segundo, o caso ilustra a importância de padrões técnicos rigorosos. O SciELO não se contentou em simplesmente disponibilizar conteúdo *online*, mas desenvolveu infraestruturas técnicas sofisticadas que competem com plataformas comerciais. Esta decisão foi fundamental para garantir credibilidade e aceitação da plataforma.

Terceiro, o modelo de financiamento do SciELO demonstra que alternativas ao sistema comercial dominante são viáveis, mas requerem compromisso político e financeiro sustentado de governos e instituições públicas. O sucesso a longo prazo de tais iniciativas depende não apenas de recursos financeiros, mas de reconhecimento político da importância de infraestruturas científicas independentes.

O SciELO também representa uma contribuição importante para o que Santos (2014) chama de “epistemologias do Sul” - formas de conhecimento que emergem de experiências de resistência à dominação epistêmica. Ao criar uma plataforma que prioriza a produção científica regional e facilita o diálogo entre pesquisadores do Sul Global, o SciELO contribui para a construção de alternativas epistêmicas ao sistema científico hegemônico.

Esta contribuição vai além da simples disponibilização de conteúdo. O SciELO criou um espaço onde pesquisas sobre problemas específicos da América Latina podem circular e ser reconhecidas, onde metodologias adaptadas a contextos regionais podem ser validadas, e onde perspectivas teóricas que emergem de experiências do Sul Global podem encontrar audiência.

Olhando para o futuro, o SciELO enfrenta o desafio de manter sua relevância em um ambiente científico em rápida transformação.

A crescente importância de preprints, dados abertos, e novas formas de comunicação científica requer adaptação contínua da plataforma. Simultaneamente, pressões por internacionalização e integração com sistemas globais devem ser balanceadas com o compromisso de servir às necessidades específicas da comunidade científica regional.

O modelo SciELO também está sendo adaptado e replicado em outras regiões do Sul Global, incluindo iniciativas na África e Ásia. Estes desenvolvimentos sugerem que o modelo pode ter relevância além da América Latina, oferecendo um framework para o desenvolvimento de infraestruturas científicas regionais independentes.

A experiência do SciELO demonstra que é possível construir alternativas viáveis ao sistema dominante de comunicação científica, mas que tais alternativas requerem compromisso sustentado, recursos adequados, e reconhecimento da importância de diversidade epistêmica no sistema científico global. Suas lições continuam relevantes à medida que comunidades científicas ao redor do mundo buscam formas de democratizar o acesso ao conhecimento e desafiar estruturas hegemônicas de produção e circulação científica.

5.3. CASO 3: UNIVERSIDADES E BIG TECH - A DEPENDÊNCIA DE PLATAFORMAS PRIVADAS NA EDUCAÇÃO PÚBLICA

A crescente dependência de universidades públicas de plataformas digitais controladas por grandes corporações tecnológicas representa uma das manifestações mais visíveis do colonialismo de dados na educação superior. Este caso examina como essa dependência se desenvolveu, quais são suas implicações para a autonomia institucional e a democracia educacional, e que alternativas estão sendo exploradas.

Mais do que uma simples questão de conveniência tecnológica, esta dependência representa uma transformação fundamental na natureza da educação pública e nas relações de poder que a governam.

A dependência de universidades de plataformas *Big Tech* desenvolveu-se gradualmente ao longo das últimas duas décadas, seguindo um padrão que Zuboff (2019) identifica como característico do capitalismo de vigilância: a oferta inicial de serviços “gratuitos” que gradualmente criam dependência e permitem extração de dados em escala massiva.

Inicialmente, muitas universidades adotaram serviços como Gmail, Google Drive, e Microsoft Office 365 como soluções convenientes e aparentemente econômicas para comunicação e colaboração. A atratividade destes serviços não estava apenas em seu custo zero aparente, mas em sua facilidade de implementação e manutenção. Para universidades com recursos de TI limitados, a possibilidade de terceirizar infraestruturas complexas para empresas especializadas parecia uma solução racional.

No entanto, esta adoção inicial criou dependências que se aprofundaram ao longo do tempo. À medida que dados institucionais, processos de trabalho, e competências técnicas se tornaram integrados a essas plataformas, o custo de migração para alternativas aumentou exponencialmente. Esta dinâmica, conhecida como “*vendor lock-in*”, é uma estratégia deliberada das empresas de tecnologia para garantir dependência a longo prazo.

A pandemia de COVID-19 intensificou dramaticamente essa dependência, forçando universidades a rapidamente adotar plataformas de videoconferência, sistemas de gestão de aprendizagem, e outras ferramentas digitais para manter operações educacionais. Plataformas como Zoom, Microsoft Teams, e Google Classroom tornaram-se essenciais para a continuidade educacional, criando uma situação onde a educação pública tornou-se fundamentalmente dependente de infraestruturas privadas.

A dependência de universidades de plataformas *Big Tech* não se limita a ferramentas de comunicação e colaboração. Ela se estende a múltiplas dimensões da atividade acadêmica, criando uma rede complexa de dependências que são difíceis de desfazer.

Na pesquisa, universidades dependem crescentemente de plataformas como Google Scholar para descoberta de literatura, GitHub para compartilhamento de código, e serviços de computação em nuvem da Amazon, Microsoft, e Google para processamento de dados. Esta dependência é particularmente problemática em áreas como inteligência artificial e *Big Data*, onde o acesso a recursos computacionais massivos é frequentemente necessário para pesquisa competitiva.

Na administração, sistemas de gestão estudantil, recursos humanos, e finanças são crescentemente migrados para plataformas em nuvem controladas por grandes corporações. Esta migração é frequentemente justificada em termos de eficiência e redução de custos, mas cria vulnerabilidades significativas e transfere controle sobre dados institucionais sensíveis para entidades privadas.

No ensino, a dependência vai além de plataformas de videoconferência para incluir sistemas de gestão de aprendizagem, ferramentas de avaliação, e recursos educacionais digitais. Esta dependência é particularmente preocupante porque afeta diretamente a experiência educacional dos estudantes e pode influenciar pedagogias e práticas de ensino.

A dependência de plataformas *Big Tech* tem implicações profundas para a autonomia institucional de universidades públicas. Como argumenta Norris (2022), esta dependência representa uma forma de “cooptação” da educação pública por interesses corporativos privados, minando a função democrática da educação.

Primeiro, cria vulnerabilidades operacionais significativas. Interrupções de serviço, mudanças em políticas corporativas, ou

decisões de descontinuar produtos podem afetar drasticamente operações educacionais. Universidades que dependem de plataformas proprietárias têm pouco controle sobre essas decisões e podem encontrar-se em situações onde precisam rapidamente migrar para alternativas com custos e disrupções significativas.

Segundo, essa dependência facilita a extração de dados sobre atividades educacionais e de pesquisa em escala sem precedentes. Empresas como Google e Microsoft coletam dados extensivos sobre como estudantes e professores utilizam suas plataformas, incluindo padrões de comunicação, hábitos de estudo, interesses de pesquisa, e redes sociais acadêmicas. Estes dados são valiosos não apenas para desenvolvimento de produtos comerciais, mas também para análise de mercado e direcionamento de publicidade.

Terceiro, a dependência de plataformas proprietárias pode limitar a inovação educacional e a experimentação pedagógica. Quando universidades dependem de ferramentas desenvolvidas para mercados comerciais gerais, elas podem não ter acesso a funcionalidades específicas que seriam úteis para contextos educacionais particulares. Além disso, a necessidade de compatibilidade com sistemas proprietários pode limitar a adoção de ferramentas inovadoras desenvolvidas pela comunidade acadêmica.

Um aspecto particularmente preocupante da dependência de *Big Tech* na educação é a extensão da extração de dados e vigilância que ela facilita. Como documenta a UNESCO (2023), a dependência de tecnologias educacionais durante a pandemia resultou em coleta massiva de dados sobre estudantes, frequentemente sem consentimento adequado ou transparência sobre como esses dados são utilizados.

Estes dados incluem não apenas informações acadêmicas tradicionais, mas também dados comportamentais detalhados sobre como estudantes interagem com tecnologias educacionais. Padrões de cliques, tempo gasto em diferentes atividades, sequências de navegação, e até mesmo dados biométricos (através de câmeras e microfones) podem ser coletados e analisados.

A análise destes dados pode revelar informações sensíveis sobre estudantes, incluindo dificuldades de aprendizagem, problemas de saúde mental, situação socioeconômica, e preferências políticas. Quando estes dados são combinados com outras informações disponíveis comercialmente, eles podem criar perfis detalhados que podem ser utilizados para discriminação ou manipulação.

Além disso, a coleta de dados educacionais levanta questões sobre privacidade intergeracional. Dados coletados sobre estudantes podem ter implicações para suas oportunidades futuras de emprego, crédito, seguros, e outras áreas da vida. A criação de “dossiês digitais” permanentes sobre indivíduos desde idade precoce representa uma forma de vigilância que pode ter consequências duradouras.

A dependência de *Big Tech* na educação também tem implicações mais amplas para a função democrática da educação pública. Como argumenta Norris (2022), a educação pública tem um papel fundamental na formação de cidadãos críticos e na promoção de valores democráticos. Quando esta função é mediada por plataformas controladas por corporações privadas, ela pode ser comprometida de várias formas.

Primeiro, algoritmos proprietários podem influenciar que informações estudantes e professores veem, criando “bolhas de filtro” que limitam exposição a perspectivas diversas. Quando plataformas educacionais utilizam algoritmos de recomendação similares àqueles utilizados em redes sociais comerciais, elas podem inadvertidamente promover polarização e fragmentação epistêmica.

Segundo, a dependência de plataformas comerciais pode influenciar currículos e pedagogias de formas que favorecem interesses corporativos. Quando ferramentas educacionais são desenvolvidas por empresas com interesses comerciais específicos, elas podem promover certas formas de conhecimento ou habilidades em detrimento de outras.

Terceiro, a normalização da vigilância digital na educação pode socializar estudantes para aceitar monitoramento constante como normal e desejável. Esta socialização pode minar valores democráticos de privacidade, autonomia, e resistência à autoridade que são fundamentais para sociedades democráticas saudáveis.

Apesar dos desafios significativos, algumas universidades e sistemas educacionais estão explorando alternativas à dependência de *Big Tech*. Estas iniciativas oferecem modelos importantes para como instituições educacionais podem recuperar autonomia tecnológica e promover valores democráticos através de suas escolhas tecnológicas.

Iniciativas de software livre educacional, como Moodle, BigBlueButton, e Nextcloud, oferecem alternativas abertas a plataformas proprietárias. Estas ferramentas não apenas evitam problemas de *vendor lock-in* e extração de dados, mas também permitem que instituições educacionais adaptem tecnologias às suas necessidades específicas.

Consórcios universitários também estão desenvolvendo infraestruturas compartilhadas que reduzem a dependência de fornecedores comerciais. Por exemplo, a rede GÉANT na Europa fornece infraestrutura de rede compartilhada para universidades, enquanto iniciativas como eduGAIN facilitam a federação de identidades entre instituições.

Algumas universidades estão investindo no desenvolvimento de capacidades técnicas internas que reduzem dependência de fornecedores externos. Isto inclui não apenas contratação de pessoal técnico, mas também desenvolvimento de políticas e processos que priorizam soluções abertas e sustentáveis.

No entanto, essas alternativas enfrentam desafios significativos que ilustram por que a dependência de *Big Tech* se desenvolveu em primeiro lugar. Software livre frequentemente requer mais recursos técnicos para implementação e manutenção, recursos que muitas universidades não possuem ou não estão dispostas a investir.

Soluções desenvolvidas por consórcios universitários podem não ter a mesma facilidade de uso ou funcionalidades avançadas de produtos comerciais, criando resistência de usuários acostumados com interfaces polidas e recursos sofisticados.

Além disso, a migração de sistemas existentes para alternativas abertas pode ser custosa e disruptiva, particularmente quando dados e processos estão profundamente integrados a plataformas proprietárias. O custo de migração pode ser proibitivo para instituições com recursos limitados.

Reconhecendo os riscos associados à dependência de *Big Tech* na educação, alguns governos e organizações estão desenvolvendo políticas e regulamentações que promovem maior autonomia tecnológica no setor educacional.

Na Europa, o Regulamento Geral sobre Proteção de Dados (GDPR) estabelece requisitos rigorosos para processamento de dados pessoais que afetam como plataformas educacionais podem coletar e utilizar dados de estudantes. Embora não elimine a dependência de *Big Tech*, o GDPR cria incentivos para maior transparência e controle sobre dados educacionais.

Alguns países estão desenvolvendo políticas específicas que promovem uso de software livre e padrões abertos no setor público, incluindo educação. Estas políticas reconhecem que dependência de tecnologias proprietárias pode comprometer soberania digital e autonomia institucional.

O caso da dependência de universidades de *Big Tech* oferece várias lições importantes para promoção de soberania digital no setor educacional. Primeiro, demonstra que decisões tecnológicas aparentemente técnicas têm implicações políticas e democráticas profundas. A escolha de plataformas educacionais não é neutra, mas reflete e reforça certas relações de poder e valores.

Segundo, ilustra a importância de pensar a longo prazo sobre dependências tecnológicas. Soluções que parecem convenientes e econômicas no curto prazo podem criar dependências custosas e problemáticas no longo prazo.

Terceiro, destaca a necessidade de investimento público em infraestruturas educacionais independentes. A construção de alternativas viáveis à dependência de *Big Tech* requer recursos, compromisso político, e visão estratégica que vão além de considerações de custo imediato.

Finalmente, o caso demonstra que resistência à dependência tecnológica é possível, mas requer coordenação, recursos, e compromisso sustentado. As alternativas existem, mas sua implementação bem-sucedida depende de reconhecimento da importância de autonomia tecnológica para a função democrática da educação pública.

A experiência de universidades com *Big Tech* serve como um caso paradigmático de como o colonialismo de dados opera em instituições públicas fundamentais. Suas lições são relevantes não apenas para o setor educacional, mas para qualquer contexto onde instituições públicas enfrentam pressões para adotar tecnologias comerciais que podem comprometer sua autonomia e função democrática.

5.4. CASO 4: GOVERNANÇA DE DADOS EM SAÚDE DURANTE A PANDEMIA DE COVID-19

A pandemia de COVID-19 criou uma situação única onde a necessidade urgente de compartilhamento de dados para pesquisa e saúde pública colidiu com preocupações sobre privacidade, soberania e equidade. Este caso examina como diferentes países

e organizações navegaram essas tensões e quais lições podem ser extraídas para futuras crises de saúde pública. A pandemia serviu como um laboratório global para experimentação com diferentes modelos de governança de dados, revelando tanto as possibilidades quanto os perigos da digitalização acelerada da saúde pública.

A pandemia de COVID-19 representou uma crise de saúde pública sem precedentes na era digital, criando demandas urgentes por dados, análises, e coordenação global que testaram os limites dos sistemas existentes de governança de dados. A velocidade e escala da crise forçaram governos, instituições de pesquisa, e organizações de saúde pública a tomar decisões rápidas sobre compartilhamento de dados, frequentemente sem tempo adequado para consideração completa das implicações éticas e políticas.

A natureza global da pandemia também criou pressões para coordenação internacional e compartilhamento de dados entre países, ao mesmo tempo em que nacionalismos crescentes e competição geopolítica complicavam tais esforços. Esta tensão entre necessidades globais e interesses nacionais tornou-se uma característica definidora da resposta à pandemia.

A pandemia resultou em um compartilhamento de dados científicos sem precedentes em termos de velocidade e escala. Sequências genômicas do vírus SARS-CoV-2 foram disponibilizadas publicamente em questão de dias após sua identificação, estabelecendo um novo padrão para transparência científica em situações de emergência. Plataformas como GISAID facilitaram o compartilhamento rápido de dados genômicos, permitindo que pesquisadores ao redor do mundo rastreassem a evolução do vírus em tempo real.

Preprints sobre COVID-19 proliferaram rapidamente, acelerando dramaticamente a comunicação científica. Plataformas como medRxiv e bioRxiv viram aumentos exponenciais em submissões relacionadas à COVID-19, com alguns estudos recebendo atenção global antes mesmo de passar por revisão por pares. Como documenta

Zastrow (2020), esta abertura facilitou o desenvolvimento rápido de vacinas, tratamentos, e medidas de saúde pública, demonstrando o potencial da ciência aberta para responder a crises globais.

A colaboração científica global permitiu que pesquisadores construíssem rapidamente sobre o trabalho uns dos outros, resultando em avanços científicos em velocidade recorde. O desenvolvimento de vacinas mRNA em menos de um ano, por exemplo, foi possível em parte devido ao compartilhamento rápido de dados sobre a estrutura do vírus e resultados de ensaios clínicos.

No entanto, o compartilhamento acelerado também criou novos riscos e desafios. A pressão por velocidade às vezes comprometeu a qualidade da revisão por pares, resultando em estudos com metodologias questionáveis ou conclusões prematuras sendo amplamente divulgados antes de serem adequadamente validados. Casos como o estudo sobre hidroxicloroquina publicado na *The Lancet* e posteriormente retraído ilustraram os perigos de acelerar processos científicos sem salvaguardas adequadas.

A pandemia também revelou tensões significativas entre abertura de dados e soberania nacional, demonstrando como crises podem exacerbar competições geopolíticas existentes. Alguns países restringiram o compartilhamento de dados sobre variantes virais por preocupações de segurança nacional, temendo que informações sobre a prevalência de variantes específicas pudessem afetar sua reputação internacional ou capacidade de atração turística.

Outros países limitaram o acesso a dados epidemiológicos por razões de competitividade comercial ou política. A China, por exemplo, foi criticada por limitar o acesso de pesquisadores internacionais a dados sobre os primeiros casos de COVID-19, dificultando esforços para compreender as origens da pandemia.

Essas restrições às vezes prejudicaram esforços globais de resposta à pandemia. Quando países não compartilhavam dados sobre

variantes emergentes, isso limitava a capacidade de outros países de preparar respostas adequadas, incluindo desenvolvimento de vacinas adaptadas e implementação de medidas de saúde pública apropriadas.

A experiência ilustrou como dados de saúde pública são simultaneamente recursos globais (na medida em que doenças infecciosas não respeitam fronteiras) e ativos nacionais (na medida em que refletem capacidades e vulnerabilidades específicas de países). Esta dualidade criou dilemas complexos sobre quando e como compartilhar informações sensíveis.

A pandemia também acelerou dramaticamente a adoção de tecnologias de vigilância digital para rastreamento de contatos e monitoramento de saúde pública. Aplicativos de rastreamento de contatos, sistemas de monitoramento de movimento, e outras tecnologias foram implementados rapidamente em muitos países, frequentemente com pouca consideração para implicações de longo prazo para privacidade e direitos humanos.

Como analisa Kitchin (2020), embora essas tecnologias tenham potencial para melhorar respostas de saúde pública, elas também levantam preocupações significativas sobre privacidade, discriminação, e autoritarismo digital. Muitos sistemas foram implementados sem salvaguardas adequadas, supervisão democrática, ou mecanismos de prestação de contas.

A eficácia dessas tecnologias também se mostrou questionável em muitos contextos. Aplicativos de rastreamento de contatos, por exemplo, frequentemente tiveram baixas taxas de adoção e problemas técnicos que limitaram sua utilidade para saúde pública. No entanto, a infraestrutura de vigilância criada durante a pandemia permaneceu em muitos países, levantando preocupações sobre normalização da vigilância digital.

A experiência da pandemia demonstrou como crises podem ser utilizadas para justificar a implementação de tecnologias de vigi-

lância que podem persistir além da emergência original. Como argumenta Toh (2020), isso destaca a importância de frameworks éticos robustos para governança de dados em situações de emergência, incluindo mecanismos para garantir que medidas excepcionais sejam temporárias e proporcionais.

A pandemia também revelou e exacerbou desigualdades digitais existentes, tanto dentro de países quanto entre países. Comunidades com menor acesso a tecnologias digitais frequentemente foram sub-representadas em dados de saúde pública, criando lacunas na compreensão da distribuição e impactos da doença.

Países com menor capacidade técnica enfrentaram dificuldades para coletar, processar, e compartilhar dados de saúde pública, limitando sua capacidade de participar em esforços globais de pesquisa e coordenação. Esta “divisão digital” em saúde pública teve consequências reais para capacidade de resposta à pandemia.

Além disso, o acesso desigual a tecnologias de saúde digital significou que benefícios de inovações como telemedicina e monitoramento remoto foram distribuídos de forma desigual, potencialmente exacerbando disparidades de saúde existentes.

Apesar dos desafios, a pandemia também gerou experimentação com modelos alternativos de governança de dados que oferecem lições importantes para futuras crises. Algumas iniciativas demonstraram como é possível combinar abertura científica com proteção de privacidade e soberania.

O consórcio COVID-19 Host Genetics Initiative, por exemplo, desenvolveu protocolos para compartilhamento de dados genéticos que permitiam análises colaborativas sem transferência de dados brutos, protegendo privacidade individual e soberania nacional enquanto facilitava pesquisa global.

Iniciativas como a COVID-19 Data Portal do European Bioinformatics Institute criaram infraestruturas para compartilhamento de dados que incorporavam princípios FAIR (Findable, Accessible, Interoperable, Reusable) enquanto respeitavam restrições éticas e legais específicas.

A experiência da pandemia oferece várias lições importantes para governança de dados em futuras emergências de saúde pública. Primeiro, demonstra a importância de ter frameworks éticos e legais estabelecidos antes de crises ocorrerem. Tentativas de desenvolver tais frameworks durante emergências frequentemente resultam em soluções inadequadas ou inconsistentes.

Segundo, ilustra a necessidade de equilibrar transparência científica com proteção de privacidade e soberania. Este equilíbrio não é apenas uma questão técnica, mas requer consideração cuidadosa de valores, contextos culturais, e relações de poder.

Terceiro, destaca a importância de inclusão e equidade em sistemas de dados de saúde pública. Dados que não representam adequadamente populações diversas não apenas são cientificamente inadequados, mas podem perpetuar ou exacerbar desigualdades de saúde.

À medida que o mundo se prepara para futuras pandemias, as lições da COVID-19 sobre governança de dados são particularmente relevantes. A próxima pandemia provavelmente será ainda mais digitalizada, com maior dependência de dados, inteligência artificial, e coordenação digital.

Preparação adequada requer não apenas desenvolvimento de capacidades técnicas, mas também estabelecimento de frameworks éticos e legais que possam funcionar em situações de emergência. Isso inclui protocolos para compartilhamento de dados que equilibrem necessidades de saúde pública com proteção de direitos, mecanismos para garantir que medidas excepcionais sejam temporárias, e sistemas para promover equidade no acesso a benefícios de inovações digitais.

A experiência da pandemia também tem implicações importantes para conceitos emergentes de soberania de dados em saúde. Como demonstra a iniciativa da África CDC (2025) para desenvolver um Framework Continental de Governança de Dados de Saúde, há reconhecimento crescente de que países e regiões precisam de maior controle sobre dados de saúde para garantir que benefícios de pesquisa e inovação sejam distribuídos equitativamente.

Este framework reconhece que dados de saúde são simultaneamente recursos globais e ativos nacionais, e busca criar mecanismos que facilitem compartilhamento benéfico enquanto protegem interesses nacionais e regionais. A abordagem enfatiza governança ética, confiável, e coordenada que se alinha com princípios de equidade e direitos.

Finalmente, a experiência da pandemia levanta questões fundamentais sobre democracia e saúde pública digital. A velocidade necessária para responder a crises de saúde pública pode conflitar com processos democráticos de deliberação e prestação de contas. No entanto, como a pandemia demonstrou, respostas que ignoram considerações democráticas podem ser tanto ineficazes quanto prejudiciais a longo prazo.

O desafio para o futuro é desenvolver formas de governança de dados que possam responder rapidamente a emergências enquanto mantêm compromissos com transparência, prestação de contas, e participação democrática. Isso requer não apenas inovação técnica, mas também inovação institucional e política.

A pandemia de COVID-19 serviu como um teste de estresse para sistemas globais de governança de dados, revelando tanto possibilidades quanto perigos da digitalização acelerada da saúde pública. Suas lições continuarão relevantes à medida que sociedades navegam tensões entre eficiência tecnológica e valores democráticos, entre necessidades globais e soberania nacional, e entre inovação científica e proteção de direitos humanos.



6

**EXPERIÊNCIAS
BRASILEIRAS
E LATINO-AMERICANAS**

O Brasil e a América Latina ocupam uma posição singular no debate global sobre Ciência Aberta e colonialismo de dados. A região combina uma história de iniciativas pioneiras de democratização do conhecimento com desigualdades sociais e dependência tecnológica que dificultam a apropriação plena de tecnologias digitais. Esses contrastes conferem ao contexto latino-americano o papel de laboratório social: as políticas de acesso aberto buscam quebrar barreiras históricas de exclusão, enquanto a pressão para aderir a padrões globais – geralmente estabelecidos por agências e plataformas do Norte global – coloca a região diante do risco de *subordinação epistêmica*. A seguir analisam-se políticas públicas, infraestrutura informacional e experiências de cidadania científica, destacando conquistas, contradições e desafios.

6.1 POLÍTICAS DE CIÊNCIA ABERTA NO BRASIL

A principal base normativa para a transparência pública é a Lei de Acesso à Informação (Lei 12.527/2011). A lei concretiza o direito constitucional de todos receberem informações do Estado e estabelece que a publicidade é regra e o sigilo, exceção (IFCE, 2011). Ela distingue transparência ativa — publicação espontânea de dados de interesse coletivo — e transparência passiva, quando o órgão responde a pedidos de informação, com prazo de 20 dias prorrogáveis por mais dez (IFCE, 2011). Embora não trate especificamente de dados científicos, a LAI abriu caminho para políticas de dados abertos na administração pública. Nesse contexto surgiram o Portal Brasileiro de Dados Abertos (2012) e a Política Nacional de Dados Abertos (Decreto 8.777/2016), que formalizou compromissos governamentais de transparência e incluiu disposições sobre dados científicos produzidos com recursos públicos.

O ordenamento jurídico brasileiro situa-se num campo de tensões entre abertura e proteção de dados. De um lado, o Estado incentiva a publicação de dados e a ciência aberta como mecanismos de *accountability* e inovação; de outro, normas como a Lei Geral de Proteção de Dados Pessoais – LGPD – impõem limites ao compartilhamento de informações pessoais (ver §6.3).

As principais agências de fomento brasileiras desenvolveram diretrizes específicas que vão além de mandatos genéricos de acesso aberto. A FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) foi pioneira ao aprovar em 2008 uma política de acesso aberto às publicações resultantes de auxílios e bolsas, atualizada pela Portaria CTA n.º 37/2021, que fixa prazo de até 12 meses para a disponibilização de artigos em repositórios institucionais. O Conselho Técnico-Administrativo também aprovou diretrizes preliminares para planos de gestão de dados em 2022. O CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) incluiu nos editais pontuação adicional para projetos que disponibilizam dados e código, estimulando reprodutibilidade. A CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) determinou que teses e dissertações financiadas com recursos públicos sejam depositadas em repositórios abertos, ampliando a visibilidade da produção da pós-graduação.

Apesar de avanços normativos, a implementação dessas políticas enfrenta obstáculos estruturais. Muitas universidades carecem de infraestrutura técnica para repositórios de dados e sistemas de gestão de pesquisa, e faltam profissionais capacitados em curadoria de dados e documentação de código aberto. Além disso, a avaliação acadêmica ainda privilegia artigos em periódicos de alto impacto (geralmente de acesso restrito) em detrimento de práticas de compartilhamento. Consequentemente, parte dos pesquisadores resiste a publicar seus dados por temor de perder prioridade na autoria ou sofrer uso indevido, como apontou estudo do grupo de Ciência Aberta da Fiocruz (SciELO Brasil; Fiocruz, 2021).

Organizações da sociedade civil e coletivos de tecnologia foram fundamentais para traduzir o discurso de acesso aberto em políticas concretas. Iniciativas como o projeto “Open Access Brasil” (2011) e redes comunitárias de repositórios pressionaram o Estado a adotar padrões de interoperabilidade e licenças livres. A Fundação Oswaldo Cruz (Fiocruz) promoveu amplo debate interno e produziu relatório comparativo sobre políticas de dados abertos de diferentes países (SciELO Brasil; Fiocruz, 2021). Essas ações catalisaram a adoção, pela Fiocruz, de repositórios institucionais e de um programa de formação em gestão de dados, demonstrando como instituições de pesquisa podem liderar movimentos de abertura.

6.2 REPOSITÓRIOS INSTITUCIONAIS E INICIATIVAS REGIONAIS

O Brasil dispõe de uma das mais extensas redes de repositórios institucionais da América Latina. A maioria das universidades públicas mantém plataformas de acesso aberto para teses, dissertações e artigos científicos, geralmente baseadas em softwares livres como DSpace. O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) coordena esses esforços, oferecendo capacitação e definindo diretrizes de interoperabilidade. Um marco é a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), que integra sistemas de informação de instituições de ensino superior e incentiva a publicação eletrônica; o portal ressalta que a BDTD possibilita à comunidade científica brasileira publicar e disseminar teses e dissertações, ampliando a visibilidade da produção nacional (IBICT, 2025). Com mais de 500 mil documentos, a BDTD é hoje uma das maiores coleções de pós-graduação em acesso aberto no mundo.

Além de iniciativas nacionais, o Brasil participa de redes regionais que buscam superar a fragmentação informacional na

América Latina. A Rede Federada de Repositórios Institucionais de Publicações Científicas (LA Referencia) integra repositórios de nove países latino-americanos. De acordo com o portal Land Portal, a LA Referencia é uma rede latino-americana de repositórios de acesso aberto que apoia estratégias nacionais de acesso aberto por meio de uma plataforma com padrões de interoperabilidade, compartilhando e dando visibilidade à produção científica de instituições de ensino superior e pesquisa (Land Portal Foundation, 2025). A rede integra artigos, teses de doutorado e mestrado de mais de cem universidades e centros de pesquisa dos países membros (Land Portal Foundation, 2025). Ela baseia-se em acordos técnicos e organizacionais entre agências públicas de ciência e tecnologia e a rede RedCLARA, e surgiu a partir do Acordo de Buenos Aires de 2012, que expressa a vontade de tornar a produção científica latino-americana um bem público regional (Land Portal Foundation, 2025).

No campo da saúde, o Brasil colabora com países latino-americanos por meio da Biblioteca Virtual em Saúde (BVS). A BVS Fiocruz é descrita como um espaço de integração de fontes de informação em uma rede da América Latina e do Caribe (Fiocruz, 2025). Trata-se de parte de um projeto internacional para promover parcerias, facilitar o acesso e o uso de informações, reunindo mais de 30 mil documentos e 13 bibliotecas temáticas coordenadas pelo ICICT (Fiocruz, 2025). A iniciativa resulta de parceria entre a Organização Pan-Americana da Saúde (OPAS) – por meio do Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde (Bireme) – e o Ministério da Saúde (Fiocruz, 2025), reforçando o caráter cooperativo da produção e disseminação de conhecimento em saúde.

A ampliação de repositórios institucionais impõe questões de sustentabilidade financeira e tecnológica. Muitos repositórios dependem de orçamentos limitados e correm risco diante de cortes. Manter plataformas seguras e interoperáveis requer técnicos especializados e planejamento de preservação digital. Outra dificuldade é a heterogeneidade de metadados, que compromete a integração regional.

A LA Referencia enfrenta desafios para harmonizar padrões nacionais e garantir a atualização constante dos registros, principalmente em países com recursos escassos.

6.3 DESAFIOS DA IMPLEMENTAÇÃO DA LGPD NA PESQUISA

A Lei Geral de Proteção de Dados (LGPD) (Lei 13.709/2018, em vigor desde 2020) criou um marco para o tratamento de dados pessoais no Brasil. A lei aplica-se ao tratamento de dados por pessoas naturais ou jurídicas, públicas ou privadas, inclusive por meios digitais (Brasil, 2018). Entre os fundamentos da LGPD estão o respeito à privacidade, a autodeterminação informativa, a liberdade de expressão, a inviolabilidade da intimidade, o desenvolvimento econômico e a inovação (Brasil, 2018). O art. 4º estabelece exceções, entre elas o tratamento de dados realizado para fins exclusivos de pesquisa acadêmica, observadas as normas de ética e proteção aplicáveis (Brasil, 2018).

A LGPD teve forte impacto na pesquisa científica, especialmente em projetos que envolvem dados de pessoas. Pesquisadores de áreas como saúde, educação e ciências sociais precisam obter consentimento informado específico para cada uso de dados, garantir transparência sobre a finalidade do tratamento e adotar medidas de segurança. A reutilização de dados originalmente coletados para outros fins tornou-se mais complexa, e a exigência de consentimento específico dificulta estudos longitudinais ou colaborativos. Além disso, técnicas de anonimização e pseudonimização nem sempre são suficientes para proteger privacidade, sobretudo no caso de dados genéticos, que são intrinsecamente identificáveis e podem revelar informações sobre familiares.

Instituições públicas e privadas têm investido em compliance para atender à LGPD. Universidades criaram comitês de proteção de

dados, elaboraram políticas internas, desenvolveram sistemas de gestão de consentimento e capacitaram pesquisadores. Alguns institutos implementaram plataformas seguras para armazenamento e compartilhamento de dados pessoais, com trilhas de auditoria e procedimentos de eliminação. Apesar desses esforços, persistem incertezas jurídicas sobre a adequação de bancos de dados pré-existent à nova lei, e há receio de que burocracias excessivas desestimulem a ciência aberta.

6.4 MOVIMENTOS DE SOFTWARE LIVRE E TECNOLOGIAS SOCIAIS

Desde o início dos anos 2000, o governo brasileiro adotou políticas para reduzir dependências de fornecedores estrangeiros e fomentar o uso de software livre. O Portal do Software Público Brasileiro, criado em abril de 2007, serve como plataforma para compartilhamento de soluções de código aberto desenvolvidas pela administração pública. O portal cataloga softwares utilizados em setores variados e é utilizado também em países como Uruguai, Argentina, Portugal, Venezuela, Chile e Paraguai; ele se consolidou como ambiente de compartilhamento de software e promove a racionalização de recursos de TI, a expansão de parcerias e o fortalecimento da política de software livre (Governo Federal do Brasil, 2025). Além de possibilitar economia de custos — estatísticas oficiais apontam que os gastos de TI do governo federal cresceram 227 % entre 2003 e 2009 — o portal incentiva a independência de fornecedores, já que licenças de código aberto permitem modificar e adequar programas, aumentando a competição e a segurança (Governo Federal do Brasil, 2025).

O movimento de software livre no Brasil articula programadores, comunidades hacker, universidades e agências governamentais. Eventos como o Fórum Internacional de Software Livre (FISL),

realizado desde 2000 em Porto Alegre, tornaram-se espaços de articulação política e técnica. Essa tradição de defesa da liberdade de código dialoga com a ciência aberta e com a soberania tecnológica: a adoção de ferramentas abertas nos repositórios e nas infraestruturas de pesquisa reduz custos, permite auditoria de algoritmos e estimula a formação de capacidades locais.

O conceito de tecnologia social foi formulado no Brasil para designar soluções desenvolvidas em interação com a comunidade, replicáveis e voltadas à transformação social. A Pró-Reitoria de Extensão da Universidade Federal de Sergipe (UFS) resume que tecnologias sociais compreendem *“produtos, técnicas e/ou metodologias reaplicáveis, desenvolvidas na interação com a comunidade e que representam efetivas soluções de transformação social”* (UFS, 2025). Além de definirem uma metodologia participativa, essas tecnologias têm caráter transversal, abrangendo diversas áreas e promovendo inclusão. A UFS ressalta que a aplicação de tecnologias sociais impacta diretamente o desenvolvimento humano, pois constitui instrumento poderoso para efetivação de direitos sociais e ampliação de acesso a políticas públicas, como geração de renda, melhoria das condições de saúde, segurança alimentar, construção de habitações de baixo custo e saneamento (UFS, 2025).

No Brasil, a Rede de Tecnologias Sociais (RTS), coordenada pelo Ministério da Ciência, Tecnologia e Inovação e pela Fundação Banco do Brasil, mapeia e divulga experiências de tecnologias sociais. Exemplos incluem filtros de água de baixo custo, sistemas de captação de água de chuva em comunidades rurais e programas de alfabetização desenvolvidos com a participação de educadores populares. Ao priorizarem conhecimento local e autogestão, essas iniciativas desafiam modelos de inovação centrados em grandes corporações e aproximam a ciência aberta das demandas sociais.

6.5 CIÊNCIA CIDADÃ E PARTICIPAÇÃO COMUNITÁRIA

A ciência cidadã – participação de cidadãos não especialistas em projetos científicos – tem se expandido como componente da ciência aberta na região. Projetos de monitoramento ambiental e saúde comunitária ilustram sua relevância, ao mesmo tempo em que evidenciam desafios de inclusão digital e equidade.

Um dos casos mais longevos de ciência cidadã no Brasil é o programa Observando os Rios, mantido pela Fundação SOS Mata Atlântica. O portal do projeto explica que ele *reúne comunidades e as mobiliza em torno da qualidade da água de rios, córregos e outros corpos d'água das localidades onde vivem* (SOS Mata Atlântica, 2025). A iniciativa é aberta à população: qualquer pessoa pode participar de grupos de monitoramento existentes ou criar novos grupos próximos a escolas, igrejas e centros comunitários (SOS Mata Atlântica, 2025). O monitoramento da água é realizado com um kit de análise desenvolvido pela ONG; os grupos medem a qualidade da água mensalmente e enviam os resultados pela internet (SOS Mata Atlântica, 2025). Segundo a plataforma Cívis, a metodologia utiliza o Índice de Qualidade da Água (IQA) e combina indicadores laboratoriais e de percepção; os dados de todos os grupos são compilados em relatórios técnicos anuais, formando um retrato da qualidade das águas de 17 estados da Mata Atlântica (Cívis, 2025). Desde 1991, o programa mobilizou mais de 20 mil pessoas em 149 grupos de monitoramento (SOS Mata Atlântica, 2025), demonstrando a escala de engajamento possível.

Outro exemplo é o monitoramento participativo da qualidade do ar e do ruído em áreas urbanas coordenado por universidades brasileiras. Moradores usam sensores de baixo custo e aplicativos de código aberto para registrar dados, que são integrados a plataformas de pesquisa. Esses projetos aproximam ciência, educação ambiental e mobilização

social, mas enfrentam desafios de sustentabilidade e de representação: participantes tendem a ser grupos urbanos escolarizados, enquanto comunidades periféricas têm menor acesso a tecnologias digitais.

Na área da saúde, o Brasil também desenvolve ferramentas de ciência cidadã. O Guardiões da Saúde é um aplicativo de vigilância participativa criado pelo Ministério da Saúde em 2007 para monitorar eventos de massa. Segundo a Associação Brasileira de Profissionais de Epidemiologia de Campo (ProEpi), o aplicativo foi usado na Copa do Mundo FIFA 2014; em 2016, foi novamente utilizado nos Jogos Olímpicos e Paralímpicos do Rio de Janeiro, graças a uma parceria entre a Skoll Global Threats Fund, o Ministério da Saúde, a startup EpiTrack e a própria ProEpi (ProEpi, 2025). A estratégia baseia-se na população relatar voluntariamente sua condição de saúde: registros de sintomas são compilados e analisados para verificar incidências de doenças e contribuir para a construção de cenários epidemiológicos (ProEpi, 2025). Em 2020, em parceria com a Sala de Situação em Saúde da Universidade de Brasília, o aplicativo foi atualizado para a Covid-19 e passou a ser usado também por estudantes universitários; em março de 2021 já contava com 26 mil usuários cadastrados (ProEpi, 2025). Essa experiência demonstra o potencial de ferramentas móveis para vigilância em tempo real, mas também ressalta desafios como a representatividade dos usuários e a proteção de dados pessoais (problemas acentuados pela LGPD).

Além do Guardiões da Saúde, outras iniciativas de saúde comunitária envolvem a coleta participativa de sintomas de arboviroses, monitoramento de focos de mosquito *Aedes aegypti* e aplicação de testes rápidos por agentes comunitários. Esses projetos reforçam a importância de incorporar comunidades na construção de políticas de saúde e de valorizar o conhecimento local.

Embora a ciência cidadã promova democratização da pesquisa, ela também reproduz exclusões. A participação em projetos costuma exigir acesso à internet, dispositivos móveis e alfabetização

digital, recursos nem sempre disponíveis em comunidades rurais ou periferias urbanas. Programas como Observando os Rios oferecem kits analógicos e treinamento presencial para reduzir essas barreiras, mas a escalabilidade permanece limitada. Outro desafio é a representatividade: voluntários tendem a ser pessoas com maior escolaridade e interesse ambiental, o que pode gerar viés nos dados coletados. Para superar essas limitações, é necessário combinar plataformas digitais com estratégias de educação popular e políticas de universalização da internet.

6.6 REDES DE COLABORAÇÃO SUL-SUL

A cooperação entre países do Sul global é estratégica para reduzir dependências tecnológicas e epistemológicas. O Brasil participa de redes como os BRICS, que possuem um *Framework* de cooperação em ciência, tecnologia e inovação priorizando áreas como energia, saúde e tecnologia da informação. No âmbito latino-americano, o país integra organizações como o Mercosul e a União de Nações Sul-Americanas (UNASUL), que promovem intercâmbio de pesquisadores, projetos colaborativos e compartilhamento de infraestrutura científica. Redes setoriais – como a LA Referencia e a BVS discutidas acima – evidenciam que a integração regional em bases abertas é possível quando há vontade política e investimentos em infraestrutura.

Apesar do potencial, as redes Sul-Sul enfrentam assimetrias em capacidades técnicas e recursos. Diferenças no nível de digitalização e no financiamento público de ciência tornam difícil manter compromissos de interoperabilidade. Contudo, a cooperação regional oferece oportunidades para elaborar padrões de ciência aberta adaptados às realidades locais, evitando a imposição de modelos provenientes do Norte global. Exemplos incluem o desenvolvimento de softwares livres para repositórios, a capacitação conjunta de gestores de dados e a elaboração de parâmetros éticos regionais para uso de dados sensíveis.

Parte

IV

**PROPOSTAS
E PERSPECTIVAS
FUTURAS**

7

**PROPOSTAS E SOLUÇÕES
PARA UMA GOVERNANÇA
DE DADOS JUSTA E ÉTICA**

A construção de um futuro em que a Ciência Aberta cumpra suas promessas democratizantes exige uma transformação na forma como os dados são governados, as tecnologias são desenvolvidas e o conhecimento circula. Esta parte propõe um conjunto de princípios e ações para promover soberania digital, equidade e sustentabilidade, reconhecendo que a governança de dados é inseparável das lutas por justiça social e epistêmica.

7.1 SOBERANIA DE DADOS E MARCOS REGULATÓRIOS

O conceito de soberania de dados amplia a ideia de privacidade individual para incorporar dimensões coletivas e políticas. Para povos indígenas e comunidades historicamente marginalizadas, o controle sobre dados é parte da autodeterminação. Os Princípios CARE – desenvolvidos pela Global Indigenous Data Alliance – complementam a estrutura FAIR ao enfatizar quatro valores: benefício coletivo, autoridade para controlar, responsabilidade e ética. Esses princípios propõem que os ecossistemas de dados devem gerar benefícios tangíveis para as comunidades, reconhecer seu direito de decidir quais dados são coletados e como são usados, exigir que pesquisadores demonstrem que o uso de dados apoia a autodeterminação e garantir que direitos e bem-estar sejam a principal preocupação em todas as etapas (ResearchHub, 2024).

Legislações nacionais têm buscado incorporar noções de soberania. O Regulamento Geral de Proteção de Dados da União Europeia (GDPR) estabelece que a transferência de dados pessoais para países terceiros só pode ocorrer quando forem cumpridas condições que garantam um nível de proteção adequado (União Europeia, 2016). A Lei Geral de Proteção de Dados do Brasil reproduz princípios da GDPR e adiciona requisitos de localização de dados em alguns

casos, reforçando o controle nacional sobre fluxos de dados (ver §6.3). Países como a Índia discutem projetos de lei que classificam dados segundo sua sensibilidade e exigem armazenamento local de dados críticos, enquanto nações africanas elaboram marcos que combinam proteção da privacidade com soberania coletiva dos dados.

A governança internacional continua fragmentada. Diversas iniciativas tentam preencher o vazio regulatório: a proposta de Declaração Universal dos Direitos Humanos na Era Digital insiste que a governança da internet deve ser orientada pelos princípios da Declaração de 1948. O texto argumenta que a formulação de políticas digitais deve começar pela identificação de valores fundamentais – como dignidade, igualdade e liberdade – e que os direitos humanos se aplicam tanto em espaços físicos quanto virtuais (Lanfranco; Stoll, 2020). Organizações como ONU, UNESCO e OCDE elaboram frameworks de inteligência artificial e governança de dados, mas esses processos são frequentemente dominados por países e empresas do Norte global; é necessário que países do Sul participem ativamente para incluir perspectivas locais.

Para além desses debates internacionais, diferentes comunidades reivindicam formas de soberania que valorizam suas cosmovisões. Por exemplo, no contexto da Nova Zelândia, bibliotecários da Universidade de Waikato explicam que dados sobre iwi, hapū e whānau (famílias e subtribos maori) são considerados taonga – tesouros culturais – e devem ser governados pelas próprias comunidades. O guia recomenda que pesquisadores armazenem dados localmente em servidores no país, evitando serviços em nuvem sediados no exterior, pois leis estrangeiras podem permitir acesso governamental aos dados (University of Waikato, 2025). Essa abordagem ressalta a necessidade de práticas de armazenamento e licenciamento que respeitem a autodeterminação dos povos indígenas e cria um precedente para países que buscam proteger seus dados de legislações extraterritoriais.

Outras propostas inovadoras incluem a criação de data trusts e commons digitais como modelos de governança coletiva. Data trusts são arranjos legais em que uma entidade fiduciária administra conjuntos de dados em nome de grupos ou comunidades, definindo regras para coleta, uso e repartição de benefícios. Esses mecanismos asseguram que dados de comunidades ou de setores estratégicos sejam explorados de modo transparente e em consonância com os interesses dos titulares. Os commons digitais, inspirados na teoria dos bens comuns de Elinor Ostrom, propõem que recursos informacionais sejam geridos coletivamente por comunidades de usuários, com regras claras de acesso, manutenção e compartilhamento que previnam a apropriação privada. Experiências em países como Canadá e Alemanha têm testado modelos de *data trusts* para regular dados urbanos e de mobilidade, enquanto cooperativas agrícolas criaram plataformas de dados nas quais agricultores compartilham informações sobre safras e mercados sob controle coletivo. Essas iniciativas sugerem caminhos para superar tanto a lógica de propriedade privada quanto a de controle estatal centralizado.

7.2 INFRAESTRUTURAS ABERTAS E TECNOLOGIAS DECOLONIAIS

Reduzir dependências estruturais passa por construir infraestruturas abertas e tecnologias orientadas pela liberdade e pela colaboração. O movimento de software livre oferece um modelo poderoso: a Free Software Foundation define software livre como aquele que respeita a liberdade dos usuários, permitindo que qualquer pessoa execute o programa para qualquer propósito, estude seu código, modifique-o e redistribua cópias originais ou modificadas (Free Software Foundation, 2025). Essas quatro liberdades permitem que pesquisadores adaptem ferramentas às suas necessidades,

promovendo transparência e reprodutibilidade. No entanto, comunidades do Sul global ainda participam pouco na governança de grandes projetos de código aberto; ampliar essa participação exige investimentos em formação técnica e apoio a iniciativas locais de desenvolvimento, como hackathons e cooperativas de programadores.

Para escapar da concentração de dados em plataformas corporativas, é preciso explorar tecnologias descentralizadas. O InterPlanetary File System (IPFS), por exemplo, é um sistema de armazenamento e compartilhamento de arquivos construído sobre uma rede de operadores que hospedam partes dos dados; ele identifica cada arquivo por um hash e permite que qualquer par na rede recupere o conteúdo. Ao não depender de servidores centrais, o IPFS fornece um meio resiliente de preservação de conhecimento e de distribuição de dados científicos, especialmente útil em contextos com riscos de censura ou interrupção de serviços.

Outra inspiração vem das tecnologias apropriadas, conceito popularizado por E. F. Schumacher. Em 1965, Schumacher defendia que o desenvolvimento deveria abandonar a imposição de tecnologias de grande escala e adotar “tecnologias intermediárias” apropriadas aos contextos locais e às habilidades das pessoas (Practical Action, 2021). Esse princípio continua relevante para a era digital: soluções tecnológicas devem ser acessíveis, energeticamente eficientes e viáveis de manter com recursos locais. Iniciativas como computadores de baixo custo e redes comunitárias de internet demonstram que é possível construir infraestrutura digital que respeite limites ambientais e culturais.

Além de software, a independência tecnológica exige *hardware* aberto. Padrões como RISC-V, um conjunto de instruções de processador open source, permitem que universidades e pequenas empresas projetem microchips sem pagar royalties às gigantes da indústria. A disponibilização de projetos de circuitos, impressoras 3D e outros artefatos sob licenças abertas cria ecossistemas de fabricação distribuída.

Redes de produção baseadas em fabricação digital e maker spaces possibilitam que comunidades adaptem dispositivos às suas necessidades, reduzindo a dependência de importações.

Outro exemplo são as redes comunitárias de telecomunicações. Na Argentina, o projeto *Altermundi* desenvolve redes *mesh* baseadas no protocolo LibreRouter, que permite que comunidades rurais construam sua própria infraestrutura de internet com equipamentos de baixo custo e software livre. Essas redes descentralizadas demonstram que é possível ampliar a conectividade em regiões negligenciadas por operadoras comerciais, fortalecendo a soberania digital e a inclusão.

Para que essas infraestruturas abertas sejam ambientalmente sustentáveis, é preciso considerar o consumo de energia e a pegada de carbono. Tecnologias de computação de borda (*edge computing*), nas quais os dados são processados localmente em dispositivos ou micros servidores próximos à fonte, reduzem a necessidade de enviar grandes volumes de informação a data centers distantes, economizando energia e largura de banda. A adoção de servidores energeticamente eficientes e o uso de energia renovável em repositórios científicos podem diminuir o impacto ambiental da ciência aberta. Projetos de micro-servidores alimentados por painéis solares em comunidades rurais demonstram que é possível hospedar repositórios e websites de forma autônoma e ecológica, combinando soberania digital com sustentabilidade.

7.3 PRINCÍPIOS ÉTICOS E INCLUSÃO

A governança de dados justa exige um arcabouço ético robusto. A informação é parte constitutiva da identidade de cada pessoa; para Luciano Floridi, a proteção da privacidade informacional é fundamental porque cada indivíduo é “constituído por sua

informação” e violações de privacidade correspondem a agressões à identidade (Bawden & Robinson, 2019). A ética da informação demanda que instituições protejam tanto a privacidade individual quanto a coletiva e que criem mecanismos de consentimento, transparência e responsabilidade. Entrevistas com Rafael Capurro ressaltam que uma ética contemporânea da informação deve problematizar as normas morais da comunicação e refletir criticamente sobre como a digitalização cria novas regras e valores, de modo a evitar que prevaleça a lei do mais forte ou o simples cálculo de custos e benefícios (Lara, 2014).

No debate filosófico, a britânica Miranda Fricker cunhou o termo injustiça epistêmica para descrever injustiças relacionadas ao conhecimento: elas incluem exclusão e silenciamento de determinadas pessoas, distorção das contribuições e discriminação na atribuição de autoridade. Fricker distingue injustiça testemunhal – quando o testemunho de alguém é desvalorizado por preconceitos de raça, gênero ou classe – e injustiça hermenêutica – quando um grupo carece de conceitos para expressar suas experiências. A ciência aberta deve combater essas injustiças valorizando saberes diversos, evitando que algumas vozes sejam sistematicamente ignoradas.

Para democratizar a produção e circulação do conhecimento é preciso promover bibliodiversidade. Esse conceito, originado na América Latina, aplica a ideia de diversidade cultural ao mundo do livro. A bibliodiversidade defende uma variedade equilibrada de publicações e critica a concentração das editoras e a “bestsellerização” do mercado, que privilegia poucos títulos lucrativos em detrimento de obras experimentais e de autores marginalizados. Editoras independentes e coletivos de pesquisa desempenham papel crucial na preservação dessa diversidade; políticas públicas podem apoiar editoras pequenas, revistas comunitárias e repositórios alternativos para ampliar vozes sub-representadas.

Outro requisito essencial é a inclusão digital. Conforme descreve um guia sobre inclusão digital, o termo refere-se ao processo de garantir que todas as pessoas, independentemente de sua condição socioeconômica, tenham acesso às tecnologias de informação e comunicação (TIC); a inclusão digital busca eliminar a exclusão causada por barreiras econômicas, geográficas ou educacionais e criar um ambiente equitativo e acessível a todos (Inova Promos, 2025). As políticas de ciência aberta devem ir além do acesso físico a equipamentos e incluir formação em literacia digital, desenvolvimento de interfaces acessíveis e oferta de conteúdo em múltiplos idiomas. Sem enfrentar o digital divide, a abertura de dados corre o risco de beneficiar apenas os grupos que já dispõem de infraestrutura e competências.

Outra preocupação ética é o viés algorítmico. Como evidenciado por Helm *et al.* (2023), algoritmos podem reproduzir padrões discriminatórios porque o design ou os dados utilizados para treiná-los estão enviesados; esse fenômeno manifesta-se em motores de busca e redes sociais e pode gerar desvantagens sistemáticas para certos grupos (Helm *et al.*, 2023). O problema pode resultar tanto de escolhas de projeto quanto da seleção de dados e está associado ao colonialismo digital, pois modelos de linguagem treinados majoritariamente em inglês reproduzem perspectivas anglo-americanas e marginalizam outras línguas e culturas (Helm *et al.*, 2023). Regulamentos como o GDPR europeu e a proposta de Lei de Inteligência Artificial da UE buscam mitigar esses vieses exigindo transparência, auditoria e responsabilidade sobre algoritmos. Para a ciência aberta, isso implica desenvolver ferramentas que incorporem dados multilíngues e representem a diversidade cultural, além de instituir auditorias de IA e fortalecer a participação de comunidades no desenho de sistemas.

7.4 RECOMENDAÇÕES DE POLÍTICAS PÚBLICAS E INICIATIVAS COMUNITÁRIAS

Governos devem elaborar marcos regulatórios que equilibrem abertura e proteção. Isso inclui mandatos de acesso aberto para pesquisas financiadas com recursos públicos, acompanhados por financiamento de repositórios e capacitação para curadores de dados. A legislação de proteção de dados deve considerar exceções para pesquisa científica, mas exigir transparência e consentimento. Políticas de inclusão digital – como programas de banda larga comunitária e centros de acesso público – são fundamentais para que todos possam usufruir de dados abertos. Além disso, é necessário apoiar editoras independentes e cooperativas de software, por meio de incentivos fiscais e compras públicas, para fortalecer a bibliodiversidade e o desenvolvimento tecnológico local.

Instituições acadêmicas podem reformular sistemas de avaliação para valorizar contribuições à ciência aberta, como compartilhamento de dados, software e participações em ciência cidadã. Devem estabelecer escritórios de gestão de dados e promover treinamento interdisciplinar que combine competências técnicas com noções de ética e justiça epistêmica. Projetos de pesquisa devem ser co-desenhados com comunidades locais, garantindo que agendas científicas respondam a necessidades sociais e que os benefícios do uso de dados retornem às comunidades.

Organizações da sociedade civil desempenham papel duplo de advocacy e inovação. Elas podem pressionar por marcos regulatórios justos, monitorar a implementação de políticas e promover campanhas de alfabetização digital e direitos digitais. No âmbito local, coletivos podem desenvolver projetos de ciência cidadã e tecnologias sociais (ver §6.4) que sirvam de modelos replicáveis. Movimentos sociais têm o poder de mobilizar a opinião pública e disputar narrativas sobre quem controla os dados e para que fim eles são usados.

Outra recomendação é a criação de comitês de governança de dados comunitários e códigos de conduta. Esses comitês, formados por representantes de comunidades, pesquisadores e autoridades públicas, podem deliberar sobre propostas de pesquisa, definir requisitos de consentimento e de auditoria e supervisionar a partilha de benefícios. Códigos de conduta acordados socialmente estabelecem obrigações de respeito, reciprocidade e devolução de resultados, contribuindo para institucionalizar a justiça epistêmica no uso dos dados. Projetos de sequenciamento genômico com povos indígenas do Canadá e programas de dados urbanos em cidades alemãs demonstram que conselhos comunitários podem co-definir as regras de coleta, armazenamento e destruição de amostras biológicas ou dados sensíveis.

7.5 MODELOS ALTERNATIVOS DE FINANCIAMENTO E SUSTENTABILIDADE

A transição para uma ciência aberta ética exige modelos de financiamento que não reproduzam a dependência de grandes editoras. Financiamento público e cooperativo pode sustentar plataformas de acesso aberto e repositórios sem transferir custos aos autores; consórcios internacionais como o SCOAP³ demonstram a viabilidade dessa abordagem. O consórcio, liderado pelo CERN, reuniu bibliotecas, agências de fomento e centros de pesquisa de dezenas de países para converter revistas de física de acesso fechado em acesso aberto, transferindo o custo de publicação dos leitores ou autores para contribuições proporcionais à produção científica de cada país; os artigos permanecem sob licença CC BY, com autores mantendo os direitos autorais (Wikipedia, 2025). Esse modelo ilustra como a cooperação multilateral pode financiar a comunicação científica sem taxas de processamento de artigos (APCs).

Outra via é a economia solidária. Estudo recente sobre economia solidária e bancos comunitários define a economia solidária como um modelo de produção local baseado na autogestão e na inclusão; ela surge como alternativa para enfrentar as consequências negativas dos modelos de negócios convencionais e se apoia em práticas baseadas na confiança e na solidariedade (Oliveira *et al.*, 2023). No campo da ciência, plataformas cooperativas de publicação e bancos de tempo, em que pesquisadores trocam revisão por acesso a serviços, podem reduzir custos e criar redes de apoio. Moedas complementares e sistemas de crédito mutuo têm sido experimentados em comunidades para financiar projetos sociais; adaptá-los à ciência aberta pode incentivar contribuições que não se traduzem imediatamente em lucro, mas têm impacto social.

Outra alternativa é o acesso aberto diamante, modelo em que nem autores nem leitores pagam taxas. Estimativas apontam que existem entre 17 mil e 29 mil periódicos que operam nesse regime, representando a maioria dos títulos registrados em diretórios de periódicos de acesso aberto, mas ainda pouco indexados em bases internacionais (Wikipedia, 2025). Na América Latina, plataformas como SciELO e Redalyc adotam esse modelo há décadas, financiadas por consórcios de universidades e agências públicas, demonstrando que é possível manter revistas sem cobrar APCs. Apoiar o acesso diamante reforça a bibliodiversidade e reduz desigualdades de publicação.

A filantropia e o apoio de fundações podem complementar esses mecanismos. Fundos privados podem sustentar repositórios temáticos, plataformas de pré-prints e programas de ciência cidadã sem exigir retorno financeiro. Além disso, modelos de membresia cooperativa – em que bibliotecas, universidades e indivíduos contribuem com pequenas quantias para manter uma revista – têm sido testados em iniciativas de código aberto e poderiam ser adaptados para financiar infraestruturas de ciência aberta.

Modelos híbridos – combinando financiamento público, privado e comunitário – também são necessários. Parcerias público-privadas podem apoiar infraestruturas abertas desde que salvaguardados princípios de transparência e acesso universal. Campanhas de *crowdfunding* para projetos específicos e sistemas de compensação baseados no impacto social das pesquisas podem diversificar as fontes de recursos. No entanto, é preciso evitar que a lógica de mercado imponha critérios de rentabilidade sobre áreas de pesquisa de interesse público.

7.6 EDUCAÇÃO E CAPACITAÇÃO PARA SOBERANIA DIGITAL

Para que indivíduos e comunidades exerçam soberania sobre seus dados, a educação deve ir além de ensinar a usar ferramentas. Literacia digital crítica envolve compreender como algoritmos e plataformas funcionam, quem controla os dados e quais são os impactos sociais da tecnologia. Programas educacionais devem capacitar estudantes a questionar modelos de negócios das plataformas e reconhecer violações de privacidade. A alfabetização científica pode incorporar projetos de ciência cidadã, integrando ensino e pesquisa e cultivando pensamento crítico.

A educação científica participativa convida comunidades a co-definir currículos e agendas de pesquisa. Incluir conhecimentos indígenas, saberes tradicionais e perspectivas locais no ensino de ciências amplia o pluralismo epistêmico e combate injustiças. Metodologias como aulas em campo, oficinas de tecnologia social e laboratórios cidadãos permitem que participantes desenvolvam experimentos relevantes para suas realidades.

A capacitação técnica é igualmente vital. Universidades e centros de pesquisa devem formar administradores de sistemas, curadores de dados e desenvolvedores de *software* com sensibilidade ética. Programas de mentoria e intercâmbio entre instituições do Norte e do Sul podem acelerar a transferência de conhecimento e reduzir desigualdades. Investir em formação técnica também reduz a dependência de consultorias externas e fortalece a capacidade dos países de desenvolver e manter infraestruturas abertas.

Programas de formação para servidores públicos e legisladores também são essenciais. Políticas de soberania digital e ciência aberta requerem que gestores e reguladores compreendam as implicações técnicas e éticas do uso de dados; sem esse conhecimento, leis podem ser inadequadas ou de difícil aplicação. Bibliotecas públicas e escolares podem atuar como polos de capacitação, oferecendo oficinas sobre privacidade, segurança digital, licenciamento aberto e produção colaborativa de conhecimento. Parcerias com movimentos de *software* livre e coletivos de mídia podem estender essas iniciativas a áreas rurais e periferias urbanas, reforçando a cidadania digital.



8

**PERSPECTIVAS
FUTURAS E AGENDA
DE PESQUISA**

À medida que avançamos na terceira década do século XXI, as tensões entre Ciência Aberta e colonialismo de dados continuam a evoluir, moldadas por desenvolvimentos tecnológicos emergentes, mudanças geopolíticas, e crescente consciência sobre questões de justiça e equidade. Este capítulo examina tendências emergentes e delinea uma agenda de pesquisa para os próximos anos.

A Ciência Aberta, consolidada como um movimento global que visa democratizar o acesso, a produção e os benefícios do conhecimento científico (UNESCO, 2021), encontra-se em um ponto de inflexão crítico. Impulsionada pela revolução digital que facilitou o compartilhamento de dados e publicações em uma escala sem precedentes, a Ciência Aberta agora se depara com uma nova vaga de transformações tecnológicas que apresentam um conjunto complexo e ambivalente de oportunidades e desafios. A mesma tecnologia que promete acelerar descobertas e aprofundar a colaboração global também ergue novas barreiras, concentra poder e ameaça perpetuar desigualdades históricas sob novas formas.

O futuro da Ciência Aberta, portanto, não é um caminho único e predeterminado, mas uma negociação constante e delicada. De um lado, reside a promessa de uma colaboração científica sem fronteiras, transparente e eficiente, capaz de responder a desafios planetários urgentes. Do outro, emerge a necessidade imperativa de construir soberania digital e prevenir a consolidação de um novo colonialismo na era do *Big Data*, onde o valor extraído não são mais os recursos naturais, mas os dados gerados em grande parte pelo Sul Global. Este capítulo argumenta que a sustentabilidade e a legitimidade do movimento da Ciência Aberta dependem fundamentalmente da nossa capacidade de navegar essa tensão.

Para dissecar essa encruzilhada, enfocaremos em como as tecnologias emergentes que estão redefinindo a prática científica, com destaque para a Inteligência Artificial (IA) e a Ciência Descentralizada (DeSci). Em seguida, partiremos para a análise do

dilema central que opõe a abertura irrestrita à soberania digital e ao risco do colonialismo de dados. Por fim, com base nessa análise crítica, será proposta uma agenda de pesquisa robusta, destinada a guiar a comunidade científica, gestores e formuladores de políticas na construção de um ecossistema científico futuro que seja, simultaneamente, aberto, equitativo e soberano.

8.1 AS NOVAS FRONTEIRAS TECNOLÓGICAS DA PESQUISA CIENTÍFICA

A prática científica está sendo profundamente remodelada por tecnologias disruptivas que alteram não apenas as ferramentas disponíveis, mas o próprio ciclo de vida da pesquisa, desde a formulação de hipóteses até a validação e disseminação dos resultados. Uma análise crítica dessas novas fronteiras é essencial para compreender como os princípios da Ciência Aberta podem ser potencializados ou, inversamente, subvertidos. Duas tendências se destacam por seu potencial transformador: a Inteligência Artificial e a Ciência Descentralizada.

8.2 INTELIGÊNCIA ARTIFICIAL (IA) COMO CATALISADORA E DESAFIO

A Inteligência Artificial, especialmente em suas vertentes de aprendizado de máquina e modelos de linguagem generativos, emerge como uma força poderosa para acelerar a produtividade científica. Ferramentas de IA podem analisar vastos conjuntos de dados abertos para identificar padrões, correlacionar informações de

diferentes domínios e formular novas hipóteses, superando a capacidade humana em escala e velocidade (OECD, 2023; ScienceDirect, 2021). A automação de tarefas repetitivas, como a revisão de literatura ou a análise de imagens, libera os pesquisadores para se concentrarem em questões mais complexas e criativas, potencializando os ideais de eficiência e progresso acelerado da Ciência Aberta.

Contudo, a integração da IA na ciência não é isenta de riscos significativos. O principal desafio reside na tensão entre a natureza muitas vezes opaca dos algoritmos e o pilar da reprodutibilidade científica. A chamada “caixa-preta” da IA, onde os processos de tomada de decisão do modelo não são facilmente interpretáveis, colide frontalmente com a necessidade de transparência metodológica (Castelvecchi, 2016). Além disso, questões éticas são proeminentes. Modelos generativos podem produzir “alucinações” — informações plausíveis, mas factualmente incorretas — que, se não verificadas, podem poluir a literatura científica (Alkaisi & McFarlane, 2023). Vieses presentes nos dados de treinamento podem ser amplificados pelos algoritmos, perpetuando preconceitos e desigualdades. Diante disso, a comunidade científica enfrenta o desafio de desenvolver uma governança robusta para o uso ético e transparente da IA, garantindo que ela sirva como uma ferramenta para a abertura, e não para a ofuscação (Rockembach, 2024).

8.3 CIÊNCIA DESCENTRALIZADA (DESCI) E A PROMESSA DA BLOCKCHAIN

Paralelamente à IA, emerge a Ciência Descentralizada (DeSci), um movimento que utiliza tecnologias da Web3, como *blockchain*, contratos inteligentes e Organizações Autônomas Descentralizadas (DAOs), para construir um ecossistema científico mais aberto,

colaborativo e resistente à censura. A DeSci desafia diretamente o modelo tradicional, onde editoras comerciais e agências de fomento centralizam o poder de validação, publicação e financiamento da ciência (Ulam.io, 2025).

O potencial da DeSci se manifesta em aplicações práticas inovadoras. DAOs, como a VitaDAO, que foca em pesquisa sobre longevidade, permitem que comunidades globais financiem coletivamente projetos de pesquisa, democratizando as decisões de fomento. No Brasil, uma iniciativa notável é o desenvolvimento do identificador persistente descentralizado dARK (decentralized Archival Resource Key) pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict). Utilizando *blockchain*, o dARK visa garantir a rastreabilidade e o acesso perene a objetos digitais (artigos, dados, etc.) sem a dependência de uma autoridade central e com custos reduzidos, um passo importante para a autonomia infraestrutural (Observatório *Blockchain*, 2025). Outras aplicações incluem a criação de registros de pesquisa imutáveis e a tokenização da propriedade intelectual, permitindo que pesquisadores mantenham maior controle sobre suas criações.

Apesar da promessa, a DeSci enfrenta seus próprios desafios. A escalabilidade da tecnologia *blockchain*, os custos de transação em algumas redes e a complexidade da experiência do usuário ainda são barreiras para a adoção em massa. Mais crucialmente, a ausência de modelos de governança maduros levanta questões sobre como garantir a qualidade, a revisão por pares e a integridade da pesquisa em um ambiente totalmente descentralizado. A transição de um modelo baseado na reputação de instituições para um modelo baseado em protocolos tecnológicos requer uma cuidadosa reflexão para não substituir um conjunto de problemas por outro.

8.4 O DILEMA CENTRAL: ABERTURA GLOBAL, SOBERANIA DIGITAL E O ESPECTRO DO COLONIALISMO DIGITAL

No cerne das discussões sobre o futuro da Ciência Aberta reside uma tensão fundamental, uma espécie de trilema entre os ideais de abertura global, a necessidade de soberania digital e o risco iminente de um novo colonialismo de dados. A forma como esse dilema for resolvido definirá se a Ciência Aberta cumprirá sua promessa de equidade ou se tornará um mecanismo que, inadvertidamente, aprofunda as assimetrias de poder globais.

O ideal da Ciência Aberta é o de um bem público global. A resposta científica à pandemia de COVID-19, por exemplo, demonstrou como o compartilhamento rápido e aberto de dados genômicos e resultados de pesquisas entre laboratórios de todo o mundo foi decisivo para o desenvolvimento acelerado de vacinas e tratamentos (UNESCO, 2021). Esse é o poder da colaboração sem fronteiras. No entanto, essa visão otimista deve ser contrastada com a realidade geopolítica da infraestrutura digital que a sustenta. As plataformas de computação em nuvem, os repositórios de dados em larga escala e os algoritmos de IA mais avançados são dominados por um pequeno número de corporações multinacionais (*Big Techs*), concentradas majoritariamente no Norte Global. Essa concentração de poder cria uma dependência estrutural.

Na prática científica, esse fenômeno se manifesta quando pesquisadores do Sul Global geram dados valiosos — sobre biodiversidade, saúde pública, questões sociais — que são, por necessidade ou conveniência, armazenados e processados em infraestruturas estrangeiras. Sobre essas infraestruturas, eles não possuem governança ou controle. Essa dinâmica perpetua uma divisão desigual do trabalho intelectual: alguns produzem a matéria-prima (dados), enquanto outros

detêm os meios de produção (tecnologia) e colhem os benefícios da análise avançada, da inovação e da propriedade intelectual. Como aponta Charles Tilly, o controle sobre o conhecimento se organiza em torno de fronteiras que definem quem tem direitos e quem não tem, e a produção de conhecimento quase inevitavelmente reserva vantagens para seus produtores e financiadores (Tilly, in SciELO). A “abertura” sem soberania corre o risco de ser uma via de mão única.

Como contraponto, emerge o conceito de Soberania Digital. Longe de ser um apelo ao isolacionismo, a soberania digital representa o direito e a capacidade de uma nação ou comunidade de governar seu próprio futuro digital. Isso inclui a gestão de seus dados, o desenvolvimento de infraestruturas próprias e a formulação de políticas tecnológicas alinhadas a seus interesses estratégicos e valores (FGV, 2023; Polido, 2024). A soberania digital é, portanto, uma condição necessária para uma participação justa e equitativa na Ciência Aberta global. Iniciativas como a criação de repositórios de dados nacionais, redes federadas e plataformas soberanas, como a Rede Moara desenvolvida pelo Ibict para conectar ciência e software livre (Gov.br, 2025), são passos concretos nessa direção.

A solução para este dilema não reside em uma escolha binária entre uma “soberania fechada” e uma “abertura colonizada”. O caminho mais promissor parece ser o de uma “soberania cooperativa” ou “abertura federada”. O conceito de *Data Spaces*, por exemplo, propõe a criação de ecossistemas onde diferentes entidades podem compartilhar dados de forma segura e governada, sem ceder o controle sobre seus ativos de informação (Aquare.la, 2025). Isso requer uma mudança cultural profunda, superando a visão de que dados são “feudos privados” de laboratórios e instituições, em direção a uma lógica sistêmica de compartilhamento responsável, baseada em padrões de interoperabilidade e governança institucional clara, como defende Thiago Lima Nicodemo (Nicodemo, 2025). Trata-se de construir pontes, não muros, mas garantir que essas pontes sejam construídas sobre fundações sólidas de equidade e respeito mútuo.

8.5 AGENDA DE PESQUISA PARA UMA CIÊNCIA ABERTA, EQUITATIVA E SOBERANA

A construção de uma agenda de pesquisa para uma ciência aberta, equitativa e soberana representa um dos desafios mais complexos e urgentes do século XXI. Em um contexto global marcado por profundas assimetrias na produção e distribuição do conhecimento científico, a simples adoção de práticas de abertura, sem uma reflexão crítica sobre suas implicações, pode inadvertidamente perpetuar e até mesmo amplificar as desigualdades existentes no ecossistema científico mundial.

Esta agenda de pesquisa emerge da necessidade premente de repensar os fundamentos epistemológicos, metodológicos e políticos da ciência contemporânea. Não se trata apenas de tornar os dados e publicações mais acessíveis, mas de questionar fundamentalmente as estruturas de poder que determinam quem produz conhecimento, onde esse conhecimento é validado, como é distribuído e quem se beneficia de seus resultados. A ciência aberta, quando implementada de forma acrítica, corre o risco de se tornar um mecanismo sofisticado de reprodução das hierarquias globais do conhecimento, transformando pesquisadores e instituições do Sul Global em meros fornecedores de dados brutos para agendas de pesquisa definidas nos centros hegemônicos do Norte.

A proposta aqui apresentada articula-se em torno de três eixos estruturantes que se complementam e se reforçam mutuamente. O primeiro eixo dedica-se aos fundamentos para uma ciência equitativa, focando no diagnóstico rigoroso das assimetrias globais e na proposição de mecanismos concretos para combatê-las. O segundo eixo concentra-se na construção prática da soberania científica e digital, explorando as dimensões tecnológicas, políticas e

regulatórias necessárias para que nações e comunidades científicas possam exercer controle efetivo sobre seus dados, infraestruturas e prioridades de pesquisa. O terceiro eixo examina criticamente o papel das tecnologias emergentes, particularmente *blockchain* e inteligência artificial, como potenciais ferramentas de emancipação ou de aprofundamento da dependência tecnológica.

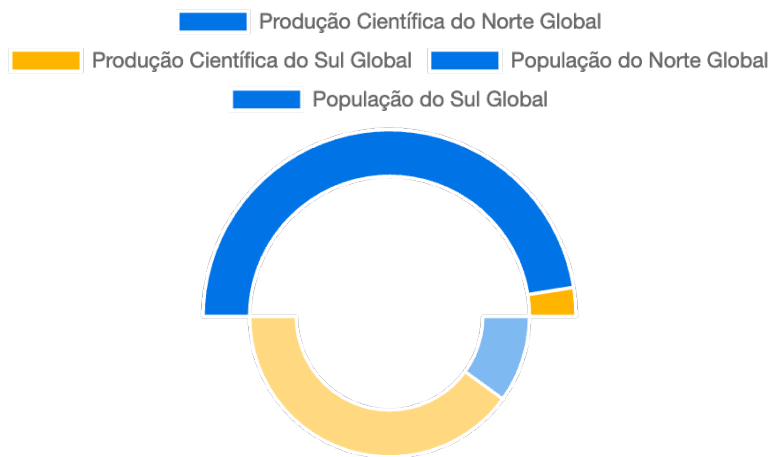
Perpassando esses três eixos, identificamos desafios transversais que demandam atenção especial: a ética na coleta e uso de dados, a necessária mudança cultural na academia e a urgência de programas de letramento digital crítico. Estes desafios não podem ser tratados como questões secundárias ou meramente técnicas, mas devem ser reconhecidos como elementos centrais de qualquer transformação substantiva do sistema científico global.

EIXO 1: FUNDAMENTOS PARA UMA CIÊNCIA EQUITATIVA: MAPEANDO E COMBATENDO AS DESIGUALDADES

Uma implementação acrítica da Ciência Aberta corre o risco de reforçar, em vez de mitigar, as desigualdades existentes no ecossistema científico global. Portanto, o primeiro pilar de uma agenda de pesquisa robusta deve ser dedicado a diagnosticar, compreender e propor soluções para essas assimetrias. O objetivo é garantir que a abertura não se torne um mecanismo que aprofunda divisões históricas.

A primeira linha de investigação é o mapeamento rigoroso da “distribuição desigual da atividade científica” (ANNAN, 2003, citado em TILLY, s.d.). Dados históricos já apontavam que cerca de 95% da nova ciência era criada em países que abrigavam apenas um quinto da população mundial. É crucial atualizar e aprofundar essa análise no contexto digital, investigando a concentração de publicações em periódicos de alto impacto, o direcionamento de financiamentos internacionais e, fundamentalmente, a localização e o controle das infraestruturas de dados (servidores, repositórios, plataformas).

Figura 1 - Ilustração da disparidade global na produção científica versus população, baseada em dados históricos



Fonte: Elaborado pelo autor com base em Annan (2003).

A análise contemporânea das assimetrias científicas globais deve transcender os indicadores tradicionais de produção acadêmica e incorporar uma compreensão mais sofisticada das múltiplas dimensões da desigualdade no conhecimento. O fenômeno da concentração científica não pode ser compreendido apenas através de métricas quantitativas de publicações ou citações, mas requer uma investigação profunda das estruturas sistêmicas que perpetuam essas disparidades.

A geografia da produção científica revela padrões persistentes de concentração que refletem e reproduzem as hierarquias geopolíticas globais. Enquanto os países do Norte Global continuam a dominar não apenas a produção de conhecimento, mas também os mecanismos de sua validação e disseminação, observamos um fenômeno paradoxal: o aumento da participação quantitativa de

países do Sul Global na produção científica não tem sido acompanhado por uma redistribuição proporcional do poder de definir agendas de pesquisa, estabelecer prioridades temáticas ou controlar as infraestruturas críticas do sistema científico.

Esta concentração manifesta-se em múltiplas camadas interconectadas. Na camada da infraestrutura física, a localização geográfica dos servidores que hospedam os principais repositórios científicos, bases de dados e plataformas de publicação concentra-se predominantemente nos países desenvolvidos, criando dependências tecnológicas que transcendem a mera questão do acesso. Na camada institucional, as editoras científicas de maior prestígio, as agências de indexação mais influentes e os sistemas de avaliação de impacto mais utilizados permanecem sob controle de organizações sediadas no Norte Global, estabelecendo critérios de qualidade e relevância que frequentemente privilegiam perspectivas, metodologias e temáticas alinhadas com os interesses e paradigmas dominantes nessas regiões.

A dimensão financeira das assimetrias científicas revela-se particularmente complexa no contexto da ciência aberta. Embora o discurso da abertura prometa democratizar o acesso ao conhecimento, a realidade é que os custos de publicação em acesso aberto têm criado novas barreiras para pesquisadores e instituições com recursos limitados. O modelo de Article Processing Charges (APCs), amplamente adotado por periódicos de prestígio, transfere os custos da publicação dos leitores para os autores, criando um sistema onde a capacidade de pagar determina não apenas o acesso ao conhecimento, mas também a possibilidade de contribuir para sua produção e disseminação.

A pesquisa deve ir além dos números e analisar qualitativamente como as políticas de Ciência Aberta impactam pesquisadores e instituições com menos recursos. Investigações são necessárias para verificar se, na prática, cientistas do Sul Global estão sendo

relegados ao papel de meros “consumidores de tecnologia” ou fornecedores de dados brutos para agendas de pesquisa definidas no Norte (OUTRAS PALAVRAS, 2022). Adicionalmente, é preciso analisar as “desigualdades categóricas” (TILLY, s.d.) dentro do próprio Sul Global, examinando como recortes de gênero, raça e localização geográfica influenciam o acesso e a capacidade de participar plenamente do ecossistema científico aberto.

O conceito de “desigualdades categóricas” proposto por Tilly oferece um framework analítico particularmente útil para compreender como diferentes formas de exclusão se interseccionam e se reforçam mutuamente no contexto científico. Estas desigualdades não são meramente resultado de diferenças individuais de capacidade ou recursos, mas emergem de sistemas organizados de distinção social que criam e mantêm fronteiras duráveis entre grupos. No contexto da ciência global, essas fronteiras manifestam-se através de múltiplos mecanismos: linguísticos (dominância do inglês como língua científica), institucionais (prestígio diferencial de universidades e centros de pesquisa), metodológicos (privilégio de certas abordagens epistemológicas) e tecnológicos (acesso desigual a infraestruturas digitais avançadas).

A análise das desigualdades de gênero na ciência aberta revela dinâmicas particularmente complexas. Embora as mulheres tenham aumentado significativamente sua participação na produção científica global nas últimas décadas, persistem disparidades substanciais em termos de visibilidade, reconhecimento e acesso a posições de liderança. No contexto da ciência aberta, essas disparidades podem ser tanto atenuadas quanto amplificadas, dependendo de como as práticas de abertura são implementadas. Por um lado, a transparência dos processos de revisão por pares e a disponibilização aberta de dados podem reduzir alguns vieses implícitos que historicamente prejudicaram a participação feminina. Por outro lado, a intensificação da competição por visibilidade em plataformas digitais pode reproduzir e até mesmo amplificar dinâmicas de exclusão baseadas em gênero.

As desigualdades raciais e étnicas na ciência global constituem outra dimensão crítica que requer investigação aprofundada. A sub-representação de pesquisadores negros, indígenas e de outras minorias étnicas em posições de prestígio no sistema científico global não pode ser compreendida apenas como resultado de diferenças de acesso à educação, mas deve ser analisada como produto de sistemas históricos de exclusão que continuam a operar através de mecanismos sutis, mas eficazes. A ciência aberta, ao tornar mais visíveis os processos de produção e validação do conhecimento, oferece oportunidades únicas para identificar e combater esses vieses, mas também pode inadvertidamente reproduzi-los se não for acompanhada de políticas específicas de inclusão e equidade.

A dimensão geográfica das desigualdades científicas dentro do próprio Sul Global merece atenção especial. A concentração da produção científica em grandes centros urbanos e universidades de elite reproduz, em escala nacional e regional, as mesmas dinâmicas de exclusão observadas em nível global. Pesquisadores e instituições localizados em regiões periféricas, mesmo dentro de países considerados emergentes na produção científica, enfrentam barreiras adicionais de acesso a recursos, infraestruturas e redes de colaboração que limitam sua capacidade de participar plenamente do ecossistema científico aberto.

Equidade nas práticas de colaboração: repensando as dinâmicas de poder na produção colaborativa do conhecimento

A colaboração é um pilar da Ciência Aberta, mas pode mascarar dinâmicas de poder desiguais. A agenda de pesquisa deve, portanto, focar no desenvolvimento e avaliação de modelos de colaboração (Norte-Sul e Sul-Sul) que garantam justiça e reciprocidade. Isso implica investigar práticas de coautoria, o reconhecimento equitativo de contribuições (não apenas a coleta de dados, mas também a análise e conceitualização) e a capacitação mútua, de modo a

evitar o “extrativismo de dados”. Uma linha de pesquisa promissora é o impacto de métricas de avaliação alternativas, que se afastem da obsessão por publicações em periódicos de alto custo e difícil acesso, e passem a valorizar a colaboração genuína, a formação de recursos humanos e o impacto local da pesquisa (EUA, 2022).

A retórica da colaboração científica internacional frequentemente obscurece relações de poder profundamente assimétricas que reproduzem, no âmbito da produção do conhecimento, as mesmas dinâmicas de dominação observadas em outras esferas das relações internacionais. O discurso da “parceria” e da “cooperação” pode servir como verniz ideológico para práticas que, na realidade, perpetuam formas sofisticadas de extrativismo epistêmico, onde pesquisadores e instituições do Sul Global fornecem dados, amostras, contextos locais e força de trabalho, enquanto a conceituação teórica, a análise interpretativa e o reconhecimento acadêmico concentram-se nos centros hegemônicos do Norte.

A análise crítica das práticas de coautoria revela padrões sistêmicos que merecem investigação aprofundada. A ordem dos autores em publicações científicas, longe de ser uma questão meramente técnica, reflete e reproduz hierarquias de poder que têm implicações concretas para as carreiras acadêmicas e o reconhecimento profissional. Estudos preliminares sugerem que, em colaborações Norte-Sul, pesquisadores do Norte Global tendem a ocupar posições de primeiro ou último autor (tradicionalmente associadas à liderança intelectual do projeto) com frequência desproporcional à sua contribuição efetiva para a coleta e análise de dados. Esta dinâmica é particularmente problemática em áreas como saúde global, estudos ambientais e pesquisa social, onde o conhecimento local e o acesso a contextos específicos são fundamentais para a qualidade da pesquisa.

O fenômeno do “extrativismo de dados” representa uma das manifestações mais perversas das assimetrias na colaboração científica. Este conceito, inspirado na crítica pós-colonial ao extrativismo

econômico, descreve práticas onde pesquisadores e instituições com maior poder econômico e simbólico extraem dados, amostras e conhecimentos de contextos locais sem estabelecer relações genuinamente recíprocas com as comunidades e pesquisadores que tornam essa extração possível. O extrativismo de dados não se limita à apropriação de recursos materiais, mas inclui também a apropriação de conhecimentos tradicionais, saberes locais e perspectivas culturais que são posteriormente recontextualizados e reinterpretados segundo marcos teóricos e metodológicos dominantes.

A superação do extrativismo de dados requer o desenvolvimento de novos modelos de colaboração baseados em princípios de reciprocidade, transparência e empoderamento mútuo. Estes modelos devem ir além da simples inclusão de pesquisadores do Sul Global em projetos concebidos e liderados no Norte, para promover formas genuínas de co-criação onde as agendas de pesquisa, as metodologias empregadas e os marcos interpretativos são desenvolvidos colaborativamente desde o início do processo.

A capacitação mútua emerge como elemento central de qualquer modelo genuinamente equitativo de colaboração científica. Diferentemente dos modelos tradicionais de “transferência de tecnologia” ou “desenvolvimento de capacidades”, que pressupõem um fluxo unidirecional de conhecimento dos centros “avançados” para as “periferias”, a capacitação mútua reconhece que todos os participantes de uma colaboração possuem conhecimentos, habilidades e perspectivas valiosas que podem enriquecer o processo de pesquisa. Este reconhecimento não é apenas uma questão de cortesia política, mas reflete uma compreensão epistemológica mais sofisticada sobre a natureza plural e contextual do conhecimento científico.

A implementação de práticas de capacitação mútua requer mudanças substanciais nas estruturas institucionais que governam a colaboração científica internacional. Programas de financiamento devem ser redesenhados para incluir recursos específicos para atividades de

capacitação bilateral, reconhecendo que o investimento em formação de recursos humanos é tão importante quanto a aquisição de equipamentos ou a coleta de dados. Universidades e centros de pesquisa devem desenvolver políticas institucionais que incentivem e recompensem a participação em atividades de capacitação mútua, reconhecendo-as como contribuições legítimas para a produção científica.

As métricas de avaliação alternativas representam uma fronteira crucial para a promoção de práticas mais equitativas de colaboração científica. O sistema atual de avaliação acadêmica, baseado predominantemente em indicadores bibliométricos como fator de impacto de periódicos e número de citações, tende a privilegiar formas de produção científica que se alinham com os padrões e interesses dos centros hegemônicos. A proposta de métricas alternativas, mencionada no documento original (EUA, 2022), aponta para a necessidade de desenvolver sistemas de avaliação que valorizem dimensões da atividade científica tradicionalmente negligenciadas: a formação de recursos humanos, o impacto social local, a contribuição para políticas públicas, a inovação metodológica e a promoção da equidade.

O desenvolvimento de métricas alternativas não é apenas uma questão técnica, mas envolve escolhas políticas e epistemológicas fundamentais sobre o que constitui conhecimento valioso e como esse valor deve ser mensurado. A valorização da “colaboração genuína” requer o desenvolvimento de indicadores capazes de capturar a qualidade das relações estabelecidas entre pesquisadores, a reciprocidade dos benefícios gerados e o grau de empoderamento das partes envolvidas. Estes indicadores devem ser suficientemente sofisticados para distinguir entre colaborações meramente instrumentais e parcerias genuinamente transformadoras.

A promoção do “impacto local da pesquisa” como critério de avaliação representa um desafio particular, pois requer o desenvolvimento de metodologias capazes de mensurar transformações sociais, econômicas e políticas que frequentemente se manifestam em escalas

temporais e espaciais diferentes daquelas tradicionalmente consideradas na avaliação acadêmica. O impacto local não pode ser reduzido a indicadores quantitativos simples, mas deve incorporar avaliações qualitativas que considerem a perspectiva das comunidades e atores locais sobre a relevância e utilidade da pesquisa realizada.

8.6 INCLUSÃO E CIÊNCIA CIDADÃ: DEMOCRATIZANDO A PRODUÇÃO DO CONHECIMENTO ALÉM DA ACADEMIA

A democratização do conhecimento não se limita à comunidade científica. A pesquisa deve avaliar a eficácia e os desafios éticos de projetos de ciência cidadã, que engajam o público na coleta e análise de dados. Tais projetos têm o potencial de conectar a ciência às “necessidades da sociedade” (UNESCO, 2021), mas exigem estudos sobre como garantir a qualidade dos dados, o reconhecimento dos participantes e a proteção contra a exploração. Além disso, uma ciência verdadeiramente aberta deve investigar como “outros sistemas de conhecimento”, como saberes tradicionais, indígenas e populares, podem dialogar e ser integrados de forma ética e respeitosa, sem apropriação ou descaracterização, enriquecendo o ecossistema científico global (UNESCO, 2021).

A ciência cidadã representa uma das manifestações mais promissoras e, simultaneamente, mais problemáticas da democratização do conhecimento científico. Embora o discurso da participação pública na ciência seja frequentemente celebrado como um avanço inequívoco em direção a formas mais inclusivas e socialmente relevantes de produção do conhecimento, a realidade revela tensões e contradições que demandam análise crítica cuidadosa.

O potencial democratizante da ciência cidadã reside em sua capacidade de romper com o monopólio institucional da produção científica, criando oportunidades para que indivíduos e comunidades sem credenciais acadêmicas formais possam contribuir significativamente para o avanço do conhecimento. Esta abertura é particularmente valiosa em contextos onde o conhecimento local, a experiência vivida e a observação cotidiana podem fornecer insights cruciais para a compreensão de fenômenos complexos. Projetos bem-sucedidos de ciência cidadã têm demonstrado que a participação pública pode não apenas ampliar a escala e o escopo da coleta de dados, mas também enriquecer a qualidade da pesquisa através da incorporação de perspectivas e conhecimentos que permaneceriam inacessíveis aos pesquisadores acadêmicos.

No entanto, a implementação da ciência cidadã também pode reproduzir e até mesmo amplificar as mesmas dinâmicas de exclusão e exploração observadas na ciência tradicional. Quando projetos de ciência cidadã são concebidos e controlados exclusivamente por instituições acadêmicas, existe o risco de que os participantes públicos sejam reduzidos a meros coletores de dados, executando tarefas definidas por outros sem participar efetivamente dos processos de conceituação, análise e interpretação. Esta dinâmica reproduz, em nova roupagem, as relações hierárquicas tradicionais entre “produtores” e “consumidores” de conhecimento, limitando o potencial genuinamente transformador da participação pública na ciência.

A questão da qualidade dos dados em projetos de ciência cidadã tem sido objeto de intenso debate, revelando tensões fundamentais entre diferentes concepções de rigor científico e inclusão social. Críticos da ciência cidadã argumentam que a participação de não-especialistas inevitavelmente compromete a precisão e confiabilidade dos dados coletados, enquanto defensores sustentam que protocolos adequados de treinamento e validação podem garantir padrões de qualidade comparáveis àqueles obtidos através de métodos tradicionais. Esta controvérsia não é meramente técnica,

mas reflete concepções mais profundas sobre a natureza do conhecimento científico e os critérios de validação epistêmica.

A pesquisa sobre ciência cidadã deve desenvolver metodologias sofisticadas para avaliar a qualidade dos dados que vão além de simples comparações com padrões estabelecidos por especialistas. É necessário reconhecer que diferentes tipos de conhecimento podem ter critérios de qualidade específicos e que a diversidade de perspectivas pode, em muitos casos, enriquecer a compreensão de fenômenos complexos mesmo quando introduz variabilidade nos dados coletados. O desafio consiste em desenvolver frameworks de avaliação que sejam simultaneamente rigorosos e inclusivos, capazes de distinguir entre variabilidade problemática (resultado de erros ou inconsistências) e variabilidade informativa (resultado de diferentes perspectivas ou contextos).

O reconhecimento dos participantes em projetos de ciência cidadã constitui outra dimensão crítica que requer atenção cuidadosa. A tradição acadêmica de atribuição de crédito através de autorias e citações não se adapta facilmente a contextos onde centenas ou milhares de indivíduos contribuem para a coleta e análise de dados. É necessário desenvolver novos modelos de reconhecimento que sejam simultaneamente práticos e justos, capazes de valorizar adequadamente as contribuições dos participantes sem burocratizar excessivamente os processos de participação.

A proteção contra a exploração em projetos de ciência cidadã requer o desenvolvimento de frameworks éticos específicos que considerem as particularidades da participação pública na pesquisa. Os códigos de ética tradicionais, desenvolvidos para regular a pesquisa com seres humanos como “sujeitos” de investigação, não se aplicam diretamente a contextos onde os indivíduos participam como “colaboradores” na produção do conhecimento. É necessário desenvolver novos princípios éticos que garantam que a participação seja verdadeiramente voluntária, informada e mutuamente benéfica.

A integração de “outros sistemas de conhecimento” representa talvez o desafio mais complexo e politicamente sensível na construção de uma ciência verdadeiramente inclusiva. A menção da UNESCO (2021) aos saberes tradicionais, indígenas e populares aponta para a necessidade de superar a falsa dicotomia entre conhecimento “científico” e “não-científico”, reconhecendo que diferentes tradições epistêmicas podem oferecer contribuições valiosas para a compreensão de fenômenos naturais e sociais.

A integração respeitosa de diferentes sistemas de conhecimento requer, em primeiro lugar, o reconhecimento de que estes sistemas possuem suas próprias metodologias, critérios de validação e formas de organização que não podem ser simplesmente subsumidas aos padrões da ciência ocidental moderna. O diálogo entre diferentes tradições epistêmicas deve ser baseado em princípios de reciprocidade e respeito mútuo, evitando tanto a apropriação descontextualizada quanto a romantização acrítica de saberes tradicionais.

O risco de apropriação cultural e epistêmica é particularmente agudo quando conhecimentos tradicionais são incorporados em projetos de pesquisa sem o consentimento adequado ou sem o reconhecimento apropriado das comunidades que os desenvolveram e preservaram. A pesquisa sobre integração de sistemas de conhecimento deve desenvolver protocolos rigorosos para garantir que esta integração seja baseada em relações de colaboração genuína, onde as comunidades detentoras de conhecimentos tradicionais participam não apenas como fornecedoras de informação, mas como parceiras na definição de objetivos, metodologias e usos dos resultados da pesquisa.

A descaracterização de saberes tradicionais através de sua tradução para linguagens e frameworks científicos dominantes representa outro risco que deve ser cuidadosamente evitado. A integração respeitosa requer o desenvolvimento de metodologias capazes de preservar a integridade e especificidade de diferentes tradições epistêmicas, mesmo quando estas são colocadas em diálogo com abordagens

científicas convencionais. Isto pode exigir a criação de novos formatos de publicação e disseminação que sejam capazes de acomodar diferentes formas de organização e apresentação do conhecimento.

EIXO 2: CONSTRUINDO A SOBERANIA CIENTÍFICA E DIGITAL NA PRÁTICA

A soberania na era digital não significa isolamento, mas sim a capacidade de uma nação ou comunidade científica de governar seus próprios dados, infraestruturas e prioridades de pesquisa. Este eixo da agenda de pesquisa foca nos elementos práticos — tecnológicos, políticos e regulatórios — necessários para construir essa autonomia no contexto da Ciência Aberta.

Infraestruturas nacionais e soberanas: arquiteturas tecnológicas para a autonomia científica

A dependência de plataformas estrangeiras é um dos maiores riscos à soberania. A pesquisa deve, portanto, analisar a viabilidade, os custos e os modelos de governança para a criação de uma “infraestrutura digital nacional e soberana” (NICODEMO, 2025). Isso inclui não apenas repositórios de dados, mas também plataformas de publicação, redes de computação de alto desempenho e sistemas de identificação persistente. Iniciativas brasileiras como a Rede Moara, que conecta ciência, *software* livre e soberania tecnológica, são objetos de estudo cruciais (REDE MOARA, 2025). O conceito de “Data Spaces”, que propõe ecossistemas federados onde os participantes compartilham dados de forma segura e governada sem ceder o controle, deve ser investigado como um modelo alternativo às arquiteturas centralizadas das Big Techs (AQUARE.LA, 2025). Um desafio central para a pesquisa é como garantir a interoperabilidade entre esses sistemas soberanos, para que a autonomia não se converta em fragmentação e isolamento científico.

A construção de infraestruturas digitais nacionais e soberanas para a ciência representa um dos desafios mais complexos e estratégicos da contemporaneidade. A questão transcende aspectos meramente técnicos para envolver dimensões geopolíticas, econômicas e epistemológicas fundamentais sobre o futuro da produção e circulação do conhecimento científico. A dependência de infraestruturas controladas por corporações transnacionais ou governos estrangeiros não constitui apenas um risco técnico ou econômico, mas uma vulnerabilidade estratégica que pode comprometer a capacidade de nações e comunidades científicas de definir suas próprias prioridades de pesquisa e controlar o uso de seus dados.

A análise da viabilidade de infraestruturas nacionais soberanas deve começar por uma compreensão clara dos diferentes componentes que constituem o ecossistema digital da ciência contemporânea. Estes componentes incluem, mas não se limitam a: repositórios de dados de pesquisa, plataformas de publicação científica, sistemas de identificação persistente, redes de computação de alto desempenho, infraestruturas de preservação digital, plataformas de colaboração científica, sistemas de gestão de identidade digital e ferramentas de análise e visualização de dados. Cada um destes componentes apresenta desafios específicos em termos de desenvolvimento, implementação e manutenção, mas também oportunidades únicas para a construção de alternativas soberanas.

Os repositórios de dados científicos constituem talvez o componente mais crítico de qualquer infraestrutura científica soberana. A capacidade de armazenar, organizar, preservar e disponibilizar dados de pesquisa de forma segura e sustentável é fundamental para a autonomia científica. No entanto, a construção de repositórios verdadeiramente soberanos vai muito além da simples instalação de servidores em território nacional. É necessário desenvolver capacidades técnicas locais para a gestão de grandes volumes de dados, implementar padrões internacionais de metadados e interoperabilidade, estabelecer políticas claras de acesso e uso, e garantir a sustentabilidade financeira e institucional a longo prazo.

A experiência internacional oferece lições valiosas sobre os desafios e oportunidades associados ao desenvolvimento de repositórios nacionais. Países como o Reino Unido, com o UK Data Service, a França, com o PROGEDO, e a Alemanha, com o GESIS, desenvolveram infraestruturas robustas que combinam investimento público significativo, parcerias institucionais sólidas e padrões técnicos avançados. No entanto, a simples replicação destes modelos em contextos diferentes pode não ser apropriada, sendo necessário desenvolver abordagens que considerem as especificidades locais em termos de recursos disponíveis, prioridades de pesquisa e marcos regulatórios.

As plataformas de publicação científica representam outro componente crucial de qualquer infraestrutura soberana. O domínio exercido por grandes editoras comerciais sobre os canais de disseminação do conhecimento científico constitui uma das formas mais perversas de dependência no sistema científico global. A construção de alternativas soberanas requer não apenas o desenvolvimento de plataformas técnicas robustas, mas também a criação de novos modelos de negócio, sistemas de avaliação por pares e mecanismos de reconhecimento acadêmico que possam competir efetivamente com as estruturas estabelecidas.

A iniciativa brasileira da Rede Moara, mencionada no documento original (REDE MOARA, 2025), representa um exemplo particularmente interessante de como diferentes dimensões da soberania tecnológica podem ser articuladas de forma integrada. A conexão entre ciência, software livre e soberania tecnológica proposta pela Rede Moara sugere um modelo de desenvolvimento que vai além da simples substituição de tecnologias proprietárias por alternativas nacionais, para promover uma transformação mais profunda nas relações entre tecnologia, conhecimento e poder.

O software livre emerge como elemento fundamental de qualquer estratégia de soberania digital, não apenas por razões econômicas (redução de custos de licenciamento), mas principalmente

por razões políticas e epistemológicas. O controle sobre o código-fonte das ferramentas utilizadas na produção científica é essencial para garantir transparência, auditabilidade e adaptabilidade às necessidades específicas de diferentes comunidades científicas. Além disso, o modelo de desenvolvimento colaborativo característico do software livre oferece oportunidades únicas para a construção de capacidades técnicas locais e para o estabelecimento de redes de cooperação Sul-Sul.

Os sistemas de identificação persistente constituem uma dimensão frequentemente negligenciada, mas crucial, da soberania científica digital. A dependência de sistemas como o DOI (Digital Object Identifier), controlados por organizações internacionais com sede nos países desenvolvidos, cria vulnerabilidades que podem comprometer o acesso a longo prazo aos recursos científicos. O desenvolvimento de alternativas soberanas, como o sistema dARK mencionado posteriormente no documento, representa uma linha de pesquisa e desenvolvimento particularmente promissora.

As redes de computação de alto desempenho representam talvez o componente mais desafiador em termos de recursos necessários para o desenvolvimento de capacidades soberanas. A pesquisa científica contemporânea, particularmente em áreas como modelagem climática, bioinformática, física de partículas e inteligência artificial, depende crescentemente de capacidades computacionais que requerem investimentos substanciais em hardware, *software* e recursos humanos especializados. A construção de capacidades nacionais nesta área requer estratégias de longo prazo que combinem investimento público, parcerias internacionais e desenvolvimento de recursos humanos.

O conceito de "Data Spaces", mencionado no documento original (AQUARE.LA, 2025), oferece uma perspectiva particularmente interessante para superar a falsa dicotomia entre centralização e fragmentação. Os Data Spaces propõem arquiteturas federadas onde diferentes

organizações podem compartilhar dados e recursos de forma controlada, mantendo a governança local sobre seus próprios ativos digitais. Este modelo oferece oportunidades para a construção de ecossistemas científicos que sejam simultaneamente soberanos e interoperáveis, evitando tanto a dependência de plataformas centralizadas quanto o isolamento resultante de sistemas completamente autônomos.

A implementação de Data Spaces para a ciência requer o desenvolvimento de padrões técnicos, protocolos de governança e marcos legais que permitam o compartilhamento seguro e controlado de dados entre diferentes instituições e jurisdições. Isto inclui questões complexas como autenticação e autorização federada, controle de acesso granular, auditoria de uso de dados e resolução de disputas. A pesquisa nesta área deve investigar tanto os aspectos técnicos quanto os aspectos políticos e legais necessários para a implementação bem-sucedida de tais sistemas.

A questão da interoperabilidade representa talvez o desafio mais complexo na construção de infraestruturas soberanas. Por um lado, a interoperabilidade é essencial para evitar que a busca por autonomia resulte em fragmentação e isolamento. Por outro lado, a adoção acrítica de padrões desenvolvidos por e para os sistemas dominantes pode reproduzir formas sutis de dependência. A pesquisa deve investigar estratégias para o desenvolvimento de padrões de interoperabilidade que sejam simultaneamente abertos, inclusivos e respeitosos da diversidade de contextos e necessidades.

O desenvolvimento de capacidades técnicas locais emerge como pré-requisito fundamental para qualquer estratégia de soberania digital. Não basta instalar infraestruturas desenvolvidas em outros contextos; é necessário desenvolver capacidades endógenas para a concepção, implementação, manutenção e evolução dessas infraestruturas. Isto requer investimentos substanciais em educação, formação profissional e pesquisa e desenvolvimento, bem como a criação de incentivos para a retenção de talentos e o desenvolvimento de carreiras técnicas especializadas.

Políticas públicas e marcos regulatórios: construindo o arcabouço legal para a soberania científica

A tecnologia por si só não garante a soberania; ela precisa ser sustentada por um arcabouço legal e político robusto. Uma linha de pesquisa urgente é a formulação de políticas públicas que definam dados de pesquisa financiados com recursos públicos como bens públicos, estabelecendo diretrizes claras para sua coleta, gestão, preservação e acesso (NICODEMO, 2025). É imperativo realizar análises comparativas de marcos regulatórios internacionais, como o CLOUD Act dos EUA, que permite que autoridades americanas acessem dados armazenados por empresas de tecnologia do país, mesmo que esses dados estejam em servidores fora dos EUA. A pesquisa deve elucidar os riscos que tais leis representam para a soberania de dados de pesquisas brasileiras hospedadas em provedores estrangeiros (PENSO, s.d.). Finalmente, são necessários estudos aprofundados sobre modelos de governança de dados que consigam equilibrar os princípios da abertura com a segurança, a privacidade dos participantes da pesquisa e os interesses estratégicos nacionais (JOTA, 2025).

A construção de marcos regulatórios adequados para a soberania científica digital representa um dos desafios mais complexos e politicamente sensíveis da agenda de pesquisa proposta. A intersecção entre direito, tecnologia e política científica cria um campo de tensões onde diferentes interesses e valores frequentemente entram em conflito. A formulação de políticas públicas eficazes requer uma compreensão sofisticada não apenas dos aspectos técnicos envolvidos, mas também das dinâmicas políticas, econômicas e sociais que moldam o desenvolvimento e a implementação de marcos regulatórios.

A definição de dados de pesquisa financiados com recursos públicos como bens públicos, mencionada por Nicodemo (2025), representa uma mudança paradigmática fundamental nas concepções tradicionais sobre propriedade intelectual e acesso ao conhecimento.

Esta proposta desafia diretamente os modelos de negócio baseados na privatização e comercialização de resultados de pesquisa, criando tensões com interesses estabelecidos tanto no setor privado quanto em certas instituições acadêmicas que dependem de receitas geradas pela exploração comercial de propriedade intelectual.

A implementação efetiva de políticas que tratem dados de pesquisa pública como bens públicos requer o desenvolvimento de frameworks legais sofisticados que sejam capazes de distinguir entre diferentes tipos de dados, considerar questões de privacidade e segurança, e estabelecer mecanismos apropriados de governança e acesso. Isto inclui a definição clara de critérios para determinar quando dados devem ser considerados públicos, o estabelecimento de exceções apropriadas para dados sensíveis ou estratégicos, e a criação de instituições responsáveis pela implementação e fiscalização dessas políticas.

A análise comparativa de marcos regulatórios internacionais revela a diversidade de abordagens adotadas por diferentes países para equilibrar os imperativos da abertura científica com considerações de soberania e segurança nacional. O CLOUD Act americano, mencionado no documento original (PENSO, s.d.), representa um exemplo particularmente problemático de como legislações nacionais podem ter implicações extraterritoriais que comprometem a soberania de dados de outros países. A capacidade das autoridades americanas de acessar dados armazenados por empresas de tecnologia americanas, independentemente da localização física dos servidores, cria vulnerabilidades significativas para países que dependem de provedores de serviços baseados nos EUA.

A pesquisa sobre os riscos representados pelo CLOUD Act e legislações similares deve ir além de análises puramente legais para investigar as implicações práticas dessas leis para a condução de pesquisas científicas. Isto inclui questões como: em que circunstâncias dados de pesquisa podem ser acessados por autoridades estrangeiras? Quais são as implicações para a confidencialidade de

dados de participantes de pesquisa? Como essas leis afetam a capacidade de pesquisadores de cumprir compromissos éticos e legais com participantes de pesquisa? Quais são as consequências para a colaboração científica internacional?

O desenvolvimento de alternativas soberanas aos provedores de serviços dominados por empresas americanas requer não apenas capacidades técnicas, mas também marcos regulatórios que incentivem e protejam essas alternativas. Isto pode incluir políticas de compras públicas que privilegiem provedores nacionais, incentivos fiscais para o desenvolvimento de capacidades locais, e regulamentações que estabeleçam requisitos de soberania de dados para certas categorias de pesquisa.

A questão da governança de dados representa talvez o aspecto mais complexo do desenvolvimento de marcos regulatórios para a soberania científica. A necessidade de equilibrar os princípios da abertura com considerações de segurança, privacidade e interesses estratégicos nacionais, mencionada no documento original (JOTA, 2025), requer o desenvolvimento de modelos de governança sofisticados que sejam capazes de acomodar múltiplos valores e interesses frequentemente conflitantes.

Os modelos de governança de dados devem ser capazes de operar em múltiplas escalas e contextos, desde a governança de projetos individuais de pesquisa até a governança de infraestruturas nacionais de dados. Isto requer o desenvolvimento de frameworks institucionais que incluam representação apropriada de diferentes stakeholders (pesquisadores, instituições, governo, sociedade civil), mecanismos transparentes de tomada de decisão, e procedimentos claros para a resolução de conflitos e disputas.

A dimensão internacional da governança de dados científicos adiciona camadas adicionais de complexidade. A ciência é inerentemente internacional, e qualquer framework de governança

deve ser capaz de acomodar a necessidade de colaboração e compartilhamento de dados através de fronteiras nacionais. Isto requer o desenvolvimento de acordos internacionais, protocolos de reciprocidade e mecanismos de coordenação que respeitem a soberania nacional enquanto facilitam a cooperação científica.

A proteção da privacidade dos participantes de pesquisa representa uma consideração fundamental em qualquer framework de governança de dados científicos. As leis de proteção de dados, como o GDPR europeu e a LGPD brasileira, estabelecem requisitos rigorosos para o processamento de dados pessoais que têm implicações significativas para a condução de pesquisas científicas. A pesquisa deve investigar como esses requisitos podem ser cumpridos de forma que proteja efetivamente a privacidade dos indivíduos sem inviabilizar a pesquisa científica legítima.

O conceito de “interesses estratégicos nacionais” em relação aos dados científicos requer definição cuidadosa e implementação equilibrada. Por um lado, é legítimo que países busquem proteger dados que possam ter implicações para a segurança nacional, a competitividade econômica ou outros interesses estratégicos. Por outro lado, a definição excessivamente ampla de dados estratégicos pode comprometer os princípios da ciência aberta e limitar desnecessariamente a colaboração científica internacional.

A pesquisa deve investigar critérios apropriados para a classificação de dados científicos como estratégicos, procedimentos para a revisão e reclassificação dessas determinações, e mecanismos para equilibrar considerações de segurança com os benefícios da abertura científica. Isto pode incluir o desenvolvimento de sistemas de classificação graduada que permitam diferentes níveis de acesso e compartilhamento dependendo da sensibilidade dos dados e da natureza dos usuários.

A implementação efetiva de marcos regulatórios para a soberania científica digital requer também o desenvolvimento de capacidades institucionais apropriadas. Isto inclui a criação ou fortalecimento de agências regulatórias especializadas, o desenvolvimento de expertise técnica e legal dentro do governo, e o estabelecimento de mecanismos de coordenação entre diferentes órgãos e níveis de governo. A pesquisa deve investigar modelos institucionais apropriados para diferentes contextos nacionais, considerando fatores como estrutura federativa, capacidades administrativas existentes e tradições regulatórias.

A dimensão internacional da regulação de dados científicos requer também atenção especial ao desenvolvimento de capacidades de negociação e participação em fóruns internacionais. A capacidade de influenciar o desenvolvimento de padrões e normas internacionais é crucial para garantir que os interesses de países em desenvolvimento sejam adequadamente representados nos frameworks globais de governança de dados científicos.

EIXO 3: TECNOLOGIAS EMERGENTES COMO FERRAMENTAS DE SOBERANIA OU DOMINAÇÃO

Tecnologias como *Blockchain* e Inteligência Artificial (IA) estão no centro da transformação digital e apresentam uma dualidade fascinante: podem tanto aprofundar a dependência tecnológica quanto fornecer ferramentas inéditas para a construção da soberania. Este eixo de pesquisa se dedica a analisar criticamente esse potencial ambivalente.

Ciência Descentralizada (DeSci) e Blockchain: promessas e limitações da descentralização científica

a Ciência Descentralizada (DeSci) surge como um movimento que utiliza a tecnologia *blockchain* para criar ecossistemas científicos alternativos, com a promessa de revolucionar o financiamento,

a publicação e a revisão por pares (OBSERVATORIO BLOCKCHAIN, 2025). A pesquisa precisa investigar o potencial real da DeSci para criar sistemas mais transparentes, justos e resistentes à censura. Casos de uso concretos, como o identificador persistente descentralizado dARK, desenvolvido pelo Ibict no Brasil, devem ser analisados como alternativas a sistemas de PID (Persistent Identifier) como o DOI, que são centralizados e possuem custos de manutenção (IBICT, 2025). Contudo, a pesquisa também deve abordar com seriedade os desafios da DeSci: a escalabilidade das redes *blockchain*, os custos de transação (as “gas fees”), a dificuldade de garantir o controle de qualidade em um ambiente descentralizado e a necessidade de um amplo letramento digital para que a comunidade científica possa adotar essas ferramentas de forma eficaz (ULAM.IO, 2025).

A emergência da Ciência Descentralizada (DeSci) representa uma das propostas mais radicais e potencialmente transformadoras para a reorganização do sistema científico global. Fundamentada nos princípios da tecnologia *blockchain* e inspirada pelos ideais de descentralização que caracterizam o movimento Web3, a DeSci promete não apenas resolver problemas específicos do sistema científico atual, mas reimaginar fundamentalmente as relações de poder, os mecanismos de incentivo e as estruturas de governança que moldam a produção e disseminação do conhecimento científico.

A análise crítica da DeSci deve começar por uma compreensão clara de suas proposições fundamentais e das tecnologias que as sustentam. A tecnologia *blockchain*, em sua essência, oferece um mecanismo para a criação de registros distribuídos, imutáveis e verificáveis sem a necessidade de autoridades centralizadas. Aplicada ao contexto científico, esta tecnologia promete resolver problemas como a manipulação de dados, a censura de resultados inconvenientes, a concentração de poder em editoras comerciais e a falta de transparência nos processos de financiamento e revisão por pares.

O potencial revolucionário da DeSci manifesta-se em múltiplas dimensões do sistema científico. No âmbito do financiamento, propostas como os Science DAOs (Decentralized Autonomous Organizations) sugerem mecanismos alternativos de alocação de recursos que poderiam democratizar o acesso ao financiamento científico e reduzir a influência de vieses institucionais e geográficos. Estes sistemas permitiriam que comunidades científicas globais participassem diretamente das decisões de financiamento, potencialmente priorizando projetos com base em critérios mais diversos e inclusivos do que aqueles tradicionalmente empregados por agências de fomento centralizadas.

Na esfera da publicação científica, a DeSci propõe alternativas radicais ao sistema tradicional de periódicos acadêmicos. Plataformas descentralizadas de publicação poderiam eliminar as barreiras de acesso impostas por editoras comerciais, reduzir os custos de publicação e criar novos modelos de revisão por pares baseados em incentivos criptoecônômicos. A imutabilidade dos registros *blockchain* poderia garantir a integridade dos dados publicados e criar trilhas de auditoria transparentes para todo o processo de produção científica.

O sistema dARK, desenvolvido pelo IbiCT no Brasil e mencionado no documento original (IBICT, 2025), representa um exemplo concreto de como a tecnologia *blockchain* pode ser aplicada para criar alternativas soberanas a sistemas centralizados. Como alternativa ao sistema DOI, o dARK oferece não apenas independência técnica, mas também autonomia política e econômica em relação às organizações internacionais que controlam os sistemas tradicionais de identificação persistente. A análise deste sistema pode fornecer insights valiosos sobre os desafios práticos e as oportunidades associadas à implementação de infraestruturas científicas descentralizadas.

No entanto, a avaliação crítica da DeSci deve também reconhecer e analisar rigorosamente as limitações e desafios significativos que esta abordagem enfrenta. A questão da escalabilidade representa talvez o obstáculo mais fundamental para a adoção ampla de sistemas baseados em *blockchain* na ciência. As redes *blockchain* existentes, particularmente aquelas que utilizam mecanismos de consenso baseados em prova de trabalho, enfrentam limitações severas em termos de throughput de transações e consumo energético que podem torná-las inadequadas para aplicações científicas de grande escala.

Os custos de transação, conhecidos como “gas fees” no ecossistema Ethereum, mencionados no documento original (ULAM.IO, 2025), representam uma barreira prática significativa para a adoção da DeSci, particularmente por pesquisadores e instituições com recursos limitados. A volatilidade destes custos, que podem variar dramaticamente dependendo da congestão da rede, cria incertezas que podem inviabilizar o uso rotineiro de sistemas descentralizados para atividades científicas cotidianas.

A questão do controle de qualidade em ambientes descentralizados apresenta desafios epistemológicos e práticos complexos. O sistema tradicional de revisão por pares, apesar de suas limitações bem documentadas, oferece mecanismos estabelecidos para a avaliação da qualidade científica. A transição para sistemas descentralizados requer o desenvolvimento de novos mecanismos de controle de qualidade que sejam simultaneamente rigorosos, transparentes e resistentes a manipulações. Isto inclui questões como a prevenção de ataques Sybil (onde atores maliciosos criam múltiplas identidades falsas), a garantia de expertise adequada dos revisores e a manutenção de padrões de qualidade em um ambiente onde não existem autoridades centralizadas.

A necessidade de letramento digital amplo, mencionada no documento original (ULAM.IO, 2025), representa outro desafio significativo para a adoção da DeSci. A complexidade técnica dos

sistemas *blockchain* e das ferramentas associadas pode criar novas formas de exclusão digital que reproduzem ou até mesmo amplificam as desigualdades existentes no sistema científico. A democratização prometida pela descentralização pode, paradoxalmente, beneficiar principalmente aqueles que já possuem recursos e conhecimentos técnicos avançados.

A governança de sistemas descentralizados apresenta desafios particulares que merecem investigação aprofundada. Embora a descentralização prometa reduzir a concentração de poder, a realidade é que muitos sistemas *blockchain* são dominados por pequenos grupos de desenvolvedores, mineradores ou detentores de tokens com grande influência sobre as decisões de governança. A pesquisa deve investigar como garantir que a governança de sistemas DeSci seja genuinamente democrática e inclusiva, evitando a reprodução de hierarquias de poder sob novas formas.

A sustentabilidade ambiental dos sistemas *blockchain* representa uma preocupação crescente que não pode ser ignorada na avaliação da DeSci. O consumo energético massivo de algumas redes *blockchain* levanta questões sobre a compatibilidade desta tecnologia com objetivos de sustentabilidade ambiental que são centrais para muitas áreas da pesquisa científica. A pesquisa deve investigar alternativas mais sustentáveis, como mecanismos de consenso baseados em prova de participação, e avaliar seus trade-offs em termos de segurança e descentralização.

A interoperabilidade entre diferentes sistemas *blockchain* e entre sistemas descentralizados e infraestruturas científicas tradicionais representa outro desafio técnico e político significativo. A fragmentação do ecossistema *blockchain* em múltiplas redes incompatíveis pode reproduzir os problemas de silos de dados que a DeSci pretende resolver. A pesquisa deve investigar estratégias para garantir que a adoção de tecnologias descentralizadas não resulte em nova fragmentação do sistema científico.

*Inteligência artificial (IA) na ciência aberta:
oportunidades transformadoras e riscos epistêmicos*

A IA promete acelerar as descobertas científicas de forma sem precedentes, mas seu uso levanta sérias questões. A agenda de pesquisa deve priorizar os desafios éticos e de transparência, especialmente no uso de modelos generativos na escrita científica (ROCKEMBACH, 2024). É crucial investigar o problema da “caixa-preta” da IA, onde os processos de tomada de decisão do algoritmo são opacos (CASTELVECCHI, 2016), e o risco de “alucinações” — a geração de informações factualmente incorretas ou citações inexistentes — que podem minar a integridade da pesquisa (ALKAISSI & MCFARLANE, 2023). Uma frente de pesquisa fundamental é o desenvolvimento e uso de modelos de IA que sejam, eles mesmos, abertos, auditáveis e alinhados aos princípios da Ciência Aberta, como uma contrapartida à crescente dependência de modelos proprietários de grandes empresas. Finalmente, é vital investigar como a IA, ao ser treinada com dados históricos, pode reproduzir e amplificar vieses existentes, aprofundando uma forma de “colonialismo algorítmico” que marginaliza ainda mais as perspectivas do Sul Global (FIOCRUZ, 2025).

A integração da Inteligência Artificial nos processos de produção científica representa simultaneamente uma das oportunidades mais promissoras e um dos desafios mais complexos para a construção de uma ciência aberta, equitativa e soberana. A capacidade da IA de processar volumes massivos de dados, identificar padrões complexos e acelerar processos de descoberta oferece possibilidades transformadoras para o avanço do conhecimento científico. No entanto, a implementação acrítica dessas tecnologias pode reproduzir e amplificar as mesmas assimetrias e vieses que caracterizam o sistema científico global, criando novas formas de exclusão e dependência.

A promessa de aceleração das descobertas científicas através da IA manifesta-se em múltiplas dimensões da atividade científica. Na análise de dados, algoritmos de aprendizado de máquina podem identificar padrões e correlações que permaneceriam invisíveis para análises tradicionais, potencialmente revelando insights científicos inovadores. Na modelagem e simulação, a IA pode acelerar dramaticamente processos computacionais complexos, permitindo investigações que seriam impraticáveis com métodos convencionais. Na descoberta de literatura, sistemas de IA podem processar volumes massivos de publicações científicas, identificando conexões e lacunas que poderiam escapar a revisões manuais.

No entanto, a aplicação da IA na ciência também introduz riscos epistêmicos fundamentais que demandam investigação cuidadosa. O problema da “caixa-preta”, mencionado por Castelvechchi (2016), refere-se à opacidade dos processos de tomada de decisão em muitos sistemas de IA, particularmente em redes neurais profundas. Esta opacidade é particularmente problemática no contexto científico, onde a transparência e a reprodutibilidade são valores fundamentais. Quando algoritmos de IA são utilizados para análise de dados ou geração de hipóteses, a incapacidade de compreender como essas conclusões foram alcançadas pode comprometer a validação e replicação dos resultados.

A questão das “alucinações” em modelos generativos, destacada por Alkaissi e McFarlane (2023), representa um risco particular para a integridade da pesquisa científica. A capacidade desses modelos de gerar informações factualmente incorretas ou citações inexistentes, apresentadas de forma convincente e aparentemente autoritativa, pode introduzir erros sistemáticos na literatura científica. Este problema é agravado pela tendência de alguns pesquisadores de utilizar modelos generativos como ferramentas de escrita sem verificação adequada das informações produzidas.

A crescente dependência de modelos proprietários desenvolvidos por grandes corporações tecnológicas representa uma forma particularmente insidiosa de dependência que pode comprometer a soberania científica. Quando pesquisadores e instituições dependem de sistemas de IA controlados por empresas privadas, eles perdem controle sobre aspectos fundamentais de seus processos de pesquisa, incluindo a transparência dos algoritmos utilizados, a privacidade dos dados processados e a continuidade do acesso às ferramentas. Esta dependência pode ser especialmente problemática para pesquisadores e instituições do Sul Global, que podem ter menor poder de negociação com fornecedores de tecnologia.

O desenvolvimento de modelos de IA abertos, auditáveis e alinhados aos princípios da Ciência Aberta emerge como uma prioridade estratégica fundamental. Estes modelos devem ser desenvolvidos de forma transparente, com código-fonte aberto, dados de treinamento documentados e processos de desenvolvimento auditáveis. A pesquisa nesta área deve investigar não apenas os aspectos técnicos do desenvolvimento de modelos abertos, mas também os modelos de governança, financiamento e sustentabilidade necessários para garantir que essas alternativas possam competir efetivamente com soluções proprietárias.

A questão dos vieses algorítmicos representa talvez o desafio mais complexo e politicamente sensível na aplicação da IA à ciência. Quando modelos de IA são treinados com dados históricos que refletem desigualdades e preconceitos existentes, eles tendem a reproduzir e amplificar esses vieses em suas previsões e recomendações. No contexto científico, isto pode resultar na perpetuação de perspectivas dominantes, na marginalização de abordagens alternativas e na reprodução de desigualdades estruturais.

O conceito de “colonialismo algorítmico”, mencionado no documento original (FIOCRUZ, 2025), oferece um framework analítico particularmente útil para compreender como os vieses da IA

podem reproduzir dinâmicas de dominação epistêmica. Quando modelos de IA são desenvolvidos predominantemente por organizações sediadas no Norte Global, utilizando dados que refletem principalmente as perspectivas e prioridades dessas regiões, eles podem funcionar como mecanismos sofisticados de imposição de paradigmas dominantes sobre comunidades científicas periféricas.

A investigação dos vieses algorítmicos na ciência deve ir além da simples identificação de problemas para desenvolver metodologias concretas de mitigação. Isto inclui técnicas de debiasing durante o treinamento de modelos, desenvolvimento de datasets mais diversos e representativos, e criação de métricas para avaliar a equidade e inclusividade de sistemas de IA aplicados à ciência. A pesquisa deve também investigar como diferentes comunidades científicas podem participar ativamente do desenvolvimento e avaliação de sistemas de IA, garantindo que suas perspectivas e necessidades sejam adequadamente representadas.

A questão da transparência algorítmica na ciência requer atenção especial ao desenvolvimento de metodologias de explicabilidade que sejam adequadas ao contexto científico. Diferentemente de outras aplicações da IA, onde a precisão pode ser mais importante que a interpretabilidade, a ciência requer que os pesquisadores possam compreender e validar os processos através dos quais conclusões são alcançadas. Isto exige o desenvolvimento de técnicas de IA explicável que sejam simultaneamente rigorosas tecnicamente e acessíveis para pesquisadores sem formação técnica avançada em IA.

A privacidade e segurança dos dados em sistemas de IA científica representam outra dimensão crítica que requer investigação aprofundada. Quando dados de pesquisa sensíveis são processados por sistemas de IA, particularmente aqueles baseados em nuvem, surgem questões complexas sobre proteção de privacidade, segurança de dados e soberania informacional. A pesquisa deve investigar técnicas como aprendizado federado, computação

homomórfica e privacidade diferencial como alternativas que permitam o aproveitamento dos benefícios da IA sem comprometer a privacidade e segurança dos dados.

A avaliação e validação de sistemas de IA aplicados à ciência requerem o desenvolvimento de novos frameworks metodológicos que considerem as especificidades do contexto científico. Os métodos tradicionais de avaliação de sistemas de IA, focados em métricas como precisão e recall, podem não ser adequados para avaliar aspectos como reprodutibilidade, interpretabilidade e impacto científico. A pesquisa deve desenvolver metodologias de avaliação que sejam capazes de capturar essas dimensões mais complexas da qualidade científica.

A formação e capacitação de pesquisadores para o uso crítico e responsável da IA representa um desafio educacional fundamental. A integração efetiva da IA na ciência requer que pesquisadores desenvolvam não apenas competências técnicas para utilizar essas ferramentas, mas também capacidades críticas para avaliar suas limitações, identificar vieses e interpretar resultados de forma apropriada. Isto exige o desenvolvimento de programas de formação que combinem aspectos técnicos, éticos e epistemológicos do uso da IA na ciência.

Pontos-chave da agenda de pesquisa

Equidade: Investigar e corrigir as assimetrias na produção científica, financiamento e infraestrutura, promovendo colaborações justas.

Soberania: Fomentar o desenvolvimento de infraestruturas digitais nacionais e políticas públicas que garantam o controle sobre os dados de pesquisa.

Tecnologias Emergentes: Analisar criticamente o potencial da DeSci e da IA como ferramentas para a autonomia, investigando seus riscos éticos e de dependência.

Estes pontos-chave representam mais do que diretrizes programáticas; constituem princípios estruturantes que devem orientar não apenas a condução de pesquisas específicas, mas a transformação fundamental do sistema científico global. A equidade não pode ser tratada como um objetivo secundário ou cosmético, mas deve ser reconhecida como pré-requisito para a legitimidade e eficácia de qualquer sistema científico que pretenda servir aos interesses da humanidade como um todo. A soberania não deve ser confundida com isolacionismo ou nacionalismo científico, mas compreendida como a capacidade de comunidades científicas de exercer controle democrático sobre suas próprias prioridades, recursos e processos. As tecnologias emergentes não devem ser adotadas acriticamente com base em promessas de eficiência ou inovação, mas avaliadas rigorosamente em termos de seu potencial para promover ou comprometer os objetivos de equidade e soberania.

A interconexão entre estes três pontos-chave é fundamental para compreender a natureza sistêmica dos desafios enfrentados. A promoção da equidade sem atenção à soberania pode resultar em formas paternalistas de inclusão que mantêm as estruturas fundamentais de dependência. A busca por soberania sem consideração pela equidade pode reproduzir exclusões e hierarquias em escala nacional ou regional. A adoção de tecnologias emergentes sem atenção aos imperativos de equidade e soberania pode amplificar as desigualdades existentes e criar novas formas de dependência tecnológica.

Desafios transversais e a ética dos dados

Além dos eixos estruturantes, uma agenda de pesquisa completa deve abordar desafios que perpassam todas as áreas, focando nos dilemas éticos e na profunda mudança cultural necessária para que uma Ciência Aberta, equitativa e soberana, floresça.

*A ética na coleta e uso de dados:
construindo frameworks para uma ciência responsável*

A abertura não pode ser um pretexto para negligenciar a ética. É preciso consolidar a “Ética de Dados” (Data Ethics) como uma nova área do conhecimento, que transcenda a mera conformidade com leis de proteção de dados (como a LGPD no Brasil). A pesquisa nesta área deve se aprofundar em questões complexas como a discriminação algorítmica, a justiça no acesso aos benefícios da pesquisa e a prevenção de vieses (étnicos, raciais, de gênero) que podem ser embutidos nos conjuntos de dados (PRP UNICAMP, s.d.). A transparência sobre como os dados são coletados e analisados, o consentimento informado e a proteção rigorosa de dados sensíveis devem ser pilares inegociáveis de qualquer prática de ciência aberta (EDITVERSE, s.d.).

A consolidação da Ética de Dados como disciplina autônoma representa uma das necessidades mais urgentes para a construção de uma ciência verdadeiramente responsável e sustentável. Esta nova área do conhecimento deve ir muito além da simples aplicação de princípios éticos tradicionais ao contexto digital, para desenvolver frameworks conceituais, metodológicos e práticos que sejam capazes de lidar com as complexidades específicas da era dos *Big Data* e da inteligência artificial.

A emergência da Ética de Dados como campo de estudo reflete o reconhecimento crescente de que os dados não são entidades neutras ou objetivas, mas construções sociais que incorporam valores, perspectivas e relações de poder. A coleta, processamento, análise e uso de dados envolvem escolhas que têm implicações éticas profundas, desde a definição de quais dados coletar até a interpretação e aplicação dos resultados obtidos. Estas escolhas não são meramente técnicas, mas refletem concepções sobre o que constitui conhecimento válido, quais questões merecem investigação e quais grupos e perspectivas devem ser priorizados.

A questão da discriminação algorítmica representa uma das fronteiras mais críticas da Ética de Dados aplicada à ciência. Quando algoritmos são utilizados para processar dados de pesquisa, tomar decisões sobre financiamento, selecionar participantes de estudos ou avaliar propostas de pesquisa, existe o risco de que vieses implícitos nos dados ou nos algoritmos resultem em discriminação sistemática contra grupos específicos. Esta discriminação pode ser particularmente insidiosa porque é frequentemente invisível e pode ser racionalizada como resultado de processos “objetivos” e “baseados em dados”.

A investigação da discriminação algorítmica na ciência requer o desenvolvimento de metodologias sofisticadas para identificar, mensurar e mitigar vieses em diferentes estágios do processo de pesquisa. Isto inclui técnicas para auditar datasets em busca de vieses de representação, métodos para avaliar a equidade de algoritmos de aprendizado de máquina e frameworks para monitorar os impactos de sistemas automatizados sobre diferentes grupos. A pesquisa deve também investigar como diferentes definições de equidade (igualdade de oportunidades, paridade demográfica, equidade individual) se aplicam ao contexto científico e como trade-offs entre diferentes concepções de equidade devem ser negociados.

A justiça no acesso aos benefícios da pesquisa constitui outra dimensão fundamental da Ética de Dados que requer atenção especial. A ciência aberta promete democratizar o acesso ao conhecimento, mas a realidade é que os benefícios da pesquisa científica continuam a ser distribuídos de forma desigual. Comunidades que fornecem dados para pesquisas frequentemente não se beneficiam dos resultados obtidos, enquanto os lucros e reconhecimentos concentram-se em instituições e pesquisadores com maior poder econômico e simbólico.

O desenvolvimento de princípios de justiça distributiva aplicados aos dados de pesquisa requer uma reflexão cuidadosa sobre como os benefícios e riscos da pesquisa devem ser distribuídos

entre diferentes stakeholders. Isto inclui questões como: que tipo de retorno as comunidades que fornecem dados devem receber? Como garantir que pesquisas conduzidas em contextos de vulnerabilidade social não reproduzam ou ampliem essas vulnerabilidades? Quais são as obrigações dos pesquisadores em relação às comunidades que estudam?

A prevenção de vieses étnicos, raciais e de gênero em conjuntos de dados, mencionada no documento original (PRP UNICAMP, s.d.), representa um desafio técnico e político complexo que requer abordagens multidisciplinares. Os vieses em dados podem originar-se de múltiplas fontes: sub-representação de grupos específicos, uso de categorias inadequadas ou problemáticas, aplicação de métodos de coleta que privilegiam certas perspectivas, ou interpretação de resultados através de lentes culturalmente específicas.

A identificação e correção de vieses em dados requer o desenvolvimento de metodologias que sejam simultaneamente rigorosas tecnicamente e sensíveis às dimensões sociais e políticas dos dados. Isto inclui técnicas estatísticas para detectar vieses de seleção e representação, métodos qualitativos para compreender como diferentes grupos experienciam e interpretam os fenômenos estudados, e abordagens participativas que envolvem as comunidades afetadas na definição de categorias e interpretação de resultados.

A transparência na coleta e análise de dados, mencionada no documento original (EDITVERSE, s.d.), representa um princípio fundamental que deve ser operacionalizado através de práticas concretas e verificáveis. A transparência não pode ser reduzida à simples disponibilização de dados brutos, mas deve incluir documentação detalhada sobre métodos de coleta, critérios de seleção, processos de limpeza e transformação, e pressupostos analíticos. Esta documentação deve ser acessível não apenas para especialistas técnicos, mas também para as comunidades e stakeholders afetados pela pesquisa.

O consentimento informado em contextos de *Big Data* e análise automatizada apresenta desafios únicos que requerem repensar os modelos tradicionais de consentimento desenvolvidos para pesquisas convencionais. Quando dados são coletados em larga escala, processados através de algoritmos complexos e utilizados para múltiplos propósitos, torna-se difícil ou impossível informar adequadamente os participantes sobre todos os usos potenciais de seus dados. Isto requer o desenvolvimento de novos modelos de consentimento que sejam simultaneamente práticos e respeitosos da autonomia dos participantes.

A proteção de dados sensíveis em contextos de ciência aberta apresenta tensões fundamentais entre os imperativos da transparência e da privacidade. Embora a abertura de dados seja frequentemente apresentada como um bem inequívoco, a realidade é que a disponibilização de certos tipos de dados pode expor indivíduos e comunidades a riscos significativos. A pesquisa deve investigar técnicas de anonimização e pseudonimização que sejam robustas contra ataques de re-identificação, bem como modelos de acesso controlado que permitam o compartilhamento de dados sensíveis para fins de pesquisa legítima.

A mudança cultural na academia: transformando incentivos e práticas institucionais

Um dos maiores obstáculos é a “resistência cultural de pesquisadores” (NICODEMO, 2025). A cultura acadêmica tradicional, focada no artigo científico como produto final e principal métrica de sucesso, desincentiva a partilha de dados brutos, metodologias e processos. A agenda de pesquisa deve, portanto, propor e avaliar novos sistemas de incentivo e avaliação de pesquisadores. É preciso estudar como as agências de fomento e as universidades podem reformar seus critérios para valorizar práticas de ciência aberta, como a curadoria de dados, o desenvolvimento de *software* de código aberto e o trabalho coletivo, reconhecendo-os como produção científica legítima e de alto valor (EUA, 2022).

A transformação cultural necessária para a implementação efetiva de uma ciência aberta, equitativa e soberana representa talvez o desafio mais complexo e de longo prazo de toda a agenda de pesquisa proposta. A resistência cultural mencionada por Nicodemo (2025) não deve ser compreendida simplesmente como obstinação ou conservadorismo individual, mas como resultado de estruturas institucionais profundamente enraizadas que moldam comportamentos, incentivos e práticas através de mecanismos sutis, mas poderosos.

A cultura acadêmica tradicional, construída ao longo de séculos, estabeleceu o artigo científico publicado em periódicos de prestígio como a unidade fundamental de produção e avaliação acadêmica. Este sistema, embora tenha contribuído significativamente para o avanço do conhecimento científico, criou também uma série de distorções e incentivos perversos que são incompatíveis com os princípios da ciência aberta. A ênfase na publicação como produto final desencoraja a partilha de dados brutos, metodologias detalhadas e processos de pesquisa que não resultem diretamente em publicações, criando uma cultura de sigilo e competição que é antitética aos valores da colaboração e transparência.

A análise da resistência cultural deve reconhecer que os pesquisadores individuais operam dentro de sistemas de incentivos que frequentemente penalizam comportamentos alinhados com a ciência aberta. Em um ambiente onde promoções, financiamentos e reconhecimento profissional dependem primariamente de métricas como número de publicações e fator de impacto de periódicos, é racional que pesquisadores priorizem atividades que maximizem essas métricas, mesmo que isso signifique negligenciar práticas de abertura e colaboração que poderiam beneficiar a ciência como um todo.

A transformação destes sistemas de incentivos requer intervenções coordenadas em múltiplos níveis institucionais. As agências de fomento desempenham um papel crucial nesta transformação, pois controlam recursos significativos e podem influenciar comportamentos através de seus critérios de avaliação e financiamento.

A reforma dos critérios de avaliação de propostas de pesquisa para incluir considerações sobre abertura de dados, colaboração e impacto social pode criar incentivos poderosos para mudanças de comportamento.

As universidades e institutos de pesquisa têm responsabilidades igualmente importantes na promoção da mudança cultural. Os processos de contratação, promoção e avaliação de desempenho de pesquisadores devem ser reformulados para reconhecer e valorizar contribuições que vão além das publicações tradicionais. Isto inclui atividades como curadoria de dados, desenvolvimento de software científico, mentoria de jovens pesquisadores, engajamento público com a ciência e colaboração interdisciplinar.

A valorização da curadoria de dados como atividade científica legítima representa uma mudança paradigmática fundamental. A curadoria de dados de alta qualidade requer expertise técnica significativa, conhecimento profundo do domínio científico e investimento substancial de tempo e recursos. No entanto, estas atividades são frequentemente invisíveis nos sistemas tradicionais de avaliação acadêmica, criando desincentivos para que pesquisadores invistam na criação e manutenção de recursos de dados de alta qualidade.

O desenvolvimento de *software* científico de código aberto representa outra categoria de contribuição que merece reconhecimento e valorização. O *software* tornou-se uma infraestrutura crítica para a pesquisa científica contemporânea, mas o desenvolvimento e manutenção de *software* de qualidade requer habilidades especializadas e investimento contínuo que raramente são adequadamente reconhecidos nos sistemas de avaliação acadêmica. A criação de carreiras e trajetórias profissionais que valorizem a contribuição para o desenvolvimento de infraestruturas digitais científicas é essencial para a sustentabilidade da ciência aberta.

O trabalho coletivo e colaborativo, mencionado no documento original (EUA, 2022), representa talvez a mudança mais fundamental necessária na cultura acadêmica. A ciência contemporânea

é crescentemente colaborativa e interdisciplinar, mas os sistemas de avaliação e reconhecimento permanecem focados em contribuições individuais e disciplinares. Isto cria tensões entre as necessidades da pesquisa científica moderna e os incentivos institucionais, frequentemente resultando em sub-valorização de contribuições colaborativas e interdisciplinares.

A implementação de novos sistemas de avaliação que reconheçam adequadamente o trabalho colaborativo requer o desenvolvimento de metodologias sofisticadas para atribuir crédito e reconhecimento em contextos de autoria múltipla e contribuições diversificadas. Isto pode incluir sistemas de taxonomia de contribuições que permitam identificar e valorizar diferentes tipos de participação em projetos de pesquisa, desde a conceituação inicial até a disseminação de resultados.

A resistência à mudança cultural não se limita aos pesquisadores individuais, mas inclui também resistência institucional de organizações que se beneficiam do sistema atual. Editoras comerciais, empresas de métricas acadêmicas e outras organizações que lucram com o sistema tradicional de publicação científica têm incentivos para resistir a mudanças que possam ameaçar seus modelos de negócio. A superação desta resistência requer estratégias coordenadas que incluam pressão política, desenvolvimento de alternativas viáveis e criação de coalizões de instituições comprometidas com a transformação.

A dimensão internacional da mudança cultural apresenta desafios adicionais, particularmente para países e instituições que dependem de reconhecimento internacional para financiamento e prestígio. A adoção unilateral de novos sistemas de avaliação pode resultar em desvantagens competitivas se esses sistemas não forem reconhecidos internacionalmente. Isto requer esforços coordenados para promover mudanças culturais em escala global, incluindo o desenvolvimento de padrões internacionais e a criação de redes de instituições comprometidas com a transformação.

Letramento digital e capacitação: educação crítica para a autonomia tecnológica

Finalmente, nenhuma infraestrutura ou política será eficaz sem pessoas capacitadas para utilizá-la de forma crítica e consciente. A pesquisa deve orientar o desenvolvimento de programas de capacitação que vão além do treinamento instrumental. É necessária uma “educação digital freireana” (OUTRAS PALAVRAS, 2022), que habilite os pesquisadores a usar as ferramentas digitais e, ao mesmo tempo, a compreender suas implicações políticas, econômicas e sociais. Esse letramento crítico é a base para que a comunidade científica possa fazer escolhas informadas sobre quais tecnologias adotar e como moldá-las para servir aos interesses da ciência e da sociedade, em vez de se tornar refém delas.

A questão do letramento digital crítico representa o elemento fundamental que determina o sucesso ou fracasso de qualquer agenda de transformação da ciência na era digital. Sem pesquisadores e comunidades científicas adequadamente capacitados para compreender, avaliar e utilizar criticamente as tecnologias digitais, mesmo as infraestruturas mais sofisticadas e as políticas mais bem intencionadas podem resultar em novas formas de dependência e exclusão.

O conceito de “educação digital freireana”, mencionado no documento original (OUTRAS PALAVRAS, 2022), oferece um *framework* particularmente poderoso para repensar os objetivos e métodos da capacitação digital na ciência. Inspirada na pedagogia crítica de Paulo Freire, essa abordagem reconhece que a educação não é um processo neutro de transmissão de conhecimentos técnicos, mas uma prática política que pode tanto reproduzir quanto transformar relações de poder existentes.

A educação digital freireana aplicada ao contexto científico deve começar pelo reconhecimento de que as tecnologias digitais não são ferramentas neutras, pois incorporam valores, interesses

e relações de poder específicos. A capacitação de pesquisadores deve, portanto, envolver o desenvolvimento de competências técnicas para utilizar essas ferramentas, assim como a capacidade crítica para compreender como essas tecnologias são desenvolvidas, por quem são controladas e a quais interesses servem.

O desenvolvimento de programas de capacitação que vão além do treinamento instrumental requer uma compreensão sofisticada das diferentes dimensões do letramento digital. O letramento técnico, embora importante, é apenas uma das dimensões necessárias. Igualmente importantes são o letramento crítico (capacidade de avaliar criticamente as tecnologias e suas implicações), o letramento ético (compreensão das dimensões éticas do uso de tecnologias digitais) e o letramento político (reconhecimento das dimensões de poder envolvidas no desenvolvimento e uso de tecnologias).

A dimensão técnica do letramento digital na ciência inclui competências como programação, análise de dados, uso de ferramentas de visualização, compreensão de algoritmos de aprendizado de máquina e familiaridade com infraestruturas de computação em nuvem. No entanto, estas competências técnicas devem ser desenvolvidas dentro de um framework crítico que permita aos pesquisadores compreender as limitações, vieses e implicações políticas dessas ferramentas.

A dimensão crítica do letramento digital requer o desenvolvimento de capacidades para questionar e avaliar as tecnologias digitais. Isto inclui a capacidade de identificar vieses em algoritmos, compreender as implicações de diferentes modelos de dados, avaliar a confiabilidade de diferentes fontes de informação digital e reconhecer como diferentes tecnologias podem reproduzir ou desafiar relações de poder existentes.

A dimensão ética do letramento digital na ciência é particularmente importante dado o potencial das tecnologias digitais para impactar indivíduos e comunidades de formas significativas. Os pesquisadores

devem ser capacitados para reconhecer e navegar dilemas éticos complexos relacionados à privacidade de dados, consentimento informado, justiça algorítmica e responsabilidade social da pesquisa.

A dimensão política do letramento digital requer que os pesquisadores compreendam como as tecnologias digitais se inserem em estruturas mais amplas de poder político e econômico. Isto inclui compreensão sobre modelos de negócio de empresas de tecnologia, implicações geopolíticas da dependência tecnológica, e como diferentes escolhas tecnológicas podem promover ou comprometer objetivos de equidade e soberania.

O desenvolvimento de programas de capacitação eficazes requer também atenção às diferentes necessidades e contextos de diferentes grupos de pesquisadores. Pesquisadores em início de carreira podem ter necessidades diferentes de pesquisadores seniores, assim como pesquisadores de diferentes disciplinas podem requerer abordagens específicas. A capacitação deve ser adaptada para considerar essas diferenças, evitando abordagens universais que podem não ser apropriadas para todos os contextos.

A sustentabilidade dos programas de capacitação representa outro desafio crítico. O desenvolvimento de competências digitais é um processo contínuo que requer atualização constante dado o ritmo acelerado de mudança tecnológica. Isto requer o desenvolvimento de modelos de capacitação que sejam sustentáveis a longo prazo e que possam se adaptar a mudanças tecnológicas futuras.

A dimensão coletiva da capacitação é igualmente importante. Além de desenvolver competências individuais, é necessário construir capacidades coletivas dentro de comunidades científicas, incluindo redes de apoio mútuo, comunidades de prática e infraestruturas institucionais que suportem o uso crítico e responsável de tecnologias digitais.

A avaliação da eficácia dos programas de capacitação requer o desenvolvimento de métricas que vão além de indicadores tradicionais de aprendizagem. É necessário avaliar não apenas se os participantes adquiriram competências técnicas específicas, mas também se desenvolveram capacidades críticas para avaliar e utilizar tecnologias de forma responsável e se estão contribuindo para a construção de comunidades científicas mais equitativas e soberanas.

A dimensão internacional da capacitação digital apresenta oportunidades e desafios únicos. Por um lado, as tecnologias digitais facilitam o compartilhamento de recursos educacionais e a criação de redes globais de aprendizagem. Por outro lado, existe o risco de que programas de capacitação desenvolvidos em contextos específicos reproduzam vieses culturais e epistemológicos quando aplicados em outros contextos. A pesquisa deve investigar como desenvolver abordagens de capacitação que sejam simultaneamente globalmente relevantes e localmente apropriadas.

CONSIDERAÇÕES FINAIS: RUMO A UMA TRANSFORMAÇÃO SISTÊMICA DA CIÊNCIA GLOBAL

A agenda de pesquisa apresentada neste capítulo não constitui meramente um conjunto de diretrizes técnicas ou recomendações políticas, mas uma proposta de transformação sistêmica fundamental do sistema científico global. A construção de uma ciência aberta, equitativa e soberana requer mais do que reformas incrementais ou ajustes pontuais; demanda uma reimaginação radical das estruturas, processos e valores que moldam a produção e circulação do conhecimento científico na contemporaneidade.

Os três eixos estruturantes desta agenda — fundamentos para uma ciência equitativa, construção da soberania científica e digital, e análise crítica das tecnologias emergentes — não devem ser compreendidos como domínios separados ou independentes, mas como dimensões interconectadas de um projeto político e epistemológico mais amplo. A equidade sem soberania pode resultar em formas paternalistas de inclusão que mantêm as estruturas fundamentais de dependência. A soberania sem equidade pode reproduzir exclusões e hierarquias em escala nacional ou regional. A adoção de tecnologias emergentes sem atenção aos imperativos de equidade e soberania pode amplificar as desigualdades existentes e criar novas formas de dependência tecnológica.

Os desafios transversais identificados, ética dos dados, mudança cultural na academia e letramento digital crítico, perpassam todos os eixos da agenda e constituem pré-requisitos fundamentais para qualquer transformação substantiva. Sem frameworks éticos robustos, a abertura científica pode resultar em novas formas

de exploração e violação de direitos. Sem mudanças culturais profundas nas instituições acadêmicas, mesmo as políticas mais bem intencionadas podem ser neutralizadas por resistências sistêmicas. Sem capacitação crítica adequada, as comunidades científicas podem se tornar reféns das tecnologias que deveriam servir aos seus interesses.

A implementação desta agenda requer reconhecer que a transformação da ciência não é apenas uma questão técnica ou administrativa, mas um processo político que envolve redistribuição de poder, recursos e reconhecimento. As resistências a esta transformação não são meramente resultado de inércia institucional ou conservadorismo individual, mas refletem interesses estabelecidos que se beneficiam do sistema atual. A superação destas resistências requer estratégias coordenadas que combinem pressão política, desenvolvimento de alternativas viáveis e construção de coalizões amplas comprometidas com a transformação.

A dimensão internacional desta agenda apresenta desafios e oportunidades únicos. Por um lado, a ciência é inerentemente global, e qualquer transformação substantiva requer coordenação e cooperação internacional. Por outro lado, as assimetrias de poder no sistema científico global significam que iniciativas de transformação podem ser capturadas ou neutralizadas pelos interesses dominantes. A construção de alternativas soberanas requer, portanto, estratégias que sejam simultaneamente localmente enraizadas e globalmente articuladas.

O sucesso desta agenda depende fundamentalmente da capacidade de diferentes atores, como pesquisadores, instituições, governos e organizações da sociedade civil, de construir visões compartilhadas sobre o futuro desejável da ciência e de desenvolver estratégias coordenadas para alcançar essa visão. Isso requer mais do que competências técnicas e recursos materiais. Exige também imaginação política e coragem para desafiar estruturas estabelecidas.

A urgência dessa transformação é amplificada pelos desafios globais contemporâneos, como mudanças climáticas, pandemias, desigualdades crescentes e crises democráticas, que demandam respostas científicas que sejam tecnicamente competentes, socialmente responsáveis e politicamente legítimas. A construção de uma ciência aberta, equitativa e soberana não é só um imperativo ético, mas uma necessidade prática para enfrentar os desafios do século XXI.

A agenda aqui proposta deve ser compreendida como um convite ao diálogo e à ação coletiva. Sua implementação requer adaptação às especificidades de diferentes contextos, disciplinas e comunidades científicas. Mais importante ainda, requer o reconhecimento de que a transformação da ciência é um processo contínuo que demanda vigilância constante, avaliação crítica e disposição para aprender e adaptar-se diante de novos desafios e oportunidades.

O futuro da ciência não está predeterminado pelas trajetórias tecnológicas ou pelas dinâmicas de mercado, mas será moldado pelas escolhas políticas e éticas que fazemos hoje. A construção de uma ciência verdadeiramente aberta, equitativa e soberana é possível, mas requer que assumamos a responsabilidade coletiva de trabalhar para torná-la realidade.

REFERÊNCIAS BIBLIOGRÁFICAS

ABALKINA, Anna; BISHOP, Dorothy V. M. The Threat of Paper Mills to Biomedical and Social Science Journals. **Preprint**, 2025.

ABCD USP. UNESCO publica recomendações sobre Ciência Aberta. *In*: ABCD USP.: ABCD USP, [s.d.]. Disponível em: <https://www.abcd.usp.br/noticias/unesco-publica-recomendacoes-sobre-ciencia-aberta/>. Acesso em: 4 ago. 2025.

ABRAMSON, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. **Nature**, v. 630, p. 493–500, 8 maio 2024.

AFRICA CDC. Africa Sets Course to Strengthen and Harmonise Health Data Governance. **Africa CDC**, Adis Abeba, 23 jul. 2025. Disponível em: <https://africacdc.org/news-item/africa-sets-course-to-strengthen-and-harmonise-health-data-governance/>. Acesso em: 31 jul. 2025.

AFRICAN UNION. **Data Policy Framework**. Adis Abeba: African Union Commission, 2022.

AJL – ALGORITHMIC JUSTICE LEAGUE. **Fighting Algorithmic Bias**. 2022. Disponível em: <https://www.ajl.org>. Acesso em: 30 set. 2025.

AQUARE.LA. Data Spaces: O Futuro da Organização de Dados na Era Digital. **Aquare.la**, 7 maio 2025. Disponível em: <https://aquare.la/data-spaces-o-futuro-da-organizacao-de-dados-na-era-digital/>. Acesso em: 05 ago. 2025.

ALKAISSI, H.; MCFARLANE, S. I. Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. **Cureus**, v. 15, n. 2, e35179, 2023. Disponível em: <https://www.cureus.com/articles/35179-artificial-hallucinations-in-chatgpt-implications-in-scientific-writing>. Acesso em: 05 ago. 2025.

ALBORNOZ, Denisse *et al.* Approaching openness: A critical analysis of open science policies. *In*: ARZA, Valeria; FRESSOLI, Mariano (org.). **Open Science in Latin America**. Buenos Aires: CENIT, 2020.

ALL EUROPEAN ACADEMIES. **The European Code of Conduct for Research Integrity**. Berlim: ALLEA, 2017. Disponível em: <https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf>. Acesso em: 29 jun. 2025.

ALL EUROPEAN ACADEMIES. **The European Code of Conduct for Research Integrity**. Edição revisada. Berlim: ALLEA, 2023. Disponível em: https://www.allea.org/wp-content/uploads/2023/06/ECOC_Research_Integrity_2023.pdf. Acesso em: 29 jun. 2025.

AN, Yonghong; WILLIAMS, Michael A.; XIAO, Mo. High Prices and Market Power of Academic Publishing Reduce Article Citations. **ProMarket**, 24 abr. 2024. Disponível em: <https://www.promarket.org/2024/04/24/high-prices-and-market-power-of-academic-publishing-reduce-article-citations/>. Acesso em: 5 ago. 2025.

ANDERSON, Benedict. **Comunidades imaginadas**: reflexões sobre a origem e a difusão do nacionalismo. São Paulo: Companhia das Letras, 2008.

ANDERSON, Jane; CHRISTEN, Kimberly. *Decolonizing Methodologies: Data Sovereignty and TK Labels*. 2017

ANGWIN, Julia *et al.* Machine Bias. **ProPublica**, 23 maio 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 30 set. 2025.

ANSARI, Ahmed; HUI, Yuk. Decolonizing Philosophy of Technology: Learning from Bottom-Up Practices. **Philosophy ; Technology**, v. 35, n. 1, p. 1-25, 2022. DOI: 10.1007/s13347-021-00489-w.

ANTUNES, Aline *et al.* Percepções dos investigadores face ao ecossistema da Ciência Aberta. **Ibersid: revista de sistemas de información y documentación**, v. 15, n. 2, 2021. DOI: 10.54886/ibersid.v15i2.4759.

APA – AMERICAN PSYCHOLOGICAL ASSOCIATION. **Social Media and Mental Health: Highlights from the APA**. 2017. Disponível em: <https://www.apa.org/news/press/releases/stress/2017/technology-social-media>. Acesso em: 30 set. 2025.

ARUN, Chinmayi. Data colonialism and the political economy of big tech. **Lawfare**, 2025. Disponível em: <https://www.lawfaremedia.org/article/data-colonialism-and-the-political-economy-of-big-tech>. Acesso em: 10 jun. 2025.

ARUN, Chinmayi. Revisão crítica sobre capitalismo de vigilância e desigualdades geopolíticas. **Lawfare**, mar. 2025. Disponível em: [fonte]. Acesso em: 30 jul. 2025.

ASKYAZI. **WhatsApp Usage Across Africa: Key Statistics & Insights for 2025**. 2025. Disponível em: <https://www.askyazi.com/useful-data-sources-for-africa/whatsapp-usage-across-africa-key-statistics-insights-for-2025>. Acesso em: 5 ago. 2025.

BABBAGE, Charles. **Reflections on the Decline of Science in England, and on Some of Its Causes**. Londres: B. Fellowes, 1830. Disponível em: <https://archive.org/details/reflectionsonth00babbgoog>. Acesso em: 29 jun. 2025.

BABINI, Dominique; MACHIN-MASTROMATTEO, Juan D. Latin American science is meant to be open: Arguments and experiences. *In*: CHAN, Leslie; LOIZIDES, Fernando (ed.).

Expanding Perspectives on Open Science: Communities, Cultures and Diversity in Concepts and Practices. Amsterdã: IOS Press, 2017. p. 31-56.

BAKER, Monya. 1,500 scientists lift the lid on reproducibility. **Nature**, v. 533, p. 452-454, 2016. Disponível em: <https://www.nature.com/articles/533452a>. Acesso em: 30 jul. 2025.

BACKLINKO. **WhatsApp Monthly Active Users Worldwide**. 2024. Disponível em: <https://backlinko.com/whatsapp-users>. Acesso em: 5 ago. 2025.

BALDÉ, Cornelis P. *et al.* **The Global E-waste Monitor 2020: Quantities, Flows and the Circular Economy Potential**. Genebra: ITU, United Nations University, ISWA, 2020. Disponível em: <https://www.itu.int/en/ITU-D/Environment/Pages/E-waste>. Acesso em: 30 set. 2025.

BAWDEN, David; ROBINSON, Lyn. "Essentially made of information": concepts and implications of informational privacy. **Information Research**, v. 24, n. 4, 2019. Disponível em: <https://informationr.net/ir/24-4/colis/colis1913.html>. Acesso em: 4 ago. 2025.

BAYRAU, Bethel. From Parachute Science to Community Engagement and Advocacy. **Global Health Now**, 6 jun. 2023.

BEALL, Jeffrey. Predatory publishers are corrupting open access. **Nature**, v. 489, p. 179, 2012.

BENDER, Ashley *et al.* Feminist perspectives on FAIR data: Critical reflections on the data sharing movement in science. **Journal of Information Science**, v. 47, n. 5, p. 595-606, 2021.

BIAGIOLI, Mario; LIPPMAN, Alexandra (org.). **Scientific Authorship: Credit and Intellectual Property in Science**. Londres: Routledge, 2020.

BIAŁAS, Szymon. How Decentralized Science (DeSci) Improves Research. *In*: ULAM LABS BLOG.: Ulam Labs, 15 jan. 2025. Disponível em: <https://www.ulam.io/blog/how-decentralized-science-is-revolutionizing-research>. Acesso em: 4 ago. 2025.

BIK, Emma M.; BOYER, Jennifer S.; KURLAND, Brian J. The prevalence of inappropriate image duplication in biomedical research publications. **mBio**, v. 7, n. 3, p. e00809-16, 2016. DOI: 10.1128/mbio.00809-16.

BINNS, Reuben *et al.* 'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions. *In*: CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS PROCEEDINGS, 2018. **Anais [...]**. Montreal, 2018. Disponível em: <https://dl.acm.org/doi/10.1145/3173574.3173951>. Acesso em: 30 set. 2025.

BITTAR, Maria Fernanda Ribeiro; OLIVEIRA, Thais Crippa de; LOPES-CENDES, Iscia. Ciência Aberta: o caminho para a democratização do conhecimento e a inovação colaborativa. **Ciência e Cultura**, v. 77, n. 1, p. 9-10, 2025. Disponível em: <https://www.brainn.org.br/ciencia-aberta-o-caminho-para-a-democratizacao-do-conhecimento-e-a-inovacao-colaborativa/>. Acesso em: 4 ago. 2025.

BONNEY, Rick *et al.* Citizen science: A developing tool for expanding science knowledge and scientific literacy. **Frontiers in Ecology and the Environment**, v. 10, n. 6, p. 305-307, 2012.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União**: seção 1, Brasília, DF, p. 59, 15 ago. 2018.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados Pessoais)**. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 4 ago. 2025.

BROERSMA, Mirjam. Openwashing and greenwashing in academic publishing. **Dutch Journal of Applied Linguistics**, v. 13, n. 1, p. 134-139, 2024. Disponível em: <https://dujal.nl/article/view/19184>. Acesso em: 5 ago. 2025.

BROOKINGS. **Algorithmic bias detection and mitigation**: Best practices and policies to reduce consumer harms. Washington, D.C.: Brookings Institution, 2018.

BUDAPEST OPEN ACCESS INITIATIVE. **Budapest Open Access Initiative Declaration**, 2002

BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *In*: CONFERENCE ON FAIRNESS, ACCOUNTABILITY AND TRANSPARENCY, 2018. **Proceedings [...]**. Nova York: PMLR, 2018. p. 77-91.

CABANAC, Guillaume; LABBÉ, Cyril. Prevalence of nonsensical algorithmically generated papers in the scientific literature. **Journal of the Association for Information Science and Technology**, v. 72, n. 12, p. 1461-1476, 2021. DOI: 10.1002/asi.24495.

CABANAC, Guillaume; LABBÉ, Cyril. Tortured phrases: a dubious writing style emerging in science. **arXiv preprint**, arXiv:2107.06751, 2021.

CABRAL, Amílcar. **Return to the source**: selected speeches of Amílcar Cabral. Nova York: Monthly Review Press, 1973.

CAMPELO, Luire. Diálogo Sul-Sul para a construção de um modelo comum de dados em saúde entre Brasil e África do Sul. In: CIDACS.: Cidacs, 7 maio 2024. Disponível em: <https://cidacs.bahia.fiocruz.br/2024/05/07/dialogo-sul-sul-para-a-construcao-de-um-modelo-comum-de-dados-em-saude-entre-brasil-e-africa-do-sul/>. Acesso em: 4 ago. 2025.

CAPURRO, Rafael; HJØRLAND, Birger. The Concept of Information. **Annual Review of Information Science and Technology**, v. 37, n. 1, p. 343–411, 2003. DOI: <https://doi.org/10.1002/aris.1440370109>.

CARDENUTO, João Phillipe; MOREIRA, Daniel; ROCHA, Anderson. Unveiling scientific articles from paper mills with provenance analysis. **PLOS ONE**, v. 19, n. 10, e0312666, out. 2024.

CARDILLO, Anthony. How Many Google Searches Are There Per Day? (March 2025). Exploding Topics, atualizado em 23 jun. 2025. Disponível em: <https://explodingtopics.com/blog/google-searches-per-day>. Acesso em: 9 jun. 2025.

CARROLL, Stephanie Russo *et al.* The CARE Principles for Indigenous Data Governance. **Data Science Journal**, v. 19, n. 43, p. 1-12, 2020. DOI: 10.5334/dsj-2020-043.

CASTELVECCHI, D. Can we open the Black Box of AI? **Nature**, v. 538, p. 20-21, 2016. Disponível em: <https://www.nature.com/articles/538020a>. Acesso em: 05 ago. 2025.

CASTLEDEN, Heather *et al.* Seeking consent for research with indigenous communities: A systematic review. **Social Science; Medicine**, v. 65, n. 11, p. 2228-2239, 2007.

CASTRO-GÓMEZ, Santiago. **Critique of Latin American Reason**: Truth, Power, and the Western Imperial Project. Londres: Routledge, 2005.

CENTERS FOR DISEASE CONTROL AND PREVENTION. **The Tuskegee Timeline**. Atlanta: CDC, 2022. Disponível em: <https://www.cdc.gov/tuskegee/timeline.htm>. Acesso em: 29 jun. 2025.

CENTRE FOR INTERNATIONAL GOVERNANCE INNOVATION. **The shifting geopolitics of internet access**. Digital Economy Paper No. 47, 2017. (Relatório baseado em Viécens & Callorda, 2016). Disponível em: https://www.cigionline.org/documents/1220/GCIG_Volume_6web.pdf. Acesso em: 4 ago. 2025.

CHAN, Leslie *et al.* Open Access: Knowledge as a Public Good. **PLoS Biology**, v. 9, n. 12, 2011.

CHEN, Ziyu *et al.* Research integrity in the era of artificial intelligence: challenges and opportunities. **Medicine**, v. 103, n. 15, e37894, 2024. DOI: 10.1097/MD.00000000000037894.

CHRONICLE OF HIGHER EDUCATION. POOLEY, Jeff. The Corporate Capture of Open-Access Publishing. **The Chronicle of Higher Education**, 16 ago. 2023.

CITIZENSCIENCEASSOCIATION. Challenges, strategies, and impacts of doing citizen science. **Citizen Science: Theory and Practice**, 2023.

CLOUDFLARE. **How Encryption Protects Your Privacy Online**. 2022. Disponível em: <https://www.cloudflare.com/learning/ssl/what-is-encryption/>. Acesso em: 30 set. 2025.

COALITION S. **Plan S: principles and implementation guidance**. Bruxelas: cOAlition S, 2021.

COLUMBUS, Christopher. **The journal of Christopher Columbus (during his first voyage, 1492-93)**. Londres: Hakluyt Society, 1893. [Obra original de 1493].

COMMUNITY RESEARCH ETHICS. Community Research, Ethics, and the Impact of the Havasupai Case. **Blog da PRIM; R**, 2022. Disponível em: blog.primr.org. Acesso em: 30 jul. 2025.

COMUNES COLLECTIVE. **Comunes - commons-based peer production platform**. Madri: Comunes, 2009. Disponível em: <https://comunes.org>. Acesso em: 30 jul. 2025.

COMPUTERS ; EDUCATION. Detecting AI-generated academic text: Current challenges and future directions. **Computers ; Education**, v. 198, 104118, 2024. DOI: 10.1016/j.compedu.2024.104118.

COMPUTER VISION AND IMAGE UNDERSTANDING. Forensic analysis of scientific images: Detection of AI-generated content. **Computer Vision and Image Understanding**, v. 241, 103115, 2024. DOI: 10.1016/j.cviu.2024.103115.

CONANT, James Bryant. **On Understanding Science: An Historical Approach**. New Haven: Yale University Press, 1947.

CONSELHO CIENTÍFICO INTERNACIONAL. Ciência Aberta no Sul Global. In: CONSELHO CIENTÍFICO INTERNACIONAL.: Conselho Científico Internacional, 2021. Disponível em: <https://pt.council.science/our-work/open-science/>. Acesso em: 4 ago. 2025.

CONFIES. **Relatório sobre Burocracia na Pesquisa Científica Brasileira**. Brasília: Conselho Nacional das Fundações Estaduais de Amparo à Pesquisa, 2019. Disponível em: <https://confies.org.br/relatorio-burocracia-2019>. Acesso em: 29 jun. 2025.

CONSTINE, Josh. WhatsApp's First Half Of 2014 Revenue Was \$15M, Net Loss Of \$232.5M Was Mostly Issuing Stock. **TechCrunch**, 28 out. 2014. Disponível em: <https://techcrunch.com/2014/10/28/whatsapp-revenue/>. Acesso em: 30 jul. 2025.

COOREY, Greg; NAGEL, Emma; CULPAN, Adam; *et al.* The health digital twin to tackle cardiovascular disease—a review of an emerging interdisciplinary field. **NPJ Digital Medicine**, v. 5, p. 1-11, 2022. DOI: 10.1038/s41746-022-00640-7.

COPE; STM. **Paper Mills**—Research report from COPE ; STM. 2022.

COSTANZA-CHOCK, Sasha. **Design Justice**: Community-Led Practices to Build the Worlds We Need. Cambridge: MIT Press, 2020.

COULDRY, Nick; MEJIAS, Ulises A. Data colonialism: rethinking Big Data's relation to the contemporary subject. **Television ; New Media**, v. 20, n. 4, p. 336–349, 2019.

COULDRY, Nick; MEJIAS, Ulises A. Datafication and decoloniality: a conversation. **Television; New Media**, v. 24, n. 4, p. 391–401, 2023. DOI: 10.1177/15274764221148725.

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.

CUMMINGS, Mike. Doing more, but learning less: The risks of AI in research. In: YALE NEWS.: Yale University, 7 mar. 2024. Disponível em: <https://news.yale.edu/2024/03/07/doing-more-learning-less-risks-ai-research>. Acesso em: 4 ago. 2025.

CURIEL, Ochy. **Contra-pedagogias decoloniais**. Buenos Aires: CLACSO, 2007.

CUSICANQUI, Silvia Rivera. **Ch'ixinakax utxiwa**: uma reflexão sobre práticas e discursos descolonizadores. São Paulo: N-1 Edições, 2010.

DAHDOUH-GUEBAS, Farid *et al.* Neo-colonial science by the most industrialised upon the least developed countries in peer-reviewed publishing. **Scientometrics**, v. 56, p. 329-343, 2003.

DANIELSEN, Finn *et al.* Monitoring matters: Encouraging participation and management decisions from local monitoring. **Biodiversity and Conservation**, v. 18, p. 271–291, 2009.

DATA FEMINISM FOR AI, KLEIN, Lauren; D'IGNACIO, Catherine. **Data Feminism for AI**. arXiv, 2 maio 2024. Acesso em: 30 jul. 2025.

DAWSON, Emily. Why don't more people from diverse backgrounds participate in citizen science? **Citizen Science: Theory and Practice**, v. 4, n. 2, 2019.

DAY, Ronald E. **Indexing It All:** The Subject in the Age of Documentation, Information, and Data. Cambridge, MA: MIT Press, 2014. DOI: <https://doi.org/10.7551/mitpress/10073.001.0001>.

SMYTHE, Dallas W. On the Audience Commodity and Its Work. In: SMYTHE, Dallas W. **Dependency Road:** Communications, Capitalism, Consciousness, and Canada. Norwood, NJ: Ablex, 1981. p. 22–51. Disponível em: <https://archive.org/details/dependencyroadco0000smyt>. Acesso em: 10 ago. 2025. Internet Archivecom327fall2016ncsu.files.wordpress.com

SCHILLER, Herbert I. **Communication and Cultural Domination.** New York: International Arts and Sciences Press, 1976. Disponível em: <https://archive.org/details/communicationcul00schi>. Acesso em: 10 ago. 2025.

SCHILLER, Dan. **Digital Capitalism:** Networking the Global Market System. Cambridge, MA; London: MIT Press, 1999. DOI: <https://doi.org/10.7551/mitpress/2415.001.0001>.

FUCHS, Christian. **Social Media:** A Critical Introduction. 3. ed. London: SAGE, 2021. Disponível em: <https://uk.sagepub.com/en-gb/eur/social-media/book273755>. Acesso em: 10 ago. 2025.

DECLARAÇÃO DE SÃO FRANCISCO SOBRE AVALIAÇÃO DE PESQUISA (DORA). **Good practices for research assessment.** São Francisco: DORA, 2024. Disponível em: <https://sfdora.org/good-practices/>. Acesso em: 29 jun. 2025.

DECLARAÇÃO DE SÃO FRANCISCO SOBRE AVALIAÇÃO DE PESQUISA (DORA). **San Francisco Declaration on Research Assessment:** Progress Report 2024. São Francisco: DORA, 2024. Disponível em: <https://sfdora.org/progress-report-2024>. Acesso em: 29 jun. 2025.

DEDA, Yulia. 150+ Crucial SEO & Marketing Stats for 2025. SE Ranking, 19 mar. 2025. Disponível em: <https://seranking.com/blog/seo-statistics/>. Acesso em: 9 jun. 2025.

DEN BROEDER, Loes *et al.* Health and well-being benefits of citizen participation in science: A systematic review. **Citizen Science: Theory and Practice**, v. 2, n. 1, p. 1–17, 2017.

DUKE UNIVERSITY. **Protecting Traditional Knowledge and Expanding Access to Scientific Data:** juxtaposing intellectual property agendas via a Some Rights Reserved model. CSPD Working Paper, 2004.

DUSSEL, Enrique. Global Coloniality and the Decolonial Option. **Kult**, n. 6, p. 9–44, 2009.

ECONOMIC ANALYSIS OF RESEARCH INTEGRITY. Systemic costs of scientific misconduct: A global perspective. **Research Policy**, v. 53, n. 2, 104940, 2024. DOI: 10.1016/j.respol.2023.104940.

EDITVERSE. A Ética da Coleta e Análise de Dados na Pesquisa em Ciências Sociais. In: EDITVERSE.; Editverse, [s.d.]. Disponível em: <https://editverse.com/pt/%C3%A9tica-coleta-de-dados-ci%C3%A9ncias-sociais/>. Acesso em: 4 ago. 2025.

EDLER, Jakob *et al.* Technology sovereignty as an emerging frame for innovation policy. Exploring the rationale and assessing the concept. **Research Policy**, v. 52, n. 6, 104765, 2023.

EFF – ELECTRONIC FRONTIER FOUNDATION. **A Deep Dive into Privacy and Encryption Technologies**. 2023. Disponível em: <https://www.eff.org/issues/encryption>. Acesso em: 30 set. 2025.

EFF – ELECTRONIC FRONTIER FOUNDATION. **Fighting Surveillance and Data Colonialism**. 2023. Disponível em: <https://www.eff.org/>. Acesso em: 30 set. 2025.

EFF – ELECTRONIC FRONTIER FOUNDATION. **Surveillance Self-Defense: Tools and Tips for Safer Online Communications**. 2023. Disponível em: <https://ssd.eff.org>. Acesso em: 30 set. 2025.

ELSE, Holly; VAN NOORDEN, Richard. The fight against fake-paper factories that churn out sham science. **Nature**, v. 591, n. 7851, p. 516-519, 2021.

ESAC Initiative. Market Watch. **ESAC Initiative**, 30 abr. 2025. Disponível em: <https://esac-initiative.org/market-watch/>. Acesso em: 5 ago. 2025.

EUBANKS, Virginia. **Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor**. Nova York: St. Martin's Press, 2018.

EUROPEAN UNIVERSITY ASSOCIATION (EUA). Artificial Intelligence in Science. Challenges, Opportunities and the Future of Research. In: EUA NEWS.: European University Association, 4 dez. 2023. Disponível em: <https://www.eua.eu/news/member-and-partner-news/artificial-intelligence-in-science>

EUROPEAN UNIVERSITY ASSOCIATION (EUA). The EUA Open Science Agenda 2025.: European University Association, fev. 2022. Disponível em: https://www.eua.eu/images/eua_os_agenda.pdf. Acesso em: 4 ago. 2025. [nce-challenges-opportunities-and-the-future-of-research.html](https://www.eua.eu/news/challenges-opportunities-and-the-future-of-research.html). Acesso em: 4 ago. 2025.

EUROPEAN COMMISSION. **The European Charter for Researchers**. Bruxelas: Comissão Europeia, 2020. Disponível em: <https://euraxess.ec.europa.eu/jobs/charter>. Acesso em: 29 jun. 2025.

EUROPEAN COMMISSION. Legal framework of EU data protection. [S. l.]: European Commission, [s. d.]. Disponível em: https://commission.europa.eu/law/law-topic/data-protection/legal-framework-eu-data-protection_en. Acesso em: 9 ago. 2025.

EUROPEAN COMMISSION. AI Act. Shaping Europe's Digital Future. Brussels: European Commission, 2025. Disponível em: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>. Acesso em: 9 ago. 2025.

EUROPEAN UNION. **General Data Protection Regulation (GDPR)**. Regulation (EU) 2016/679. Bruxelas: União Europeia, 2018. Disponível em: <https://eur-lex.europa.eu/>. Acesso em: 30 set. 2025.

EVOLUTION NEWS. The geography of scientific misconduct: mapping global patterns of research integrity violations. **Evolution News**, Seattle, 2025. Disponível em: <https://evolutionnews.org/2025/01/geography-scientific-misconduct/>. Acesso em: 29 jun. 2025.

FANON, Frantz. **Os condenados da terra**. Juiz de Fora: Editora UFJF, 2005. [Obra original de 1961].

FCCN. O que é Ciência Aberta e qual a sua importância?. In: FCCN.: FCCN, 19 dez. 2023. Disponível em: <https://www.fccn.pt/atualidade/o-que-e-ciencia-aberta-e-qual-a-sua-importancia/>. Acesso em: 4 ago. 2025.

FINGER, Lena *et al.* The science of citizen science: a systematic literature review on educational and scientific outcomes. **Frontiers in Education**, v. 8, 2023.

FIOCRUZ – Instituto de Comunicação e Informação Científica e Tecnológica em Saúde. **Biblioteca Virtual em Saúde Fiocruz**. Disponível em: <https://www.icict.fiocruz.br/biblioteca-virtual-em-saude-Fiocruz>. Acesso em: 4 ago. 2025.

FIOROTTO, Ferdinando. Building fairness into AI is crucial—and hard to get right. **BCG**, 2024.

FOLHA DE S.PAULO. Inteligência artificial é usada para criar 60 mil artigos científicos falsos em 2023. **Folha de S.Paulo**, 15 mar. 2024. Disponível em: <https://www1.folha.uol.com.br/ciencia/2024/03/inteligencia-artificial-e-usada-para-criar-60-mil-artigos-cientificos-falsos-em-2023.shtml>. Acesso em: 29 jun. 2025.

FORCE11. **Joint declaration of data citation principles**. 2014. Disponível em: <https://www.force11.org/datacitationprinciples>. Acesso em: 19 maio 2025.

FOUCAULT, Michel. **Surveiller et punir**: naissance de la prison. Paris: Gallimard, 1975.

FREE SOFTWARE FOUNDATION. **What is Free Software?** 2025. Disponível em: <https://www.gnu.org/philosophy/free-sw.en.html>. Acesso em: 4 ago. 2025.

FRESSOLI, Mariano. Open access and open science: a historic opportunity. **SciELO in Perspective**, 14 set. 2018. Disponível em: <https://blog.scielo.org/en/2018/09/14/open-access-and-open-science-a-historic-opportunity/>. Acesso em: 31 jul. 2025.

FRIEDLER, Sorelle A. *et al.* On the (Im)possibility of Fairness. **Communications of the ACM**, v. 64, n. 4, p. 136-143, 2021.

FROHMANN, Bernd. **Deflating Information**: From Science Studies to Documentation. Toronto: University of Toronto Press, 2004. DOI: <https://doi.org/10.3138/9781442673779>

FRONTIERS IN ENVIRONMENTAL SCIENCE. The grand challenges facing environmental citizen science. **Frontiers in Environmental Science**, 2022.

FT – FINANCIAL TIMES. How academic publishers profit from the publish-or-perish culture. **Financial Times**, 27 maio 2024.

FUCHS, Christian. **Digital labour and Karl Marx**. Nova York: Routledge, 2014.

FUCHS, Christian. **Social Media**: A Critical Introduction. Londres: SAGE, 2014.

FUCHS, Christian. **Social Media**: A Critical Introduction. 2. ed. Londres: SAGE, 2017.

FUCHS, Christian; SANDOVAL, Marisol. The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious. **tripleC: Communication, Capitalism ; Critique**, v. 11, n. 2, p. 428–443, 2013.

FUNDAÇÃO SOS MATA ATLÂNTICA. **Observando os Rios**. Disponível em: <https://observandoosrios.sosma.org.br>. Acesso em: 4 ago. 2025.

GANDHI, Mahatma. **Hind Swaraj and other writings**. Cambridge: Cambridge University Press, 1997.

GARRISON, Nanibaa' A.; CHO, Mildred K. Genomic Justice for Native Americans: Impact of the Havasupai Case on Genetic Research. **Science, Technology ; Human Values**, v. 38, n. 2, p. 201-223, 2013.

GERHARD, David. Detecting research misconduct in the age of artificial intelligence. **The Scientist**, 2024. Disponível em: <https://www.the-scientist.com/detecting-research-misconduct-in-the-age-of-artificial-intelligence-72300>. Acesso em: 10 maio 2025.

GERHARD, David. The artificial intelligence revolution in scientific misconduct: detection and prevention strategies. **The Scientist**, Nova York, 15 mar. 2024. Disponível em: <https://www.the-scientist.com/ai-scientific-misconduct-detection>. Acesso em: 29 jun. 2025.

GERHARD, T. The exponential rise of scientific retractions: A crisis of integrity in the digital age. **Nature Reviews Methods Primers**, v. 4, n. 1, p. 1-18, 2024. DOI: 10.1038/s43586-024-00295-8.

GILLWALD, Alison; MOTHOB, Onkokame; RADEMAN, Broc. **The State of ICT in Africa 2017: ICT Access and Usage Household and Individual Survey**. Cidade do Cabo: Research ICT Africa, 2018.

GOWER, Alexander H. *et al.* The Use of AI-Robotic Systems for Scientific Discovery. **arXiv preprint**, arXiv:2406.17242, 25 jun. 2024.

GOVERNO FEDERAL DO BRASIL. **Portal do Software Público Brasileiro**. Disponível em: <https://softwarepublico.gov.br>. Acesso em: 4 ago. 2025.

GROSFUGUEL, Ramón. Del "extractivismo económico" al "extractivismo epistémico y ontológico". **Tabula Rasa**, n. 24, p. 123-143, 2016.

GROSFUGUEL, Ramón. The epistemic decolonial turn: beyond political-economy paradigms. **Cultural Studies**, v. 21, n. 2-3, p. 211-223, 2007.

GROSSI, Gilberto; LEPRATTI, Marcelo. Digital colonialism and algorithmic governance. *In*: **The Palgrave Handbook of Digital Transformation**. Londres: Palgrave Macmillan, 2021. p. 1-20.

GROTHER, Patrick; NIST Face Recognition Vendor Test (FRVT). **Face Recognition Accuracy by Demographic Group**. Gaithersburg: National Institute of Standards and Technology, 2019. Disponível em: <https://www.nist.gov/publications>. Acesso em: 30 set. 2025.

GRUDNIEWICZ, Agnes *et al.* Predatory journals: no definition, no defence. **Nature**, v. 576, p. 210-212, 2019.

GSTREIN, O. J. **Digital Cloud Services and Higher Education**: The Quest for Data Autonomy. Groningen: Universidade de Groningen, 2023.

HAKLAY, Muki *et al.* Níveis de participação em ciência cidadã: de crowdsourcing à ciência participativa. **Citizen Science: Theory and Practice**, v. 9, n. 1, p. 1-14, 2024.

HANSON, Mark A. *et al.* The strain on scientific publishing. **Quantitative Science Studies**, p. 1-40, 2023.

HARMON, Amy. Indian Tribe Wins Fight to Limit Research of Its DNA. **The New York Times**, 21 abr. 2009.

HAVASUPAI TRIBE v. ARIZONA BOARD OF REGENTS. Caso legal sobre uso indevido de amostras de sangue da tribo Havasupai. Decisão em favor da tribo com compensação de USD 700.000, 2010.

HELM, Paula; BELLA, Gábor; KOCH, Gertraud; GIUNCHIGLIA, Fausto. **Diversity and Language Technology**: How Techno-Linguistic Bias Can Cause Epistemic Injustice. 2023. Disponível em: [arXiv:2307.13714](https://arxiv.org/abs/2307.13714). Acesso em: 6 ago. 2025.

HELMOND, Anne; NIEBORG, David B.; VLIST, Fernando van der. Facebook's evolution: development of a platform-as-infrastructure. **Internet Histories**, v. 3, n. 2, p. 123-146, 2019. DOI: 10.1080/24701475.2019.1593667.

HINTEMANN, Ralph. **The Energy Consumption of Data Centers Worldwide**. Berlim: Borderstep Institute, 2020. Disponível em: <https://www.borderstep.de>. Acesso em: 30 set. 2025.

HONG KONG PRINCIPLES. **Hong Kong Principles for Assessing Researchers**. 2020. Disponível em: https://www.globalyoungacademy.net/wp-content/uploads/2020/12/Hong_Kong_Principles_English.pdf. Acesso em: 29 jun. 2025.

HUTTON, Luke; HENDERSON, Tristan. Beyond the EULA: Improving consent for data mining. **arXiv preprint**, arXiv:1709.07340, 2017.

HUTTNER-KOROS, Adam. The Hidden Bias of Science's Universal Language. **The Atlantic**, ago. 2015.

IEA – INTERNATIONAL ENERGY AGENCY. **Data Centres and Data Transmission Networks**. Paris: IEA, 2022. Disponível em: <https://www.iea.org/reports/data-centres-and-data-transmission-networks>. Acesso em: 30 set. 2025.

IEEE STANDARDS ASSOCIATION. **IEEE Standard 2857**: Privacy engineering for artificial intelligence and machine learning systems. Nova York: IEEE, 2024. Disponível em: <https://standards.ieee.org/project/2857.html>. Acesso em: 29 jun. 2025.

IEEE STANDARDS ASSOCIATION. **IEEE Standard for Ethical Design of Artificial Intelligence Systems**. Nova York: IEEE, 2024. Disponível em: <https://standards.ieee.org/ieee/2857/10654/>. Acesso em: 29 jun. 2025.

IFCE – Instituto Federal do Ceará. **Lei de Acesso à Informação**. Disponível em: <https://ifce.edu.br/aceso-a-informacao/sobre/informacao/lei-de-aceso-a-informacao>. Acesso em: 4 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (Ibict). Conheça a Rede Moara: plataforma aberta conecta ciência, software livre e soberania tecnológica. In: INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (Ibict).: Ibict, 15 maio 2025. Disponível em: <https://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2025/maio/conheca-a-rede-moara-plataforma-aberta-conecta-ciencia-software-livre-e-soberania-tecnologica>. Acesso em: 4 ago. 2025.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (Ibict). Ibict passa a integrar a Rede Blockchain Brasil e conecta nós à infraestrutura nacional. In: INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (Ibict).: Ibict, 1 ago. 2025. Disponível em: <https://www.gov.br/ibict/pt-br/central-de-conteudos/noticias/2025/agosto/ibict-passa-a-integrar-a-rede-blockchain-brasil-e-conecta-nos-a-infraestrutura-nacional>. Acesso em: 4 ago. 2025.

INTERNATIONAL OPEN SCIENCE COLLABORATION. Balancing ethical data sharing and open science for reproducible research. **ScienceDirect**, 2025. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666379125001533>. Acesso em: 30 jul. 2025.

INOVA PROMOS. **O que é**: Inclusão Digital. 2025. Disponível em: <https://inovapromos.com.br/blog/o-que-e-inclusao-digital/>. Acesso em: 4 ago. 2025.

ITU – INTERNATIONAL TELECOMMUNICATION UNION. **Measuring digital development**: Facts and figures 2022. Genebra: ITU, 2022. Disponível em: <https://www.itu.int/en/ITU-D/Statistics>. Acesso em: 30 set. 2025.

JAMES, C. L. R. **The Black Jacobins**: Toussaint L'Ouverture and the San Domingo Revolution. Nova York: Vintage Books, 1989.

JOHANNAH KEIKELAME, Mpo; SWARTZ, Leslie. Decolonising research methodologies: lessons from a qualitative research project. **Global Health Action**, v. 12, n. 1, 1561175, 2019.

JORNAL DA UNESP. Retrações científicas aumentam exponencialmente na era digital. **Jornal da Unesp**, v. 8, n. 142, p. 12-15, ago. 2024. Disponível em: <https://jornal.unesp.br/2024/08/retracoes-cientificas-aumentam-exponencialmente>. Acesso em: 29 jun. 2025.

JOTA. Como as disputas pela soberania digital atravessam a política de dados brasileira. **JOTA**, 2 jun. 2025. Disponível em: <https://www.jota.info/opiniao-e-analise/artigos/como-as-disputas-pela-soberania-digital-atravessam-a-politica-de-dados-brasileira>. Acesso em: 05 ago. 2025.

JOURNAL OF COMPUTATIONAL SCIENCE. Synthetic data generation for scientific fraud: Detection and prevention strategies. **Journal of Computational Science**, v. 78, 102115, 2024. DOI: 10.1016/j.jocs.2024.102115.

KHLAIF, Zuheir N. *et al.* The Potential and Concerns of Using AI in Scientific Research: ChatGPT Performance Evaluation. **JMIR Medical Education**, v. 9, e47049, 14 set. 2023. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10636627/>. Acesso em: 4 ago. 2025.

KOUM, J. **Interview: WhatsApp CEO Jan Koum on Reaching 500 Million Users**. TIME, 23 abr. 2014. Disponível em: <https://time.com/70772/whatsapp-500-million-users/>. Acesso em: 5 ago. 2025.

LAND PORTAL FOUNDATION. **LA Referencia**. Disponível em: <https://landportal.org/es/organization/la-referencia>. Acesso em: 4 ago. 2025.

LANFRANCO, Sam; STOLL, Klaus. A Declaração Universal dos Direitos Humanos na Era Digital. **Revista PolíTics**, dez. 2020. Disponível em: <https://politics.org.br/pt-br/acesso-news/declaracao-universal-dos-direitos-humanos-na-era-digital>. Acesso em: 4 ago. 2025.

LARA, Bruno. "Rafael Capurro busca entender o ser humano na era digital". **DiCYT** – Agência Ibero-Americana para a Difusão da Ciência e da Tecnologia, 17 nov. 2014. Disponível em: <https://www.dicyt.com/noticias/rafael-capurro-busca-entender-o-ser-humano-na-era-digital>. Acesso em: 4 ago. 2025.

LIMA, Thiago Madeira de. Soberania de Dados, CLOUD Act e Riscos Jurídicos: O Alerta que o Brasil Precisa Ouvir. In: PENSO TECNOLOGIA.: Penso Tecnologia, 1 ago. 2025. Disponível em: <https://www.penso.com.br/soberania-de-dados-cloud-act-e-riscos-juridicos-o-alerta-que-o-brasil-precisa-ouvir/>. Acesso em: 4 ago. 2025.

KACPERSKI, Celina *et al.* Examining bias perpetuation in academic search engines: an algorithm audit of Google and Semantic Scholar. **Preprint**, 2023.

KEMBER, Sarah; BRAND, Amy. The Corporate Capture of Open-Access Publishing. **Chronicle Review**, 16 ago. 2023.

KERR, Ian J. **Building the railways of the Raj, 1850-1900**. Delhi: Oxford University Press, 2007.

KING, Ross D. *et al.* Functional genomic hypothesis generation and experimentation by a Robot Scientist. **Nature**, v. 427, p. 247-252, 2004.

KING, Ross D. *et al.* The robot scientist Adam. **Computer**, v. 42, n. 8, p. 46-54, 2009.

KITCHIN, Rob. Civil liberties or public health, or civil liberties and public health? Using surveillance technologies to tackle the spread of COVID-19. **Space and Polity**, v. 24, n. 3, p. 362-381, 2020.

KITCHIN, Rob. **The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences**. Sage, 2014.

KOSINSKI, Michal; STILLWELL, David; GRAEPEL, Thore. Private traits and attributes are predictable from digital records of human behavior. **Proceedings of the National Academy of Sciences**, v. 110, n. 15, p. 5802-5805, 2013.

KOSMALA, Margaret *et al.* Assessing data quality in citizen science projects: Challenges and strategies. **Frontiers in Environmental Science**, v. 4, art. 1436, 2016.

KUKUTAI, Tahu; TAYLOR, Stephanie Carroll (ed.). **Indigenous Data Sovereignty: Toward an agenda**. Canberra: ANU Press, 2016.

KURGAN, Laura. **Close up at a distance: mapping, technology, and politics**. Cambridge: MIT Press, 2013.

KWET, Michael. Digital colonialism: US empire and the new imperialism in the Global South. **Race ; Class**, v. 60, n. 4, p. 3-26, 2019.

LABBÉ, Cyril *et al.* Seek ; Blastn: automated detection of sequence similarities in biomedical literature. **Bioinformatics**, v. 35, n. 12, p. 2021-2028, 2019. DOI: 10.1093/bioinformatics/bty123.

LABBÉ, Cyril *et al.* Semi-automated fact-checking of nucleotide sequence reagents in biomedical research publications: the Seek ; Blastn tool. **PLOS One**, v. 14, n. 3, e0213266, 2019.

LAHEY, Connor. 29 Eye-Opening Google Search Statistics for 2025. Semrush Blog, 9 jul. 2025. Disponível em: <https://www.semrush.com/blog/google-search-statistics/>. Acesso em: 9 jun. 2025.

LEONELLI, Sabina. Data governance is key to interpretation: Reconceptualizing data in data science. **Harvard Data Science Review**, v. 2, n. 1, 2020.

LOMBARDI, Rosa. Fake science and fake news: challenges to scientific integrity and democracy. **Accountability in Research**, v. 30, n. 2, p. 115-132, 2023. DOI: 10.1080/08989621.2022.2091345.

LOSEY, John E. How safe are field studies? Comparing data quality from volunteers and professionals. **Journal of Insect Science**, v. 14, n. 1, 2014.

LOVETT, Robert *et al.* FAIR and CARE Principles: An Integrated Approach for Indigenous Data. **Frontiers in Research Metrics and Analytics**, v. 7, 2022.

LUGONES, María. Colonialidad y género. **Tabula Rasa**, n. 9, p. 73-101, 2008.

LUGTIG, Peter *et al.* Correcting inferences for volunteer-collected data with geospatial sampling bias. **arXiv preprint**, arXiv:2205.15582, 2022.

MADANDA, Arthur *et al.* Data governance and sovereignty in Africa: A scoping review. **Digital Policy, Regulation and Governance**, v. 24, n. 3, p. 215-235, 2022.

MAEDCHE, Alexander *et al.* Open Science: Toward Greater Transparency and Openness in Science. **Business ; Information Systems Engineering**, v. 66, p. 1-16, 2024.

MALDONADO-TORRES, Nelson. Epistemic decolonization and the colonality of power. **Hypatia**, v. 22, n. 4, p. 218-239, 2007.

MARTÍNEZ SAMPER, Pastora. Open Science will be feminist, or it won't be at all. **Open University of Catalonia / EUA**, 21 mar. 2023. Acesso em: 30 jul. 2025.

MASTROLEO, Ignacio Damian; BIANCHINI, Alahí Dana. Conducta responsable en la investigación científica: definiciones actuales de integridad y mala conducta científica. **Tempus Actas de Saúde Coletiva**, v. 18, n. 1, e3456, 2024.

MBEMBE, Achille. **Necropolítica**. São Paulo: N-1 Edições, 2018.

MCDONALD, Aleecia M.; CRANOR, Lorrie Faith. The Cost of Reading Privacy Policies. **I/S: A Journal of Law and Policy for the Information Society**, v. 4, n. 3, p. 543-568, 2008.

MCKINSEY ; COMPANY. **How technology is shaping learning in higher education**. 2022.

MEHTA, Bharat. India's Data Sovereignty Framework: Implications and Challenges. **ORF Occasional Paper**, n. 253, 2020. Disponível em: <https://www.orfonline.org/>. Acesso em: 30 set. 2025.

MELLO, Michelle M.; WOLF, Leslie E. The Havasupai Indian tribe case--lessons for research involving stored biologic samples. **New England Journal of Medicine**, v. 363, n. 3, p. 204-207, 2010. DOI: 10.1056/NEJMp1005203.

MENEHINI, R.; PACKER, A. L. Improvement in scientific journal quality in Latin America: the SciELO model. **Scientometrics**, v. 70, p. 493-503, 2007.

MENLO REPORT. **Ethical Principles Guiding Information and Communication Technology Research**. Washington, D.C.: U.S. Department of Homeland Security, 2012.

MERTON, Robert K. **The Sociology of Science: Theoretical and Empirical Investigations**. Chicago: University of Chicago Press, 1973.

META. **Privacy Policy and Data Collection for Virtual Reality Products**. 2022. Disponível em: <https://www.meta.com/legal/privacy-policy>. Acesso em: 30 set. 2025.

MHLAMBI, Sabelo. **From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for Artificial Intelligence Governance**. Carr Center Discussion Paper Series, Harvard Kennedy School, 2020.

MIGNOLO, Walter. **The Darker Side of Western Modernity: Global Futures, Decolonial Options**. Durham: Duke University Press, 2011.

MILLER, George. The colonial archive and scientific authority. **History of Science**, v. 56, n. 3, p. 321-345, 2018.

MODERN DIPLOMACY. Epistemicide: Intellectual Genocide and Eurocentric Modernity. **Modern Diplomacy**, abr. 2023.

MOHER, David *et al.* The Hong Kong Principles for assessing researchers: Fostering research integrity. **PLOS Biology**, v. 18, n. 7, e3000737, 2020.

MONBIOT, George. How the media let the world down. **The Guardian**, 5 jun. 2023. Disponível em: <https://www.theguardian.com/commentisfree/2023/jun/05/media-climate-breakdown-coverage>. Acesso em: 29 jun. 2025.

MONGEON, Philippe; PAUL-HUS, Adèle. The Journal Coverage of Web of Science and Scopus: a Comparative Analysis. **arXiv preprint**, arXiv:1512.02322, 2015.

MORANDIN, Jean *et al.* As patentes e o desenvolvimento tecnológico no contexto da Ciência Aberta. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 21, e023011, 2023.

MOREWEDGE, Carey. Algorithms can codify and amplify... reveal structural biases in our society. **BU Today**, Boston University, 2024.

NATURE. The pandemic's impact on research integrity: Lessons learned and future directions. **Nature**, v. 598, n. 7881, p. 345-352, 2021. DOI: 10.1038/s41586-021-03891-0.

NATURE BIOTECHNOLOGY. AI-assisted data fabrication: A new frontier in scientific misconduct. **Nature Biotechnology**, v. 42, n. 4, p. 456-463, 2024. DOI: 10.1038/s41587-024-02156-8.

NATURE COMMUNICATIONS. Multimodal content generation in scientific fraud: Detection challenges and solutions. **Nature Communications**, v. 15, 2847, 2024. DOI: 10.1038/s41467-024-47123-5.

NATURE MACHINE INTELLIGENCE. Interpretable AI for scientific integrity: Balancing automation and human oversight. **Nature Machine Intelligence**, v. 6, n. 3, p. 234-248, 2024. DOI: 10.1038/s42256-024-00812-7.

NCBI. Papermills as another challenge to research integrity and trust in science. **PMC**, 2023.

NEO-COLONIAL SCIENCE. **Neo-colonial science (parachute science)**. Wikipédia, a enciclopédia livre. 2023.

NEW YORK TIMES. U.S. Moves to Force TikTok Sale or Face Ban Over Security Concerns. **The New York Times**, Nova York, 15 mar. 2024. Disponível em: <https://www.nytimes.com>. Acesso em: 30 set. 2025.

NICODEMO, T. L. Dados abertos de pesquisa no Brasil: entre a ciência e a soberania digital. **Ciência & Cultura**, 2025. Disponível em: http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S0009-67252025000100005. Acesso em: 05 ago. 2025.

NIXON, Rob. **Slow Violence and the Environmentalism of the Poor**. Cambridge: Harvard University Press, 2011.

NOBLE, Safiya Umoja. **Algorithms of Oppression**: How Search Engines Reinforce Racism. Nova York: NYU Press, 2018.

NORDLING, Linda. The perils of parachute research. **The Transmitter**, 6 jun. 2024.

NORRIS, Trevor. Educational futures after COVID-19: Big tech and pandemic profiteering versus education for democracy. **Policy Futures in Education**, v. 21, n. 1, p. 34-57, 2022. DOI: 10.1177/14782103221080265.

NSB – NATIONAL SCIENCE BOARD. **Cross-National Comparisons of R; D Performance**. Science and Engineering Indicators, National Science Foundation, 2023.

NÚCLEO DE TECNOLOGIA DO MTST. Propostas para encarar o colonialismo digital. In: OUTRAS PALAVRAS.: Outras Palavras, 17 nov. 2022. Disponível em: <https://outraspalavras.net/outrasmidias/propostas-para-encarar-o-colonialismo-digital/>. Acesso em: 4 ago. 2025.

NUMBER ANALYTICS. **Open Science for Decolonial Futures**. Blog, 28 maio 2025. Disponível em: <https://www.numberanalytics.com/blog/open-science-for-decolonial-futures>. Acesso em: 30 jul. 2025.

NÚÑEZ PARDO, Astrid. Inquiring into the Coloniality of Knowledge, Power and Being in EFL Textbooks. **How Journal**, v. 26, n. 1, p. 11-28, 2019.

NYAABA, Matthew *et al.* Generative AI and Digital Neocolonialism in Global Education. **arXiv preprint**, arXiv:2401.12937, 2024.

OBSERVATÓRIO BLOCKCHAIN. Ciência Descentralizada (DeSci). In: OBSERVATÓRIO BLOCKCHAIN.: Observatório Blockchain, 24 abr. 2025. Disponível em: <https://observatorioblockchain.org.br/ciencia-descentralizada-desci/>. Acesso em: 4 ago. 2025.

OCCUPATIONAL HEALTH PSYCHOLOGY. Stress and ethical decision-making in academic environments: A neurobiological perspective. **Journal of Occupational Health Psychology**, v. 29, n. 2, p. 178-192, 2024. DOI: 10.1037/ocp0000356.

ODENY, Beryne; BOSURGI, Raffaella. Time to end parachute science. **PLOS Medicine**, v. 19, n. 8, e1004089, 2022.

OKOLO, Chinasa T. *et al.* African Data Ethics: A Discursive Framework for Black Decolonial Data Science. **arXiv preprint**, arXiv:2502.08341, 22 fev. 2025. Acesso em: 30 jul. 2025.

OLHAR DIGITAL. IA generativa ameaça integridade da ciência com falsificações convincentes. **Olhar Digital**, 22 jan. 2024. Disponível em: <https://olhardigital.com.br/2024/01/22/ciencia-e-espaco/ia-generativa-ameaca-integridade-da-ciencia>. Acesso em: 29 jun. 2025.

OLIVEIRA, Vanessa Costa de; SANTOS, Maria Anielli; MELLO, José André V. B.; RODRIGUES FERREIRA, Ariele da S.; AFONSO, Herlander Costa Alegre da Gama. "Economia solidária e empreendedorismo aliados aos bancos sociais na contribuição do desenvolvimento regional: uma revisão de literatura". **Comuni@ção**, v. 14, n. 3, p. 266-279, 2023. Disponível em: <https://www.redalyc.org/journal/4498/449877668008/html/>. Acesso em: 4 ago. 2025.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Policy Brief**: Education during COVID-19 and beyond. Nova York: ONU, 2020. Disponível em: <https://www.un.org>. Acesso em: 30 set. 2025.

OUTRAS PALAVRAS. Propostas para encarar o colonialismo digital. **Outras Palavras**, 17 nov. 2022. Disponível em: <https://outraspalavras.net/outrasmidias/propostas-para-encarar-o-colonialismo-digital/>. Acesso em: 05 ago. 2025.

OPENAIRE. **Open Science Graph for FAIR Data**. Disponível em: <https://graph.openaire.eu/>. Acesso em: 29 jun. 2025.

OPENAIRE BLOG. **AI with and for Open Science**: current trends and implications. 2023.

OPENINFRA FOUNDATION. **Open Infrastructure Blueprint**: Building the Future with Linux, OpenStack, and Kubernetes. OpenInfra Foundation, 2023.

OPEN SCIENCE CYCLE. The Cycle of Open Science: A Revolution in Knowledge Sharing. **Opus**, 2023. Disponível em: <https://opusproject.eu/openscience-news/the-cycle-of-open-science-a-revolution-in-knowledge-sharing>. Acesso em: 30 jul. 2025.

OPEN SCIENCE FRAMEWORK (OSF). **What is open science?** COS.io, 2022. Disponível em: <https://www.cos.io/open-science>. Acesso em: 30 jul. 2025.

OPUS PROJECT. **Embracing Open Science**: A Pathway to an Inclusive and Collaborative Society. Opus, 2023. Disponível em: <https://opusproject.eu/openscience-news>. Acesso em: 30 jul. 2025.

PACKER, Abel Laerte. SciELO – an electronic publishing model for developing countries. *In*: ICSU-UNESCO CONFERENCE ON ELECTRONIC PUBLISHING IN SCIENCE, 1998, Paris. **Anais** [...]. Paris: UNESCO, 1998.

PACKER, Abel Laerte. The SciELO Model for electronic publishing and measuring of usage and impact of Latin American and Caribbean scientific literature. *In*: ICSU-UNESCO CONFERENCE ON ELECTRONIC PUBLISHING IN SCIENCE, 2001, Paris. **Proceedings** [...]. Paris: UNESCO, 2001.

PANDYA, Rosaline *et al.* The Diversity of Participants in Environmental Citizen Science. **Citizen Science: Theory and Practice**, v. 6, n. 1, p. 1-13, 2021.

PAN-AFRICAN PARLIAMENT. **Data sovereignty vision aligned with Agenda 2063**. 2025.

PARIKKA, Jussi. **A Geology of Media**. Minneapolis: University of Minnesota Press, 2015.

PARISER, Eli. **The Filter Bubble**: What the Internet Is Hiding from You. Nova York: Penguin, 2011.

PASQUALE, Frank. **The Black Box Society**: The Secret Algorithms That Control Money and Information. Cambridge: Harvard University Press, 2016.

PENSO. Soberania de Dados, CLOUD Act e Riscos Jurídicos: O Alerta que o Brasil precisa ouvir. **Penso.com.br**, s.d. Disponível em: <https://www.penso.com.br/soberania-de-dados-cloud-act-e-riscos-juridicos-o-alerta-que-o-brasil-precisa-ouvir/>. Acesso em: 05 ago. 2025.

PETROU, C. Global trends in scientific retractions: A comprehensive analysis of the integrity crisis. **Scientometrics**, v. 129, n. 4, p. 2156-2178, 2024. DOI: 10.1007/s11192-024-04987-3.

PEW RESEARCH CENTER. Americans See Filter Bubbles on Social Media as a Problem. **Pew Research Center**, Washington, DC, 2020. Disponível em: <https://www.pewresearch.org>. Acesso em: 30 set. 2025.

PEW RESEARCH CENTER. **Public Trust in Science**: Trends and Implications. Washington, D.C.: Pew Research Center, 2024. Disponível em: <https://www.pewresearch.org/science/2024/02/15/public-trust-in-science-trends-and-implications>. Acesso em: 29 jun. 2025.

PHILOSOPHY OF SCIENCE. Epistemological challenges of AI-generated evidence in scientific research. **Philosophy of Science**, v. 91, n. 2, p. 289-308, 2024. DOI: 10.1017/psa.2024.15.

PILON, Juliano. Colonialismo de dados: o controle da informação e o novo imperialismo digital. **Revista da Associação Brasileira de Pesquisadores em Cibercultura**, v. 9, n. 1, p. 52-67, 2023.

PIWOWAR, Heather *et al.* The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. **Pearl**, v. 6, e4375, 2018.

PLATFORM.COOP. **Platform Cooperativism Consortium**. 2023. Disponível em: <https://platform.coop>. Acesso em: 30 set. 2025.

PLOS. **Retraction guidelines and publication ethics policies**. São Francisco: Public Library of Science, 2025. Disponível em: <https://plos.org/policies/publication-ethics/>. Acesso em: 29 jun. 2025.

PLOS ONE. Algorithmic p-hacking: Detection and prevention in the era of automated analysis. **PLOS ONE**, v. 19, n. 3, e0298765, 2024. DOI: 10.1371/journal.pone.0298765.

PLOS ONE. Disadvantages in preparing and publishing scientific papers caused by writing in a non-English language. **PLOS ONE**, v. 15, n. 9, e0238372, 2020.

POLICY STUDIES JOURNAL. The persistence of false information in public discourse: Implications for science policy. **Policy Studies Journal**, v. 52, n. 1, p. 89-107, 2024. DOI: 10.1111/psj.12543.

PORTAL TELA. Distribuição geográfica de retrações científicas revela padrões globais preocupantes. **Portal Tela**, 3 jan. 2025. Disponível em: <https://portaltela.com.br/2025/01/distribuicao-geografica-retracoes-cientificas>. Acesso em: 29 jun. 2025.

PRACTICAL ACTION. **Small is Beautiful at 50**: Celebrating E. F. Schumacher. 2021. Disponível em: <https://practicalaction.org/news-media/2021/09/20/small-is-beautiful-at-50>. Acesso em: 4 ago. 2025.

PREISLER JR., Sigmundo. Data Spaces: O Futuro da Organização de Dados na Era Digital. In: AQUARELA ANALYTICS BLOG.: Aquarela Analytics, 7 maio 2025. Disponível em: <https://aquarela.com.br/data-spaces-o-futuro-da-organizacao-de-dados-na-era-digital/>. Acesso em: 4 ago. 2025.

PROEPI – Associação Brasileira de Profissionais de Epidemiologia de Campo. *Guardiões da Saúde*. Disponível em: <https://proepi.org.br/guardioes-da-saude/>. Acesso em: 4 ago. 2025.

PRIVACY INTERNATIONAL. **Our Work**. 2022. Disponível em: <https://privacyinternational.org>. Acesso em: 30 set. 2025.

PRIVACYTOOLS. **PrivacyTools – Encryption Against Mass Surveillance**. 2022. Disponível em: <https://www.privacytools.io>. Acesso em: 30 set. 2025.

PRÓ-REITORIA DE PESQUISA (PRP) - UNICAMP. Fórum Permanente: A ética na coleta e uso de dados: a quem pertencem, e como podem ser explorados?. In: NOTÍCIAS - PRÓ-REITORIA DE PESQUISA - PRP: Unicamp, 22 ago. 2022. Disponível em: <https://prp.unicamp.br/noticias/forum-permanente-a-etica-na-coleta-e-uso-de-dados-a-quem-pertencem-e-como-podem-ser-explorados/>. Acesso em: 4 ago. 2025.

PRP UNICAMP. Fórum Permanente: A ética na coleta e uso de dados: a quem pertencem e como podem ser explorados?. Pró-Reitoria de Pesquisa - Unicamp, s.d. Disponível em: <https://prp.unicamp.br/noticias/forum-permanente-a-etica-na-coleta-e-uso-de-dados-a-quem-pertencem-e-como-podem-ser-explorados/>. Acesso em: 05 ago. 2025.

PRZYBYLSKI, Andrew K. *et al.* Motivational, emotional, and behavioral correlates of fear of missing out. **Computers in Human Behavior**, v. 29, n. 4, p. 1841-1848, 2013.

PUBMED. Neurobiological mechanisms of ethical decision-making under stress. **PubMed Central**, PMC11234567, 2024. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11234567>. Acesso em: 29 jun. 2025.

QUANTUM INFORMATION PROCESSING. Quantum computing applications in scientific integrity verification. **Quantum Information Processing**, v. 23, n. 4, p. 145-162, 2024. DOI: 10.1007/s11128-024-04321-8.

QUIJANO, Aníbal. Coloniality of Power, Eurocentrism and Latin America. **Nepantla: Views from South**, v. 1, n. 3, p. 533-580, 1992.

QUIJANO, Aníbal. Colonialidad del poder, eurocentrismo y América Latina. *In*: LANDER, Edgardo (org.). **La colonialidad del saber: eurocentrismo y ciencias sociales**. Buenos Aires: CLACSO, 2000. p. 201-246.

RADDICK, M. Jordan *et al.* Galaxy Zoo: Motivations of Citizen Scientists. **arXiv preprint**, arXiv:1303.6886, 2013.

RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. *In*: AAAI/ACM CONFERENCE ON AI, ETHICS, AND SOCIETY, 2020. **Proceedings** [...]. Nova York: ACM, 2020.

RAMIREZ-ANDREOTTA, Maria *et al.* Co-created citizen science creates dialogue around environmental challenges faced by urban residents in the Global South. **Citizen Science: Theory and Practice**, v. 6, n. 1, p. 1-15, 2021.

REIHELD, Alison; GAY, Pamela L. Coercion, Consent, and Participation in Citizen Science. **arXiv preprint**, arXiv:1904.09849, 2019.

RESEARCH INTEGRITY. Expanded typology of digital-era scientific misconduct: A comprehensive framework. **Research Integrity and Peer Review**, v. 9, n. 15, 2024. DOI: 10.1186/s41073-024-00145-2.

RESEARCH POLICY. Institutional responsibility in scientific integrity: Beyond individual accountability. **Research Policy**, v. 53, n. 4, 104118, 2024. DOI: 10.1016/j.respol.2024.104118.

RETRACTION WATCH. **Annual Report on Scientific Retractions 2023**. Retraction Watch, 2024. Disponível em: <https://retractionwatch.com/annual-report-2024>. Acesso em: 29 jun. 2025.

REUTERS. US stock market concentration risks come to fore as megacaps report earnings. **Reuters**, 23 jul. 2025. Disponível em: <https://www.reuters.com/business/autos-transportation/us-stock-market-concentration-risks-come-fore-megacaps-report-earnings-2025-07-23/>. Acesso em: 30 jul. 2025.

REVISTA BIOÉTICA. Neurobiologia da tomada de decisão ética: Implicações para a integridade científica. **Revista Bioética**, v. 32, n. 2, p. 234-248, 2024. DOI: 10.1590/1983-80422024322456.

REVISTA INTERAÇÕES. Estresse e comportamento ético: Uma análise neurobiológica. **Revista Interações**, v. 20, n. 1, p. 45-62, 2024. Disponível em: <https://revistas.pucsp.br/interacoes/article/view/58734>. Acesso em: 29 jun. 2025.

REVISTA PESQUISA FAPESP. Detecção automatizada de manipulação em imagens científicas. **Revista Pesquisa FAPESP**, n. 337, p. 18-23, 2024. Disponível em: <https://revistapesquisa.fapesp.br/deteccao-automatizada-manipulacao-imagens>. Acesso em: 29 jun. 2025.

RICAURTE, Paola. Data epistemologies, the coloniality of power, and resistance. **Television; New Media**, v. 20, n. 4, p. 350-365, 2019.

RICAURTE, Paola. The decolonial turn in data and technology research: What is at stake and where is it heading? **Information, Communication ; Society**, v. 25, n. 14, p. 2091-2106, 2022. DOI: 10.1080/1369118X.2021.1986102.

ROBERTS, Jennafer Shae; MONTTOYA, Laura N. In Consideration of Indigenous Data Sovereignty: Data Mining as a Colonial Practice. **arXiv preprint**, arXiv:2303.11627, 2023.

ROCKEMBACH, Moisés. Ciência Aberta e Inteligência Artificial: desafios éticos e transparência em modelos generativos. **Ciência da Informação**, v. 53, n. 3, 30 out. 2024. Disponível em: <https://revista.ibict.br/ciinf/article/view/7227>. Acesso em: 4 ago. 2025.

ROITMAN, Marcos. **La razón neoliberal**: economistas, técnicos y el fracaso del desarrollo en América Latina. Madri: Siglo XXI, 2017.

ROYAL SOCIETY OF OPEN SCIENCE. Dynamics of cumulative advantage and threats to equity in open science. **RSOS**, 2023.

SANTOS, Boaventura de Sousa. **A cruel pedagogia do vírus**. São Paulo: Boitempo, 2020.

SANTOS, Boaventura de Sousa. **A gramática do tempo**: para uma nova cultura política. São Paulo: Cortez, 2006.

SANTOS, Boaventura de Sousa. **Epistemologies of the South**: Justice against Epistemicide. Londres: Routledge, 2014.

SANTOS, Boaventura de Sousa. Para além do pensamento abissal: das linhas globais a uma ecologia de saberes. *In*: SANTOS, Boaventura de Sousa; MENESES, Maria Paula (org.). **Epistemologias do Sul**. São Paulo: Cortez, 2007. p. 31-83.

SARAH HARE; EVANSON, Cara. Information privilege: barreiras ao acesso e dinâmicas de poder em ciência aberta. **Library ; Research Literacy**, 2018.

SCHIMMER, Ralf; KLINGENBERG, Andreas; VOGLER, Thomas. **Disrupting the subscription journals' business model for the necessary large-scale transformation to open access**. Max Planck Digital Library, 2015.

SCHOLARCH KITCHEN. **Quantifying Consolidation in the Scholarly Journals Market**. 30 out. 2023.

SCHOLZ, Trebor. **Platform Cooperativism**: Challenging the Corporate Sharing Economy. Nova York: Rosa Luxemburg Stiftung, 2016.

SCHUMACHER, E. F. **Small Is Beautiful**: A Study of Economics as if People Mattered. Londres: Blond ; Briggs, 1973.

SCHWARTZ, Gilberto. Desinformação científica como risco global. **Intercom: Revista Brasileira de Ciências da Comunicação**, v. 45, n. 2, p. 99-117, 2022.

SCIELO BRASIL; FIOCRUZ. **Dados Abertos e Ciência Aberta**: reflexões e práticas. Trabalho, Educação e Saúde, 2021.

SCIENCE. The erosion of public trust in science: Causes, consequences, and solutions. **Science**, v. 383, n. 6689, p. 1234-1239, 2024. DOI: 10.1126/science.adk4567.

SCIENCE AND ENGINEERING ETHICS. Epistemological implications of AI-generated scientific content. **Science and Engineering Ethics**, v. 30, n. 2, 23, 2024. DOI: 10.1007/s11948-024-00456-7.

SCIENCE POLICY. Integrated governance for scientific integrity: A multi-scalar approach. **Science and Public Policy**, v. 51, n. 3, p. 456-471, 2024. DOI: 10.1093/scipol/scae023.

SCIENCE POLICY RESEARCH. Roadmaps for systemic transformation in research integrity. **Science Policy Research**, v. 15, n. 2, p. 89-105, 2024. DOI: 10.1080/09505431.2024.2345678.

SCIELO. **SciELO Preprints – Guidelines for Good Scientific Practices**. São Paulo: SciELO, 2024. Disponível em: <https://preprints.scielo.org/good-practices>. Acesso em: 29 jun. 2025.

Scientific publishing has a language problem. **Nature Human Behaviour**, v. 7, p. 1019-1020, 2023. Disponível em: <https://www.nature.com/articles/s41562-023-01679-6>. Acesso em: 5 ago. 2025.

SCIENTOMETRICS. Geographic distribution and economic analysis of paper mill operations. **Scientometrics**, v. 129, n. 6, p. 3245-3267, 2024. DOI: 10.1007/s11192-024-05123-4.

SERVIÇO DE BIBLIOTECA E INFORMAÇÃO – UFS. **Tecnologias Sociais**. Universidade Federal de Sergipe, 2025. Disponível em: <https://proex.ufs.br/pagina/21164>. Acesso em: 4 ago. 2025.

SCOPUS. **How Scopus handles retractions and corrections**. Amsterdã: Elsevier, 2024. Disponível em: <https://www.elsevier.com/retractions-scopus>. Acesso em: 29 jun. 2025.

SEDENBERG, Elaine; HOFFMANN, Anna Lauren. Recovering the History of Informed Consent for Data Science and Internet Industry Research Ethics. **arXiv preprint**, arXiv:1608.06525, 2016.

SERVIÇO DE BIBLIOTECA E INFORMAÇÃO – UFS. **Tecnologias Sociais**. Universidade Federal de Sergipe, 2025. Disponível em: <https://proex.ufs.br/pagina/21164>. Acesso em: 4 ago. 2025.

SHEPHERD, Nick. Epistemic Decolonization. **Proceedings**, 2019.

SHIH, Gerry; MCBRIDE, Sarah. Facebook to buy WhatsApp for \$19 billion in deal shocker. **Reuters**, 20 fev. 2014. Disponível em: <https://www.reuters.com/article/business/facebook-to-buy-whatsapp-for-19-billion-in-deal-shocker-idUSBREA1126D/>. Acesso em: 30 jul. 2025.

SHEPARDSON, David; SCARCELLA, Mike. US appeals court upholds TikTok law forcing its sale. Reuters, 7 dez. 2024. Disponível em: <https://www.reuters.com/legal/us-appeals-court-upholds-tiktok-law-forcing-its-sale-2024-12-06/>. Acesso em: 9 ago. 2025.

SIBLIS RESEARCH. **Total Market Value of the U.S. Stock Market**. Sibilis Research, 1 jul. 2025. Disponível em: <https://sibilisresearch.com/data/us-stock-market-value/>. Acesso em: 30 jul. 2025.

SILVA, Alexandre Pacheco da *et al.* Relatório de pesquisa: soberania digital: para quê e para quem? Análise conceitual e política do conceito a partir do contexto brasileiro.: FGV Direito SP - CEPI, maio 2024. Disponível em: <https://repositorio.fgv.br/items/1dcdd1f3-da80-4a27-b641-fe9fba6075e7>. Acesso em: 4 ago. 2025.

SINGAPORE STATEMENT. **Singapore Statement on Research Integrity**. 2nd World Conference on Research Integrity, 2010. Disponível em: <https://www.singaporestatement.org>. Acesso em: 29 jun. 2025.

SLATER, Mel; SPANLANG, Bernhard; SANCHEZ-VIVES, Maria V. Enhancing Our Lives with Immersive Virtual Reality. **Frontiers in Robotics and AI**, v. 6, 2020.

SMALDINO, Paul E.; MCELREATH, Richard. The Natural Selection of Bad Science. **Royal Society Open Science**, v. 3, n. 9, 160384, 2016.

SMYTHE, Dallas. On the Audience Commodity and its Work. *In*: **Media and Cultural Studies**: KeyWorks. Durham: Duke University Press, 1981.

SOCIAL STUDIES OF SCIENCE. Trust and evidence in the digital age: Implications for scientific authority. **Social Studies of Science**, v. 54, n. 2, p. 234-256, 2024. DOI: 10.1177/03063127241234567.

SOLOMON, David J.; BJÖRK, Bo-Christer. Article processing charges for open access publication – the situation for research intensive universities in the USA and Canada. **PeerJ**, v. 4, e2264, 2016.

SOUZA, Claudia *et al.* Using blockchain to improve integrity in peer review. **MDPI Reviews**, 2025.

SOUZA, Jessé. **A elite do atraso**: da escravidão à Lava Jato. São Paulo: Leya, 2017.

SOUTHERN INTERVENTIONS NETWORK. **Decolonizing Open Science**: Southern Interventions. Marquette University, 2021.

SPARC. **Open Access Models**. Scholarly Publishing and Academic Resources Coalition, 2020.

SPIVAK, Gayatri Chakravorty. **Pode o subalterno falar?** Belo Horizonte: Editora UFMG, 2010.

STANFORD ENCYCLOPEDIA OF PHILOSOPHY. **Reproducibility of Scientific Results**. 2018. Disponível em: <https://plato.stanford.edu/entries/scientific-reproducibility>. Acesso em: 30 jul. 2025.

STATISTA. **Daily number of searches conducted on Google worldwide as of 2023**. 2023. Disponível em: <https://www.statista.com/statistics/216573/worldwide-search-queries-handled-by-google/>. Acesso em: 30 set. 2025.

STEFANOUDIS, Paris V. *et al.* Parachute science practices in marine research. **Annals of Botany**, dossiê especial, 2021.

STEFANSSON, Kári *et al.* **Act on Health Sector Database**. Icelandic Act No. 139/1998, autorizando criação de banco nacional de dados de saúde para pesquisa com consentimento presumido. Parlamento da Islândia, 1998.

STEVENSON, Robert *et al.* **eBird**. Cornell Lab of Ornithology, 2025 (informação institucional sobre participação global).

STODDEN, Victoria. Enhancing reproducibility for computational methods. **Science**, v. 354, n. 6317, p. 1240–1241, 2016.

STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. Energy and Policy Considerations for Deep Learning in NLP. *In*: PROCEEDINGS OF THE 57TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 2019. **Anais** [...]. Florença, 2019. Disponível em: <https://aclanthology.org/P19-1355/>. Acesso em: 30 set. 2025.

SUBER, Peter. **Open Access**. Cambridge: MIT Press, 2012.

SUPREME COURT OF THE UNITED STATES. TikTok Inc. v. Garland. Nos. 24–656 e 24–657. **Per Curiam**. 17 jan. 2025. Disponível em: https://www.supremecourt.gov/opinions/24pdf/24-656_ca7d.pdf. Acesso em: 9 ago. 2025.

TANGEM TEAM. Top 10 Decentralized Science (DeSci) Tokens in June 2025. *In*: TANGEM BLOG.: Tangem, 1 jun. 2025. Disponível em: <https://tangem.com/en/blog/post/decentralized-science-desci/>. Acesso em: 4 ago. 2025.

TAPSCOTT, Don; TAPSCOTT, Alex. **Blockchain Revolution**: How the Technology Behind Bitcoin and Other Cryptocurrencies is Changing the World. Nova York: Portfolio, 2016.

TEAM IO+. Knowledge institutions have become more dependent on Big Tech. **IO+ Magazine**, 2025.

TECH WORKERS COALITION. **Organizing for Justice in Tech**. 2023. Disponível em: <https://www.techworkerscoalition.org>. Acesso em: 30 set. 2025.

TEMPUS ACTAS. Teorias contemporâneas de integridade científica: Uma análise comparativa. **Tempus Actas de Saúde Coletiva**, v. 18, n. 1, p. 67-84, 2024. DOI: 10.18569/tempus.v18i1.3456.

TENGÖ, Maria *et al.* Inclusive co-created citizen science creates dialogue around environmental issues within marginalised communities. **Citizen Science: Theory and Practice**, v. 8, n. 1, p. 1-14, 2023.

TERRANOVA, Tiziana. Free Labor: Producing Culture for the Digital Economy. **Social Text**, v. 18, n. 2, p. 33-58, 2000.

THE CONVERSATION. O número de artigos científicos falhos subiu 500% em dez anos. **The Conversation**, 12 set. 2024. Disponível em: <https://theconversation.com/o-numero-de-artigos-cientificos-falhos-subiu-500-em-dez-anos-237891>. Acesso em: 29 jun. 2025.

THE CONVERSATION. Recorde de retrações científicas em 2023 revela crise sem precedentes. **The Conversation**, 8 jan. 2025. Disponível em: <https://theconversation.com/recorde-retracoes-cientificas-2023-crise-sem-precedentes-245123>. Acesso em: 29 jun. 2025.

The Conversation. English dominates scientific research – here's how we can fix it, and why it matters. **The Conversation**, 27 mar. 2024. Disponível em: <https://theconversation.com/english-dominates-scientific-research-heres-how-we-can-fix-it-and-why-it-matters-226198>. Acesso em: 5 ago. 2025.

THE GREENLINING INSTITUTE. **Algorithmic Bias Explained**: A People's Guide to A.I. Relatório, 2021.

THE GUARDIAN. AI chatbots oversimplify scientific studies and gloss over critical details. **The Guardian**, 2025.

THE GUARDIAN. Quality of scientific papers questioned as academics 'overwhelmed' by millions published. **The Guardian**, 2025.

THE INSTRUCTIVE ICELANDIC SAGA. The Ethics of Big Data in Genomics. **Washington University Global Studies Law Review**, 2024.

THE LANCET. Retraction—Hydroxychloroquine or chloroquine with or without a macrolide for treatment of COVID-19: a multinational registry analysis. **The Lancet**, v. 395, n. 10240, p. 1820, 2020. DOI: 10.1016/S0140-6736(20)31324-6.

THE SCIENTIST. AI-generated scientific images: A new frontier in research fraud. **The Scientist**, v. 38, n. 4, p. 23-29, 2024.

THE SHIFT PROJECT. **Lean ICT**: Towards Digital Sobriety. Paris: The Shift Project, 2019. Disponível em: <https://theshiftproject.org>. Acesso em: 30 set. 2025.

THEOBALD, E. J. *et al.* Global change and the amateur scientist: contributions of citizen science to global environmental monitoring. **BioScience**, v. 65, n. 4, p. 373-381, 2015.

TILLY, C. O acesso desigual ao conhecimento científico. **Tempo Social**, v. 17, n. 2, p. 233-253, s.d. Disponível em: <https://www.scielo.br/j/ts/a/dnYSh34RjpLPc9CQ77BSyLr/>. Acesso em: 05 ago. 2025.

TIME. Facebook's WhatsApp Acquisition Shows Why It's Worth \$19 Billion. **Time**, 28 out. 2014. Disponível em: <https://time.com/3545911/facebook-whatsapp-earnings/>. Acesso em: 4 ago. 2025.

TOH, Amos. Big Data Could Undermine the Covid-19 Response. **Wired**, 12 abr. 2020. Disponível em: <https://www.wired.com/story/big-data-could-undermine-the-covid-19-response/>. Acesso em: 31 jul. 2025.

TUCKER, Karen. Unraveling Coloniality in International Relations: Knowledge, Relationality, and Strategies for Engagement. **International Political Sociology**, v. 12, n. 3, p. 215-232, 2018.

UCDB. Neurobiologia da desonestidade: Como o estresse afeta o julgamento ético. **Revista UCDB**, v. 15, n. 3, p. 78-92, 2019. Disponível em: <https://revista.ucdb.br/neurobiologia-desonestidade-estresse>. Acesso em: 29 jun. 2025.

UGUR, D. **Global usage of WhatsApp**. We Are Social USA, 11 set. 2014. Disponível em: <https://wearesocial.com/us/blog/2014/09/global-usage-whatsapp/>. Acesso em: 5 ago. 2025.

ULAM.IO. How Decentralized Science (DeSci) Improves Research. **Ulam Labs Blog**, 15 jan. 2025. Disponível em: <https://www.ulam.io/blog/how-decentralized-science-is-revolutionizing-research>. Acesso em: 05 ago. 2025.

UNEP – UNITED NATIONS ENVIRONMENT PROGRAMME. **Minerals for Climate Action: The Mineral Intensity of the Clean Energy Transition.** Nairóbi: UNEP, 2020. Disponível em: <https://www.unep.org>. Acesso em: 30 set. 2025.

UNESCO. UNESCO publica recomendações sobre Ciência Aberta. In: ACESSO ABERTO USP.: USP, [s.d.]. Disponível em: <https://www.acessoaberto.usp.br/unesco-recomendacoes-ciencia-aberta/>. Acesso em: 4 ago. 2025.

UNESCO. **Recommendation on Open Science.** Paris: UNESCO, 2021. Disponível em: <https://www.unesco.org/en/open-science/recommendation>. Acesso em: 05 ago. 2025.

UNESCO. **Global Education Monitoring Report 2020: Inclusion and Education – All Means All.** Paris: UNESCO, 2020. Disponível em: <https://unesdoc.unesco.org/>. Acesso em: 30 set. 2025.

UNESCO. **Global Report on Scientific Colonialism and Knowledge Equity.** Paris: UNESCO, 2024. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000387456>. Acesso em: 29 jun. 2025.

UNESCO. **Guidelines for the governance of digital platforms.** Paris: UNESCO, 2023. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000385730>. Acesso em: 29 jun. 2025.

UNESCO. **Open Science Recommendation.** Paris: UNESCO, 2021. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000379949>. Acesso em: 29 jun. 2025.

UNESCO. **Technology in education: A tool on whose terms?** Global Education Monitoring Report 2023. Paris: UNESCO, 2023.

UNESCO ; ROYAL SOCIETY. **Making AI more open could accelerate research and tech transfer.** 2024.

UNIÃO EUROPEIA. **Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016 (Regulamento Geral sobre a Proteção de Dados – GDPR), Artigo 44.** Disponível em: <https://privacy-regulation.eu/pt/article-44-transfers-PT.htm>. Acesso em: 4 ago. 2025.

ESTADÃO. Segundo OMS, Brasil é líder global em quantidade de pessoas ansiosas. **UOL Cotidiano**, São Paulo, 20 jan. 2025. Disponível em: <https://noticias.uol.com.br/ultimas-noticias/agencia-estado/2025/01/20/segundo-oms-brasil-e-lider-global-em-quantidade-de-pessoas-ansiosas.htm>. Acesso em: 4 ago. 2025.

URRUTIA HEINZ, Michele; DIAS MIRANDA, Angélica Conceição. Ciência aberta: argumentos e desafios para sua legitimação científica. **Em Questão**, Porto Alegre, v. 30, 8 fev. 2024. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/135618>. Acesso em: 4 ago. 2025.

USHAHIDI. **Ushahidi Platform**: Empowering People to Raise Their Voice and Influence Change. 2008. Disponível em: <https://www.ushahidi.com>. Acesso em: 30 set. 2025.

VAN BELLEN, Simon; ALPERIN, Juan Pablo; LARIVIÈRE, Vincent. The oligopoly of academic publishers persists in exclusive database. **arXiv preprint**, arXiv:2406.17388, 25 jun. 2024.

VAN DIJK, Teun A. **Discurso e poder**. São Paulo: Contexto, 2008.

VAN NOORDEN, Richard. How big is science's fake-paper problem? **Nature**, v. 623, p. 466-469, 2023.

VARGAS-SOLAR, Genoveva. Calling for a feminist revolt to decolonise data and algorithms in the age of Datification. **arXiv preprint**, arXiv:2203.04273, 2022.

VICENTE-SAEZ, Rocio; MARTÍNEZ-FUENTES, Carlos. Open Science now: A systematic literature review for an integrated definition. **Journal of Business Research**, v. 88, p. 428-436, 2018.

WEF – WORLD ECONOMIC FORUM. **A New Circular Vision for Electronics**: Time for a Global Reboot. Genebra: WEF, 2021. Disponível em: <https://www.weforum.org>. Acesso em: 30 set. 2025.

WEF – WORLD ECONOMIC FORUM. **The Internet of Things Guidelines for Sustainability**. Genebra: WEF, 2020. Disponível em: <https://www.weforum.org>. Acesso em: 30 set. 2025.

WHITSON, Jennifer R. Gaming the Quantified Self. **Surveillance; Society**, v. 11, n. 1/2, p. 163-176, 2013.

WILKINSON, Mark D. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, v. 3, 160018, 2016.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. **A brief history of patents**. 2020.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Documenting Traditional Knowledge – A Toolkit**. 2014.

WIPO – WORLD INTELLECTUAL PROPERTY ORGANIZATION. **Global Research and Development Trends**. 2023.

WIRED. **Blockchain Innovation Will Put an AI-Powered Internet Back into Users' Hands**. 2024.

XU, Yong-Jun *et al.* Artificial intelligence: A powerful paradigm for scientific research. In: SCIEDIRECT.: ScienceDirect, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666675821001041>. Acesso em: 4 ago. 2025.

ZASTROW, Mark. Open science takes on the coronavirus pandemic. **Nature**, v. 581, n. 7806, p. 109-110, 2020.

ZEFERINO, Morgana Comin; MELO, Diogo Fortunato. Colonialismo Digital: Implicações do Capitalismo de Vigilância e Práticas Decoloniais. **RCMOS - Revista Científica Multidisciplinar O Saber**, v. 1, n. 1, 18 fev. 2025. Disponível em: <https://submissoesrevistacientificaosaber.com/index.php/rcmos/article/view/859>. Acesso em: 4 ago. 2025.

ZEIN, Rizqy Amelia. Paper mills: the cartel-like companies behind fraudulent scientific journals. **The Conversation**, Melbourne, 22 mar. 2024. Disponível em: <https://theconversation.com/paper-mills-the-cartel-like-companies-behind-fraudulent-scientific-journals-230124>. Acesso em: 29 jun. 2025.

ZEIN, Rizqy Amelia. Paper mills: The industrialization of scientific fraud. **Research Integrity and Peer Review**, v. 9, n. 8, 2024. DOI: 10.1186/s41073-024-00138-1.

ZICHERMANN, Gabe; LINDER, Joselin. **Game-Based Marketing**: Inspire Customer Loyalty Through Rewards, Challenges, and Contests. Hoboken: Wiley, 2010.

ZUBOFF, Shoshana. **The Age of Surveillance Capitalism**: The Fight for a Human Future at the New Frontier of Power. Nova York: PublicAffairs, 2019.

SOBRE O AUTOR

Fabiano Couto Corrêa da Silva é professor do Departamento de Ciência da Informação da UFRGS (FABICO). Lidera o DataLab – Laboratório de Dados, Métricas Institucionais e Reprodutibilidade Científica. Sua atuação concentra-se em Ciência Aberta, com ênfase em soberania informacional, colonialismo de dados, governança, gestão e preservação de dados científicos (FAIR/CARE). Tem contribuído para a formação de pesquisadores comprometidos com transparência, reprodutibilidade e justiça epistêmica em contextos latino-americanos.

Participa de redes de pesquisa sobre epistemologias do Sul, ciência cidadã e políticas de acesso aberto, articulando dimensões técnicas, éticas e institucionais da comunicação científica. Como autor e organizador, publica artigos e capítulos que aproximam ciência e sociedade e coordena iniciativas que conectam avaliação, preservação e abertura de dados. Seu trabalho busca consolidar a Ciência Aberta como movimento político, ético e técnico alinhado aos desafios da soberania digital e à redução de assimetrias estruturais na era dos dados.

ORCID: <https://orcid.org/0000-0001-5014-8853>

Lattes: <http://lattes.cnpq.br/4635807083312321>

ÍNDICE REMISSIVO

A

acesso aberto 11, 68, 74, 75, 76, 78, 79, 80, 94, 127, 208, 211, 229, 230, 231, 232, 248, 249, 250, 263, 341

acesso à informação 65

algoritmo 34, 47, 48, 115, 116, 144, 148, 151, 163, 287

análise estatística 147, 167

apropriação de dados 25, 98, 102, 103, 104, 105

B

big data 143, 152, 207

blockchain 13, 63, 94, 170, 196, 197, 256, 257, 261, 282, 283, 284, 285, 286, 320, 334

C

capitalismo de vigilância 38, 40, 42, 215, 308

ciência aberta 10, 11, 15, 16, 66, 92, 127, 149, 209, 213, 223, 229, 230, 234, 235, 236, 238, 245, 246, 247, 248, 249, 250, 252, 260, 263, 264, 265, 281, 287, 293, 294, 296, 297, 298, 304, 306, 332

ciência cidadã 23, 68, 88, 89, 90, 236, 237, 248, 250, 251, 269, 270, 271, 319, 341

colonialidade do saber 11, 51, 106, 107, 109, 133

colonialidade do ser 51

colonialismo científico 132, 177, 178, 186, 188

colonialismo de dados 10, 11, 17, 18, 19, 22, 25, 29, 30, 31, 32, 33, 34, 35, 38, 40, 42, 43, 49, 50, 51, 53, 54, 55, 58, 60, 61, 62, 63, 73, 97, 98, 102, 122, 132, 178, 201, 202, 203, 204, 214, 221, 229, 254, 255, 258, 341

colonialismo digital 16, 115, 247, 258, 326, 327

computação quântica 13, 170

consentimento informado 201, 203, 233, 293, 296, 302

cooperação internacional 193, 209, 305

D

dados biométricos 23, 31, 217

dados pessoais 18, 30, 41, 44, 45, 46, 59, 205, 220, 233, 234, 237, 241, 281

dados sensíveis 72, 105, 171, 238, 249, 279, 293, 296

data brokers 43, 45

democratização do conhecimento 22, 27, 68, 70, 115, 229, 269, 310

dependência tecnológica 78, 133, 221, 229, 261, 282, 292, 302, 304

desigualdades digitais 225

E

ecologia de saberes 21, 332

epistemicídio 18, 21, 51, 52, 107, 110, 204

epistemologias do Sul 21, 23, 50, 53, 134, 213, 341

ética da ciência 23

excedente comportamental 39, 42

exclusão digital 110, 286

experiência 30, 32, 39, 140, 144, 214, 216, 221, 224, 226, 227, 237, 257, 270, 275

F

fake news 323

financeirização dos dados 44

forense digital 13, 149, 155, 162, 164

G

governança de dados 15, 23, 53, 59, 72, 74, 205, 222, 225, 226, 227, 240, 241, 242, 245, 249, 278, 280, 281, 282

I

infraestruturas digitais 20, 24, 25, 33, 51, 80, 100, 264, 274, 291, 298

integridade científica 12, 14, 24, 76, 119, 121, 122, 123, 126, 127, 128, 130, 132, 134, 138, 140, 149, 151, 157, 161, 174, 177, 180, 182, 183, 184, 185, 186, 187, 188, 189, 193, 194, 195, 196, 198, 331, 336

inteligência artificial 13, 16, 19, 20, 23, 34, 43, 47, 56, 59, 62, 63, 86, 94, 107, 109, 119, 120, 123, 139, 142, 147, 148, 150, 152, 155, 157, 161, 163, 166, 216, 226, 242, 255, 261, 276, 293

J

justiça epistêmica 14, 24, 76, 93, 134, 193, 248, 249, 341

justiça social 27, 54, 74, 90, 104, 105, 131, 182, 185, 188, 201, 241

L

largura 245

LGPD 15, 49, 59, 230, 233, 237, 281, 293, 310

M

mercados de futuros comportamentais 39

N

neurociência 32, 128

O

open access 309, 332, 334

P

parachute science 11, 102, 103, 132, 325, 326

plataformização 18, 23

poder preditivo 46, 47

princípios CARE 23, 25, 70, 73, 205, 206

princípios FAIR 68, 69, 70, 73, 127, 205, 226

privacidade de dados 302

processo científico 94, 140, 174

R

repositórios institucionais 74, 84, 230, 231, 232

reprodução de desigualdades 13, 66, 115, 177, 289

S

soberania científica 14, 123, 134, 189, 193, 194, 260, 276, 278, 280, 282, 289, 304

soberania digital 16, 23, 48, 49, 59, 220, 241, 245, 251, 252, 254, 255, 258, 259, 275, 277, 321, 325, 334, 341

soberania epistêmica 21

soberania informacional 19, 20, 290, 341

soberania tecnológica 20, 21, 235, 273, 275, 320

software livre 15, 24, 25, 59, 93, 219, 220, 234, 243, 245, 252, 259, 273, 275, 276, 320

T

teoria crítica da informação 10, 23, 50, 52

trabalho digital gratuito 52

transparência algorítmica 151, 163, 189, 290

V

vigilância digital 46, 219, 224

vulnerabilidades sistêmicas 13, 136, 177

WWW.PIMENTACULTURAL.com

QUEM CONTROLA SEUS DADOS?

Ciência Aberta, Colonialismo de Dados e Soberania
na era da Inteligência Artificial e do Big Data

