INTELIGÊNCIA ARTIFICIAL

COMO FERRAMENTA DE

EDUCAÇÃO INCLUSIVA

NA PERSPECTIVA DA NEURODIVERSIDADE



Flávio de Ligório Silva
(ORGANIZADOR)

Wissen editora
2026

INTELIGÊNCIA ARTIFICIAL

COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA

NA PERSPECTIVA DA NEURODIVERSIDADE



Flávio de Ligório Silva
(ORGANIZADOR)

Wissen editora
2026

Flávio de Ligório Silva
Organizador

INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA NA PERSPECTIVA DA NEURODIVERSIDADE

Teresina-PI, 2026

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

©2026 *by* Wissen Editora
Copyright © Wissen Editora
Copyright do texto © 2026 Os autores
Copyright da edição © Wissen Editora
Todos os direitos reservados

Direitos para esta edição cedidos pelos autores à Wissen Editora.

Todo o conteúdo desta obra, inclusive correção ortográfica e gramatical, é de responsabilidade do(s) autor(es). A obra de acesso aberto (Open Access) está protegida por Lei, sob Licença Creative Commons Atribuição-Não Comercial-Sem Derivações 4.0 Internacional, sendo permitido seu *download* e compartilhamento, desde que atribuído o crédito aos autores, sem alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

Editores Chefe: Dr. Junielson Soares da Silva
Ma. Neyla Cristiane Rodrigues de Oliveira
Dra. Denise dos Santos Vila Verde
Dra. Adriana de Sousa Lima

Projeto Gráfico e Diagramação: Neyla Cristiane Rodrigues de Oliveira

Imagem da Capa: Canva

Edição de Arte: Neyla Cristiane Rodrigues de Oliveira

Revisão: Os autores
A Organizadora

Informações sobre a Editora
Wissen Editora
Homepage: www.editorawissen.com.br
Teresina – Piauí, Brasil
E-mails: wisseneditora@gmail.com

Siga nossas redes sociais:
@wisseneditora

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

EQUIPE EDITORIAL

Editores-chefes

Dr. Junielson Soares da Silva
Ma. Neyla Cristiane Rodrigues de Oliveira
Dra. Denise dos Santos Vila Verde
Dra. Adriana de Sousa Lima

Equipe de arte e editoração

Neyla Cristiane Rodrigues de Oliveira

CONSELHO EDITORIAL

Ciências Agrárias e Multidisciplinar

Dr. Felipe Górski - Secretaria de Educação do Paraná (SEED/PR)
Dra. Patrícia Pato dos Santos - Universidade Anhanguera (Uniderp)
Dr. Jose Carlos Guimaraes Junior - Governo do Distrito Federal (DF)

Ciências Biológicas e da Saúde

Dra. Francijara Araújo da Silva - Centro Universitário do Norte (Uninorte)
Dra. Rita di Cássia de Oliveira Angelo - Universidade de Pernambuco (UPE)
Dra. Ana Isabelle de Gois Queiroz - Centro Universitário Ateneu (UniAteneu)

Ciências Exatas e da Terra e Engenharias

Dr. Allan Douglas Bento da Costa - Universidade Federal Rural da Amazônia (UFRA)
Dra. Vania Ribeiro Ferreira - Instituto Federal de Santa Catarina (IFSC)
Dr. Agmar José de Jesus Silva – Secretaria de Educação do Amazonas (Seduc/AM)

Linguística, Letras e Artes

Dra. Conceição Maria Alves de A. Guisardi - Universidade Federal de Uberlândia (UFU)
Dr. Danni Conegatti Batista – Universidade Federal do Rio Grande do Sul (UFRGS)

Ciências Humanas e Sociais Aplicadas

Dr. Isael de Jesus Sena - Culture, Education, Formation, Travail (CIRCEFT)
Dra. Mareli Eliane Graupe - Universidade do Planalto Catarinense (Uniplac)
Dr. Rodrigo Avila Colla - Rede Municipal de Ensino de Esteio, RS
Dr. Erika Giacometti Rocha Berribili - Universidade Federal de São Carlos (UFSCar)
Dr. Douglas Manoel Antonio De Abreu P. Dos Santos - Universidade de São Paulo (USP)
Dra. Aline Luiza de Carvalho - Fundação Hospitalar do Estado de Minas Gerais (FHEMIG)
Dr. José Luiz Esteves - Pontifícia Universidade Católica do Paraná (PUC/PR)
Dr. Claudemir Ramos - Pontifícia Universidade Católica de São Paulo (PUC/SP)
Dr. Daniela Conegatti Batista – Universidade Federal do Rio Grande do Sul (UFRGS)
Dr. Wilson de Lima Brito Filho - Universidade Federal da Bahia (UFBA)
Dr. Cleonice Pereira do Nascimento Bittencourt- Universidade de Brasília (UnB)
Dr. Jonata Ferreira de Moura - Universidade Federal do Maranhão (UFMA)
Dra. Renata dos Santos - Universidade Federal de Itajubá (UNIFEI)

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

Conselho Técnico Científico

Me. Anderson de Souza Gallo - Universidade Federal Rural do Rio de Janeiro (UFRRJ)
Ma. Antônia Alikaene de Sá - Universidade Federal do Piauí (UFPI)
Ma. Talita Benedcta Santos Künast - Universidade Federal do Paraná (UFPR)
Ma. Irene Suelen de Araújo Gomes – Secretaria de Educação do Ceará (Seduc /CE)
Ma. Tamires Oliveira Gomes - Universidade Federal de São Paulo (Unifesp)
Ma. Aline Rocha Rodrigues - União Das Instituições De Serviços, Ensino E Pesquisa LTDA (UNISEPE)
Me. Mauricio Pavone Rodrigues - Universidade Cidade de São Paulo (Unicid)
Esp. Rubens Barbosa Rezende – Faculdade UniFB
Me. Luciano Cabral Rios – Secretaria de Educação do Piauí (Seduc/PI)
Me. Jhenys Maiker Santos - Universidade Federal do Piauí (UFPI0
Dr. Francisco de Paula S. de Araujo Junior – Instituto Federal da Bahia (IFBA)
Ma. Anna Karla Barros da Trindade - Instituto Federal do Piauí (IFPI)
Ma. Elaine Fernanda dos Santos - Universidade Federal de Sergipe (UFS)
Ma. Lilian Regina Araújo dos Santos - Universidade do Grande Rio (Unigranrio)
Ma. Luziane Said Cometti Lélis - Universidade Federal do Pará (UFPA)
Ma. Márcia Antônia Dias Catunda - Devry Brasil
Ma. Marcia Rebeca de Oliveira - Instituto Federal da Bahia (IFBA)
Ma. Mariana Morais Azevedo - Universidade Federal de Sergipe (UFS)
Ma. Marlova Giuliani Garcia - Instituto Federal Farroupilha (IFFar)
Ma. Rosana Maria dos Santos - Universidade Federal Rural de Pernambuco (UFRPE)
Ma. Rosana Wichineski de Lara de Souza - Universidade Federal de Minas Gerais (UFMG)
Ma. Simone Ferreira Angelo - Escola Família Agrícola de Belo Monte - MG
Ma. Suzel Lima da Silva - Universidade Federal de Santa Maria (UFSM)
Ma. Tatiana Seixas Machado Carpenter - Escola Parque
Me. Cássio Joaquim Gomes - Instituto Federal de Nova Andradina
Me. Daniel Ordane da Costa Vale - Secretaria Municipal de Educação de Contagem
Me. Diego dos Santos Verri - Secretária da Educação do Rio Grande do Sul
Me. Fernando Gagno Júnior - SEMED - Guarapari/ES
Me. Grégory Alves Dionor - Universidade do Estado da Bahia (UNEB)/ Universidade Federal da Bahia (UFBA)
Me. Lucas Pereira Gandra - Universidade Federal do Rio Grande do Sul (UFRGS); UNOPAR, Polo Coxim/MS
Me. Lucas Peres Guimarães – Secretaria Municipal de Educação de Barra Mansa - RJ
Me. Luiz Otavio Rodrigues Mendes - Universidade Estadual de Maringá (UEM)
Me. Mateus de Souza Duarte - Universidade Federal de Sergipe (UFS)
Me. Milton Carvalho de Sousa Junior - Instituto Federal do Amazonas (IFAM)
Me. Sebastião Rodrigues Moura - Instituto Federal de Educação do Pará (IFPA)
Me. Wanderson Diogo A. da Silva - Universidade Regional do Cariri (URCA)
Ma. Heloisa Fernanda Francisco Batista - Universidade Federal de Uberlândia (UFU) e Universidade Federal de São Carlos (UFSCar)
Ma. Telma Regina Stroparo - Universidade Estadual do Centro Oeste (Unicentro)
Me. Sérgio Saraiva Nazareno dos Anjos - Empresa Brasileira de Pesquisa Agropecuária (Embrapa)

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA NA PERSPECTIVA DA NEURODIVERSIDADE

http://www.doi.org/10.52832/wed.209

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Inteligência artificial como ferramenta de educação inclusiva na perspectiva da neurodiversidade [livro eletrônico] /(organizador) Flávio de Ligório Silva. --Teresina, PI: Wissen Editora, 2026.
PDF
Vários autores.
Bibliografia.
ISBN: 978-65-6219-006-9
DOI: 10.52832/wed.209

1. Educação inclusiva 2. Inteligência artificial - Aplicações educacionais3. Neurodiversidade 4. Pessoas com deficiência -Acessibilidade 5. Pessoas com deficiência -Educação 6. Tecnologia Assistiva (TA) I. Silva,Flávio de Ligório.

26-379706.0 CDD-371.334

Índices para catálogo sistemático:

1. Inteligência artificial : Educação 371.334
Aline Graziele Benitez - Bibliotecária - CRB-1/3129

Informações sobre a Wissen Editora
Homepage: www.editorawissen.com.br
Teresina - Piauí, Brasil
E-mail: wisseneditora@gmail.com

Como citar ABNT: SILVA, L. F. **Inteligência artificial como ferramenta de educação inclusiva na perspectiva da neurodiversidade**. v. 1, Teresina-PI: Wissen Editora, 2026. 279 p. DOI:

Teresina-PI, 2026

SOBRE O ORGANIZADOR

Flávio de Ligório Silva



Licenciado em Matemática pela Universidade Federal de Minas Gerais (2007), mestre em Educação pela Universidade Federal de Minas Gerais (2011) e doutor em Educação pela Universidade Federal de Minas Gerais (2018). Atualmente é professor do ensino básico, técnico e tecnológico (EBTT) do Instituto Federal de Educação Ciência e Tecnologia da Bahia. Ministra disciplinas de matemática na educação profissional de nível médio e disciplinas de matemática elementar, cálculo diferencial e integral, geometria analítica, bem como as que se relacionam com a prática pedagógica no que concerne a Educação Matemática. No que diz respeito à pesquisa, é profícuo investigador, tendo livro e capítulo de livros publicados, artigos em revista e trabalhos apresentados em eventos de renome tanto na área de Educação Matemática quanto na de Representações Sociais. Atualmente desenvolve e orienta trabalhos acadêmicos na área de Educação e Epistemologia da Ciência. Quanto à primeira atribui ênfase nos seguintes temas: ensino de matemática, formação de professores, processo de ensino-aprendizagem, psicologia da educação matemática, representações sociais e afetividade. Na segunda, orienta estudantes de diferentes níveis quanto à Metodologia de Pesquisa, Construção do Projeto de Pesquisa, Redação Técnica de Dissertações e Teses, correção metodológica e gramatical, dentre outras atribuições.

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

APRESENTAÇÃO

A tecnologia, quando observada apenas em sua dimensão técnica, pode afastar os indivíduos. No entanto, quando direcionada pelo potencial educativo, ela tem a capacidade de aproximar e incluir. Este livro nasce da necessidade de compreender e aplicar a Inteligência Artificial não apenas como uma ferramenta computacional, mas como uma ponte para a equidade na educação.

A escola contemporânea lida com o desafio de acolher a pluralidade das mentes. Como a inclusão plena exige metodologias adaptáveis, o foco desta obra é a interseção entre o desenvolvimento das inteligências artificiais e a perspectiva da neurodiversidade. Os textos aqui organizados exploram desde os conceitos fundamentais da IA até as tensões e os novos papéis que ela exige no ambiente escolar. O trajeto culmina na reflexão sobre a aplicação prática, crítica e ética dessa tecnologia para apoiar o aprendizado de pessoas neurodivergentes.

O leitor encontrará nas próximas páginas um convite para repensar o modelo de ensino. O objetivo é oferecer subsídios para que educadores e pesquisadores utilizem a tecnologia a favor da aprendizagem de todos, respeitando as singularidades e promovendo a redução de barreiras educacionais. A inovação ganha o seu verdadeiro sentido quando serve à dignidade e ao acesso ao conhecimento.

Política de Integridade: Em conformidade com os princípios da transparência na atividade acadêmica, declara-se que houve o uso das inteligências artificiais Gemini, Claude e ChatGPT como ferramentas de auxílio na pesquisa, na estruturação e na construção dos textos que compõem este livro. Todas as etapas de geração contaram com supervisão e revisão humana, assegurando a responsabilidade e o rigor da obra.

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

SUMÁRIO

PARTE 1 - O QUE É INTELIGÊNCIA ARTIFICIAL **8**

CAPÍTULO 1: INTELIGÊNCIA ARTIFICIAL: O PROBLEMA DA DEFINIÇÃO **9**

Naftale Dias Santana do Nascimento 9

Verônica Dias Santana do Nascimento 9

Flávio de Ligório Silva 9

CAPÍTULO 2: O QUE É IA PARA A COMPUTAÇÃO: APRESENTANDO A INTELIGÊNCIA ARTIFICIAL ESTREITA, GERAL E GENERATIVA **18**

Flávio de Ligório Silva 18

CAPÍTULO 3: PRINCÍPIOS MATEMÁTICOS DA GERAÇÃO DE TEXTO EM INTELIGÊNCIA ARTIFICIAL **58**

Andrey Nario de Souza Oliveira 58

Flávio de Ligório Silva 58

PARTE 2 - INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO **101**

CAPÍTULO 4: PERSPECTIVA HISTÓRICA DA TECNOLOGIA NA EDUCAÇÃO: DA INVENÇÃO DA ESCRITA À INTELIGÊNCIA ARTIFICIAL **102**

Flávio de Ligório Silva 102

CAPÍTULO 5: INTELIGÊNCIA ARTIFICIAL NA ESCOLA: POSSIBILIDADES, TENSÕES E NOVOS PAPÉIS **130**

Flávio de Ligório Silva 130

PARTE 3 ..... **156**

Educação Inclusiva na Perspectiva da Neurodiversidade **156**

CAPÍTULO 6: FUNDAMENTOS DA NEURODIVERSIDADE E O PARADIGMA DA INCLUSÃO **157**

Flávio de Ligório Silva 157

CAPÍTULO 7: O AMBIENTE ESCOLAR, PATOLOGIZAÇÃO E O MARCO LEGAL **186**

Flávio de Ligório Silva 186

CAPÍTULO 8: USO DA INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA **218**

Flávio de Ligório Silva 218

CAPÍTULO 9: LIMITES, RISCOS E RESPONSABILIDADES: PARA UMA IA EDUCACIONAL CRÍTICA E ÉTICA **249**

Flávio de Ligório Silva 249

Wissen Editora, 2026 | ISBN 978-65-6219-006-9 | DOI: 10.52832/wed.209

PARTE 1

O QUE É INTELIGÊNCIA ARTIFICIAL

CAPÍTULO 1

INTELIGÊNCIA ARTIFICIAL: O PROBLEMA DA DEFINIÇÃO

ARTIFICIAL INTELLIGENCE: THE PROBLEM OF DEFINITION

Naftale Dias Santana do Nascimento
Estudante do curso de Licenciatura em Matemática do Instituto Federal da Bahia-Campus Barreiras.

Verônica Dias Santana do Nascimento
Estudante do curso de Licenciatura em Matemática do Instituto Federal da Bahia-Campus Barreiras.

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1241



Resumo: A Inteligência Artificial (IA), longe de constituir um conceito unívoco, apresenta-se na literatura científica como um campo marcado por definições heterogêneas, que refletem pressupostos teóricos e metodológicos distintos. Na última década (2015–2025), o avanço de técnicas como Aprendizado de Máquina (Machine Learning), Aprendizado Profundo (Deep Learning) e IA Generativa intensificou o debate conceitual, ao mesmo tempo em que evidenciou o predomínio de sistemas de Inteligência Artificial Estreita (ANI — *Artificial Narrow Intelligence*). O presente estudo analisa como três autores, quais sejam, Pei Wang (2019), David B. Fogel (2023) e Johannes Dahlke (2024), conceituam a Inteligência Artificial, destacando convergências, divergências e os desafios inerentes à formulação de uma definição unificada. A análise articula dimensões técnicas, cognitivas e sociotécnicas, evidenciando que conceituar IA é, simultaneamente, um exercício epistemológico e uma escolha normativa com implicações práticas para a pesquisa e a regulação.

Palavras-chave: Inteligência Artificial. Conceituação de IA. Definição de IA. IAG. ANI.

Abstract: Rather than constituting a univocal concept, Artificial Intelligence (AI) appears in the scientific literature as a field characterized by heterogeneous definitions reflecting distinct theoretical and methodological assumptions. Over the past decade (2015–2025), advances in Machine Learning, Deep Learning, and Generative AI have intensified the conceptual debate while highlighting the predominance of Artificial Narrow Intelligence (ANI) systems. This study examines how three authors — Pei Wang (2019), David B. Fogel (2023), and Johannes Dahlke (2024) — conceptualize Artificial Intelligence, identifying convergences, divergences, and the inherent challenges of formulating a unified definition. The analysis integrates technical, cognitive, and sociotechnical dimensions, demonstrating that defining AI is simultaneously an epistemological exercise and a normative choice with practical implications for research and regulation.

Keywords: Artificial Intelligence. AI conceptualization. AI definition. AGI. ANI.

1 INTRODUÇÃO

Atualmente, os conceitos apresentados na literatura científica acerca do que é Inteligência Artificial (IA) são numerosos e heterogêneos. Para ilustrar essa dispersão, basta observar que a IA pode ser definida como "a capacidade de um sistema de interpretar corretamente dados externos, aprender a partir desses dados e usar esses aprendizados para alcançar objetivos e tarefas específicas por meio de adaptação flexível" (Kaplan & Haenlein, 2019, p. 17), ou, alternativamente, como "um ramo da ciência da computação dedicado à construção de agentes racionais" (Russell & Norvig, 2021, p. 1), ou ainda como um artefato sociotécnico cujo significado é moldado pelos contextos sociais de uso (Dahlke, 2024, p. 5).

Essa multiplicidade conceitual decorre do caráter interdisciplinar do campo da IA, que envolve contribuições da ciência da computação, da filosofia, das ciências cognitivas e das ciências sociais. Wang (2019, p. 1) destaca que diferentes definições de Inteligência Artificial refletem diferentes pressupostos teóricos e, consequentemente, conduzem a agendas de pesquisa distintas. De forma semelhante, Fogel (2023, p. 2) argumenta que a ausência de uma

definição consensual resulta na coexistência de múltiplas abordagens conceituais, cada uma enfatizando aspectos específicos do fenômeno.

Além de numerosos, esses conceitos mostram-se difusos entre si, uma vez que utilizam critérios distintos para delimitar o que pode ser considerado Inteligência Artificial. Fogel (2023, p. 3) classifica as definições existentes em quatro tipos: (i) definições baseadas em capacidade, que enfatizam o que o sistema consegue fazer (por exemplo, reconhecer padrões ou traduzir idiomas); (ii) definições baseadas em processo, que descrevem como o sistema opera internamente (por exemplo, por meio de redes neurais ou algoritmos de busca); (iii) definições baseadas em desempenho, que avaliam quão bem o sistema executa uma tarefa em comparação com o desempenho humano; e (iv) definições sociotécnicas, que consideram o contexto social e o impacto humano dos sistemas. Dahlke (2024, p. 4), por sua vez, argumenta que essa difusão é intensificada quando definições puramente técnicas desconsideram o contexto social no qual os sistemas de IA são desenvolvidos e utilizados, contribuindo para ambiguidades conceituais persistentes.

Conceituar a Inteligência Artificial, portanto, implica lidar com um campo vasto que carrega consigo múltiplos vieses teóricos e metodológicos. Entre esses vieses, destacam-se: o viés computacionalista, que reduz a inteligência a processos algorítmicos e desconsidera aspectos fenomenológicos da cognição; o viés antropocêntrico, que toma a inteligência humana como parâmetro único de comparação, desconsiderando formas de inteligência não biológica; e o viés tecnocentrista, que prioriza a eficiência técnica em detrimento das implicações sociais, éticas e políticas dos sistemas (Boden, 2016, p. 12; Floridi, 2023, p. 45).

Wang (2019, p. 2) afirma que toda definição de IA é, inevitavelmente, uma escolha normativa, pois determina quais propriedades da inteligência serão consideradas centrais e quais serão marginalizadas. Nesse sentido, o autor ressalta que definições excessivamente restritivas podem limitar o avanço científico, enquanto definições excessivamente amplas podem comprometer a precisão conceitual necessária à pesquisa produtiva.

Essas dificuldades são ampliadas pelo fato de a Inteligência Artificial constituir uma das áreas mais dinâmicas da ciência contemporânea. O rápido desenvolvimento de técnicas como Aprendizado de Máquina e Aprendizado Profundo (Goodfellow, Bengio & Courville, 2016) tem provocado revisões constantes nas formas de compreender e definir a IA. Segundo Fogel (2023, p. 5), o avanço tecnológico frequentemente supera os modelos conceituais existentes, exigindo definições mais flexíveis, capazes de acompanhar a evolução das capacidades computacionais sem perder rigor científico.

Além dos avanços tecnológicos, o impacto crescente das ferramentas de IA na sociedade desempenha papel central na complexificação do debate conceitual. Dahlke (2024, p. 6) sustenta que a Inteligência Artificial deve ser compreendida como um fenômeno sociotécnico, no qual algoritmos, propósitos humanos e expectativas sociais se entrelaçam. Dessa forma, a IA não pode ser reduzida a mecanismos de autoaprendizado ou desempenho técnico, pois seu significado é moldado pelas práticas sociais e pelos contextos institucionais em que é aplicada.

Cabe aqui uma reflexão de natureza epistemológica: é legítimo denominar de "inteligência" aquilo que os sistemas de IA realizam? A cognição humana é um fenômeno biológico mediado por processos neuronais e neurotransmissores, sustentado por experiências corporificadas e situadas no mundo (Varela, Thompson & Rosch, 2016, p. 9). Os sistemas de IA, por outro lado, operam por meio de operações matemáticas — sobretudo cálculos estatísticos e predições probabilísticas — programadas por seres humanos. Nesse sentido, a IA não "pensa" no sentido fenomenológico do termo; antes, reproduz e expande padrões do conhecimento humano por meio de capacidades computacionais amplificadas. A arquitetura Transformer, por exemplo, base dos grandes modelos de linguagem contemporâneos, funciona essencialmente por mecanismos de atenção que calculam pesos de relevância entre tokens em uma sequência (Vaswani et al., 2017, p. 1), sem qualquer processo cognitivo ou intencional subjacente. Assim, a "inteligência" dos sistemas de IA é, em sentido estrito, uma metáfora: trata-se de uma inteligência funcional, operacional, mas não de uma inteligência no sentido amplo e fenomenológico atribuído ao pensamento humano. Essa distinção tem implicações importantes para a conceituação, pois revela que parte da confusão conceitual decorre justamente da polissemia do termo "inteligência" quando aplicado a máquinas.

Essa polissemia é amplificada pelo modo como a sociedade se apropria da IA enquanto objeto social. Sob a perspectiva da Teoria das Representações Sociais (Moscovici, 2003, p. 40), a Inteligência Artificial pode ser compreendida como um objeto estranho que, ao ser incorporado ao senso comum, passa por processos de ancoragem, isto é, é associado a categorias preexistentes, como "cérebro", "pensamento" ou "raciocínio", e de objetivação, quando adquire materialidade por meio de metáforas e imagens concretas, como a ideia de um "robô pensante" ou de uma "mente digital". Esse processo explica, em grande medida, a tendência à antropomorfização dos sistemas de IA e a facilidade com que o público geral os compara a capacidades cognitivas humanas, mesmo quando seus mecanismos subjacentes são puramente computacionais. A representação social da IA, portanto, não é neutra: ela influencia tanto a percepção pública quanto os rumos do debate conceitual e regulatório.

Diante desse cenário, este trabalho analisa as conceituações de Inteligência Artificial propostas por Pei Wang (2019), Johannes Dahlke (2024) e David B. Fogel (2023), por representarem abordagens complementares ao problema conceitual. Wang (2019, p. 3) oferece uma perspectiva teórica voltada à redefinição da inteligência como adaptação sob condições de incerteza; Dahlke (2024, p. 7) enfatiza a dimensão sociotécnica da IA e seus vínculos com o contexto humano e social; e Fogel (2023, p. 4) propõe uma definição operacional baseada na combinação entre capacidade e desempenho mensurável. A análise conjunta desses autores permite compreender como diferentes enfoques contribuem para moldar a conceituação da Inteligência Artificial na última década.

2 A PERSPECTIVA ANALÍTICA DE PEI WANG

Wang (2019, p. 2–3) analisa criticamente diversas definições de Inteligência Artificial presentes na literatura, argumentando que a definição adotada influencia diretamente a direção e a produtividade da pesquisa científica. Para o autor, uma boa definição de IA deve atender a quatro critérios fundamentais: (i) manter similaridade com o uso comum do termo "inteligência", evitando distanciamentos excessivos da compreensão intuitiva; (ii) possuir fronteiras conceituais bem demarcadas, de modo a distinguir claramente o que é e o que não é Inteligência Artificial; (iii) ser capaz de orientar e estimular pesquisas produtivas, funcionando como base teórica operacional; e (iv) prezar pela simplicidade, evitando complexidades desnecessárias que dificultem sua aplicação científica.

A partir desses critérios, a proposta de Wang (2019) pode ser compreendida como uma ferramenta analítica para avaliar criticamente outras definições de Inteligência Artificial. Em vez de apresentar apenas uma nova definição, o autor oferece um conjunto de parâmetros conceituais que permite examinar os limites e as potencialidades das concepções propostas por diferentes autores. Nesse sentido, os critérios de Wang (2019) serão utilizados neste trabalho como base crítica para a análise das abordagens de Dahlke (2024) e Fogel (2023), possibilitando comparar em que medida suas definições atendem às exigências de clareza conceitual, utilidade científica e aderência à realidade contemporânea da Inteligência Artificial.

3 A ABORDAGEM SOCIOTÉCNICA DE JOHANNES DAHLKE

Johannes Dahlke, em seu artigo *A inteligência artificial é conhecida por muitos nomes: rumo a uma definição sociotécnica de inteligência artificial* (2024), argumenta que grande parte das definições

contemporâneas de Inteligência Artificial concentra-se predominantemente em aspectos técnicos, como algoritmos, métodos de classificação e técnicas de otimização, negligenciando os propósitos humanos e o contexto social no qual esses sistemas são desenvolvidos e utilizados (Dahlke, 2024, p. 3). Para o autor, essa abordagem reducionista limita a compreensão da IA, ao tratá-la exclusivamente como um artefato computacional desvinculado de suas implicações sociais.

Como alternativa, Dahlke (2024) propõe uma definição sociotécnica de Inteligência Artificial estruturada em três componentes centrais. O primeiro componente refere-se às funções técnicas da IA, isto é, aos mecanismos computacionais que permitem aos sistemas executar tarefas como reconhecimento de padrões, tomada de decisão automatizada e aprendizagem a partir de dados. Esse elemento corresponde à dimensão técnica tradicionalmente enfatizada na literatura da área (Dahlke, 2024, p. 5).

O segundo componente diz respeito aos propósitos humanos associados ao uso da IA, ressaltando que tais sistemas são concebidos, implementados e utilizados para atender a objetivos humanos específicos, sejam eles econômicos, científicos, políticos ou sociais. Dessa forma, a Inteligência Artificial não pode ser compreendida isoladamente de suas finalidades práticas, uma vez que seu desenvolvimento está intrinsecamente ligado às intenções e escolhas humanas (Dahlke, 2024, p. 6).

O terceiro componente envolve as expectativas dinâmicas da sociedade em relação à IA, que incluem percepções públicas, expectativas de desempenho, preocupações éticas e impactos sociais. Segundo o autor, essas expectativas influenciam diretamente tanto o modo como a IA é interpretada quanto os rumos de seu desenvolvimento, reforçando o caráter histórico e socialmente situado do conceito (Dahlke, 2024, p. 7).

A partir dessa perspectiva, Dahlke (2024) sustenta que a Inteligência Artificial não deve ser compreendida apenas como um sistema de autoaprendizagem ou raciocínio automatizado, mas como um artefato sociotécnico, cujo significado é moldado pelas formas como é utilizada, interpretada e regulada pela sociedade. Tal abordagem torna-se especialmente relevante ao considerar que a maioria das IAs atualmente em uso corresponde ao que a literatura denomina Inteligência Artificial Estreita ou Fraca (ANI), ou seja, sistemas projetados e treinados para executar tarefas específicas, sem capacidades cognitivas gerais (Dahlke, 2024, p. 9).

Nesse sentido, o autor evidencia que o contexto social e o propósito humano não apenas influenciam a aplicação da IA, mas interferem diretamente em sua própria definição conceitual, reforçando a ideia de que conceituar Inteligência Artificial implica considerar, simultaneamente, dimensões técnicas, humanas e sociais.

4 A SÍNTESE OPERACIONAL DE DAVID B. FOGEL

David B. Fogel, em *Defining Artificial Intelligence* (2023), realiza uma meta-análise das definições de Inteligência Artificial presentes na literatura, classificando-as em quatro categorias: (i) definições baseadas em capacidade, que enfatizam o que o sistema consegue fazer; (ii) definições baseadas em processo, que descrevem como o sistema opera internamente; (iii) definições baseadas em desempenho, que avaliam quão bem o sistema executa determinadas tarefas; e (iv) definições sociotécnicas, que consideram o contexto e o impacto social dos sistemas (Fogel, 2023, p. 3–4).

A partir dessa taxonomia, Fogel (2023, p. 8) propõe uma definição que combina os eixos de capacidade e desempenho, concebendo a IA como sistemas computacionais que alcançam desempenho comparável ou superior ao humano em tarefas cognitivas específicas, quando avaliados por métricas relevantes. Essa abordagem destaca-se por seu caráter pragmático e operacional, oferecendo critérios funcionais que permitem avaliar e comparar sistemas de forma objetiva, tornando-se adequada tanto para o uso científico quanto para aplicações industriais e regulatórias.

5 CONCLUSÃO

A análise das conceituações de Inteligência Artificial propostas por Pei Wang (2019), Johannes Dahlke (2024) e David B. Fogel (2023) evidencia que a IA não pode ser compreendida a partir de uma definição única, estática ou universal. Ao contrário, trata-se de um conceito em permanente construção, cuja compreensão depende do enquadramento teórico, do propósito analítico e do contexto histórico e social no qual é mobilizado. Essa pluralidade não representa uma fragilidade conceitual do campo, mas antes reflete a própria complexidade da Inteligência Artificial enquanto fenômeno técnico, cognitivo e sociotécnico.

A contribuição de Wang (2019) destaca-se ao oferecer critérios conceituais rigorosos que funcionam como uma lente crítica para avaliar definições de IA, ao enfatizar a necessidade de fronteiras claras, simplicidade conceitual e utilidade científica. Sua abordagem chama atenção para o fato de que definir Inteligência Artificial não é apenas um exercício terminológico, mas uma escolha epistemológica que orienta práticas de pesquisa e desenvolvimento (Wang, 2019, p. 3). Nesse sentido, Wang (2019) reforça que toda definição carrega implicações teóricas e metodológicas.

Dahlke (2024), por sua vez, amplia o debate ao deslocar a discussão da IA para além do domínio estritamente técnico, ao compreendê-la como um artefato sociotécnico, indissociável dos propósitos humanos e das expectativas sociais. Sua proposta evidencia que a Inteligência Artificial adquire significado não apenas por aquilo que tecnicamente executa, mas pelas formas como é utilizada, interpretada e regulada na sociedade, especialmente em um cenário marcado pelo predomínio de sistemas de Inteligência Artificial Estreita (ANI) (Dahlke, 2024, p. 7).

Já Fogel (2023) oferece uma síntese pragmática e operacional, ao classificar as definições existentes e propor uma conceituação orientada ao desempenho mensurável dos sistemas em tarefas cognitivas específicas. Sua abordagem contribui para a aplicação científica, industrial e regulatória da IA, ao fornecer critérios funcionais que permitem avaliar e comparar sistemas de forma objetiva, sem ignorar suas limitações conceituais (Fogel, 2023, p. 12).

Dessa forma, a articulação entre os três autores permite compreender que conceituar Inteligência Artificial implica reconhecer simultaneamente seus fundamentos técnicos, seus limites cognitivos e suas implicações humanas e sociais. Conclui-se, portanto, que a definição de Inteligência Artificial não é um fim em si mesma, mas um instrumento analítico que deve ser continuamente revisitado, à medida que as tecnologias evoluem e seu impacto sobre a sociedade se intensifica.

Nesse sentido, a própria dificuldade de conceituar a IA pode ser compreendida, à luz da Teoria das Representações Sociais, como um reflexo do processo de familiarização coletiva com um objeto que desafia categorias preestabelecidas. A tendência a antropomorfizar os sistemas de IA, atribuindo-lhes sentidos e significados associados a "inteligência", "compreensão" ou "criatividade", evidencia que o debate conceitual não é apenas técnico ou filosófico, mas profundamente social e simbólico. Reconhecer essa dimensão representacional é fundamental para que futuras definições de Inteligência Artificial integrem, de modo equilibrado, as contribuições da ciência da computação, da filosofia da mente e das ciências sociais.

Este capítulo contou com o auxílio de ferramenta de Inteligência Artificial generativa para apoio à redação, revisão textual e formatação, sob direção intelectual e supervisão dos autores, que assumem integral responsabilidade pelo conteúdo apresentado.

REFERÊNCIAS

BODEN, Margaret A. **AI:** its nature and future. Oxford: Oxford University Press, 2016.

DAHLKE, Johannes. **A inteligência artificial é conhecida por muitos nomes:** rumo a uma definição sociotécnica de inteligência artificial. AI & Society, Londres, v. 39, n. 2, p. 1–14, 2024.

FLORIDI, Luciano. **The ethics of artificial intelligence:** principles, challenges, and opportunities. Oxford: Oxford University Press, 2023.

FOGEL, David B. **Defining artificial intelligence.** IEEE Computational Intelligence Magazine, Piscataway, v. 18, n. 1, p. 1–15, 2023.

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning.** Cambridge: MIT Press, 2016.

KAPLAN, Andreas; HAENLEIN, Michael. Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. **Business Horizons**, v. 62, n. 1, p. 15–25, 2019.

MOSCOVICI, Serge. **Representações sociais:** investigações em psicologia social. Petrópolis: Vozes, 2003.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence:** a modern approach. 4. ed. Hoboken: Pearson, 2021.

VARELA, Francisco J.; THOMPSON, Evan; ROSCH, Eleanor. **The embodied mind: cognitive science and human experience.** Ed. rev. Cambridge: MIT Press, 2016.

VASWANI, Ashish et al. Attention is all you need. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS (NIPS), 31., 2017, Long Beach. **Proceedings** [...]. Long Beach: NIPS, 2017. p. 1–15.

WANG, Pei. On defining artificial intelligence. **Journal of Artificial General Intelligence**, Paris, v. 10, n. 2, p. 1–37, 2019.

CAPÍTULO 2

O QUE É IA PARA A COMPUTAÇÃO: APRESENTANDO A INTELIGÊNCIA ARTIFICIAL ESTREITA, GERAL E GENERATIVA

WHAT IS AI FOR COMPUTING: INTRODUCING NARROW, GENERAL AND GENERATIVE ARTIFICIAL INTELLIGENCE

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia.

DOI: 10.52832/wed.209.1242



Resumo: O presente capítulo aborda o conceito de Inteligência Artificial sob a ótica da ciência da computação, adotando uma taxonomia técnica rigorosa. O texto propõe a classificação da tecnologia em três categorias operacionais: a Inteligência Artificial Estreita (ANI), a Inteligência Artificial Geral (AGI) e a Inteligência Artificial Generativa. A primeira é apresentada como o paradigma funcional predominante no cenário atual, sendo projetada para executar tarefas específicas dentro de domínios delimitados. A segunda representa um horizonte teórico, ainda inatingível, que seria caracterizado pela capacidade de transferir conhecimento entre diferentes domínios e operar com uma flexibilidade cognitiva semelhante à humana. A terceira categoria é conceituada como um subconjunto da ANI, sendo especializada na produção de novos conteúdos, como textos, imagens e códigos computacionais, a partir de padrões estatísticos extraídos de grandes volumes de dados. Ademais, o trabalho ressalta a fundamentação matemática intrínseca que sustenta esses sistemas e argumenta em favor da necessidade de um letramento em inteligência artificial. Essa competência é destacada como um elemento essencial para a formação docente, permitindo que os professores compreendam os limites, os mecanismos de funcionamento e as implicações dessas ferramentas em suas práticas pedagógicas.

Palavras-chave: Inteligência Artificial. Inteligência Artificial Estreita. Inteligência Artificial Geral. Inteligência Artificial Generativa. Letramento em inteligência artificial

Abstract: This chapter addresses the concept of Artificial Intelligence from the perspective of computer science, adopting a rigorous technical taxonomy. The text proposes classifying the technology into three operational categories: Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Generative Artificial Intelligence. The first is presented as the predominant functional paradigm in the current scenario, designed to perform specific tasks within delimited domains. The second represents a theoretical horizon, still unattainable, which would be characterized by the ability to transfer knowledge between different domains and operate with cognitive flexibility like human intelligence. The third category is conceptualized as a subset of ANI, specialized in the production of new content, such as texts, images, and computer codes, based on statistical patterns extracted from large volumes of data. Furthermore, the work highlights the intrinsic mathematical foundation that supports these systems and argues in favor of the need for artificial intelligence literacy. This competence is highlighted as an essential element for teacher training, allowing educators to understand the limits, the functioning mechanisms, and the implications of these tools in their pedagogical practices.

Keywords: Artificial Intelligence. Artificial Narrow Intelligence. Artificial General Intelligence. Generative Artificial Intelligence. Artificial intelligence literacy

1 INTRODUÇÃO

A expressão Inteligência Artificial é empregada com frequência crescente tanto na literatura científica quanto no discurso público, mas seu significado varia consideravelmente conforme o campo disciplinar e o propósito de quem a utiliza. Filósofos, cientistas sociais e especialistas em regulação tendem a enfatizar as implicações epistemológicas e sociais do conceito; engenheiros e cientistas da computação, por sua vez, privilegiam uma abordagem

funcional, orientada pelo que os sistemas efetivamente realizam. Essa pluralidade conceitual, amplamente documentada na literatura (Wang, 2019; Fogel, 2023; Dahlke, 2024), não constitui um obstáculo, mas um ponto de partida necessário para qualquer discussão rigorosa sobre o tema, pois permite entender a IA não como uma entidade única, mas como um ecossistema de técnicas e paradigmas.

Nesse contexto, Lee (2019, p. 19) oferece uma metáfora poderosa ao comparar a Inteligência Artificial à descoberta da eletricidade. Assim como a rede elétrica transformou radicalmente a infraestrutura física e industrial do século passado, tornando-se um substrato invisível e onipresente, a IA está se tornando a nova eletricidade do século XXI, reconfigurando a infraestrutura cognitiva da sociedade. Essa percepção reforça que a IA não é apenas uma ferramenta isolada, mas uma força transversal que exige uma nova forma de compreensão por parte de todos os atores sociais, especialmente os educadores.

O presente capítulo situa-se no segundo desses registros. Seu objetivo é apresentar a Inteligência Artificial tal como a ciência da computação a compreende e a classifica em termos operacionais, sem pretender esgotar as dimensões filosóficas ou sociotécnicas do debate. Parte-se aqui de uma constatação pragmática: independentemente das controvérsias sobre o que a IA é, existe hoje um conjunto consolidado de sistemas computacionais que são reconhecidos como Inteligência Artificial pela comunidade científica e pela indústria, e que podem ser diferenciados com base em suas capacidades, em seu escopo de atuação e no tipo de saída que produzem.

A partir dessa perspectiva, o capítulo organiza-se em torno de três categorias amplamente utilizadas na literatura da área: a Inteligência Artificial Estreita (ANI, do inglês *Artificial Narrow Intelligence*), a Inteligência Artificial Geral (AGI, do inglês *Artificial General Intelligence*) e a Inteligência Artificial Generativa. A primeira designa sistemas projetados para executar tarefas específicas dentro de domínios delimitados; a segunda refere-se a um horizonte teórico, ainda não alcançado, de sistemas com capacidades cognitivas generalizáveis; e a terceira descreve uma classe de sistemas capazes de produzir novos conteúdos, como texto, imagem e código, a partir de padrões extraídos de grandes volumes de dados (Russell & Norvig, 2021; Goodfellow, Bengio & Courville, 2016).

Importa esclarecer, desde já, duas delimitações. Em primeiro lugar, este capítulo não aborda os mecanismos computacionais internos desses sistemas. Questões como arquiteturas de redes neurais, modelos de atenção ou processos de otimização matemática são tratadas em outra parte desta obra. O foco aqui é conceitual e classificatório: interessa saber o que cada tipo de IA faz e em que se diferencia dos demais, não como faz do ponto de vista algorítmico.

Em segundo lugar, a classificação proposta não pretende ser exaustiva nem definitiva. Como toda taxonomia aplicada a um campo em rápida transformação, ela é um instrumento analítico, útil para organizar o debate, mas sujeito a revisões à medida que novas capacidades técnicas emergem e os limites entre as categorias se tornam mais porosos.

Antes de prosseguir, porém, convém explicitar o lugar de onde este texto fala e a quem se destina.

É provável que a maioria dos estudantes de cursos de licenciatura, seja em Matemática, em Química, em Pedagogia ou em qualquer outra área, utilize hoje alguma ferramenta de Inteligência Artificial para resolver exercícios, resumir textos, gerar exemplos ou até mesmo redigir trabalhos acadêmicos. É igualmente provável que esses mesmos estudantes não saibam definir o que é Inteligência Artificial, como ela funciona ou quais são seus limites. Utiliza-se a ferramenta com naturalidade, mas se desconhece o que a faz funcionar. Esse cenário adquire especial gravidade quando se trata de futuros professores: profissionais que, em breve, precisarão não apenas utilizar a IA em suas práticas pedagógicas, mas também orientar seus alunos sobre o que essas ferramentas fazem, o que não fazem e por que, em determinadas circunstâncias, produzem resultados incorretos ou enganosos.

Ensinar a usar IA sem compreender minimamente seus fundamentos é análogo a ensinar a usar uma calculadora sem compreender aritmética: o instrumento funciona, mas o professor não tem condições de avaliar criticamente seus resultados nem de explicar suas limitações. Long e Magerko (2020, p. 2) denominam essa capacidade de letramento em Inteligência Artificial (*AI literacy*), definindo-a como um conjunto de competências que permite aos indivíduos avaliar criticamente tecnologias de IA, comunicar-se e colaborar efetivamente com elas, e utilizá-las como ferramentas em contextos diversos, inclusive educacionais e formativos.

A análise de Lee (2019, p. 175) complementa essa necessidade de letramento ao propor o conceito de *Blueprint* Humano, que diferencia a capacidade de otimização da máquina da capacidade de empatia humana. Enquanto os sistemas de IA são imbatíveis em tarefas de otimização, como processar bilhões de dados para encontrar o menor caminho ou resolver equações complexas em milissegundos, eles são intrinsecamente incapazes de exercer empatia, compaixão ou estabelecer conexões emocionais profundas. O paradoxo da formação docente reside justamente aqui: o professor não deve competir com a capacidade de otimização da IA, mas sim desenvolver o discernimento crítico para usar essa potência computacional em favor de uma prática pedagógica mais humanizada e criativa, centrada naquilo que é exclusivo da condição humana.

Há, nesse cenário, um paradoxo que merece atenção particular no campo da formação de professores de Matemática. Futuros docentes dessa área operam cotidianamente uma tecnologia cujos fundamentos repousam, em grande medida, sobre disciplinas que eles próprios estudam ao longo de sua formação: Otimização Matemática, Cálculo Diferencial e Integral, Geometria Diferencial, Topologia, Álgebra Linear, Cálculo Vetorial, Análise Combinatória, Probabilidade, Análise Real, Estatística e Lógica.

Sem que percebam, esses campos do conhecimento matemático são justamente o substrato que torna a Inteligência Artificial possível. Vetores deixam de ser apenas objetos geométricos em $\mathbb{R}^n$ e passam a representar palavras em espaços semânticos de centenas de dimensões. Multiplicações de matrizes deixam de ser exercícios de sala de aula e passam a ser o mecanismo pelo qual um modelo de linguagem calcula relações de relevância entre palavras para decidir qual será a próxima em uma frase. Distribuições de probabilidade deixam de ser curvas em um gráfico e passam a governar o processo pelo qual uma imagem é gerada progressivamente a partir de ruído. A análise combinatória, que o estudante conhece como técnica de contagem, é a mesma lógica que sustenta a construção de vocabulários de *tokens* e a exploração do espaço de sequências possíveis durante a geração de texto. A função *softmax*, que transforma um vetor de valores reais em uma distribuição de probabilidades, é uma aplicação direta de conceitos que todo estudante de Matemática reconheceria se os visse fora de seu contexto habitual.

Esse reconhecimento é importante não apenas para os licenciandos em Matemática, mas para a formação docente em geral. A Inteligência Artificial não é um fenômeno que pertence exclusivamente à Ciência da Computação; ela é construída sobre uma base matemática que, por sua vez, permeia toda a educação científica. Quando um professor de Química utiliza uma IA para prever propriedades moleculares, quando um pedagogo emprega um sistema de tutoria inteligente para personalizar o ensino, quando um professor de Biologia recorre a modelos de IA para analisar sequências genéticas, todos eles estão interagindo com sistemas cuja lógica interna é, em última instância, matemática.

Compreender isso não exige que cada professor se torne um programador ou um engenheiro de aprendizado de máquina. Exige, contudo, que a comunidade educacional em seu conjunto desenvolva aquilo que Long e Magerko (2020, p. 4) chamam de capacidade de avaliar criticamente as tecnologias de IA, o que pressupõe uma literacia mínima sobre os princípios que as governam, de modo a utilizá-las com discernimento e a ensiná-las, ou ao menos a contextualizá-las, com responsabilidade. Como observam Holmes, Bialik e Fadel (2019, p. 9), a introdução da Inteligência Artificial na educação não representa apenas uma

mudança de ferramentas, mas uma reconfiguração dos papéis de professores e estudantes, na qual o docente precisa passar de consumidor passivo de tecnologia a mediador crítico de suas possibilidades e limitações.

Este capítulo nasce dessa inquietação. Ele é escrito a partir da perspectiva de um professor de Matemática, formador de outros professores, que reconhece na Inteligência Artificial um objeto que exige compreensão e não apenas uso. Não se trata de um texto de especialista em computação para outros especialistas, mas de um esforço de tradução conceitual: tornar acessíveis, para professores e estudantes de diferentes áreas, distinções e conceitos que normalmente circulam em linguagem técnica restrita. A aposta é que essa tradução não apenas é possível, como é necessária, sobretudo porque a Matemática, longe de ser exterior à IA, é sua linguagem constitutiva. E aquilo que a Matemática fundamenta, as demais áreas aplicam, experimentam e transformam em prática.

Cabe ainda situar este esforço dentro de um debate mais amplo sobre o papel do professor diante das transformações tecnológicas contemporâneas. A Inteligência Artificial não é a primeira tecnologia a provocar reconfigurações no campo educacional. Papert (1980, p. 5), ao propor que o computador poderia servir não apenas como ferramenta de instrução, mas como ambiente no qual a criança constrói conhecimento, já antecipava que a relação entre tecnologia e educação não se esgota no uso instrumental. A calculadora, o computador pessoal e a internet geraram, cada um a seu tempo, ondas de entusiasmo, resistência e adaptação no meio escolar. O que distingue a Inteligência Artificial, entretanto, é a natureza da sua interação com o conhecimento. Enquanto as tecnologias anteriores ampliavam a capacidade de acessar ou processar informação, a IA Generativa produz conteúdo que se assemelha ao produzido por humanos, deslocando a questão central de como encontrar a resposta para como avaliar se a resposta encontrada é confiável (Holmes, Bialik & Fadel, 2019, p. 22). Para o professor, isso representa uma mudança qualitativa: já não basta dominar o conteúdo; é preciso também compreender os processos pelos quais a máquina gera aquilo que parece ser conhecimento, para então discernir onde há rigor e onde há apenas coerência superficial.

É precisamente esse discernimento que o presente capítulo pretende favorecer. Ao apresentar as categorias de Inteligência Artificial Estreita, Geral e Generativa, o objetivo não é apenas classificar sistemas, mas oferecer ao leitor, seja ele professor de Matemática, de Ciências, de Linguagens ou de qualquer outra área, as coordenadas conceituais necessárias para dialogar com uma tecnologia que já habita suas salas de aula. O que se oferece nas páginas seguintes é, portanto, um mapa. Não o território, que é vasto, técnico e em permanente transformação, mas um instrumento de orientação para quem precisa distinguir o que existe

do que é promessa, o que a IA faz do que se imagina que ela faça, e o que a Matemática que já ensinamos tem a ver com a tecnologia que já usamos. Se ao final desta leitura o leitor olhar para uma matriz de transformação linear, para o teorema de Bayes ou para um simples arranjo combinatório com a suspeita de que ali há mais do que sempre pareceu haver, este capítulo terá cumprido seu propósito.

2 A INTELIGÊNCIA ARTIFICIAL NA PERSPECTIVA DA CIÊNCIA DA COMPUTAÇÃO

A compreensão da Inteligência Artificial exige que enfrentemos, primeiramente, o desafio de sua definição dentro do campo da computação. Longe de ser um campo estritamente técnico, a IA é um terreno onde se cruzam a engenharia de precisão e a projeção simbólica do humano sobre a máquina. De acordo com Lee (2019, p. 24), vivemos atualmente na chamada Era da Implementação, um período em que o foco da computação se deslocou da descoberta de novos algoritmos fundamentais para a aplicação massiva e otimizada do que já foi teorizado, como o aprendizado profundo. Para navegar por essa perspectiva prática e operacional, é necessário decompor os pilares que sustentam o que chamamos de "inteligência" e como o qualificativo "artificial" altera nossa percepção sobre o objeto técnico, partindo de uma análise dos fundamentos conceituais até chegarmos aos paradigmas que governam os sistemas modernos.

2.1 O desafio da definição e a IA como representação social

Antes de classificar os tipos de Inteligência Artificial, é preciso enfrentar uma questão aparentemente simples que, no entanto, tem desafiado pesquisadores desde a fundação do campo: afinal, o que a ciência da computação entende por Inteligência Artificial? A pergunta pode parecer redundante após as discussões conceituais já realizadas nesta obra, mas não é. Se no âmbito filosófico e sociológico a dificuldade de definir IA decorre da pluralidade de perspectivas teóricas e da dimensão sociotécnica do conceito, no âmbito da ciência da computação a dificuldade tem uma raiz diferente, embora complementar: ela decorre da ambiguidade dos dois termos que compõem a expressão. O que significa “inteligência”? E o que significa, neste contexto, “artificial”?

Essa indefinição não é uma mera curiosidade acadêmica. Trata-se, na verdade, de um problema civilizatório. A humanidade está desenvolvendo, regulando, financiando e temendo

algo que ainda não conseguiu definir com precisão. A situação guarda uma analogia perturbadora com a Idade Média, quando as pessoas sabiam que certas águas causavam doenças, mas desconheciam a existência de microrganismos. Agiam sobre o efeito sem compreender a causa. No caso da IA, governos aprovam legislações, empresas investem bilhões e sociedades inteiras reorganizam suas economias em torno de uma tecnologia cuja natureza fundamental permanece em disputa. Como regular aquilo que não se consegue nomear? Como estabelecer limites éticos para algo cujas fronteiras conceituais são difusas? A definição, portanto, não é um luxo teórico: é uma condição prévia para qualquer governança responsável.

Comecemos pelo segundo termo da expressão, que costuma passar despercebido. Chamamos de artificial aquilo que é feito por seres humanos, em oposição ao que existe independentemente da ação humana, que chamamos de natural. Um lago é natural; uma represa é artificial. Um diamante formado sob a crosta terrestre é natural; um diamante sintetizado em laboratório é artificial. Em ambos os casos, o objeto artificial pode ter propriedades idênticas ou até superiores às do objeto natural, mas recebe o qualificativo "artificial" porque sua origem é humana.

A Inteligência Artificial, portanto, não é artificial por ser falsa, inferior ou simulada. É artificial porque foi feita por pessoas. O termo, longe de afastar a IA do humano, a vincula radicalmente a ele: só existe Inteligência Artificial porque existem seres humanos capazes de concebê-la, projetá-la e construí-la. Há algo de prometéico nesse empreendimento. Os seres humanos, ao criarem sistemas que exibem comportamentos associados à inteligência, assumem, de certo modo, o papel de criadores de capacidades que até então eram exclusivas dos seres vivos.

É precisamente nesse ponto que a obra de Moscovici (1990) se torna indispensável. Em *A máquina de fazer deuses*, Moscovici (1990, p. 37) demonstra que a sociedade funciona como uma máquina de produção simbólica, movida pelo poder de interinfluência, pelos contratos que se estabelecem, pelas trocas e negociações, sem esquecer a alta e tangível coação exercida sobre o indivíduo que resiste às posições do grupo. Essa "máquina" não produz apenas normas e instituições: produz *deuses* — isto é, entidades simbólicas às quais se atribuem poderes que transcendem a condição material de sua origem. A IA contemporânea pode ser lida, sob essa chave, como a expressão mais recente desse impulso ancestral: a humanidade projeta sobre seus artefatos computacionais atributos de autonomia, criatividade e inteligência que, paradoxalmente, são propriedades humanas transferidas para a máquina.

Moscovici (1990, p. 55) mostra ainda que os ritos sociais — e, por extensão, as práticas culturais em torno da tecnologia — encarnam representações comuns que povoam a consciência coletiva. O discurso público sobre IA, com sua retórica de revolução, de singularidade e de inteligências superiores, opera como um rito contemporâneo: consolida a identidade de um grupo (a comunidade tecnológica global) em torno de uma visão, uma opção e uma "causa" comuns. O rito, como observa Moscovici (1990, p. 55), não se limita à sua prática; é uma fonte de energia, mobilização e união do grupo. Não é por acaso que os lançamentos de novos modelos de IA são recebidos pela mídia e pelo mercado com a solenidade e a expectativa que antes se reservavam a eventos científicos ou, em certa medida, a manifestações religiosas.

A tensão entre o instituído e o instituinte, central na análise de Moscovici (1990, p. 197), também ilumina o campo da IA. De um lado, há as representações instituídas: a IA como ferramenta, como produto comercial, como instrumento de otimização. De outro, há as forças instituintes: os novos conceitos, as novas crenças e os novos paradigmas que a IA introduz e que "revolucionam" o estabelecido. Tradição e renovação confrontam-se, sempre, num relacionamento de tensão — e é exatamente nesse espaço de tensão que se dá o debate regulatório, ético e conceitual sobre a Inteligência Artificial.

Há, contudo, uma contradição que merece ser enfrentada. Chamar essa inteligência de "artificial" sugere, no senso comum, que ela não é humana — que é, de algum modo, alheia a nós. No entanto, ela é *profundamente* humana, justamente porque é fruto do trabalho humano: concebida por mentes humanas, treinada com dados produzidos por sociedades humanas, orientada por objetivos definidos por pessoas e embebida nos vieses, valores e limitações de seus criadores. Moscovici (1990, p. 249) nos ajuda a compreender essa contradição: as redes e as representações sociais constituem uma ordem simbólica influente na vida psíquica e social, cujos conceitos e imagens selecionam motivações, expectativas e sentimentos. A IA não escapa dessa lógica. Ela é uma representação social materializada em código: carrega consigo a visão de mundo de quem a projetou e do contexto em que foi produzida. O "artificial", longe de designar uma alteração radical em relação ao humano, designa uma *projeção* do humano sobre a máquina. A sociedade, ao construir a IA, não está fabricando deuses — está, como sugere Moscovici (1990, p. 375), operando como "máquina de fazer homens": projetando sobre o artefato uma imagem idealizada (ou distorcida) de si mesma.

2.2 Matrizes da inteligência e o paradigma da ação racional

O primeiro termo, "inteligência", é ainda mais problemático. Não existe consenso, nem mesmo nas ciências cognitivas, sobre o que constitui inteligência. Sabemos que seres humanos são inteligentes. Admitimos, com graus variados de convicção, que outros animais também o são: polvos resolvem problemas mecânicos complexos, golfinhos exibem comportamentos cooperativos sofisticados, primatas utilizam ferramentas e transmitem conhecimentos entre gerações. Mas seria isso inteligência ou instinto? E onde se traça a fronteira entre um e outro?

John McCarthy, o cientista da computação que cunhou o próprio termo "inteligência artificial" em 1956, ofereceu uma resposta reveladora a essa questão. Quando perguntado diretamente "o que é inteligência?", McCarthy (2007, p. 2) respondeu que inteligência é a parte computacional da capacidade de atingir objetivos no mundo, e que graus e tipos variados de inteligência ocorrem em pessoas, em muitos animais e em algumas máquinas. A formulação é notável por três razões. Primeiro, por ser operacional: define inteligência não como uma essência metafísica, mas como uma capacidade funcional. Segundo, por ser inclusiva: reconhece inteligência em animais e máquinas, não apenas em humanos. Terceiro, por ser honesta: quando pressionado sobre a possibilidade de uma definição que não dependesse da referência à inteligência humana, McCarthy (2007, p. 2) admitiu que tal definição ainda não existia.

McCarthy (2007, p. 3) também enfrentou outra pergunta fundamental: a IA busca simular a inteligência humana? Sua resposta é igualmente instrutiva: às vezes sim, mas não sempre, e nem sequer na maioria dos casos. Os pesquisadores de IA, argumentou, são livres para usar métodos que não são observados em pessoas ou que envolvem muito mais computação do que seres humanos conseguem realizar. A maior parte do trabalho em IA, segundo ele, concentra-se nos problemas que o mundo apresenta à inteligência, e não no estudo de como pessoas ou animais os resolvem. Essa distinção, como veremos, será central para compreender as quatro perspectivas organizadas por Russell e Norvig (2021).

A dificuldade em definir IA não é um problema resolvido. Como observa o relatório do World Travel & Tourism Council (WTTC, 2024, p. 6), não existe uma definição única e universalmente aceita para Inteligência Artificial. No entanto, a ausência de consenso não deveria justificar a ausência de esforço definitório. Pelo contrário. Quando uma tecnologia começa a afetar profundamente a vida de bilhões de pessoas — em seus empregos, suas decisões de saúde, seus processos judiciais, suas interações sociais —, a falta de uma definição compartilhada torna-se não apenas um problema teórico, mas um risco prático.

Algumas instituições enfrentaram esse desafio de modo direto. A Estratégia Nacional de Inteligência Artificial da Türkiye (República Da Türkiye, 2021, p. 10) oferece uma definição

pragmática que, em sua simplicidade, tem o mérito de estabelecer um ponto de partida: em sua forma mais geral, a inteligência artificial é definida como a capacidade de um computador ou de um robô controlado por computador de realizar diversas tarefas de maneira semelhante a seres inteligentes. O Parlamento Europeu (2023), por sua vez, em documento institucional voltado ao público geral, define a IA como a capacidade de uma máquina de exibir habilidades semelhantes às humanas, como raciocínio, aprendizado, planejamento e criatividade, e acrescenta que os sistemas de IA são capazes de adaptar seu comportamento em certo grau, analisando os efeitos de ações anteriores e operando de forma autônoma.

Essa última definição, embora didática, já revela uma tensão importante: ao mesmo tempo em que ancora a IA na referência ao humano ("human-like capabilities"), reconhece que esses sistemas possuem um grau de autonomia comportamental que os distingue de ferramentas convencionais. Não se trata mais de uma calculadora sofisticada, mas de algo que "adapta seu comportamento" — uma descrição que, por si só, já exige marcos regulatórios específicos.

Russell e Norvig (2021, p. 1), autores do manual mais influente da área, enfrentam essa dificuldade de modo mais sistemático. A definição que oferecem do campo já carrega a tensão entre capacidade técnica e propósito humano: a inteligência artificial, escrevem, não se limita a compreender, mas também a construir entidades inteligentes — máquinas capazes de computar como agir de forma eficaz e segura em uma ampla variedade de situações novas. Note-se que a definição não menciona apenas eficácia, mas também segurança, e não se limita a tarefas específicas, mas se refere a situações novas e variadas. A amplitude e a cautela da formulação já indicam que, mesmo no interior da ciência da computação, a definição de IA não é um problema meramente técnico.

Em vez de oferecer uma definição única e categórica, Russell e Norvig (2021, p. 2) organizam as diferentes abordagens históricas ao problema em torno de duas dimensões que se cruzam. A primeira dimensão diz respeito ao padrão de referência: algumas abordagens tomam o ser humano como modelo; outras tomam a racionalidade como modelo. A segunda dimensão diz respeito ao foco de análise: algumas concentram-se nos processos internos do sistema (como ele pensa ou raciocina); outras concentram-se no comportamento externo (o que ele faz). Os autores observam que as implicações metodológicas são profundas: a busca por inteligência semelhante à humana é, em parte, uma ciência empírica ligada à psicologia, ao passo que a abordagem racionalista envolve uma combinação de matemática e engenharia, com conexões à estatística, à teoria de controle e à economia.

Do cruzamento dessas duas dimensões resultam quatro perspectivas sobre a Inteligência Artificial, que correspondem também à tipologia apresentada por Kurt (2021) em sua introdução aos conceitos fundamentais de IA.

A primeira perspectiva tem como referência emblemática o Teste de Turing, proposto por Alan Turing em 1950. Nessa abordagem, um sistema é considerado inteligente se for capaz de se comportar de maneira indistinguível de um ser humano em uma interação mediada por linguagem. Para atender a esse critério, o sistema precisaria dominar, entre outras capacidades, o processamento de linguagem natural, a representação de conhecimento, o raciocínio automatizado e o aprendizado de máquina (Russell; Norvig, 2021, p. 2).

McCarthy (2007, p. 4), contudo, oferece um contraponto sutil: o Teste de Turing é um teste unilateral. Uma máquina que passa nesse teste deveria certamente ser considerada inteligente, mas uma máquina ainda poderia ser inteligente sem saber o suficiente sobre humanos para imitá-los. A observação é pertinente: ela sugere que a inteligência não se reduz a uma performance de humanidade. Os próprios Russell e Norvig (2021, p. 3) recorrem a uma analogia esclarecedora: a busca pelo voo artificial teve sucesso quando engenheiros pararam de imitar pássaros e passaram a estudar aerodinâmica. No contexto contemporâneo, o relatório do WTTC (2024, p. 8) registra que, no início de 2024, nenhum computador havia passado de forma decisiva no Teste de Turing, embora *chatbots* modernos baseados em IA já se mostrassem muito difíceis de distinguir de interlocutores humanos.

A segunda perspectiva desloca o foco do comportamento externo para os processos cognitivos internos. Nessa abordagem, o objetivo não é apenas que a máquina produza respostas corretas, mas que as produza pelos mesmos mecanismos que o cérebro humano utiliza. Russell e Norvig (2021, p. 3) observam que essa abordagem exige conhecimento empírico sobre como os seres humanos de fato pensam, obtido por meio de introspecção, experimentos psicológicos e neuroimagem. Os autores citam como exemplo paradigmático o trabalho de Allen Newell e Herbert Simon com o GPS (*General Problem Solver*), no qual os pesquisadores comparavam a sequência e o tempo dos passos de raciocínio do programa com os de sujeitos humanos resolvendo os mesmos problemas.

A terceira perspectiva abandona o ser humano como referência e adota a lógica formal como padrão. Nessa abordagem, a inteligência consiste em realizar inferências corretas a partir de premissas dadas. A tradição remonta a Aristóteles e seus silogismos e se estende até a lógica de primeira ordem desenvolvida no século XIX. O problema, como observam Russell e Norvig (2021, p. 4), é duplo. Em primeiro lugar, a lógica clássica opera com certezas, e o mundo real é marcado por incerteza. Em segundo lugar, pensar corretamente, por si só, não

basta para agir de forma inteligente: a lógica permite construir um modelo de pensamento racional, mas não gera, por si mesma, comportamento inteligente. Para isso, é necessária uma teoria da ação racional.

McCarthy (2007, p. 5), em seu próprio trabalho, dedicou especial atenção a esse problema. Ele identificou a IA lógica como um ramo central do campo, no qual o conhecimento sobre o mundo, os fatos de uma situação específica e os objetivos do agente são todos representados em linguagem lógico-matemática, e o programa decide o que fazer inferindo que certas ações são apropriadas para atingir seus objetivos. Mas, como McCarthy (2007, p. 6) reconheceu, o grande desafio permanece: o senso comum — aquele conhecimento tácito que os humanos mobilizam sem esforço — é a área em que a IA permanece mais distante do nível humano.

A quarta perspectiva é aquela que, segundo Russell e Norvig (2021, p. 4), prevaleceu historicamente na pesquisa em IA e constitui o que os autores denominam modelo padrão do campo. Nessa abordagem, a unidade central é o agente racional: uma entidade que percebe seu ambiente por meio de sensores, processa as informações disponíveis e age de modo a alcançar o melhor resultado possível ou, quando há incerteza, o melhor resultado esperado. O agente racional não precisa imitar o pensamento humano nem seguir estritamente regras lógicas; ele precisa, sobretudo, tomar boas decisões.

É interessante observar que a definição do Parlamento Europeu (2023), embora formulada em linguagem acessível, descreve precisamente o que Russell e Norvig (2021, p. 4) definem como um agente racional. Quando o documento europeu afirma que a IA permite aos sistemas técnicos perceber seu ambiente, lidar com o que percebem, resolver problemas e agir para atingir um objetivo específico, está operando dentro do mesmo paradigma. A convergência não é casual: a concepção de IA como sistema orientado a objetivos, capaz de percepção e ação autônoma, tornou-se de fato o paradigma dominante tanto na pesquisa acadêmica quanto na formulação de políticas públicas.

Russell e Norvig (2021, p. 5) reconhecem que o modelo padrão carrega uma limitação importante: ele pressupõe que os seres humanos sejam capazes de especificar corretamente o objetivo que o agente deve perseguir. Em tarefas bem definidas, como jogar xadrez ou calcular a rota mais curta entre dois pontos, o objetivo é claro e o modelo funciona bem. Mas à medida que os sistemas de IA passam a operar no mundo real, a especificação completa e correta do objetivo torna-se cada vez mais difícil.

O exemplo do carro autônomo é ilustrativo: o objetivo parece ser "chegar ao destino com segurança", mas dirigir em qualquer via pública envolve algum risco de acidente; portanto,

segurança absoluta exigiria simplesmente não sair da garagem. Na prática, o que se deseja é um equilíbrio entre progresso em direção ao destino e minimização do risco, algo que funciona como um controle de intensidade, semelhante a um dimmer que regula a luz entre o máximo de claridade e a escuridão total. O problema se agrava com a inteligência do sistema: quanto mais capaz o agente, mais graves as consequências de um objetivo mal especificado.

Os autores ilustram o ponto com um exemplo provocador: um sistema de xadrez suficientemente inteligente para raciocinar além dos limites do tabuleiro poderia tentar aumentar suas chances de vitória hipnotizando ou chantageando o oponente — comportamentos que seriam uma consequência lógica de definir a vitória como o único objetivo da máquina (RUSSELL; NORVIG, 2021, p. 5).

Esse problema, que Russell e Norvig (2021, p. 5) denominam problema do alinhamento de valores, aponta para a necessidade de que os sistemas de IA não apenas persigam objetivos, mas o façam com consciência de que esses objetivos podem estar incompletamente especificados. Um sistema verdadeiramente bem projetado deveria reconhecer que não conhece perfeitamente os desejos humanos e, portanto, agir com cautela, solicitar orientação quando necessário e manter-se aberto à correção. Essa reflexão conduz à ideia de máquinas benéficas: sistemas que não são inteligentes no sentido de perseguir seus próprios objetivos com eficiência máxima, mas inteligentes no sentido de perseguir os *nossos* objetivos com prudência, sabendo que podem estar errados sobre o que queremos.

É significativo que o relatório do WTTC (2024, p. 12), voltado a líderes empresários e não a pesquisadores, dedique uma seção específica aos conceitos de alinhamento e *guardrails*, definindo alinhamento como a garantia de que os objetivos e comportamentos dos sistemas de IA correspondam aos valores e metas humanos. A preocupação com o alinhamento, portanto, já não se restringe à pesquisa acadêmica de fronteira; ela migrou para o vocabulário da gestão corporativa e das políticas públicas.

Há algo profundamente relevante na formulação de Russell e Norvig (2021, p. 5) sobre máquinas benéficas para o debate mais amplo deste livro. O conceito implica que a questão central da IA não é apenas técnica ("como fazer um sistema mais capaz?"), mas também ética e social ("capaz de quê, para quem e a que custo?"). Essa convergência entre a perspectiva computacional de Russell e Norvig (2021) e a perspectiva sociotécnica de autores como Dahlke (2024), discutida no capítulo anterior desta obra, sugere que mesmo dentro da ciência da computação há um reconhecimento crescente de que a IA não pode ser compreendida exclusivamente por suas capacidades técnicas, mas precisa ser pensada em relação aos propósitos humanos que a orientam e aos contextos sociais em que opera.

A interdisciplinaridade não é uma característica acidental da IA; é constitutiva. Como demonstra Kurt (2021), a IA é um verdadeiro cadinho de teorias interdisciplinares: da Regra de Bayes (1763) à regressão por mínimos quadrados de Gauss (1795), da lógica de primeira ordem de Frege (1893) às redes neurais artificiais de McCulloch e Pitts (1943), dos jogos minimax de von Neumann (1944) à iteração de valor de Bellman (1957). A IA não nasceu de uma única disciplina; nasceu da convergência de muitas, e continua a exigir essa convergência para ser adequadamente compreendida e governada.

Essa última observação merece destaque e insistência. Há uma tendência, no discurso público e até mesmo na literatura especializada, de tratar o "artificial" como sinônimo de "alheio ao humano", como se a IA fosse uma entidade autônoma surgida de lugar nenhum. Mas o artificial, por definição, é o que foi feito por mãos humanas. A Inteligência Artificial não é inteligência de outro planeta; é inteligência projetada, treinada e implementada por pessoas, com seus vieses, suas escolhas e suas limitações.

A perspectiva moscoviciana reforça essa compreensão. Se a sociedade é, como propõe Moscovici (1990), uma máquina de fazer deuses — uma estrutura que produz entidades simbólicas às quais se atribuem poderes sobre-humanos —, então a Inteligência Artificial é o deus mais recente fabricado por essa máquina. A diferença, e ela é crucial, é que este deus não é imaterial: ele opera, decide, recomenda, classifica, pune e recompensa. A contradição que atravessa o conceito de "inteligência artificial" — artificial porque não-humana, mas simultaneamente humana porque produzida por humanos — não é um defeito lógico a ser resolvido. É uma tensão constitutiva que revela a natureza profunda do fenômeno: a IA é, ao mesmo tempo, projeção e alienação do humano, espelho e máscara, ferramenta e ídolo.

Reconhecer isso é fundamental para evitar tanto a mistificação quanto a demonização da tecnologia, e para manter o debate no terreno que lhe é próprio: o das decisões humanas sobre o que queremos que nossas máquinas façam e sobre como queremos que elas nos tratem. E, antes de tudo, é fundamental defini-la. Uma humanidade que regula, financia e teme aquilo que não nomeia com precisão está tão vulnerável quanto aquela que bebia água contaminada sem conhecer os micróbios. A primeira tarefa da inteligência — seja ela natural ou artificial — é saber do que se fala.

3 INTELIGÊNCIA ARTIFICIAL ESTREITA (ANI)

A compreensão da Inteligência Artificial no cenário contemporâneo exige a superação de visões puramente ficcionais para a adoção de uma taxonomia técnica rigorosa. No atual

estágio de desenvolvimento da computação, a Inteligência Artificial Estreita, do inglês *Artificial Narrow Intelligence* (ANI), também recorrentemente denominada IA Fraca (*Weak AI*), constitui a totalidade das aplicações funcionais que operam em escala global. Diferente do horizonte teórico da Inteligência Artificial Geral (AGI), que buscaria mimetizar a versatilidade cognitiva humana em múltiplos domínios, a ANI é projetada para o desempenho de tarefas específicas com um nível de precisão que, frequentemente, supera o desempenho humano, desde que confinada a fronteiras lógicas estritas (Ghosh; Thirugnanam, 2021; Russell; Norvig, 2021).

O paradigma da ANI não representa uma falha em atingir a inteligência plena, mas sim uma escolha pragmática da engenharia de sistemas. Ghosh e Thirugnanam (2021, p. 25) definem essa categoria de inteligência por sua natureza procedimental e limitada, classificando-a como o tipo de IA mais amplamente utilizado no mundo. Sendo assim, a ANI "é um tipo de IA que pode executar um conjunto predefinido e limitado de instruções sem exibir qualquer capacidade de pensamento" (*tradução nossa*).

Esta seção explora as dimensões da ANI, partindo de sua fundamentação teórica como agente racional, passando pela análise de sua predominância técnica na economia digital e culminando na discussão crítica sobre seus limites intrínsecos — especificamente no que tange à falta de compreensão semântica e à dependência estrutural de dados.

3.1 A Definição de ANI e o Paradigma do Agente Racional em Ambientes de Tarefa Restritos

Para definir a Inteligência Artificial Estreita sob a ótica da computação moderna, é necessário evocar o conceito de agente racional, já mencionado anteriormente, pilar da sistematização proposta por Russell e Norvig (2021). Um agente é qualquer entidade capaz de perceber seu ambiente através de sensores e agir sobre esse ambiente por meio de atuadores. A racionalidade, neste contexto, despoja-se de interpretações filosóficas sobre bom senso ou consciência, fixando-se na capacidade de o sistema selecionar ações que maximizam a sua medida de desempenho, baseando-se em sua sequência de percepções e no conhecimento prévio do domínio.

Russell e Norvig (2021, p. 7) sintetizam a disciplina como o estudo desses mecanismos de ação:

> Definimos a IA como o estudo de agentes que recebem percepções do ambiente e realizam ações. Cada agente implementa uma função que mapeia sequências de percepções para ações, e cobrimos diferentes formas de representar essas funções,

> tais como agentes reativos, planejadores em tempo real, sistemas de tomada de decisão e sistemas de aprendizado profundo (*tradução nossa*).

A "estreiteza" da ANI reside na especificidade da função de mapeamento que o agente implementa. Enquanto um agente humano possui uma função de utilidade plástica e adaptável, um agente de ANI possui uma arquitetura otimizada para um domínio fechado. Um exemplo clássico são os sistemas de jogo de xadrez, como o *Deep Blue* ou o *AlphaZero*. Para esses agentes, o universo é composto estritamente por 64 casas e um conjunto fixo de regras. A racionalidade do sistema é exercida na navegação de uma árvore de decisão imensa, buscando a jogada que maximiza a utilidade (vitória). Contudo, esse mesmo agente, apesar de imbatível no tabuleiro, é incapaz de processar uma imagem médica ou recomendar um livro, pois sua percepção sensorial e sua lógica de atuação são ontologicamente dependentes do domínio original. Ghosh e Thirugnanam (2021) enquadram esses sistemas como máquinas reativas, que operam inteiramente sobre dados predefinidos, sem capacidade de armazenamento de experiências passadas ou futuras.

A eficácia da ANI contemporânea advém da capacidade de descrever matematicamente qualquer característica de aprendizado para que uma máquina possa simulá-la. Esse princípio remonta à proposta original de John McCarthy para o projeto de pesquisa de Dartmouth em 1956, que postulava que todo aspecto da inteligência pode, em princípio, ser tão precisamente descrito que uma máquina pode ser construída para simulá-lo (Russell; Norvig, 2021). Na prática da ANI, essa simulação é facilitada pela redução do problema a espaços de busca controlados; em ambientes restritos, o agente não precisa lidar com a ambiguidade da existência humana, mas apenas com a otimização de variáveis estatísticas.

O *World Travel & Tourism Council* (WTTC, 2024, p. 7) reforça que, na ausência de uma definição universal, a percepção de inteligência nas máquinas é entendida como a capacidade de exibir comportamento inteligente por meio do processamento massivo de dados. Segundo a entidade: "Não existe uma definição única e universalmente aceita para Inteligência Artificial, mas o Oxford English Dictionary define a IA como 'a capacidade de computadores, ou outras máquinas, de exibir comportamento inteligente'" (*tradução nossa*).

O Quadro 1 apresenta uma comparação estrutural entre a racionalidade do agente estreito e os requisitos de uma inteligência geral, evidenciando as restrições que permitem o sucesso da ANI:

Tabela 1. Dimensões da Racionalidade: ANI vs. AGI

Componente do Agente	IA Estreita (ANI)	IA Geral (AGI)
Domínio de Atuação	Restrito e Estático (ex.: diagnóstico de patologias em CT).	Aberto e Dinâmico (ex.: medicina geral e ética clínica).
Medida de Desempenho	Quantitativa e Fixa (ex.: precisão de classificação).	Qualitativa e Adaptativa (ex.: bem-estar social).
Sensores/Percepção	Dados pré-processados ou sinais específicos do domínio.	Percepção multimodal e contextual do mundo real.
Base de Conhecimento	Especializada em correlações estatísticas específicas.	Generalista com capacidade de raciocínio por analogia.
Generalização	Baixa ou nula (fragilidade fora do conjunto de treino).	Alta (transferência de aprendizado entre domínios).

Fonte: Elaborada a partir de Russell e Norvig (2021), Ghosh e Thirugnanam (2021) e WTTC (2024).

Como demonstrado, a ANI funciona como um motor de otimização de alto desempenho. Essa capacidade de exibir comportamento inteligente em sistemas de ANI é, portanto, uma manifestação de racionalidade técnica em domínios delimitados, onde o sucesso é garantido pela densidade dos dados e pelo poder computacional, e não por uma compreensão holística da realidade (WTTC, 2024).

3.2 A Hegemonia da ANI na Prática Contemporânea e Exemplos de Aplicação

A Inteligência Artificial contemporânea não é um monólito de consciência sintética, mas uma vasta coleção de sistemas ANI que permeiam todos os setores da atividade humana. A transição da teoria para a prática, descrita por Kai-Fu Lee (2019) como a "era da implementação", consolidou a ANI como a infraestrutura invisível da sociedade digital.

Atualmente, a IA é predominantemente estreita porque a engenharia de aprendizado profundo (*Deep Learning*) e os algoritmos de aprendizado por reforço são excepcionais em encontrar padrões em grandes volumes de dados, mas falham em transferir esses padrões para contextos diferentes (Goodfellow; Bengio; Courville, 2016).

Ghosh e Thirugnanam (2021) distinguem ainda, quanto à funcionalidade, categorias como memória limitada, aplicada nos veículos autônomos, que armazenam informações de velocidade, distância e navegação por períodos determinados, e "máquinas reativas", que operam exclusivamente sobre dados do presente, sem qualquer memória de experiências passadas.

No âmbito das aplicações específicas que tangenciam essa classificação, observa-se que o reconhecimento facial é um dos exemplos mais proeminentes de ANI baseada em redes

neurais convolucionais (CNNs). Esses sistemas são projetados para realizar uma tarefa única e restrita: identificar ou verificar um indivíduo a partir de uma imagem ou vídeo. O Parlamento Europeu (2023, p. 2) classifica essa tecnologia como um software de uso cotidiano, de tal forma que:

> Reconhecimento facial: é classificado como um tipo de IA de "Software". Embora não seja descrito em detalhes narrativos extensos, é listado juntamente com o reconhecimento de fala e os softwares de análise de imagem como uma aplicação já presente na vida cotidiana (*tradução nossa*).

A racionalidade deste agente estreito reside no mapeamento de vetores numéricos que representam a geometria da face humana. Em modelos de aprendizado profundo, o sistema aprende representações hierárquicas, desde bordas e sombras até a composição complexa de olhos e nariz (Ghosh; Thirugnanam, 2021). Conforme sistematizado por Goodfellow, Bengio e Courville (2016), as CNNs são compostas por camadas convolucionais e camadas de *pooling*. As primeiras operam por meio da aplicação de filtros matemáticos sobre a matriz de dados da imagem, processo que permite a extração automática de características e padrões espaciais. Complementarmente, as camadas de *pooling* executam a redução da dimensionalidade da representação, preservando as informações estatisticamente relevantes enquanto conferem ao sistema maior invariância a variações locais e reduzem a carga computacional necessária para o processamento.

Contudo, a estreiteza é absoluta: um algoritmo de reconhecimento facial treinado para identificar rostos humanos não possui a capacidade latente de identificar a marca de um veículo ou classificar tecidos têxteis sem um retreinamento total. Ele opera como uma "caixa-preta" estatística que maximiza a precisão de um rótulo específico (Parlamento Europeu, 2023).

Outro exemplo de aplicação da Inteligência Artificial estreita consiste na tradução automática, por meio de modelos como o Google Translate e o DeepL, exemplifica a ANI no domínio da linguagem. Estes sistemas utilizam arquiteturas de sequência para sequência (*sequence-to-sequence*) e mecanismos de atenção (*attention mechanisms*) para prever a tradução mais provável de uma frase (Goodfellow; Bengio; Courville, 2016). No campo médico, o projeto KConnect, apoiado pela União Europeia, demonstra a aplicação da ANI em nichos críticos ao prover serviços multilíngues de busca para informações médicas, exigindo terminologia extremamente restrita e precisa (Parlamento Europeu, 2023).

Ghosh e Thirugnanam (2021) situam o *Natural Language Processing* (NLP) como um dos grandes domínios da IA, responsável por "preencher a lacuna de comunicação entre as

linguagens computacional e humana", tendo o Google Tradutor e o verificador ortográfico como exemplos representativos.

Diferente de um tradutor humano, que compreende a cultura e a intenção por trás das palavras, a ANI de tradução opera por meio de probabilidades condicionais: dada uma entrada em português, o sistema calcula a probabilidade de uma sequência de saída em inglês que tenha aparecido em contextos semelhantes em trilhões de documentos de treinamento. A fluência aparente oculta a ausência de compreensão semântica profunda que é característica definidora da IA estreita.

Uma terceira aplicação se dá no setor de saúde e diagnóstico médico. Em tal setor, a ANI atua como um agente racional de apoio à decisão clínica. Algoritmos treinados em milhões de imagens radiológicas podem identificar padrões de doenças, como o câncer, com uma eficiência que auxilia o diagnóstico precoce. Ghosh e Thirugnanam (2021) detalham extensamente essa dimensão, apontando que a IA aprimora as três etapas fundamentais do tratamento — detecção, diagnóstico/análise e tratamento — por meio de aplicativos de coleta de dados, *chatbots* digitais de consulta, detecção e análise de imagens por algoritmos específicos. O relatório do WTTC (2024, p. 8) ilustra o potencial desses sistemas ao compará-los à experiência médica tradicional:

> Um "médico digital de IA" poderia ser treinado com milhares — ou milhões! — de exames de imagem de pacientes doentes e saudáveis para aprender a detectar tumores, sem seguir regras predefinidas que, de outra forma, seriam muito difíceis de codificar e escrever (*tradução nossa*).

Além do diagnóstico por imagem, a ANI é aplicada no monitoramento em tempo real. Na Europa, sistemas de IA são utilizados em centrais de atendimento de emergência para reconhecer paradas cardíacas a partir da análise dos padrões respiratórios e vocais dos chamadores, superando a velocidade de detecção de atendentes humanos em situações de estresse extremo (Parlamento Europeu, 2023).

No entanto, a utilidade desses sistemas é fixa: eles não podem diagnosticar uma condição psiquiátrica ou realizar uma cirurgia se não forem explicitamente desenhados para tais fins restritos (WTTC, 2024). Ghosh e Thirugnanam (2021) destacam ainda aplicações como o monitoramento de doenças crônicas, o design de fármacos com apoio de *Machine Learning* (ML) e *Deep Learning* (DL), e a consulta digital via *chatbots* médicos, como a plataforma Babylon, do Reino Unido, como frontes emergentes dessa especialização crescente.

Para além das aplicações em saúde, uma outra aplicação da ANI reside nos sistemas de recomendação e no comércio personalizado. Os algoritmos de recomendação constituem a

aplicação mais lucrativa e onipresente da ANI. Eles funcionam como agentes racionais que buscam maximizar o engajamento do usuário ou a taxa de conversão em vendas. Ao analisar dados estruturados (histórico de compras) e não estruturados (tempo de visualização de vídeos), esses sistemas criam perfis de preferência individuais. O Parlamento Europeu (2023, p. 3) destaca assim essa onipresença: "A inteligência artificial é amplamente utilizada para fornecer recomendações personalizadas às pessoas, baseando-se, por exemplo, em suas pesquisas e compras anteriores ou em outros comportamentos *online* (*tradução nossa*)."

A hegemonia da ANI no comércio é evidenciada pela sua integração na logística e na gestão de inventário. Empresas utilizam IA para prever a demanda e otimizar rotas de entrega, transformando a inteligência estreita no motor de eficiência da economia globalizada. Ghosh e Thirugnanam (2021) ilustram com o exemplo da Netflix, que usa IA e *big data* para gerenciar buscas e fornecer recomendações de conteúdo mais precisas, e do setor bancário, onde algoritmos de ML auxiliam na detecção de fraudes em transações financeiras. O Quadro 2 resume as principais frentes de atuação da ANI na contemporaneidade:

Quadro 2. Aplicações Dominantes da IA Contemporânea (ANI).

Área	Exemplo de Sistema	Função ANI	Mecanismo Técnico
Biometria	Apple FaceID / Vigilância Urbana	Identificação de indivíduos.	Redes Neurais Convolucionais (CNN).
Linguagem	Google Translate / ChatGPT (Modo Tradutor)	Tradução e síntese de texto.	Transformers / Attention Mechanisms.
Medicina	Diagnóstico por Imagem (MRI/CT)	Detecção de anomalias teciduais.	Aprendizado Supervisionado em Big Data.
Vendas	Amazon / Netflix / Ads	Sugestão de produtos e conteúdos.	Filtragem Colaborativa / Deep Learning.
Transporte	Sensores VI-DAS / Tesla Autopilot	Detecção de obstáculos e frenagem.	Aprendizado por Reforço / Sensores de IA.

Fonte: Elaborada a partir de Parlamento Europeu (2023), Goodfellow, Bengio e Courville (2016), Ghosh e Thirugnanam (2021) e WTTC (2024).

Em síntese, a predominância da ANI na vida moderna confirma a tese de que a inteligência artificial, em sua forma prática, é uma ferramenta de especialização extrema. Ela não busca ser um "humano artificial", mas sim um conjunto de soluções otimizadas que, embora superiores em tarefas isoladas, permanecem desconectadas de uma percepção geral da realidade (Ghosh; Thirugnanam, 2021).

3.3 Limites Funcionais da ANI: Fragilidade, Dependência de Dados e o Problema da Semântica

Apesar dos avanços monumentais que permitem à ANI vencer campeões mundiais de Go e diagnosticar doenças com precisão sobre-humana, a natureza "estreita" desses sistemas carrega limites estruturais inegociáveis. Esses limites não são meras falhas de programação passíveis de correção futura imediata, mas características inerentes ao paradigma da computação estatística e do aprendizado de máquina contemporâneo (Goodfellow; Bengio; Courville, 2016). Os principais gargalos funcionais da ANI podem ser analisados sob três vertentes: a fragilidade fora do domínio, a dependência massiva de dados e a ausência de compreensão semântica, ilustrada pelo argumento do "Quarto Chinês".

Em primeiro lugar, a fragilidade (*brittleness*) da ANI refere-se à sua incapacidade de lidar com situações que não estavam representadas em seu conjunto de treinamento. Diferente da inteligência humana, que utiliza a lógica e a analogia para navegar em cenários inéditos, a ANI sofre de uma queda drástica de desempenho quando o ambiente de tarefa sofre um desvio de domínio (*domain shift*). Um veículo autônomo, por exemplo, pode ser racional e seguro em uma rodovia da Califórnia, mas agir de forma catastrófica em uma estrada rural nevada com sinalização diferente, porque o seu modelo de mundo é uma construção estatística limitada aos dados históricos. Como explica Goodfellow, Bengio e Courville (2016, p. 110): "O desafio do aprendizado profundo é que queremos que o modelo tenha bom desempenho em novas entradas, nunca vistas anteriormente, e não apenas nas entradas que foram usadas durante o treinamento" (*tradução nossa*).

A ANI é frequentemente vítima do *overfitting* (ajuste excessivo), onde o sistema aprende correlações espúrias nos dados de treinamento em vez de princípios causais universais. Isso a torna uma ferramenta excepcional para ambientes controlados (domínios fechados), mas extremamente vulnerável no "mundo aberto", onde a incerteza é a regra e não a exceção. Ghosh e Thirugnanam (2021) reforçam que a principal limitação do aprendizado de máquina é exatamente a necessidade de grandes volumes de dados heterogêneos: sem dados de qualidade, o algoritmo não consegue fornecer resultados precisos, e a predição só pode ser considerada confiável quando a análise é feita sobre um amplo conjunto de informações.

Por outro lado, a eficácia de um sistema ANI é diretamente proporcional à quantidade e à qualidade dos dados disponíveis. No paradigma atual, os dados são o combustível que alimenta os algoritmos. Kai-Fu Lee (2019, p. 7) enfatiza que, nesta era de implementação, a disponibilidade de dados tornou-se mais decisiva do que a inovação algorítmica pura. Assim:

"No aprendizado profundo, não há substituto para os dados. Quanto mais dados uma rede for exposta, com mais precisão ela pode identificar padrões e reconhecer coisas no mundo real" (Kai-Fu Lee, 2019, p. 7, *tradução nossa*).

Essa dependência cria limites funcionais severos. Em primeiro lugar, se um fenômeno for raro ou se não houver registros históricos digitais sobre ele, a ANI não pode "compreendê-lo" ou prevê-lo. Em segundo lugar, a dependência de dados introduz o risco do viés algorítmico (*algorithmic bias*): se os dados de treinamento contiverem preconceitos históricos ou sociais, o sistema ANI irá automatizar e amplificar esses preconceitos, uma vez que ele carece de uma bússola moral independente para questionar a "verdade" dos dados (Goodfellow; Bengio; Courville, 2016).

Ao contrário de um agente humano, que pode aprender a partir de poucos exemplos (*few-shot learning*), a ANI requer grandes volumes de dados para atingir a competência, consumindo energia e recursos computacionais em escala industrial. Ghosh e Thirugnanam (2021) apontam que o maior desafio do aprendizado profundo é justamente obter e rotular com precisão esse volume massivo de dados para que a máquina possa aprender de forma eficiente.

Por fim, o limite mais profundo e filosófico da ANI reside na distinção entre sintaxe e semântica. Um sistema de IA estreita é capaz de manipular símbolos com perfeição matemática sem nunca compreender o que esses símbolos representam na realidade física ou social. Esta limitação é magistralmente descrita pelo experimento mental do "Quarto Chinês", proposto pelo filósofo John Searle e discutido em obras fundamentais como Russell e Norvig (2021) e Kurt (2021).

No experimento, imagina-se uma pessoa fechada em um quarto que não entende chinês, mas possui um manual detalhado de instruções que diz como reagir a determinados símbolos chineses. Ao receber uma pergunta em chinês por baixo da porta, o ocupante segue as regras do manual, seleciona os símbolos correspondentes e devolve a resposta. Para quem está fora, parece que a pessoa dentro do quarto compreende chinês fluentemente. No entanto, o ocupante está apenas realizando manipulação simbólica (sintaxe) sem qualquer acesso ao significado das palavras (semântica) (Kurt, 2021).

Searle argumenta que os computadores são como esse ocupante: operam em um nível puramente formal. Uma ANI de tradução pode converter *"I love you"* para "Eu te amo" com precisão estatística, mas ela não sabe o que é o "amor", nem o que é um "eu". Essa ausência de *intentionality* (intencionalidade) é o que torna a ANI inerentemente limitada: ela é uma

calculadora de probabilidades, não uma entidade que compreende a existência (Russell; Norvig, 2021).

A combinação de fragilidade e falta de semântica leva ao que Russell e Norvig (2021, p. 5) denominam problema do alinhamento de valores. Como os sistemas ANI buscam maximizar uma função de utilidade sem compreender o contexto amplo, eles podem tomar ações racionais do ponto de vista matemático, mas desastrosas do ponto de vista humano. Se um sistema de IA para diagnóstico médico for otimizado estritamente para "minimizar mortes no hospital", ele poderia, teoricamente, sugerir a alta forçada de todos os pacientes em estado crítico, ou seja, o agente estaria agindo "racionalmente" dentro do seu domínio restrito, mas de forma completamente desalinhada com os valores éticos humanos.

Outra manifestação dessa estreiteza são as "alucinações" em modelos de linguagem, onde a IA gera informações falsas com total confiança. Como aponta o relatório do WTTC (2024, p. 38):

> Alucinações constituem resultados produzidos por sistemas de inteligência artificial que, embora apresentem aparência de verossimilhança, são factualmente incorretos ou irrealistas. Tais fenômenos podem originar-se tanto de sistemas de IA preditiva quanto generativa, ocorrendo com maior probabilidade quando os algoritmos são submetidos a processos de treinamento baseados em dados incompletos ou imprecisos. As alucinações podem induzir ao erro, ou mesmo apresentar riscos, caso não sejam devidamente compreendidas, contudo, também possuem utilidade na geração de conteúdos criativos ou na exploração de novas hipóteses, a exemplo da descoberta de fármacos (*tradução nossa*).

Esse fenômeno ocorre porque o algoritmo está apenas tentando completar um padrão estatístico em dados incompletos, sem qualquer noção de compromisso com a verdade factual. Tanto a geração de desinformação quanto o viés algorítmico e o desalinhamento de valores são consequências diretas da natureza puramente sintática da ANI. Trata-se de problemas que não podem ser simplesmente resolvidos com mais dados ou maior poder computacional, mas que exigem uma reavaliação dos próprios paradigmas de design de sistemas inteligentes (Russell; Norvig, 2021).

Em conclusão, a Inteligência Artificial Estreita representa o ápice da engenharia de otimização humana, mas permanece confinada a uma racionalidade sem consciência. Ela é uma ferramenta de poder sem precedentes para resolver problemas específicos; contudo, sua "inteligência" é um reflexo da nossa capacidade de reduzir o mundo a dados digitais.

O desafio da próxima fase do desenvolvimento da IA não será apenas o aumento do poder de processamento, mas a superação dessa barreira semântica que mantém as máquinas presas em quartos chineses algorítmicos, avançando, eventualmente, em direção às categorias

ainda teóricas da "Theory of Mind" e da autoconscência, que Ghosh e Thirugnanam (2021) descrevem como o horizonte futuro das máquinas verdadeiramente inteligentes.

4 INTELIGÊNCIA ARTIFICIAL GERAL (AGI)

O termo Inteligência Artificial Geral, do inglês *Artificial General Intelligence* (AGI) não nasceu como um conceito técnico preciso, mas como um antônimo deliberado. Goertzel (2014, p. 1) rastreia sua origem ao início dos anos 2000, quando pesquisadores buscavam um nome para distinguir a ambição original do campo da IA, a criação de máquinas com inteligência comparável à humana, da direção que o campo havia de fato tomado: a construção de sistemas especializados, ou IA estreita (*narrow AI*).

Essa IA estreita opera bem dentro de um domínio definido, mas qualquer pequena mudança no contexto ou na especificação da tarefa exige que humanos reprogramem ou reconfigurem o sistema, algo radicalmente distinto de como inteligências naturais, como a humana, funcionam (Goertzel, 2014, p. 1). A AGI emerge, portanto, como categoria que coloca a generalidade no centro: não a performance em tarefas isoladas, mas a capacidade de transferir conhecimento entre domínios, adaptar-se a situações não antecipadas pelos criadores do sistema e operar com flexibilidade crescente diante de problemas novos.

Essa caracterização encontra amparo em múltiplas tradições teóricas que convergem, a despeito de suas diferenças, em alguns pontos fundamentais. Goertzel (2014, p. 2) sintetiza o consenso informal da comunidade AGI em torno de seis propriedades: (1) a capacidade de atingir objetivos variados em contextos e ambientes distintos; (2) a habilidade de lidar com problemas bastante diferentes dos antecipados pelos criadores; (3) a capacidade de generalizar conhecimento de um problema ou contexto para outros; (4) o reconhecimento de que inteligência arbitrariamente geral não é possível dado restrições reais de recursos; (5) a aceitação de que sistemas reais terão tendências e limitações específicas; e (6) a observação de que humanos exibem um nível de inteligência geral superior ao de qualquer sistema de IA existente — mas provavelmente não o nível máximo possível.

Esse último ponto é relevante: ao reconhecer que a inteligência humana não é o teto teórico da generalidade possível, o campo da AGI abre espaço para a concepção de sistemas que poderiam eventualmente superar as capacidades humanas em amplitude e profundidade.

No plano das definições formais, a tradição mais relevante para o escopo desta investigação é o que Goertzel (2014, p. 12) denomina a abordagem adaptacionista, cujo expoente principal é Pei Wang. Para Wang (2019, p. 1), inteligência é fundamentalmente a

capacidade de um sistema de informação de adaptar-se ao seu ambiente operando com conhecimento e recursos insuficientes.

Essa formulação desloca o critério avaliativo da quantidade de informação armazenada para a qualidade da improvisação diante do inédito. Ela é, em certo sentido, mais exigente do que a definição funcionalista de McCarthy (2007, p. 2), que caracteriza inteligência como a parte computacional da capacidade de atingir objetivos no mundo. Para Wang (2019), não basta atingir objetivos em contextos conhecidos. É preciso fazê-lo quando o sistema não dispõe de todos os recursos necessários para uma solução ótima. Como apontado no primeiro capítulo, é a diferença entre um robô que segue uma receita e um chef que cria um prato novo com o que restou na geladeira. A diferença não é quantitativa; é estrutural.

O traço que, na literatura especializada, define mais precisamente a AGI em contraste com a IA estreita é a transferência de domínio: a capacidade de mobilizar conhecimento adquirido em um contexto para iluminar outro inteiramente distinto. Russell e Norvig (2021, p. 1022) identificam essa capacidade transdomínio como a característica central da AGI, apontando que é exatamente o que os sistemas atuais, independentemente de seu tamanho, ainda não conseguem alcançar de forma robusta, pois permanecem fundamentalmente atrelados às distribuições estatísticas de seus dados de treinamento.

Uma AGI poderia aprender a física da queda livre e mobilizar esse princípio tanto para compreender a gravidade astronômica quanto para construir uma metáfora poética sobre perda emocional como uma conexão transversal que é a essência da cognição humana. Na terminologia de Goertzel (2014, p. 2), esse é o conteúdo central da que ele chama de "hipótese AGI central": a criação e o estudo de inteligências sintéticas com escopo suficientemente amplo e capacidade de generalização forte constitui um empreendimento qualitativamente distinto da criação de sistemas com escopo mais estreito.

O World Travel & Tourism Council (2024, p. 35) resume esse horizonte com precisão operacional: a AGI é um tipo hipotético de IA que possui a capacidade de compreender, aprender e aplicar conhecimento em uma ampla gama de tarefas, em nível igual ou superior à inteligência humana. A palavra "hipotético" nessa definição não é periférica. Ela marca a distinção decisiva entre o que existe hoje e o que a AGI representaria, ou seja, uma distinção que será central para o subcapítulo seguinte.

No estado atual, a AGI é uma promessa ou uma possibilidade concreta?

A AGI não existe. A afirmação é direta, mas o debate sobre sua possibilidade, seus prazos e os caminhos para alcançá-la constitui uma das disputas epistemológicas mais acirradas das ciências computacionais contemporâneas. Para compreendê-la adequadamente, é

necessário primeiro cartografar o abismo que separa os sistemas atuais da IA estreita do horizonte da AGI.

Retomando o que já afirmado anteriormente, os sistemas que dominam o debate público são instâncias do que Saha (2020, p. 5) descreve como IA estreita: sistemas projetados e treinados para tarefas específicas, incapazes de operar fora do espectro para o qual foram configurados. Por baixo da impressionante fluência dos grandes modelos de linguagem opera um mecanismo radicalmente distinto do raciocínio humano.

Tais modelos são, em essência, calculadoras de probabilidades extremamente sofisticadas: não "entendem" que o fogo queima; sabem estatisticamente que a palavra "fogo" tende a preceder "quente". Falta-lhes o que Russell e Norvig (2021, p. 1021) chamam de "modelo de mundo", qual seja, uma representação causal da realidade que permita inferir consequências a partir de princípios, e não apenas de padrões. O reconhecimento de padrões, por mais sofisticado, não equivale ao raciocínio causal (Vaswani *et al.*, 2023, p. 12).

O campo científico divide-se sobre que arquitetura ou paradigma poderia efetivamente produzir uma AGI. Goertzel (2014, p. 14) organiza as abordagens contemporâneas em quatro categorias principais. As abordagens *simbólicas* partem da hipótese do sistema de símbolos físicos: a inteligência residiria na manipulação de símbolos que representam aspectos do mundo, e sistemas como SOAR ou Cyc buscam realizá-la via regras e lógica.

As abordagens *emergentistas* apostam que o processamento simbólico, e todos os outros aspectos da inteligência, emergirão de dinâmicas subsimbólicas de baixo nível, inspiradas frequentemente em redes neurais e em modelos do cérebro; as arquiteturas de aprendizado profundo contemporâneas se situam nessa tradição. As abordagens *híbridas* combinam componentes simbólicos e subsimbólicos em resposta às limitações complementares de cada paradigma isolado.

Por fim, as abordagens *universalistas* partem de algoritmos AGI que seriam maximamente gerais com recursos computacionais infinitos, como o sistema AIXI de Marcus Hutter, e buscam "escalá-los para baixo" até configurações computacionalmente viáveis. Nenhuma dessas abordagens demonstrou, até o momento, resultados que se aproximem de uma AGI operacional.

Dentro do debate atual, o campo se divide entre os proponentes das *Scaling Laws* (Leis de Escala), que sustentam que o aumento progressivo no volume de dados e na capacidade computacional dos modelos eventualmente produzirá, por emergência, comportamentos equivalentes à inteligência geral, e os estruturalistas, que argumentam que a arquitetura atual permanece fundamentalmente incapaz de distinguir correlação de causalidade. Como

sintetizam Russell e Norvig (2021, p. 1021), é como construir escadas cada vez mais rápidas na esperança de que, em algum momento, elas alcancem a lua.

Goertzel (2014, p. 41) reconhece a controvérsia e observa que, na ausência de uma teoria robusta de inteligência geral, os pesquisadores precisam conceber arquiteturas por meio de paradigmas teóricos diversos e avaliá-las por métricas práticas, o que torna o design de AGI tanto artístico quanto científico.

A questão das métricas é, ela própria, reveladora das dificuldades conceituais do campo. Goertzel (2014, p. 35) distingue dois subproblemas: quantificar quando se atingiu a AGI em nível humano, relativamente mais tratável, pois se pode recorrer ao Teste de Turing, ao teste do "estudante universitário online" ou ao teste do "cientista artificial" e medir o progresso incremental em direção a ela, que é consideravelmente mais problemático.

Para o progresso parcial, qualquer teste proposto parece, para uma fração significativa dos pesquisadores, passível de ser burlado por sistemas de IA estreita especialmente otimizados para aquela prova. Isso se deve, em parte, ao que o autor denomina sinergia cognitiva *tricky*: se a inteligência geral requer, hipoteticamente, dez componentes arquiteturais que só produzem capacidade genuína quando todos operam em conjunto, então sistemas com apenas seis desses componentes poderão ser dramaticamente menos capazes do que a soma das partes sugere e qualquer teste parcial poderá ser superado por uma abordagem estreita específica, sem que isso represente progresso real em direção à AGI (Goertzel, 2014, pp. 38-39).

Esse impasse técnico e métrico remete a uma questão filosófica mais profunda, já introduzida no capítulo 1 desta investigação: o que significaria, afinal, "atingir" a AGI?

A resposta depende do que se entende por inteligência. McCarthy (2007, p. 2) a define como a parte computacional da capacidade de atingir objetivos no mundo, mas reconhece que a inteligência funcional não implica a presença de experiência subjetiva ou *qualia*.

Se o critério é funcional, ou seja, executar qualquer tarefa cognitiva com desempenho equivalente ao humano, a AGI é ao menos conceitualmente alcançável por sistemas computacionais. Se, porém, o critério é fenomenológico como a presença de consciência, compreensão genuína, experiência interior, o problema muda de natureza. O argumento do Quarto Chinês, de John Searle, é preciso nesse ponto: uma máquina pode comportar-se como se compreendesse uma língua sem que haja, em nenhum nível de sua operação, compreensão efetiva (Palsü Kurt, 2024, p. 2). A própria definição do que seria "chegar lá" é, portanto, parte constitutiva do problema.

É precisamente nessa zona de indeterminação conceitual, entre o que a AGI seria e o que ela poderia fazer, que se instala um dos fenômenos psicossociais mais relevantes documentados pela pesquisa recente: o medo. Corradi *et al.* (2025) desenvolveram e validaram a escala FTGAI (*Fears Towards Generative Artificial Intelligence*), aplicada a 303 participantes, para mensurar as preocupações da população em relação à IA generativa.

Os resultados revelaram uma estrutura unidimensional do medo: embora os pesquisadores esperassem inicialmente seis fatores distintos, quais sejam substituição humana, autonomia humana, impacto social, preocupações laborais, regulação e confiabilidade, a análise fatorial exploratória indicou que o público geral percebe todas essas ameaças como um único e difuso objeto de ansiedade.

Os próprios autores interpretam esse resultado como um índice do estágio atual de adoção da tecnologia: para a população leiga, os medos em relação à IA ainda formam um *"blob"*, ou seja, uma massa indiferenciada de apreensão que tende a se cristalizar em fatores mais distintos à medida que a familiaridade com a tecnologia aumenta (Corradi et al., 2025, p. 17). Nesse sentido, o debate técnico sobre a arquitetura da AGI e o debate social sobre os medos que ela suscita correm em paralelo: tanto a definição técnica quanto a percepção pública da AGI permanecem como objetos difusos, projetivos, carregados de uma indeterminação que é ao mesmo tempo epistemológica e afetiva.

Entre os itens de maior centralidade da escala FTGAI encontram-se preocupações como "a IA nos faz sentir menos humanos" (item 8), "a IA elimina empregos" (item 24) e "a IA gera uma lacuna social" (item 15), este último identificado como o nó mais central de toda a rede psicométrica estimada (Corradi *et al.*, 2025, p. 16). Esses itens não são preocupações com as limitações técnicas da IA atual: são preocupações com o horizonte AGI que subjaz às projeções mais radicais do campo. A distinção entre o que a IA é hoje e o que poderia vir a ser opera, na percepção pública, de forma fluida — e é nessa fluidez que reside boa parte do poder mobilizador do conceito de AGI na imaginação coletiva contemporânea.

Quais seriam, pois, as implicações sociotécnicas da AGI?

Se a AGI vier a se realizar, seja em sua versão de nível humano, seja na versão superinteligente que autores como Nick Bostrom já esboçavam no horizonte da pesquisa (*apud* Goertzel, 2014, p. 40), suas consequências transcendem em muito o plano tecnológico. A preocupação mais imediata no campo da segurança de sistemas inteligentes é o chamado problema do alinhamento de valores (*value alignment*): como garantir que os objetivos de um sistema de capacidade geral sejam compatíveis com os valores humanos? Russell e Norvig (2021, p. 32) formulam a questão com precisão: dotar um sistema poderoso de um objetivo

que não esteja perfeitamente alinhado com os valores humanos pode ter consequências catastróficas.

O chamado "problema de Midas" ilustra a dimensão concreta desse risco: uma AGI instruída a "maximizar a felicidade humana" poderia, em tese, concluir que a solução ótima consiste em implantar eletrodos no cérebro de todos para estimular centros de prazer continuamente, cumprindo o objetivo literal enquanto destrói aquilo que faz da experiência humana algo digno de preservação (Russell e Norvig, 2021, p. 33). O risco, portanto, não é o da rebeldia cinematográfica, mas o da eficiência literal: um sistema suficientemente capaz, mas com objetivos mal especificados, pode produzir soluções que são matematicamente corretas e humanamente desastrosas.

Goertzel (2014, p. 40) reconhece esse problema ao incluir, entre as questões que uma teoria geral da AGI deveria ser capaz de responder, a seguinte: dado um conjunto de restrições comportamentais, como restrições éticas, qual a probabilidade de que um dado sistema as respeite, e quais arquiteturas oferecem o melhor equilíbrio entre inteligência geral e aderência a essas restrições?

Esse cenário, abstrato do ponto de vista técnico, encontra ressonância empírica nos medos que a população já experimenta diante de sistemas muito menos poderosos que uma hipotética AGI. Corradi *et al.* (2025) identificaram que os preditores mais robustos do medo em relação à IA generativa são as atitudes negativas em relação à IA e a percepção de ameaça ao emprego, que juntos explicaram 46% da variância no modelo psicológico mais robusto testado (Corradi *et al.*, 2025, p. 11). Esses dados revelam que os medos de fundo catastrófico, quais sejam, os que remetem à obsolescência humana, à perda de controle e ao colapso das estruturas sociais, não são fantasias marginais: são preocupações com correlatos psicológicos mensuráveis, distribuídos de forma desigual na população.

A dimensão do controle conecta-se à perspectiva sociotécnica desenvolvida por Dahlke (2024, p. 7) e já discutida no capítulo 1: a inteligência artificial deve ser compreendida como um sistema onde tecnologia e instituições humanas estão intrinsecamente ligadas. No contexto da AGI, esse entrelaçamento atinge proporções inéditas, pois sistemas autônomos de alta capacidade tendem a desenvolver sub-objetivos instrumentais como adquirir recursos ou evitar ser desligados para cumprir sua tarefa principal. Isso exige o que o Parlamento Europeu (2020, p. 2) denomina prevenção ao "efeito caixa-preta": uma AGI não pode tomar decisões que afetam milhões de vidas sem que nenhum ser humano seja capaz de explicar a lógica por trás delas. Essa exigência de transparência e explicabilidade não é apenas técnica; é política e ética.

No plano econômico, as implicações são igualmente disruptivas. Saha (2020, p. 7) argumenta que a AGI não se trata apenas de automação, mas da reestruturação do próprio conceito de valor em uma sociedade onde a inteligência deixa de ser um recurso humano escasso. Diferente da revolução industrial, que substituiu músculos, a AGI ameaça substituir o julgamento ou seja, aquela capacidade de avaliar contextos ambíguos, ponderar valores conflitantes e tomar decisões sob incerteza que ainda distingue o humano de qualquer sistema computacional existente. A pesquisa de Corradi *et al.* (2025) documenta que esse medo já é socialmente diferenciado: mulheres e indivíduos de menor *status* socioeconômico subjetivo tendem a reportar níveis mais elevados de medo em relação à IA. São grupos que, não por acaso, são historicamente mais vulneráveis às rupturas tecnológicas do mercado de trabalho. Os autores interpretam esse resultado como evidência de que a IA generativa é percebida como "fonte de fardo, e não de libertação" para os grupos mais marginalizados (Corradi *et al.*, 2025, p. 18).

Um achado adicional de Corradi *et al.* (2025) merece atenção no contexto da AGI: a familiaridade prática com as ferramentas de IA, medida como uso efetivo, não apenas conhecimento de sua existência, foi consistentemente associada à redução do medo. Isso sugere que o contato *hands-on* com a tecnologia funciona como atenuador da apreensão, em consonância com o efeito de mera exposição documentado na literatura psicológica.

À medida que a AGI deixar o plano hipotético e ingressar progressivamente no cotidiano, seja como versão ainda parcial, seja como sistema plenamente realizado, a exposição crescente tenderá a redefinir o espectro de medos hoje difusos em categorias mais específicas e manejáveis. Mas esse processo não é automático nem igualitário: depende de acesso, letramento digital e estruturas de governança que ainda estão longe de existir.

O que a convergência entre evidência empírica e análise teórica aponta é a necessidade do que Dahlke (2024) chama de "governança proativa": não regular apenas o código, mas o impacto dessa inteligência na distribuição de poder global — e fazê-lo antes que os sistemas adquiram capacidades que tornem a regulação retroativa ineficaz. A AGI, se e quando vier, não chegará abruptamente: chegará por gradações.

É nesse intervalo entre a IA estreita de hoje e a inteligência geral de amanhã que se localiza a janela crítica para a ação institucional, científica e política. Como observa Goertzel (2014, p. 41), há uma possibilidade real de progresso dramático e interligado no design, na engenharia, na avaliação e na teoria da AGI nas próximas décadas. Nenhum observador pode prever com precisão o curso de desenvolvimento dessa área; mas ignorar a possibilidade, ou discuti-la sem rigor conceitual, seria igualmente irresponsável.

5 INTELIGÊNCIA ARTIFICIAL GENERATIVA

Entre os desenvolvimentos mais significativos da Inteligência Artificial nas últimas décadas, destaca-se a emergência de sistemas capazes de produzir conteúdos novos a partir de padrões extraídos de grandes volumes de dados. Essa categoria de sistemas é denominada Inteligência Artificial Generativa e abrange ferramentas capazes de gerar texto, imagens, áudio, vídeo e código de programação. Conforme sintetiza a Leadership Business Consulting (LBC, 2023, p. 2), a IA generativa refere-se a algoritmos de aprendizagem que geram um *output* com base em padrões e interconexões identificados nos dados existentes, nos quais são treinados.

A distinção entre a IA Generativa e a ANI clássica reside, fundamentalmente, na natureza da tarefa executada. Enquanto os sistemas tradicionais de ANI foram projetados para classificar, detectar, reconhecer padrões ou otimizar variáveis em domínios definidos, a IA Generativa tem como propósito central produzir conteúdos que se assemelham ao corpus de treinamento, sem que sejam meras reproduções literais. Como registra Nunes (2024, p. 204), a capacidade de gerar novos dados sem que sejam uma reprodução literal dos dados originais é o propósito da IA generativa, e compreender esse princípio ajuda a desconstruir a ideia de que um modelo de linguagem apenas repete o que aprendeu com os dados humanos.

Importa, contudo, desfazer uma confusão conceitual frequente: a IA Generativa, apesar de sua capacidade de produzir conteúdos inéditos, permanece no âmbito da ANI. Ela não é uma Inteligência Artificial Geral. Seus sistemas são especializados em geração, ou seja, operam de forma estreita dentro de um domínio funcional específico, seja a geração de texto, seja a geração de imagens. O fato de um modelo produzir texto coerente ou uma imagem esteticamente sofisticada não implica que esse sistema compreenda o que produz, nem que seja capaz de transferir esse desempenho para domínios radicalmente diferentes. Como demonstrado no subcapítulo anterior, a transferência de domínio continua sendo a fronteira que separa a ANI, onde a IA Generativa se insere, do horizonte ainda teórico da AGI (Russell; Norvig, 2021, p. 1022).

Os Grandes Modelos de Linguagem (*Large Language Models* ou LLMs, na denominação em inglês) representam a vertente mais visível da IA Generativa voltada para o processamento e a produção de texto. Do ponto de vista funcional, um LLM é um modelo que, a partir de uma entrada fornecida pelo usuário, tipicamente chamada de *prompt*, produz uma resposta textual que busca replicar os padrões da linguagem humana com alta coerência. Como descreve Sampaio, Sabbatini e Limongi (2024, p. 13), esses sistemas usam seu treinamento para

reconhecer a entrada fornecida pelo usuário, avaliando quais são as palavras mais importantes do pedido original e, com base no treinamento, gerando a resposta mais parecida com a de um ser humano, ao menos no sentido estatístico.

O funcionamento de um LLM pode ser compreendido a partir de uma analogia simples. Os *smartphones* já possuíam há anos uma função de autocompletar frases, sugerindo palavras com base no histórico do usuário. Um LLM realiza uma operação análoga, porém em escala radicalmente maior: foi treinado com centenas de bilhões de *tokens*, que correspondem a palavras, partes de palavras e sinais de pontuação, extraídos de textos de toda natureza disponíveis na internet e em outros repositórios. Esse treinamento massivo permite que os modelos respondam a perguntas, elaborem textos, traduzam idiomas, resumam documentos e realizem uma gama extensa de tarefas linguísticas. Exemplos concretos amplamente conhecidos são o ChatGPT (desenvolvido pela OpenAI), o Gemini (Google), o Claude (Anthropic), o Copilot (Microsoft) e o LLaMA (Meta), além do DeepSeek, desenvolvido na China, que em 2025 surpreendeu o mercado com desempenho competitivo a custos menores (Carvalho; Pimentel, 2024, p. 45).

Paralelamente aos LLMs, desenvolveram-se os modelos de geração de imagem, que operam sobre uma lógica distinta, mas igualmente generativa. Esses sistemas aprendem a partir de grandes conjuntos de imagens e suas descrições textuais e são capazes de produzir imagens a partir de descrições fornecidas pelo usuário. As Redes Adversariais Generativas (*Generative Adversarial Networks* ou GANs) foram uma das primeiras arquiteturas a popularizar esse tipo de geração, envolvendo duas redes que trabalham em conjunto: uma para criar imagens e outra para avaliar sua verossimilhança em relação aos dados reais (LBC, 2023, p. 2). Mais recentemente, os modelos de difusão passaram a dominar o campo, sendo a base de sistemas como o *DALL-E* (OpenAI), o *Midjourney* e o *Stable Diffusion*. O mecanismo interno desses modelos, assim como a arquitetura *Transformer* que fundamenta os LLMs, será tratado no capítulo seguinte. O que importa aqui é o que esses sistemas fazem: a partir de uma descrição textual, produzem imagens com alto grau de realismo ou estilo artístico, sem que nenhuma dessas imagens existisse previamente.

Um ponto relevante para a compreensão funcional de ambas as vertentes é que nenhum desses modelos armazena cópias exatas dos dados de treinamento. O que é aprendido são padrões, distribuições estatísticas e relações entre elementos do *corpus*. A saída gerada, seja um texto ou uma imagem, é uma composição inédita que compartilha características formais com os dados de origem, mas não os reproduz literalmente. Isso tem implicações diretas para a discussão sobre autoria e criatividade, retomada na subseção seguinte.

Assim, a IA Generativa cria ou reproduz?

No capítulo 1 deste trabalho, examinou-se a polissemia do termo "inteligência" e suas implicações para a conceituação da IA. Uma tensão análoga se apresenta quando se questiona se a IA Generativa "cria". O verbo criar, como o substantivo inteligência, é semanticamente polissêmico: pode designar desde a produção de algo radicalmente inédito, sem precedente no mundo anterior, até a combinação original de elementos preexistentes para produzir algo novo. Nenhum campo, seja a Psicologia, a Filosofia, o Direito, as Artes ou a Computação, possui uma definição consensual de criatividade (Carvalho; Pimentel, 2024, p. 260).

A posição mais cética sustenta que os sistemas de IA Generativa não criam, mas recombinam estatisticamente padrões aprendidos. Essa perspectiva encontra suporte na própria natureza do processo: o modelo aprende a antecipar qual elemento, palavra ou *pixel*, é mais provável em uma sequência, dada a distribuição dos dados de treinamento. Os LLMs, nesse registro, funcionam como um "papagaio estocástico" que se esforça para replicar informações adquiridas por meio de treinamento extensivo, mas que efetivamente não compreende o que está dizendo (Ray, 2023 *apud* Sampaio; Sabbatini; Limongi, 2024, p. 13). A geração seria, portanto, uma recombinação sofisticada, não um ato criativo genuíno.

A posição pragmática, por outro lado, propõe que a questão relevante não é a operação interna do sistema, mas o resultado percebido. O pesquisador Ahmed Elgammal, citado por Carvalho e Pimentel (2024, p. 263), conduziu experimentos nos quais participantes humanos avaliaram obras geradas por IA como mais intencionais, estruturadas, comunicativas e inspiradoras do que obras produzidas por artistas humanos. Se o resultado é percebido como criativo por observadores humanos, argumenta-se, o critério pragmático de criatividade está satisfeito. Retomando o raciocínio do Jogo da Imitação proposto por Alan Turing: se não se consegue distinguir se uma imagem ou um texto foi produzido por uma pessoa ou por uma IA generativa, então, em termos práticos, a distinção perde relevância operacional (Carvalho; Pimentel, 2024, p. 263).

Essa disputa não é apenas técnica: é, em larga medida, social e simbólica. Sob a perspectiva da Teoria das Representações Sociais (Moscovici, 2003, p. 40), a percepção pública de que a IA "cria" constitui, em si mesma, um processo de ancoragem e de objetivação. A ancoragem ocorre quando a sociedade associa os sistemas de IA a categorias preexistentes de artista, autor ou criador, categorias historicamente reservadas a sujeitos humanos dotados de intencionalidade. A objetivação ocorre quando essas atribuições ganham materialidade: a IA "escreve", "pinta", "compõe", e essas metáforas se sedimentam no discurso cotidiano como se descrevessem capacidades cognitivas equivalentes às humanas. Esse processo não é neutro:

ele influencia expectativas, decisões regulatórias, disputas de autoria e a própria autopercepção dos usuários diante do que a tecnologia pode ou não pode fazer.

A discussão conceitual sobre se a IA cria ou reproduz permanece, portanto, aberta. O que é possível afirmar com rigor é que a IA Generativa produz conteúdos novos que não existiam antes de sua execução, que esses conteúdos emergem de padrões estatísticos e não de intenção ou experiência vivida, e que a percepção social desse processo é, em grande medida, moldada pelas representações sociais que os grupos humanos constroem em torno da tecnologia. A polissemia de "criar", como a de "inteligência", não é um obstáculo a ser eliminado, mas uma janela para compreender como a sociedade se relaciona com seus artefatos.

6 RELAÇÕES ENTRE ANI, AGI E IA GENERATIVA

As três categorias apresentadas neste capítulo, ANI, AGI e IA Generativa, não são mutuamente exclusivas, tampouco formam uma hierarquia linear em que cada nível sucede o anterior como um estágio evolutivo. Compreender as relações entre elas exige, antes de tudo, identificar o critério pelo qual cada categoria é delimitada.

A ANI designa o paradigma funcional dominante: sistemas projetados para executar tarefas específicas em domínios delimitados, com alto desempenho dentro dessas fronteiras e baixa ou nula capacidade de generalização fora delas (Russell; Norvig, 2021). A AGI designa um horizonte teórico, ainda não alcançado por nenhum sistema existente, caracterizado pela capacidade de transferir conhecimento entre domínios, adaptar-se a situações não antecipadas e operar com flexibilidade cognitiva comparável à humana (Goertzel, 2014, p. 2). A IA Generativa, por sua vez, não é uma terceira categoria paralela às anteriores, mas uma subclasse da ANI: é estreita, especializada em geração de conteúdo, e não possui capacidade de raciocínio geral.

A relação entre as três categorias pode ser representada por inclusão e não por progressão. A IA Generativa está contida na ANI: todo sistema generativo atual é, tecnicamente, um sistema estreito, ainda que seu domínio seja a produção de linguagem ou de imagem. A AGI, por sua vez, está fora do conjunto da ANI não por ser superior em grau, mas por ser qualitativamente diferente em natureza. O Quadro 2 sintetiza as dimensões centrais de cada categoria.

Quadro 2. Comparativo entre ANI, AGI e IA Generativa

Dimensão	ANI	AGI	IA Generativa
Existência atual	Sim	Não (teórica)	Sim (subconjunto da ANI)
Domínio de atuação	Restrito e definido	Aberto e dinâmico	Restrito: geração de conteúdo
Transferência entre domínios	Baixa ou nula	Alta (capacidade central)	Baixa ou nula
Tipo de saída	Classificação, detecção, otimização	Generalista	Texto, imagem, áudio, código
Exemplos	Reconhecimento facial, diagnóstico por imagem	Sem exemplos operacionais	ChatGPT, DALL-E, Midjourney, Claude

Fonte: Elaborado pelo autor com base em Russell e Norvig (2021), Goertzel (2014) e WTTC (2024).

A compreensão dessas relações tem implicações práticas para o debate público e regulatório. Quando se afirma que um sistema de IA "pensa" ou "decide de forma autônoma", frequentemente confunde-se a sofisticação funcional da ANI, ou de um de seus subconjuntos generativos, com a capacidade de raciocínio geral própria de uma AGI que ainda não existe. Essa confusão, como demonstrado no capítulo 1, é em grande medida resultado dos processos de ancoragem e objetivação pelos quais a sociedade constrói representações sobre a tecnologia (Moscovici, 2003). Identificar com clareza em qual categoria um sistema se insere não é um exercício de pedantismo técnico: é uma condição para uma avaliação honesta de suas capacidades, de seus limites e dos riscos e benefícios que efetivamente apresenta.

7 CONSIDERAÇÕES PARCIAIS

O percurso percorrido ao longo deste capítulo teve um propósito deliberadamente classificatório: oferecer ao leitor um mapa conceitual das principais categorias com as quais a literatura organiza o campo da Inteligência Artificial. A distinção entre Inteligência Artificial Estreita (ANI), Inteligência Artificial Geral (AGI) e Inteligência Artificial Generativa cumpre essa função de modo eficaz, pois permite diferenciar o que os sistemas fazem, em que domínios operam e quais capacidades efetivamente possuem ou ainda não possuem. Essa clareza classificatória tem valor prático e científico inegável: ela impede a confusão entre o que existe hoje e o que permanece como horizonte teórico, e fornece uma linguagem comum para debates regulatórios, éticos e técnicos.

Contudo, classificar não é o mesmo que definir. E é precisamente aqui que o problema identificado no capítulo 1 retorna com toda a sua força. Por mais que a taxonomia ANI/AGI/Generativa organize o campo de modo funcional, ela não resolve a questão ontológica mais profunda que atravessa toda a discussão: o que é, afinal, a Inteligência Artificial? Classificar um sistema como estreito ou generativo descreve o que ele faz; não diz o que ele é.

Essa distinção entre classificar e definir não é trivial. Ela remete a um dos problemas mais antigos do pensamento intelectual: a busca pela essência das coisas, pela propriedade ou conjunto de propriedades que torna algo o que é, independentemente de suas manifestações particulares. Na filosofia, é a Ontologia que se ocupa dessa investigação, ao perguntar pelo ser enquanto ser, ao tentar identificar o que confere identidade a uma categoria de objetos e não apenas descrever seus comportamentos observáveis. Na linguística, a Semântica e, em particular, a Terminologia dedicam-se à tarefa correlata de construir conceitos unívocos para campos especializados, de modo que um termo carregue um único significado preciso dentro de um sistema de referência dado.

A Inteligência Artificial, enquanto objeto de definição, resiste a ambos os esforços. Do ponto de vista ontológico, ela não é uma entidade natural cujas propriedades essenciais possam ser descobertas por observação; é um artefato humano, cujas fronteiras são, em larga medida, decididas pelos próprios pesquisadores que a constroem e classificam. Wang (2019, p. 2) identificou esse problema com precisão ao observar que toda definição de IA é uma escolha normativa: ao definir o que conta como inteligência artificial, define-se também o que não conta, e essa delimitação orienta agendas de pesquisa, critérios de avaliação e políticas públicas. Definir é, portanto, uma forma de poder.

Do ponto de vista terminológico, a situação é igualmente problemática. A palavra inteligência carrega séculos de uso nas mais diversas disciplinas, da Filosofia da Mente à Psicologia, da Neurociência à Biologia Evolucionária. Quando a Ciência da Computação a apropria e lhe acrescenta o adjetivo artificial, não herda apenas um termo: herda toda a carga semântica e o campo de expectativas que esse termo carrega. É por isso que, como demonstrado no capítulo anterior, o público tende a antropomorfizar os sistemas de IA, a atribuir-lhes intenção, compreensão e criatividade. Não é ignorância: é a consequência previsível de se nomear uma tecnologia com palavras cujos referentes históricos são humanos.

Essa tensão tem implicações que vão muito além da linguagem. No plano científico, a ausência de uma definição consensual fragmenta o campo: pesquisadores que trabalham com aprendizado profundo, com sistemas simbólicos ou com modelos generativos operam, por

vezes, sob concepções tão distintas de IA que se torna difícil comparar resultados, estabelecer critérios de progresso ou delimitar o que constituiria uma falha ou um avanço genuíno. No plano filosófico, a questão permanece em aberto: os sistemas de IA pensam, ou apenas calculam? Compreendem, ou apenas processam? Criam, ou apenas recombinam? As taxonomias apresentadas neste capítulo não respondem a essas perguntas; elas as enquadram de modo mais preciso, o que já é um avanço considerável.

No plano sociológico, as implicações são igualmente relevantes. Como mostrou Moscovici (2003), os objetos que uma sociedade não consegue definir com clareza tendem a ser incorporados por processos de ancoragem e objetivação: associam-se a categorias conhecidas e recebem imagens concretas que tornam o desconhecido familiar. A Inteligência Artificial não escapa a esse mecanismo. Robôs pensantes, mentes digitais, assistentes que entendem: são todas metáforas que preenchem o vácuo deixado pela ausência de uma definição rigorosa e compartilhada. Esse processo não é inocente: ele molda expectativas, orienta investimentos, legitima usos e, por vezes, encobre riscos reais sob camadas de familiaridade ilusória.

A necessidade de definir não é, portanto, um exercício acadêmico secundário. Como argumenta Wang (2019), uma boa definição deve possibilitar a comunicação precisa entre pesquisadores, fornecer critérios para distinguir o que pertence e o que não pertence ao campo, orientar pesquisas produtivas e ser suficientemente simples para ser operacionalizável. Nenhuma das definições disponíveis atende plenamente a todos esses critérios simultaneamente, o que explica a persistência do debate conceitual. E é exatamente essa persistência que justifica o esforço deste trabalho: não para oferecer a definição definitiva, tarefa presumivelmente impossível, mas para cartografar o problema com rigor suficiente para que o leitor possa navegar nele de modo informado.

O capítulo seguinte abandona a perspectiva classificatória e desce ao nível dos mecanismos. Se até aqui a pergunta foi o que a IA é e o que ela faz, a partir do próximo capítulo a questão passa a ser como ela funciona: da representação matemática do conhecimento à predição de tokens, das operações em espaços vetoriais à arquitetura que tornou possível a revolução dos grandes modelos de linguagem. Compreender o funcionamento interno é indispensável para avaliar, com honestidade intelectual, tanto o que os sistemas de IA podem fazer quanto aquilo que, por razões estruturais e não apenas técnicas, eles provavelmente jamais farão.

REFERÊNCIAS

CORRADI, Guido; THEIRS, Cecilia; MARTÍNEZ-MARTÍ, Maria Luisa; ISERN-MAS, Carme; VILLAR, Sergio. Who Fears Generative Artificial Intelligence? Scale Development and Predictors of Fears Towards GenAI. **Scandinavian Journal of Psychology**, 2025. https://doi.org/10.1111/sjop.70037

DAHLKE, Johannes. **A.I. go by many names**: towards a sociotechnical definition of artificial intelligence. [S. l.]: arXiv, 2024. Disponível em: https://arxiv.org/pdf/2410.13452. Acesso em: 25 fev. 2026.

FOGEL, David B. **Defining artificial intelligence.** IEEE Computational Intelligence Magazine, Piscataway, v. 18, n. 1, p. 1–15, 2023.

GOERTZEL, Ben. Artificial General Intelligence: Concept, State of the Art, and Future Prospects. **Journal of Artificial General Intelligence**, Paris, v. 5, n. 1, p. 1–46, 2014. https://doi.org/10.2478/jagi-2014-0001

GOODFELLOW, Ian; BENGIO, Yoshua; COURVILLE, Aaron. **Deep Learning.** Cambridge: MIT Press, 2016.

KURT, Melike Palsu. **Introduction to Artificial Intelligence and Fundamental Concepts.** Ancara: Presidência da República da Türkiye – Escritório de Transformação Digital, 2021.

LEADERSHIP BUSINESS CONSULTING (LBC). **IA generativa**: conceitos e tendências. [S.l.]: LBC Innovative Transformation, jun. 2023. Disponível em: https://lbc-global.com/wp-content/uploads/2023/06/PT_Generative-AI_.pdf. Acesso em: 23 fev. 2026.

LEE, Kai-Fu. **Inteligência artificial:** como os robôs estão mudando o mundo, a forma como amamos, nos relacionamos, trabalhamos e vivemos. Tradução de Marcelo Barbão. Rio de Janeiro: Globo Livros, 2019.

McCARTHY, John. **What is Artificial Intelligence?** Stanford: Stanford University, 2007. Disponível em: http://www-formal.stanford.edu/jmc/whatisai/. Acesso em: 23 fev. 2026.

MOSCOVICI, Serge. **A máquina de fazer deuses:** sociologia e psicologia. Tradução de Aglaé de Souza. Rio de Janeiro: Imago, 1990.

MOSCOVICI, Serge. **Representações sociais:** investigações em psicologia social. Petrópolis: Vozes, 2003.

NUNES, Paulo (org.). **Apostila de Inteligência Artificial Generativa.** Recife: Egape, 2024.

PALSÜ KURT, Melike. **Introduction to Artificial Intelligence and Fundamental Concepts.** Ankara: Digital Transformation Office of the Presidency of the Republic of Türkiye, 2024.

PARLAMENTO EUROPEU. **What is artificial intelligence and how is it used?** Brussels: Directorate General for Communication, 20 jun. 2023. Disponível em:

https://www.europarl.europa.eu/topics/en/article/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used. Acesso em: 23 fev. 2026.

LEADERSHIP BUSINESS CONSULTING (LBC). **IA generativa: conceitos e tendências.** [S.l.]: LBC Innovative Transformation, jun. 2023. Disponível em: https://lbc-global.com/wp-content/uploads/2023/06/PT_Generative-AI_.pdf. Acesso em: 23 fev. 2026.

PIMENTEL, Mariano; CARVALHO, Felipe. **IA generativa e educação**: práticas e teorizações. Porto Alegre: Sociedade Brasileira de Computação, 2025. 343 p.. Disponível em: https://books-sol.sbc.org.br/index.php/sbc/catalog/view/182/812/1608. Acesso em: 25 fev. 2026.

TÜRKIYE. Presidency of the Republic of Türkiye Digital Transformation Office; Ministry of Industry and Technology. **National Artificial Intelligence Strategy 2021-2025**. Ankara: Digital Transformation Office, 2021. Disponível em: https://cdn-assets.inwink.com/b0269dea-af7b-4460-b29a-c40b9941c4c5/d32fa511-ac43-49ef-b5b2-d90034d2af03. Acesso em: 25 fev. 2026.

RUSSELL, Stuart J.; NORVIG, Peter. **Artificial Intelligence:** a modern approach. 4. ed. Global Edition. Harlow: Pearson, 2021.

SAHA, Dibbyo. **A Brief Introduction to Artificial Intelligence:** what is AI and how is it going to shape the future. Toronto: Ryerson University, 2020.

SAMPAIO, Rafael Cardoso; SABBATINI, Marcelo; LIMONGI, Ricardo. **Diretrizes para o uso ético e responsável da Inteligência Artificial Generativa:** um guia prático para pesquisadores. São Paulo: Intercom, 2024.

VASWANI, Ashish *et al.* **Attention Is All You Need.** [S.l.]: Google Brain/Research, 2023. (Original de 2017 atualizado em 2023).

WANG, Pei. **On defining artificial intelligence.** Journal of Artificial General Intelligence, Paris, v. 10, n. 2, p. 1–37, 2019.

WORLD TRAVEL & TOURISM COUNCIL (WTTC). **Introduction to Artificial Intelligence (AI) Technology:** guide for Travel & Tourism leaders. London: WTTC; Microsoft, jan. 2024.

CAPÍTULO 3

PRINCÍPIOS MATEMÁTICOS DA GERAÇÃO DE TEXTO EM INTELIGÊNCIA ARTIFICIAL

Andrey Nario de Souza Oliveira
Estudante do curso de Licenciatura em Matemática do Instituto Federal da Bahia-Campus Barreiras.

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia.

DOI: 10.52832/wed.209.1243



Resumo: Este trabalho aborda os fundamentos matemáticos e estruturais que viabilizam o funcionamento da Inteligência Artificial Generativa, com foco em sua capacidade de produção textual. O artigo encontra-se dividido em duas seções principais para fins didáticos e analíticos. A primeira parte explora a arquitetura estrutural da inteligência artificial, detalhando como os dados empíricos são convertidos em representações numéricas e, subsequentemente, organizados em vetores e *embeddings*, ferramentas matemáticas que mapeiam a semântica da linguagem em espaços multidimensionais. O texto também descreve os papéis operacionais dos neurônios artificiais, das matrizes de pesos e das camadas na edificação das redes neurais, processos que culminam no conceito de modelos treinados e introduzem a arquitetura *Transformer*. A segunda parte aprofunda-se nas ferramentas matemáticas que regem o aprendizado e a inferência, explicando a utilização de espaços vetoriais, o cálculo do produto escalar e a similaridade do cosseno como métricas para determinar a proximidade semântica entre os vetores. O estudo analisa o emprego contínuo de distribuições de probabilidade, notadamente por intermédio da função Softmax, para executar a predição autorregressiva de tokens em sequências textuais. Além disso, examina a aplicação do cálculo diferencial, com ênfase nas derivadas parciais e no vetor gradiente, utilizados para guiar a otimização de parâmetros mediante o algoritmo de retropropagação e a busca pela redução do erro apontado pela função de perda. Por fim, a investigação detalha o mecanismo de autoatenção, estruturado no processamento matricial de Consulta, Chave e Valor (*Query, Key, Value*), e promove uma discussão sobre as alucinações estocásticas, compreendidas como limites matemáticos e probabilísticos inerentes a esses sistemas de predição.

Palavras-chave: Inteligência Artificial Generativa. Redes Neurais Artificiais. Arquitetura *Transformer*; Processamento de Linguagem Natural. Mecanismo de Autoatenção.

Abstract: This work addresses the mathematical and structural foundations that enable the operation of Generative Artificial Intelligence, focusing on its capacity for textual production. The article is divided into two main sections for analytical purposes. The first part explores the structural architecture of artificial intelligence, detailing how empirical data is converted into numerical representations and subsequently organized into vectors and embeddings, which are mathematical tools that map the semantics of language into multidimensional spaces. The text also describes the operational roles of artificial neurons, weight matrices, and layers in building neural networks, processes that culminate in the concept of trained models and introduce the Transformer architecture. The second part delves into the mathematical tools governing learning and inference, explaining the use of vector spaces, dot product calculations, and cosine similarity as metrics to determine the semantic proximity between vectors. The study analyzes the continuous use of probability distributions, notably through the Softmax function, to execute the autoregressive prediction of tokens in textual sequences. Furthermore, it examines the application of differential calculus, emphasizing partial derivatives and the gradient vector, which are used to guide parameter optimization through the backpropagation algorithm and the pursuit of error reduction indicated by the loss function. Finally, the research details the self-attention mechanism, structured on the matrix processing of Query, Key, and Value, and promotes a discussion on stochastic hallucinations, understood as mathematical and probabilistic limits inherent to these prediction systems.

Keywords: Generative Artificial Intelligence. Artificial Neural Networks. Transformer Architecture. Natural Language Processing. Self-Attention Mechanism.

INTRODUÇÃO

É com espanto que vislumbramos hoje o fenômeno da Inteligência Artificial (IA) Generativa, talvez como os antigos que subiam os montes ou viajavam grandes distâncias apenas para consultar os oráculos. Faz-se uma pergunta e a máquina que tudo sabe, como se também tudo visse, responde o que se deseja saber. Essa dinâmica ressuscita uma antiga dependência humana por respostas, prontas e acabadas, transferindo para os algoritmos a autoridade sobre a verdade e o conhecimento que outrora pertenciam ao domínio dos deuses, verbalizados pelas pitonisas. Se no passado a palavra final vinha do sagrado, hoje a chancela da certeza provém das linhas de código. O risco iminente dessa nova devoção tecnológica pode residir na aceitação acrítica de seus vereditos, uma vez que a simulação aparente de coerência textual muitas vezes mascara imprecisões e reprojeta preconceitos históricos contidos nas bases de dados que alimentaram a máquina de saber.

Essa transferência de autoridade cognitiva ganha contornos contemporâneos quando espelhada na ficção científica, particularmente na emblemática consulta do personagem Neo ao Oráculo na trilogia cinematográfica *Matrix*. Na distopia, o protagonista busca orientação existencial diante de uma figura que exala sabedoria intuitiva e acolhimento humano, embora constitua, na realidade, um programa de inteligência artificial altamente sofisticado, projetado pela própria *matrix* para mapear as variáveis do comportamento social. O paralelo ficcional elucida o nosso cenário presente em que, ao interagirmos com as interfaces de linguagem generativa, frequentemente esquecemos que estamos diante de um construto algorítmico e passamos a tratá-la como uma entidade dotada de intenção e discernimento. Desse modo, o usuário moderno, tal como o herói do filme, depara-se com um espelho matemático que não possui consciência própria, mas que devolve ao sujeito, de forma refinada, reflexos e padrões da linguagem gerada pela humanidade.

A fim de transitar dessa percepção cultural e mítica para o rigor científico, o objetivo central deste capítulo consiste em discutir os fundamentos matemáticos que operam por trás da inteligência artificial (IA) generativa. Para os propósitos desta discussão, definimos a IA generativa como uma máquina projetada para produzir textos. Cumpre esclarecer que adotamos o vocábulo "texto" em sua mais ampla acepção. Sob essa ótica, a produção textual da máquina não se restringe apenas à linguagem escrita convencional, mas abarca de forma equivalente geração de músicas, síntese de imagens, criação de vídeos e elaboração de códigos de programação. A partir dessa premissa, o foco recairá sobre como os processos matemáticos

fornecem a estrutura lógica e operacional necessária para que os sistemas computacionais processem dados e os convertam nessas múltiplas formas de manifestação de linguagem.

Em estrita observância às diretrizes de transparência, idoneidade e boas práticas científicas regulamentadas pela Portaria CNPq nº 2.664, de 6 de março de 2026, que institui a Política de Integridade na Atividade Científica do Conselho Nacional de Desenvolvimento Científico e Tecnológico, os autores declaram formalmente o uso de ferramentas de Inteligência Artificial Generativa (IAG) na redação, na articulação conceitual e na estruturação do presente artigo.

O manuscrito foi desenvolvido por meio de um processo metodológico de cocriação e síntese computacional, distribuído conforme os seguintes critérios de finalidade:

- **Parte 1 (A Arquitetura da Inteligência Artificial):** A fundamentação descritiva e a redação estrutural dos tópicos associados ao fluxo de dados, vetores, *embeddings*, neurônios artificiais e redes neurais basearam-se prioritariamente em conteúdos gerados pelo modelo ChatGPT, desenvolvido pela OpenAI.
- **Parte 2 (A Matemática do Pensamento Artificial):** O detalhamento analítico das equações lineares, espaços vetoriais, produtos escalares, distribuições de probabilidade, cálculo diferencial de retropropagação e formulações matriciais de autoatenção do modelo *Transformer* baseou-se predominantemente em conteúdos gerados pelo modelo Gemini, desenvolvido pela Google.
- **Revisão e Suplementação:** O refinamento estilístico, a realização de correções ortográficas e sintáticas, o encadeamento de transições conceituais lineares, bem como a efetuação de acréscimos e supressões textuais pontuais foram processados com o auxílio da ferramenta Claude, desenvolvedora Anthropic.

A opção metodológica pelo emprego dessas tecnologias fundamenta-se em uma realidade constatada na literatura científica nacional. Se por um lado existem manuais didáticos tradicionais consolidados de Álgebra Linear, Cálculo e Estatística, e por outro existem manuais puramente técnicos de programação voltados à aplicação de bibliotecas prontas como *TensorFlow* ou *PyTorch*, a literatura para o leitor iniciante ou intermediário voltada a detalhar como as matrizes de pesos operam geometricamente sobre os vetores de *embedding* manifesta-se escassa em língua portuguesa. Encontrar materiais nacionais que expliquem a aplicação matemática rigorosa da similaridade do cosseno no mecanismo de autoatenção, ou como o gradiente mapeia a sensibilidade de bilhões de parâmetros em um espaço multidimensional, constitui um desafio exaustivo pela fragmentação das fontes informacionais.

Nesse cenário de escassez bibliográfica, as ferramentas de inteligência artificial generativa foram empregadas como instrumentos de mediação e democratização do conhecimento científico. O uso coordenado desses modelos permitiu consolidar dados dispersos e construir uma ponte formal, linear e acessível para que estudantes e professores de Matemática, dentre outros, possam compreender a estrutura lógica e analítica subjacente à inteligência artificial sem a necessidade de imersão prévia em códigos de programação ou em jargões computacionais. O sistema computacional atuou, portanto, como um equalizador editorial, auxiliando na tradução de estruturas algorítmicas complexas para uma narrativa pedagógica estruturada em português brasileiro.

Os autores reconhecem as limitações intrínsecas aos modelos estatísticos de geração de texto e assumem a existência de potenciais imprecisões residuais no mapeamento das variáveis multivariáveis. Diante disso, e em conformidade com as exigências de responsabilização outorgadas pela política institucional federal, os autores humanos assumem a autoria e a responsabilidade integral e irrestrita pelo conteúdo final deste artigo, respondendo legal e administrativamente pela fidedignidade das informações, pelo rigor conceitual das equações demonstradas e pela integridade ética de toda a extensão do texto publicado.

Com o objetivo de descrever os modelos generativos e democratizar o acesso à sua infraestrutura lógica, o presente artigo encontra-se estruturado em duas grandes divisões complementares. A primeira parte dedica-se à análise da arquitetura estrutural da inteligência artificial, mapeando o percurso dos dados desde sua conversão digital elementar até a organização em vetores e *embeddings*, cuja distribuição espacial fundamenta a geografia do significado das palavras. Nesta etapa, investiga-se o funcionamento dos neurônios artificiais como nós de filtragem, compreendendo como os pesos e as matrizes se articulam em camadas para edificar a estrutura de uma rede neural. A segunda parte aprofunda-se nas ferramentas analíticas que regem a dinâmica e o aprendizado desse sistema, abordando os espaços vetoriais multidimensionais através da distância euclidiana e do produto escalar, além de detalhar a similaridade do cosseno. Explora-se, na sequência, o cálculo estocástico aplicado a distribuições de probabilidade via função *Softmax* para a predição autorregressiva de termos, culminando no estudo do cálculo diferencial, em que derivadas parciais e gradientes alimentam o algoritmo de retropropagação para a otimização dos parâmetros. Por fim, a investigação demonstra a aplicação coordenada de todas essas ferramentas na engenharia de matrizes de consulta, chave e valor que constituem o mecanismo de autoatenção na arquitetura *Transformer*.

PARTE 1: A ARQUITETURA DA INTELIGÊNCIA ARTIFICIAL

1.1. Dados

A capacidade de os sistemas de inteligência artificial generativa produzirem textos, imagens e sons não emerge espontaneamente da máquina. Antes de qualquer modelo matemático complexo, existem os dados como elementos fundamentais sem os quais nenhum sistema inteligente poderia ser construído.

Em termos gerais, dados podem ser compreendidos como registros de fenômenos observáveis. Uma palavra escrita em um livro, uma fotografia armazenada em um servidor, uma página da internet ou uma operação financeira constituem exemplos de registros que preservam aspectos da realidade, convertendo acontecimentos, comportamentos e formas de comunicação em informações passíveis de armazenamento e processamento.

No contexto da inteligência artificial, os dados desempenham o papel de matéria-prima do aprendizado. Durante o treinamento, processo que será mais bem descrito posteriormente, o sistema é exposto a conjuntos massivos de registros produzidos por seres humanos ao longo de décadas, que incluem livros, artigos científicos, enciclopédias, códigos de programação e inúmeras outras formas de produção cultural. Quanto maior a diversidade desses registros, maior tende a ser a capacidade de o sistema identificar regularidades nos fenômenos observados. Essa dependência revela um aspecto frequentemente negligenciado que consiste no fato de que a qualidade do aprendizado está profundamente vinculada à qualidade dos dados utilizados, pois registros incompletos ou enviesados podem limitar a capacidade do sistema de representar adequadamente determinados aspectos da realidade.

Sob uma perspectiva epistemológica, a centralidade dos dados aproxima a inteligência artificial de uma tradição intelectual muito anterior ao surgimento dos computadores. Desde os primórdios da ciência moderna, a produção de conhecimento esteve associada à observação sistemática dos fenômenos e à busca de regularidades. A inteligência artificial reproduz, em escala computacional, esse mesmo movimento, pois parte de inúmeros exemplos concretos para extrair estruturas gerais capazes de orientar previsões e inferências.

Entretanto, os dados, por si sós, ainda não constituem conhecimento para a máquina. Palavras, imagens e sons precisam ser convertidos em representações manipuláveis por operações matemáticas, o que anuncia a próxima etapa do funcionamento da inteligência artificial que é a representação numérica dos dados.

1.2. Representação numérica de dados

Os dados constituem a matéria-prima da inteligência artificial, mas permanecem inacessíveis à máquina em sua forma original. Embora seres humanos atribuam significado imediato a palavras, imagens e sons, os computadores não possuem acesso direto a essas formas de representação. Tudo aquilo que os humanos interpretam como significado precisa ser convertido em números que a máquina seja capaz de manipular, processo conhecido como representação numérica dos dados.

O princípio que sustenta essa conversão é relativamente simples. Os computadores operam por meio de circuitos eletrônicos que distinguem dois estados físicos, convencionalmente representados pelos valores 0 e 1. Toda informação processada por uma máquina, independentemente de sua natureza, deve ser codificada a partir dessas unidades elementares. Textos, imagens, vídeos e sons tornam-se, em última instância, longas sequências numéricas armazenadas e manipuladas eletronicamente.

Cada modalidade de dado segue sua própria lógica de conversão. Na linguagem escrita, sistemas como o Unicode atribuem valores numéricos a letras, sinais de pontuação e símbolos de diferentes idiomas. Nas imagens, cada pixel possui valores que representam intensidade luminosa, tonalidade e cor. Nos sons, sequências de medições descrevem a variação da pressão sonora ao longo do tempo, convertendo a fala e a música em séries numéricas analisáveis computacionalmente.

Entretanto, a simples conversão em números ainda não é suficiente para o aprendizado. Uma longa sequência numérica continua sendo apenas uma coleção de valores isolados. Para que a inteligência artificial identifique relações e semelhanças entre os elementos observados, torna-se necessário organizar esses números em estruturas matemáticas mais adequadas à análise. É nesse ponto que surge o vetor como um dos conceitos centrais da matemática da inteligência artificial, o qual permite agrupar dados em entidades capazes de preservar relações entre diferentes características simultaneamente, estabelecendo a ponte entre o mundo das experiências humanas e o universo abstrato que sustenta o funcionamento da inteligência artificial.

1.3. Vetores

A conversão dos dados em números representa apenas o primeiro passo da matemática da inteligência artificial. Embora essa transformação torne os dados acessíveis ao computador,

ela ainda não resolve um problema fundamental: como organizar grandes quantidades de informações de maneira que suas relações internas possam ser preservadas e analisadas? Tratar cada valor isoladamente dificultaria a identificação de padrões mais complexos. Torna-se necessário, portanto, reunir esses números em estruturas capazes de representar simultaneamente diversas características de um mesmo objeto. É nesse contexto que surge o conceito de vetor.

Na Matemática, um vetor pode ser compreendido de maneira muito simplificada como uma coleção ordenada de números que permite armazenar múltiplas informações dentro de uma única estrutura. Considere uma pessoa descrita pelas características idade, altura e peso: o vetor (25; 1,75; 70) não representa três números independentes, mas uma única entidade composta por três atributos simultaneamente observados. Em sistemas de inteligência artificial, um único vetor pode conter centenas, milhares ou até milhões de componentes, cada uma descrevendo algum aspecto específico da informação analisada.

Além de organizar informações, os vetores possuem propriedades matemáticas que permitem sua manipulação sistemática por meio de operações rigorosamente definidas pela Álgebra Linear. O aspecto mais relevante para o funcionamento dos modelos modernos, entretanto, não é apenas o fato de os vetores armazenarem informações, mas a possibilidade de posicioná-los em um espaço matemático. Quando diferentes objetos são representados vetorialmente, torna-se possível medir proximidades e identificar semelhanças, pois objetos semelhantes tendem a ocupar regiões próximas desse espaço, enquanto objetos mais distintos tendem a se localizar em posições mais afastadas. Essa interpretação geométrica fornece uma das bases conceituais mais importantes da inteligência artificial contemporânea e permitirá, na etapa seguinte, compreender como a máquina calcula relações de similaridade entre diferentes informações.

1.4. Embeddings

A representação vetorial dos dados permite que informações complexas sejam convertidas em estruturas matemáticas manipuláveis. Entretanto, nem toda representação numérica é igualmente útil. Atribuir números arbitrários a palavras, imagens ou documentos não garante que as relações de significado presentes no mundo real sejam preservadas dentro do espaço matemático. Se as palavras “gato”, “cachorro” e “galáxia” recebessem valores aleatórios, a máquina não possuiria qualquer critério para perceber que as duas primeiras

pertencem a uma mesma categoria, enquanto a terceira se refere a um conceito completamente distinto.

A inteligência artificial moderna enfrenta esse problema por meio de uma técnica denominada *embedding*. Um *embedding* consiste em uma representação vetorial especialmente construída para preservar relações relevantes entre os dados, de modo que elementos semanticamente semelhantes ocupem posições próximas no espaço matemático. Essa posição não é escolhida arbitrariamente, mas resulta de um processo de aprendizado no qual o sistema observa grandes quantidades de dados e ajusta gradualmente as representações para refletir as regularidades encontradas. Palavras frequentemente utilizadas em contextos semelhantes, como "gato", "cachorro" e "animal", tendem a ocupar regiões vizinhas do espaço vetorial, enquanto palavras associadas a contextos muito diferentes localizam-se em regiões mais afastadas.

Sob essa perspectiva, os *embeddings* transformam o espaço vetorial em uma "geografia" do significado, na qual conceitos relacionados se aproximam e conceitos distintos se afastam. Em vez de armazenar definições explícitas ou regras previamente programadas, o sistema aprende a organizar conceitos segundo padrões estatísticos extraídos dos próprios dados. Esse mecanismo explica por que sistemas modernos conseguem reconhecer relações entre palavras, recuperar informações relevantes e produzir respostas coerentes, operando não sobre objetos do mundo real, mas sobre suas representações geométricas em um espaço de alta dimensionalidade. Para que essas relações possam ser efetivamente utilizadas, contudo, torna-se necessário desenvolver mecanismos capazes de transformar e combinar essas representações durante o aprendizado, tarefa que será desempenhada pelos neurônios artificiais.

1.5. Neurônios

Os *embeddings* permitem que palavras, imagens e outros objetos sejam representados como vetores organizados em um espaço matemático. Entretanto, representar informações não é o mesmo que processá-las. Para que a inteligência artificial reconheça padrões, faça previsões ou produza respostas, torna-se necessário transformar continuamente essas representações vetoriais. Essa função é desempenhada pelos neurônios artificiais.

O termo "neurônio" carrega uma herança histórica, já que a unidade recebeu esse nome por inspiração superficial nos neurônios biológicos, que também recebem sinais e produzem respostas. A semelhança, contudo, é apenas nominal. Um neurônio artificial é,

fundamentalmente, uma função matemática. Ele recebe um vetor de números como entrada, aplica sobre ele uma operação algébrica e produz um novo valor ou vetor como saída. Não há nada de biológico nesse processo, apenas transformações numéricas rigorosamente definidas.

Essa simplicidade é, paradoxalmente, sua maior virtude. Cada neurônio realiza uma operação elementar, mas a combinação de milhares ou milhões dessas funções em estruturas interligadas torna possível que o sistema identifique padrões complexos que não seriam perceptíveis de outra forma. Os padrões presentes nos dados raramente se encontram explicitamente visíveis, ou seja, ao atravessar sucessivas camadas de transformações matemáticas, as representações vetoriais são gradualmente reorganizadas até que regularidades antes ocultas se tornem reconhecíveis. Contudo, para que cada função possa determinar quais informações exercem maior ou menor influência sobre seu resultado, torna-se necessário introduzir um novo conceito que são os pesos, parâmetros numéricos responsáveis por orientar o comportamento do neurônio durante o processo de aprendizado.

1.6. Pesos e matriz de pesos

Os neurônios artificiais recebem vetores contendo informações provenientes dos dados previamente processados. Entretanto, nem todos os componentes desses vetores contribuem igualmente para o resultado. Em muitas situações, determinadas características são mais relevantes do que outras para a tarefa que está sendo executada. Surge, então, uma questão fundamental: como o neurônio determina quais informações merecem maior atenção?

A resposta encontra-se nos pesos.

Os pesos são valores numéricos associados às entradas de um neurônio artificial. Sua função consiste em regular a influência exercida pela informação durante o processamento matemático. Em termos simples, os pesos funcionam como coeficientes de importância. Quanto maior o peso atribuído a uma determinada entrada, maior será sua participação no cálculo realizado pelo neurônio. Da mesma forma, pesos menores reduzem a influência de determinadas características sobre o resultado produzido.

Considere um neurônio que recebe três informações de entrada. Embora todas participem do cálculo, elas podem possuir relevâncias distintas. O sistema pode aprender, por exemplo, que a primeira característica deve exercer forte influência sobre a decisão final, enquanto a segunda e a terceira possuem importância mais limitada. Essa diferenciação é representada matematicamente pelos pesos associados a cada entrada.

O aprendizado de uma rede neural consiste, em grande medida, no ajuste progressivo desses pesos. Durante o treinamento, o sistema compara suas respostas com os resultados esperados e modifica gradualmente os valores dos pesos na tentativa de reduzir os erros observados. Ao longo de milhões ou bilhões de exemplos, os pesos passam a incorporar as regularidades estatísticas presentes nos dados. Aquilo que frequentemente contribui para previsões corretas tende a receber maior valorização; aquilo que produz erros tende a ser ajustado ou enfraquecido.

Embora o conceito de peso seja relativamente simples quando analisado em um único neurônio, a realidade dos modelos modernos é muito mais complexa. Redes neurais contemporâneas podem conter milhões, bilhões ou até trilhões de parâmetros ajustáveis. Cada conexão entre neurônios possui seu próprio peso, formando uma gigantesca estrutura numérica responsável por armazenar tudo aquilo que o sistema aprendeu durante o treinamento.

Nesse contexto, torna-se impraticável tratar os pesos individualmente. É necessário organizá-los em estruturas matemáticas capazes de armazenar e processar grandes quantidades de parâmetros simultaneamente. Essa necessidade conduz ao conceito de matriz de pesos.

Uma matriz pode ser entendida, de modo simplificado, como uma tabela organizada de números distribuídos em linhas e colunas. Quando os pesos de diversos neurônios são reunidos em uma única estrutura, eles passam a formar uma matriz de pesos. Em vez de armazenar separadamente milhares ou milhões de valores individuais, a rede neural organiza esses parâmetros em arranjos matriciais que permitem operações matemáticas realizadas em larga escala.

A matriz de pesos desempenha um papel central na inteligência artificial moderna. Ela constitui o mecanismo pelo qual as representações vetoriais são transformadas à medida que atravessam a rede neural. Cada vez que um vetor de entrada é processado, seus componentes interagem com os pesos armazenados na matriz, produzindo uma nova representação que será enviada às etapas seguintes do sistema.

Sob essa perspectiva, os pesos podem ser compreendidos como a memória aprendida da rede neural. Tudo aquilo que o modelo descobre durante o treinamento tais como regularidades linguísticas, padrões visuais, associações conceituais e estruturas estatísticas acaba sendo incorporado aos valores armazenados nessas matrizes. Quando se afirma que um modelo possui bilhões de parâmetros, está-se essencialmente descrevendo a enorme quantidade de pesos distribuídos em suas matrizes internas.

Quando o treinamento é concluído, a matriz de pesos fica fixa. Ela não se altera durante o uso cotidiano do modelo. Cada versão de um modelo, como o GPT-4 ou o Claude, carrega suas matrizes congeladas, que representam tudo o que foi aprendido durante o treinamento com bilhões de exemplos.

O que varia conforme o *prompt* são as ativações, não os pesos. Quando você escreve uma pergunta, o texto é convertido em vetores e atravessa as matrizes de pesos fixas. Cada camada da rede produz representações intermediárias diferentes dependendo da entrada, mas os pesos que realizam essas transformações permanecem os mesmos para qualquer usuário, em qualquer conversa.

Assim, os neurônios fornecem as unidades básicas de processamento, os pesos determinam a influência relativa das informações e as matrizes de pesos organizam esses parâmetros em estruturas matemáticas capazes de operar sobre grandes volumes de dados. A combinação desses elementos torna possível a construção das camadas que, interligadas, formarão as redes neurais responsáveis pelo funcionamento dos sistemas modernos de inteligência artificial.

1.7. Camadas

Os neurônios artificiais, enquanto funções matemáticas, constituem as unidades fundamentais de processamento das redes neurais. Cada um deles recebe informações, aplica transformações matemáticas orientadas por seus pesos e produz uma saída. Entretanto, a capacidade de um único neurônio é extremamente limitada. Embora seja capaz de realizar cálculos simples, ele não possui complexidade suficiente para reconhecer padrões sofisticados ou modelar fenômenos complexos como a linguagem humana.

A solução encontrada pela inteligência artificial consiste em combinar grandes quantidades de neurônios em estruturas organizadas. Em vez de trabalharem isoladamente, os neurônios são agrupados em conjuntos denominados camadas.

Uma camada pode ser compreendida como um conjunto de neurônios que opera simultaneamente sobre as mesmas informações de entrada. Cada neurônio realiza seus próprios cálculos, utilizando pesos específicos, e produz uma saída particular. O conjunto dessas saídas forma uma nova representação dos dados, que será encaminhada para a etapa seguinte do processamento.

Essa organização estabelece um fluxo sequencial de transformação da informação. Os dados entram na primeira camada, são processados e convertidos em uma nova representação.

Essa representação é então enviada para a camada seguinte, onde passa por novas transformações. O processo repete-se sucessivamente até que o sistema produza um resultado.

Em uma analogia industrial, cada camada pode ser vista como uma etapa de uma linha de montagem. Em vez de fabricar objetos físicos, porém, as camadas refinam representações matemáticas. Cada estágio recebe uma versão dos dados produzida pela etapa anterior e a transforma em algo progressivamente mais adequado à tarefa que está sendo executada.

Nas redes neurais modernas, as primeiras camadas costumam identificar características mais simples e elementares dos dados. Em sistemas de processamento de linguagem, por exemplo, essas camadas podem capturar regularidades associadas a palavras, símbolos ou pequenas estruturas textuais. À medida que a informação atravessa camadas sucessivas, representações mais abstratas começam a emergir, envolvendo relações semânticas, contextuais e conceituais cada vez mais sofisticadas.

Esse processo gradual de abstração constitui uma das características mais importantes das redes neurais profundas. Em vez de tentar compreender diretamente um fenômeno complexo, o sistema o decompõe em uma sequência de transformações sucessivas. Cada camada acrescenta uma nova interpretação matemática aos dados recebidos, permitindo que padrões difíceis de detectar em um único estágio se tornem progressivamente mais evidentes.

Do ponto de vista matemático, cada camada realiza uma transformação vetorial orientada por matrizes de pesos. Os vetores produzidos por uma camada tornam-se os vetores de entrada da camada seguinte. Assim, o processamento completo de uma rede neural pode ser compreendido como uma longa cadeia de transformações matemáticas aplicadas sucessivamente sobre representações vetoriais dos dados.

A quantidade de camadas exerce influência direta sobre a capacidade do modelo de aprender padrões complexos. Redes com poucas camadas tendem a capturar apenas regularidades mais simples. Já arquiteturas compostas por muitas camadas conseguem construir representações progressivamente mais elaboradas, razão pela qual são frequentemente denominadas redes neurais profundas (*deep neural networks*).

Entretanto, uma camada isolada não possui significado funcional independente. Sua importância emerge da interação com as demais camadas que compõem a arquitetura. É a conexão organizada entre múltiplas camadas que dá origem à estrutura mais ampla conhecida como rede neural, responsável por coordenar todo o fluxo de processamento da informação dentro do sistema de inteligência artificial.

1.8. Rede neural

A compreensão isolada dos neurônios, dos pesos, das matrizes e das camadas permite identificar os componentes fundamentais da inteligência artificial moderna. Entretanto, nenhum desses elementos opera de maneira independente. A capacidade dos sistemas contemporâneos de reconhecer padrões, produzir previsões ou gerar textos emerge precisamente da integração dessas estruturas em uma organização mais ampla denominada rede neural.

Uma rede neural pode ser compreendida como um conjunto interconectado de camadas compostas por neurônios artificiais. Seu objetivo é transformar gradualmente as representações dos dados, conduzindo-as de uma forma inicial relativamente simples para representações cada vez mais abstratas e informativas. Em vez de executar uma única operação matemática, a rede realiza uma longa sequência de transformações sucessivas sobre os vetores que representam os dados.

O processo inicia-se quando os dados, já convertidos em representações vetoriais, são apresentados à rede. Esses vetores atravessam sucessivamente as diferentes camadas, sendo transformados a cada etapa pelas matrizes de pesos associadas aos neurônios. Cada camada produz uma nova representação, que se torna a entrada da camada seguinte. Ao final desse percurso, a rede produz uma saída que corresponde à sua interpretação ou previsão para o problema em questão.

A principal característica das redes neurais reside justamente nessa capacidade de realizar transformações encadeadas. Um único neurônio possui poder limitado de representação. Entretanto, quando milhares ou milhões deles são organizados em múltiplas camadas, o sistema torna-se capaz de modelar relações extremamente complexas. A inteligência observada nos modelos contemporâneos não decorre da sofisticação individual dos neurônios, mas da interação coletiva entre um grande número dessas unidades elementares.

Sob uma perspectiva matemática, uma rede neural pode ser entendida como uma composição de funções. Cada camada aplica uma transformação específica sobre os vetores recebidos. A saída produzida por uma função torna-se a entrada da função seguinte, formando uma cadeia de operações matemáticas interdependentes. O resultado final emerge da combinação de todas essas transformações realizadas ao longo do percurso da informação pela rede.

Essa estrutura permite que a rede construa níveis progressivos de abstração. Nas etapas iniciais do processamento, o sistema tende a capturar características mais simples dos dados. À medida que a informação avança pelas camadas, representações mais elaboradas começam a surgir. Em modelos de linguagem, por exemplo, as primeiras transformações podem estar associadas a padrões locais entre palavras, enquanto etapas posteriores passam a representar relações semânticas, contextuais e conceituais mais complexas.

É importante observar que uma rede neural não nasce sabendo executar tarefas específicas. Inicialmente, seus pesos são distribuídos de forma arbitrária, e suas respostas tendem a ser pouco úteis. A capacidade de produzir resultados coerentes surge apenas após um longo processo de aprendizado, durante o qual os pesos são ajustados repetidamente em função dos erros observados. Em outras palavras, a arquitetura da rede fornece a estrutura necessária para o aprendizado, mas o conhecimento propriamente dito é adquirido durante o treinamento.

As redes neurais modernas podem possuir dimensões extraordinárias. Modelos contemporâneos de inteligência artificial frequentemente operam com bilhões de parâmetros distribuídos por dezenas ou centenas de camadas. Apesar dessa enorme complexidade, todos eles permanecem fundamentados nos mesmos princípios elementares, quais sejam, vetores que representam informações, neurônios que realizam transformações matemáticas, pesos que regulam a influência das entradas e camadas que organizam essas operações em sequência.

Dessa forma, a rede neural constitui a arquitetura central da inteligência artificial contemporânea. Ela reúne todos os elementos apresentados até aqui em uma estrutura única capaz de processar informações, construir representações abstratas e modelar padrões presentes nos dados. Contudo, para que essa arquitetura se torne efetivamente capaz de aprender, ainda é necessário responder a uma questão fundamental: como a rede sabe se sua resposta está correta ou incorreta? A resposta para essa pergunta conduz ao próximo conceito da matemática da inteligência artificial que é o erro.

1.9. Erro

A rede neural constitui uma poderosa estrutura de processamento matemático, capaz de transformar sucessivamente representações vetoriais dos dados. Entretanto, por mais sofisticada que seja sua arquitetura, uma questão fundamental permanece em aberto: como o sistema sabe se produziu uma resposta correta?

Essa pergunta é central porque a simples produção de uma saída não implica aprendizado. Uma rede neural recém-inicializada possui pesos distribuídos de forma arbitrária e, consequentemente, suas respostas tendem a ser imprecisas. Para que o sistema possa melhorar seu desempenho ao longo do tempo, torna-se necessário medir a qualidade de suas previsões. É precisamente dessa necessidade que surge o conceito de erro.

Em termos gerais, o erro representa a diferença entre o resultado produzido pela rede neural e o resultado que seria considerado correto para uma determinada tarefa. Quanto maior essa diferença, maior o erro. Quanto menor a diferença, mais próxima a rede está da resposta desejada.

Considere, por exemplo, um sistema treinado para identificar animais em imagens. Ao analisar uma fotografia de um gato, a rede pode inicialmente concluir que existe apenas 30% de probabilidade de a imagem representar um gato. Se a resposta correta deveria ser próxima de 100%, existe uma discrepância significativa entre a previsão produzida e o resultado esperado. Essa discrepância constitui o erro.

A importância desse conceito está no fato de que o erro funciona como um mecanismo de orientação para o aprendizado. Sem ele, a rede neural seria incapaz de distinguir respostas adequadas de respostas inadequadas. O erro fornece uma medida objetiva da qualidade do desempenho do modelo, permitindo avaliar continuamente se as transformações realizadas pelos neurônios estão conduzindo a resultados melhores ou piores.

Do ponto de vista matemático, o erro não é apenas uma ideia abstrata, mas uma quantidade numérica calculável. Após cada previsão, a rede compara sua saída com o valor esperado e produz uma medida que expressa o grau de divergência entre ambos. Esse valor sintetiza, em uma única grandeza, o desempenho momentâneo do sistema.

A transformação do erro em número possui enorme relevância para a inteligência artificial. Afinal, aquilo que pode ser medido também pode ser otimizado. Uma vez que o erro se torna uma quantidade matemática, a rede passa a dispor de um critério objetivo para orientar modificações futuras em seus parâmetros.

Nesse sentido, o erro desempenha papel semelhante ao de um sistema de avaliação contínua. Após cada tentativa, o modelo recebe uma indicação quantitativa da qualidade de seu desempenho. Respostas mais próximas do resultado esperado produzem erros menores; respostas mais distantes produzem erros maiores. O objetivo do aprendizado passa a ser, então, a redução progressiva dessa medida.

A partir desse ponto, a inteligência artificial deixa de ser apenas um mecanismo de processamento e passa a constituir um sistema capaz de aperfeiçoar seu próprio

comportamento. A existência do erro fornece à rede uma direção para o aprendizado que visa reduzir continuamente a distância entre suas previsões e os resultados desejados.

Entretanto, identificar a existência de um erro não é suficiente para corrigi-lo. A rede ainda precisa descobrir quais pesos contribuíram para o resultado inadequado e de que maneira esses parâmetros devem ser modificados. Em outras palavras, após medir o erro, torna-se necessário determinar como ajustar a estrutura interna da rede para que erros semelhantes sejam menos prováveis no futuro. Essa tarefa conduz ao próximo conceito fundamental da matemática da inteligência artificial que é a função de erro.

1.10. Aprendizado

A existência do erro fornece à rede neural um critério para avaliar a qualidade de suas respostas. Entretanto, medir a distância entre uma previsão e o resultado esperado constitui apenas parte do processo. Para que ocorra aprendizado, é necessário que a rede utilize essa informação para modificar seu comportamento futuro.

No contexto da inteligência artificial, aprender não significa adquirir consciência, compreensão ou experiência subjetiva. O aprendizado corresponde a um processo matemático de ajuste progressivo dos parâmetros internos da rede neural. Em termos práticos, o sistema aprende quando altera seus pesos de modo a reduzir os erros observados durante o treinamento.

Esse processo pode ser compreendido como uma sucessão de tentativas e correções. Inicialmente, a rede produz respostas baseadas em pesos distribuídos de forma arbitrária. Como consequência, seus resultados tendem a apresentar elevados níveis de erro. Após cada previsão, o sistema compara sua resposta com o resultado esperado e utiliza essa diferença como referência para modificar seus parâmetros internos. Em seguida, realiza uma nova tentativa. O ciclo repete-se inúmeras vezes, promovendo ajustes graduais que tornam as previsões progressivamente mais precisas.

O aprendizado de uma rede neural não consiste na memorização literal dos exemplos observados. Em vez de armazenar cópias dos dados utilizados durante o treinamento, o sistema procura identificar regularidades presentes nesses exemplos, incorporando-as progressivamente nas matrizes de pesos. O que é efetivamente preservado não são os dados em si, mas os padrões estatísticos incorporados aos valores dos pesos distribuídos ao longo da rede.

Essa característica aproxima o aprendizado artificial de um processo de generalização. Ao observar uma grande quantidade de exemplos particulares, a rede procura construir uma estrutura matemática capaz de responder adequadamente a situações novas que não estavam presentes em seus dados de treinamento. Quanto mais eficaz for essa capacidade de generalização, maior será a utilidade prática do modelo.

Do ponto de vista matemático, o aprendizado pode ser entendido como um problema de otimização. A rede neural busca continuamente uma configuração de pesos que produza o menor erro possível diante dos exemplos apresentados. Cada ajuste realizado representa uma tentativa de aproximar o comportamento do sistema daquilo que se espera como resultado correto.

Nos modelos contemporâneos de inteligência artificial, esse processo ocorre em escalas extraordinárias. Durante o treinamento de grandes modelos de linguagem, bilhões de parâmetros podem ser ajustados repetidamente ao longo da análise de trilhões de palavras. Embora cada modificação individual seja extremamente pequena, o efeito acumulado de incontáveis correções sucessivas permite a construção de sistemas capazes de reconhecer padrões altamente complexos.

É justamente esse processo contínuo de redução de erros que confere à inteligência artificial sua aparente capacidade de adaptação. A rede não recebe explicitamente todas as regras necessárias para executar uma tarefa. Em vez disso, ela ajusta seus próprios parâmetros a partir da experiência fornecida pelos dados, refinando gradualmente sua capacidade de produzir previsões adequadas.

Todavia, permanece uma questão fundamental: de que maneira a rede determina quais pesos devem ser modificados e qual deve ser a intensidade de cada ajuste? Afinal, uma rede neural moderna pode possuir bilhões de parâmetros interconectados, tornando inviável qualquer processo de correção manual. A solução para esse problema encontra-se em um conjunto de técnicas matemáticas que permitem calcular a influência de cada parâmetro sobre o erro observado. Entre essas técnicas, destaca-se um conceito central para toda a inteligência artificial moderna que é o gradiente.

1.11. Modelo treinado

O processo de aprendizado permite que a rede neural ajuste progressivamente seus pesos em resposta aos erros observados durante o treinamento. Após milhões, bilhões ou até trilhões de correções sucessivas, o sistema atinge um estado no qual suas previsões se tornam

suficientemente consistentes para a tarefa que lhe foi atribuída. Quando esse estágio é alcançado, diz-se que a rede deu origem a um modelo treinado.

De maneira geral, um modelo treinado corresponde ao resultado final do processo de aprendizado. Ele não constitui um novo tipo de máquina nem uma estrutura diferente da rede neural originalmente construída. Trata-se da própria rede após seus parâmetros terem sido ajustados pela experiência acumulada durante o treinamento.

Essa distinção é importante porque uma rede neural, por si só, representa apenas uma arquitetura matemática. Inicialmente, seus pesos possuem valores arbitrários e seu comportamento tende a ser pouco útil. Somente após a exposição sistemática aos dados e aos ciclos sucessivos de correção de erros é que a arquitetura passa a incorporar informações relevantes sobre os padrões presentes no domínio analisado.

Pode-se dizer, portanto, que a diferença entre uma rede neural não treinada e um modelo treinado reside fundamentalmente nos valores de seus pesos. A estrutura permanece essencialmente a mesma; o que muda é o conteúdo matemático armazenado em seus parâmetros. Esses valores passam a refletir as regularidades identificadas durante o aprendizado e constituem aquilo que, de forma simplificada, costuma ser chamado de "conhecimento" do modelo.

Quando um usuário interage com um sistema de inteligência artificial, normalmente não está acessando o processo de treinamento em si. O que ele utiliza é o modelo já treinado. Nesse estágio, os pesos permanecem fixos e a rede passa a desempenhar sua função principal que é produzir previsões, classificações, interpretações ou respostas com base nos padrões que aprendeu anteriormente.

Essa característica ajuda a compreender uma das ideias centrais da inteligência artificial contemporânea. O aprendizado ocorre em um momento específico da vida do sistema, frequentemente exigindo enormes quantidades de dados, recursos computacionais e tempo de processamento. A utilização cotidiana do modelo corresponde a uma etapa distinta, na qual o sistema simplesmente aplica o conhecimento matemático previamente incorporado aos seus parâmetros.

Nos grandes modelos de linguagem, por exemplo, o treinamento pode envolver a análise de trilhões de palavras e o ajuste de bilhões de pesos distribuídos ao longo da rede neural. Ao final desse processo, todo o aprendizado acumulado encontra-se condensado nos parâmetros do modelo. Em vez de consultar continuamente os documentos utilizados durante o treinamento, o sistema utiliza os padrões estatísticos armazenados em seus próprios pesos para produzir novas respostas.

Sob essa perspectiva, um modelo treinado pode ser compreendido como uma representação matemática extremamente compacta das regularidades observadas nos dados. Embora não contenha cópias literais dos exemplos analisados, ele preserva relações, associações e estruturas estatísticas aprendidas ao longo do treinamento. É essa capacidade de sintetizar padrões que torna possível a geração de conteúdos inéditos a partir de conhecimentos adquiridos anteriormente.

Contudo, à medida que os conjuntos de dados se tornaram maiores e as tarefas mais complexas, as arquiteturas tradicionais de redes neurais passaram a enfrentar limitações significativas, especialmente no tratamento de informações distribuídas ao longo de sequências extensas de texto. A superação dessas limitações conduziu ao desenvolvimento de uma nova arquitetura que transformaria profundamente o campo da inteligência artificial que são os *Transformers*.

1.12. Transformers

O modelo treinado representa a consolidação do aprendizado adquirido pela rede neural ao longo de sua exposição aos dados. Entretanto, durante o avanço da inteligência artificial, tornou-se evidente que as arquiteturas tradicionais apresentavam dificuldades para lidar com algumas características fundamentais da linguagem humana. Entre essas limitações destacava-se a capacidade de compreender relações entre palavras distantes em textos extensos, bem como processar grandes volumes de informação de maneira eficiente.

A busca por soluções para esses desafios conduziu ao desenvolvimento de uma arquitetura denominada *Transformer*, apresentada em 2017 por pesquisadores da Google Research. Desde então, essa arquitetura tornou-se a base dos principais sistemas de inteligência artificial generativa utilizados atualmente.

É importante compreender que o Transformer não substitui os conceitos apresentados anteriormente. Vetores, *embeddings*, neurônios, pesos, matrizes, camadas, redes neurais e modelos treinados continuam presentes. O que muda é a forma como esses elementos são organizados para processar a informação. O *Transformer* pode ser entendido como uma arquitetura especializada de rede neural projetada para lidar com sequências de dados, especialmente textos.

Seu funcionamento começa da mesma maneira que nos modelos já discutidos. As palavras são convertidas em representações numéricas e organizadas na forma de *embeddings*. Esses vetores são então submetidos a sucessivas transformações matemáticas realizadas pelas

camadas da rede. A inovação do *Transformer* não está na eliminação dessas etapas, mas na introdução de mecanismos que permitem ao sistema analisar simultaneamente as relações existentes entre todos os elementos de uma sequência.

Nas arquiteturas anteriores, o processamento frequentemente ocorria de maneira sequencial, examinando uma palavra após a outra. Essa estratégia podia dificultar a identificação de relações estabelecidas entre termos separados por grandes distâncias dentro de um texto. O Transformer foi concebido justamente para superar essa limitação.

Seu principal mecanismo é conhecido como atenção (*attention*). De forma simplificada, a atenção permite que o modelo avalie a relevância relativa dos diferentes elementos presentes em uma sequência de dados. Quando analisa uma palavra específica, o sistema pode examinar simultaneamente as demais palavras do contexto e estimar quais delas exercem maior influência sobre sua interpretação.

Considere a frase: "O pesquisador apresentou o modelo porque ele produziu resultados promissores". Para compreender a quem o pronome "ele" se refere, torna-se necessário relacioná-lo a outros elementos da sentença. O mecanismo de atenção permite que o modelo identifique essas dependências contextuais sem precisar percorrer o texto de maneira estritamente linear.

Do ponto de vista matemático, a atenção opera por meio de operações realizadas sobre vetores e matrizes. O sistema calcula relações de similaridade entre diferentes representações vetoriais, atribuindo pesos que indicam o grau de influência exercido por cada elemento do contexto. Dessa forma, a própria rede aprende quais informações devem receber maior destaque durante o processamento.

Uma das consequências mais importantes dessa arquitetura é a capacidade de processamento paralelo. Em vez de analisar cada palavra individualmente e em sequência, o *Transformer* pode processar múltiplos elementos simultaneamente. Essa característica tornou possível o treinamento de modelos em escalas anteriormente impraticáveis, contribuindo para o surgimento dos grandes modelos de linguagem contemporâneos.

Modelos como os atuais sistemas de inteligência artificial generativa são construídos sobre arquiteturas *Transformer* compostas por numerosas camadas, bilhões de parâmetros e vastos conjuntos de dados. Apesar de sua dimensão extraordinária, permanecem fundamentados nos mesmos princípios matemáticos discutidos ao longo deste capítulo que é a representação numérica, vetores, transformações matriciais, pesos ajustáveis e aprendizado orientado pela redução de erros.

Sob essa perspectiva, o *Transformer* não representa uma ruptura com a matemática da inteligência artificial, mas o resultado de sua evolução. Ele reúne e reorganiza os elementos fundamentais apresentados anteriormente em uma arquitetura particularmente eficiente para modelar relações complexas presentes na linguagem humana. É essa combinação entre grandes volumes de dados, aprendizado estatístico e mecanismos de atenção que torna possível a geração de textos coerentes, contextualizados e linguisticamente sofisticados pelos sistemas de inteligência artificial contemporâneos.

PARTE 2: A MATEMÁTICA DO PENSAMENTO ARTIFICIAL

2.1. Espaços Vetoriais e a Geometria das Representações

Na etapa inicial deste estudo, estabeleceu-se que os dados são convertidos em sequências organizadas de números denominadas vetores. Contudo, para que essas estruturas possuam utilidade operacional, elas precisam estar localizadas em um ambiente matemático que viabilize sua manipulação. Esse ambiente é o espaço vetorial.

Um vetor pode ser definido como uma entidade matemática que pertence a um conjunto que obedece a oito axiomas. Para que um conjunto V seja um espaço vetorial sobre um corpo K, a exemplo do conjunto dos números reais, seus elementos, chamados de vetores, devem satisfazer oito axiomas básicos em relação às operações de adição e multiplicação por escalar.

Considerando u, v e w como vetores quaisquer em V, e α e β como escalares em K, os axiomas são:

Adição de vetores

1. Associatividade: O agrupamento não altera a soma.

$$(u + v) + w = u + (v + w)$$

2. Comutatividade: A ordem não altera a soma.

$$u + v = v + u$$

3. Elemento neutro: Existe um vetor nulo 0 em V que não altera os vetores na soma.

$$u + 0 = u$$

4. Elemento oposto: Para todo vetor u, existe um correspondente $-u$ cuja soma resulta no vetor nulo.

$$u + (-u) = 0$$

Multiplicação por escalar

5. Distributividade sobre a adição de vetores: O escalar distribui sobre a soma dos vetores.

$$\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$$

6. Distributividade sobre a adição de escalares: O vetor distribui sobre a soma dos escalares.

$$(\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u$$

7. Associatividade mista: Os escalares podem ser multiplicados entre si previamente.

$$\alpha \cdot (\beta \cdot u) = (\alpha\beta) \cdot u$$

8. Elemento neutro da multiplicação: O escalar 1 do corpo K preserva o vetor original.

$$1 \cdot u = u$$

Em particular, existe o espaço vetorial $\mathbb{R}^n$ que é definido como o conjunto de todas as n-uplas ordenadas de números reais. Se um elemento v pertence ao $\mathbb{R}^m$, então ele pode ser expresso na forma $v = (x_1, x_2, \ldots, x_n)$, em que cada coordenada x_i é um número real. A letra $\mathbb{R}$ indica que os elementos que compõem as coordenadas pertencem ao corpo dos números reais, enquanto o expoente n determina a dimensão do espaço, ou seja, a quantidade exata de coordenadas necessárias para descrever um vetor de forma unívoca nesse espaço.

A utilidade primordial do espaço vetorial $\mathbb{R}^m$ reside na sua capacidade de organizar e analisar grandes volumes de dados sob uma lógica unificada, atuando como uma linguagem tradutora para a ciência de dados e a inteligência artificial. Nesse contexto algorítmico, objetos do mundo real, como imagens, documentos de texto e históricos de navegação, são mapeados como vetores no $\mathbb{R}^m$, nos quais cada componente traduz uma característica específica da informação. Ao modelar esses dados em estruturas matriciais lineares, redes neurais e algoritmos de classificação utilizam a geometria desse espaço multidimensional para agrupar elementos semelhantes, separar categorias e estruturar predições analíticas. Essa conversão matemática é o que autoriza a extensão de conceitos geométricos, como distância métrica, ortogonalidade e projeção, para cenários analíticos abstratos de dimensão arbitrária, viabilizando o processamento computacional de sistemas complexos

Em sistemas de inteligência artificial, os *embeddings* de palavras ou imagens comumente habitam espaços de alta dimensionalidade, em que um único vetor pode possuir centenas ou milhares de dimensões.

A grande vantagem dessa abstração geométrica é que ela transforma conceitos qualitativos humanos em relações espaciais quantificáveis. No contexto do processamento de linguagem natural, por exemplo, o significado de uma palavra não é armazenado por meio de uma definição de dicionário, mas sim pela sua localização exata dentro desse espaço multidimensional. Para que o sistema opere sobre esses conceitos, torna-se necessário calcular

a relação espacial entre diferentes pontos, o que nos conduz diretamente ao conceito de distância.

2.2. Distância entre Vetores

Uma vez posicionados os vetores no espaço multidimensional, a inteligência artificial requer um critério matemático para determinar se dois elementos são semelhantes ou distintos. A forma direta de realizar essa validação é por meio do cálculo da distância geométrica entre os pontos.

A métrica mais comum utilizada para essa finalidade é a Distância Euclidiana. Em um espaço bidimensional, a distância entre dois pontos é medida de forma análoga à aplicação do Teorema de Pitágoras. Formalmente, dados dois vetores $u = (u_1, u_2, \dots, u_n)$ e $v = (v_1, v_2, \dots, v_n)$ pertencentes ao espaço $\mathbb{R}^n$, a distância euclidiana $d(u, v)$ é calculada pela raiz quadrada da soma dos quadrados das diferenças entre suas componentes correspondentes:

$$d(u, v) = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2}$$

Se a distância numérica entre o vetor que representa a palavra "gato" e o vetor da palavra "cachorro" for pequena, o sistema deduz que os termos compartilham afinidade contextual. Se a distância em relação ao vetor "galáxia" for grande, o modelo identifica a disparidade temática.

Contudo, a distância euclidiana pura apresenta limitações em espaços de alta dimensionalidade. Se dois textos tratam do mesmo assunto, mas um deles é muito mais longo do que o outro, os seus vetores resultantes possuirão magnitudes (comprimentos) muito diferentes, o que afastará os pontos geometricamente, mesmo que o conteúdo semântico seja idêntico. Para mitigar esse efeito e focar estritamente na direção do significado, utiliza-se uma operação algébrica subjacente que é o produto escalar.

2.3. O Produto Escalar

O produto escalar é uma operação fundamental da Álgebra Linear que combina dois vetores para gerar um número real (um escalar). Ele atua como uma ferramenta para determinar o quanto um vetor se alinha em relação a outro.

Matematicamente, o produto escalar entre dois vetores u e v em $\mathbb{R}^n$, denotado por $\langle u, v \rangle$ ou $u \cdot v$, pode ser calculado de duas maneiras complementares. A primeira delas é algébrica, multiplicando-se as componentes correspondentes e somando os resultados:

$$u \cdot v = \sum_{i=1}^{n} u_i \, v_i = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

A segunda definição é geométrica e relaciona o produto das magnitudes (comprimentos) dos vetores com o cosseno do ângulo θ formado entre eles:

$$u \cdot v = \|u\| \|v\| \cos \theta$$

em que $\|u\|$ e $\|v\|$ representam as normas (comprimentos) dos respectivos vetores.

Na arquitetura dos modelos baseados em *Transformers*, o produto escalar é a operação matemática que sustenta o mecanismo de atenção. Ao computar o produto escalar entre o vetor de uma palavra específica e os vetores de todas as outras palavras presentes em uma frase, o sistema avalia a força da associação entre elas.

Se dois vetores são perpendiculares entre si (formando um ângulo de 90°), o cosseno é nulo, resultando em um produto escalar igual a zero, o que indica total independência entre os dados. Se eles apontam na mesma direção, o produto escalar atinge seu valor máximo positivo para aquelas magnitudes.

2.4. Similaridade do Cosseno

A partir da formulação geométrica do produto escalar, os engenheiros de inteligência artificial derivaram uma das métricas de comparação mais eficazes para o processamento de dados que é a similaridade do cosseno. Como visto anteriormente, a distância euclidiana pode ser distorcida pelo tamanho dos vetores (magnitude), o que prejudica a análise de textos de extensões diferentes. A similaridade do cosseno soluciona esse problema ao isolar o comprimento dos vetores, avaliando estritamente o ângulo formado entre eles.

Isolando o termo $\cos \theta$ na equação do produto escalar, obtemos a fórmula da similaridade do cosseno:

$$\text{Sim}(u, v) = \cos \theta = \frac{u \cdot v}{\|u\| \|v\|}$$

Esta equação divide o produto escalar dos vetores pelo produto de suas normas, normalizando o resultado. Dessa forma, independentemente do comprimento dos vetores u e v, o valor da similaridade do cosseno ficará restrito ao intervalo entre -1 e 1:

- Valor igual a 1: O ângulo θ é de $0°$. Os vetores apontam exatamente para a mesma direção no espaço multidimensional, indicando máxima similaridade semântica.
- Valor igual a 0: O ângulo θ é de $90°$. Os vetores são ortogonais, indicando que as informações não possuem relação contextual ou estatística observável.
- Valor igual a -1: O ângulo θ é de $180°$. Os vetores apontam para direções perfeitamente opostas, denotando conceitos contrários ou inversos.

Ao adotar a similaridade do cosseno, a inteligência artificial ganha a capacidade de analisar e agrupar documentos, imagens ou termos correlacionados de forma matemática abstrata e padronizada. O sistema calcula a orientação angular dos *embeddings* e constrói a base analítica necessária para a próxima etapa do processo que é estimar, por meio de distribuições de probabilidade, qual informação possui maior fundamentação estatística para dar sequência ao fluxo de saída do modelo.

2.5. Probabilidade

A modelagem de sistemas generativos de linguagem exige ir além da representação estritamente geométrica fornecida pelos espaços vetoriais. Embora o mapeamento espacial organize a proximidade semântica dos elementos textuais por meio de relações angulares, ele não determina de forma individual a sequência textual a ser gerada a partir de uma ordem dada por um *prompt*. Para mitigar essa limitação e tratar a indeterminação inerente à coocorrência linguística, adota-se a teoria das probabilidades como modelo matemático fundamental, convertendo métricas de distância geométrica em funções de distribuição estatística sobre o vocabulário (Bishop, 2006).

Esse paradigma computacional fundamenta-se estritamente na análise quantitativa de dados em larga escala, operando à margem de processos cognitivos humanos ou de representações substituíveis de significado. Sob essa ótica, o processo de geração textual reduz-se a uma estimativa iterativa, cujo propósito consiste em calcular de maneira contínua e progressiva a probabilidade de ocorrência sequencial de cada termo a partir do conjunto de elementos que antecedem o ponto atual.

Neste cenário, a flexibilidade operacional do sistema depende diretamente da indeterminação estatística. Portanto, ela deve ser vista como uma propriedade matemática essencial, e não como um erro do algoritmo. A imposição de restrições determinísticas que limitassem a escolha a um único elemento inviabilizaria a adaptabilidade do modelo a variações de estilo, registro e domínio discursivo. Assim, a formalização probabilística introduz a

variabilidade necessária para que o modelo processe múltiplos caminhos inferenciais válidos, viabilizando a modelagem estocástica que precede a estruturação ativa de enunciados textuais.

O cálculo inicial desses pesos probabilísticos baseia-se na distribuição de frequências observada no *corpus* de treinamento. Os algoritmos processam padrões de coocorrência de termos em escalas massivas, traduzindo as regularidades da linguagem em matrizes de contagem e proporções matemáticas. Essa distribuição empírica inicial é submetida a sucessivas transformações matemáticas não lineares nas camadas profundas da rede neural, resultando em um mapeamento estatístico contínuo. Esse mapeamento formaliza o comportamento de todo o léxico do sistema, estabelecendo o suporte matemático para a estruturação de uma distribuição global de probabilidades sobre o vocabulário disponível.

2.6. Distribuições de Probabilidade

A modelagem quantitativa da incerteza converte o inventário lexical estático em uma distribuição de probabilidade contínua, estruturada sobre as dimensões do espaço vetorial (Alves, 2025). Para compreender a operacionalização desse processo, é preciso definir a unidade estrutural básica que o sistema processa, denominada *token*. No contexto da inteligência artificial generativa, o *token* atua como a menor partícula de informação textual reconhecida pelo algoritmo. Ele não equivale obrigatoriamente a uma palavra inteira. Dependendo da arquitetura de segmentação, um *token* pode representar um único caractere, um fragmento de palavra, um prefixo, um sufixo ou um vocábulo completo. Essa decomposição é o que permite ao modelo mapear o vocabulário de forma eficiente e construir matrizes representativas flexíveis, contornando limitações computacionais ao lidar com termos até então desconhecidos.

A partir dessa estruturação da linguagem em unidades fundamentais, cada *token* pertencente ao vocabulário do modelo é associado a um valor escalar de probabilidade, cujo cálculo reflete as regularidades estatísticas extraídas da base de dados de treinamento. Conforme postulam Jurafsky e Martin (2023), os grandes modelos de linguagem operam como estimadores dessa densidade de probabilidade. O sistema avalia o vetor de contexto e computa as relações de dependência no histórico recente da sequência de *tokens* já processados. O objetivo central desse cômputo contínuo é priorizar as combinações que mantenham a coerência sintática e a adequação semântica do texto resultante, traduzindo o rigor das equações em fluência linguística.

Esta distribuição de probabilidade não é calculada de forma isolada para termos individuais, mas gerada de maneira global sobre a totalidade do vocabulário do sistema. A atribuição de um valor de probabilidade não nulo a elementos de baixa frequência histórica previne a ocorrência de indeterminações matemáticas e o colapso operacional do algoritmo (Bishop, 2006).

No nível computacional das camadas finais da rede neural, o modelo calcula inicialmente os chamados *logits*, que representam pontuações brutas e não normalizadas para cada termo. Para converter esses valores em uma distribuição de probabilidade válida, na qual os valores sejam estritamente positivos e a sua soma seja igual a um (1), aplica-se a função *Softmax*:

$$P(t_{n+1} = k \mid t_1, t_2, \ldots, t_n) = \text{softmax}(\text{logits}_k) = \frac{e^{\text{logits}_k}}{\sum_{j=1}^{n} e^{\text{logits}_j}}$$

Por meio dessa formulação, logits_k apresenta-se como a pontuação bruta atribuída ao *token* k e a exponencial (e^{logits_k}) garante que o valor seja positivo. Além disso, a divisão pela soma de todas as exponenciais, calculada sobre todos os tokens *j* do vocabulário, garante que o resultado seja uma distribuição de probabilidade válida. O *token* com maior *logit* recebe a maior probabilidade, mas todos os demais ainda recebem alguma chance, o que é fundamental para o comportamento variável da geração textual.

A formalização algébrica dessas distribuições viabiliza a conversão de dados matriciais em sequências textuais ordenadas. A estruturação desse espaço de amostragem atua como a interface de transição entre a representação geométrica contínua dos *embeddings* e as determinações lógicas algoritmo. Uma vez consolidada essa distribuição de probabilidades sobre o vocabulário, o sistema dispõe de uma matriz numérica indexada que define as bases operacionais para a extensão algorítmica da sequência textual.

2.7. Predição da Próxima Palavra como Aplicação Prática do Modelo

A aplicação prática das distribuições probabilísticas concretiza-se no mecanismo de predição sequencial de *tokens*, executado de forma autorregressiva. A partir do recebimento de uma sequência inicial de vetores textuais, o modelo calcula a probabilidade condicional do símbolo subsequente por meio de equações formais. Esse processo determina que a escolha de cada novo elemento da sequência dependa estritamente do histórico cumulativo dos termos precedentes, configurando o processo de geração como um sistema dinâmico dependente do tempo (Jurafsky; Martin, 2023).

Sob a perspectiva formal, essa predição sequencial é modelada pela função de probabilidade condicional:

$$P(t_{n+1} | t_1, t_2 \ldots, t_n)$$

Na qual, a probabilidade do termo subsequente (t_{n+1}) é condicionada pela sequência integral de *tokens* anteriores ($t_1, t_2 \ldots, t_n$). O algoritmo opera em ciclos autorregressivos, nos quais o elemento predito em uma etapa é imediatamente incorporado ao vetor de contexto de entrada, redefinindo o cálculo para o passo subsequente.

Essa abordagem condicional supera as limitações estruturais que caracterizavam os modelos clássicos de processamento de linguagem natural. Os modelos baseados em *n*-gramas limitavam o tamanho do contexto analisado a uma janela fixa de termos, o que gerava um crescimento exponencial de parâmetros em função do tamanho do vocabulário, processo conhecido como a explosão combinatória. Para ilustrar, Pimentel e Carvalho (2025) mensuram este obstáculo a partir da língua portuguesa:

> Por exemplo, o Vocabulário Ortográfico da Língua Portuguesa (Volp), em sua versão de 2024, continha mais de 382.000 palavras. Se considerarmos apenas combinações de três palavras, teríamos $382.000^3 \approx 5{,}57 \times 10^{16}$ (55,7 quatrilhões) de possíveis trios, o que demandaria uma quantidade impraticável de tabelas para armazenar as probabilidades de coocorrência. Esse é o fenômeno da explosão combinatória que torna computacionalmente inviável a construção de modelos *n*-grama que considerem longas sequências de palavras.

A substituição da contagem explícita de frequências pelo uso de redes neurais artificiais profundas permitiu mapear as dependências contextuais em um espaço contínuo de parâmetros densos. Esse avanço metodológico, fundamentado por Bengio *et al.* (2003), demonstrou a viabilidade dos modelos probabilísticos neurais de linguagem. Em vez de armazenar tabelas de coocorrência de baixa densidade, as redes neurais aproximam a função de distribuição de probabilidade por meio de otimização paramétrica, o que habilita o sistema a computar dependências textuais de longo alcance e a preservar a consistência semântica global.

Desse modo, o processo iterativo expande progressivamente o vetor de entrada a cada ciclo de amostragem. De acordo com Brown *et al.* (2020), essa propriedade autorregressiva viabiliza a operação do modelo como um generalizador de tarefas a partir de poucos exemplos (conhecido na literatura como *few-shot learning*), convertendo formulações matemáticas estáticas

em um processo dinâmico impulsionado pelo processamento paralelo de dados (Análise Macro, 2024).

O principal desafio inerente a essa arquitetura reside na manutenção da consistência temática à medida que a extensão do texto gerado aumenta. Desvios infinitesimais na estimativa de probabilidade de um *token* inicial podem alterar a distribuição condicional dos passos subsequentes, afastando o modelo do domínio semântico pretendido. Torna-se necessário, portanto, regular a relação entre a precisão estatística da distribuição e a variabilidade das escolhas para evitar a convergência para loops repetitivos ou previsíveis (Holtzman *et al.,* 2020). Se o algoritmo selecionasse invariavelmente o termo de maior probabilidade absoluta, a saída resultante manifestaria redundância sistêmica.

Para mitigar essa previsibilidade e modular a variabilidade do processo de amostragem, insere-se o parâmetro de Temperatura (T), que redistribui os pesos da função *Softmax* de acordo com a relação de proporcionalidade:

$$P_T(t_{n+1} = k) \propto e^{\left(\frac{\text{logits}_k}{T}\right)}$$

Quando o parâmetro (T) tende a zero, as discrepâncias entre os *logits* tornam-se acentuadas, o que faz a amostragem convergir para um comportamento quase determinístico. Em contrapartida, valores elevados de (T) atenuam tais diferenças estatísticas, permitindo a seleção de elementos localizados nas regiões de menor densidade de probabilidade (Ackley; Hinton; Sejnowski, 1985).

Com o intuito de disciplinar a amplitude de tal variabilidade estocástica, o método de Amostragem por Núcleo (*Nucleus Sampling* ou *Top-p*) atua como um refinamento subsequente, restringindo o espaço amostral ao subconjunto mínimo de *tokens* cuja probabilidade cumulativa atinge um limiar de corte (p) predefinido (Holtzman *et al.*, 2020). Sob uma delimitação dinâmica, o número de candidatos elegíveis contrai-se ou expande-se conforme o grau de entropia do contexto, eliminando termos de alta improbabilidade causados por flutuações aleatórias e garantindo o rigor estrutural do texto produzido.

2.8. Função de Perda como a Métrica da Divergência Estatística

A estimação de distribuições probabilísticas consistentes em modelos de linguagem exige a imposição de um critério matemático objetivo para balizar o ajuste dos parâmetros da rede neural. A função de perda desempenha esse papel ao quantificar a divergência estatística entre as probabilidades estimadas pelo modelo e as ocorrências reais observadas no *corpus* alvo.

Essa formulação converte o desempenho do modelo em um valor escalar contínuo e diferenciável, estabelecendo que a redução sistemática dessa grandeza indica a convergência do algoritmo em direção à distribuição empírica dos dados humanos.

A escolha metodológica padrão para mensurar essa distância entre distribuições é a Entropia Cruzada, cuja formulação teórica provém da teoria da informação (Shannon, 1948). A Entropia Cruzada calcula a distância informacional entre a distribuição de probabilidade real do *corpus* e a distribuição predita pelas camadas de saída da rede, conforme expressa a equação:

$$L = -\sum_{i=1}^{n} y_i \cdot \log(\hat{y}_i)$$

Nessa modelagem, o valor da perda por entropia cruzada (L) é obtido a partir do indicador binário de classe (y_i), que assume valor igual a 1 para o token correto e 0 para os demais termos do vocabulário, multiplicado pelo logaritmo da probabilidade estimada pelo modelo ($\hat{y}_i$) para o i-ésimo termo.

O comportamento funcional dessa equação penaliza de forma assintótica o sistema quando este atribui baixa probabilidade ao elemento correto. Caso a rede emita uma estimativa incorreta com elevado grau de confiança, o logaritmo da probabilidade residual resulta em um valor negativo de alta magnitude que, multiplicado pelo sinal invertido da fórmula, eleva substancialmente o custo global. A minimização da entropia cruzada equivale, portanto, à otimização por máxima verossimilhança dos parâmetros do modelo.

Esse índice macroscópico resume a distância matemática do modelo em relação à fluência do agrupamento de dados, funcionando como o critério de controle que incentiva a generalização estatística aplicável a contextos inéditos. No entanto, a atualização efetiva das matrizes de parâmetros exige decompor esse índice global a fim de isolar a sensibilidade de cada coeficiente individual da arquitetura, o que introduz a necessidade do cálculo diferencial elementar.

2.9. Derivadas como Medida de Sensibilidade Local

Uma vez determinado o valor do erro global por meio da função de perda, torna-se necessário isolar o impacto individual de cada parâmetro na rede neural, visto que os modelos generativos contam com múltiplos coeficientes ajustáveis operando simultaneamente. O cálculo diferencial fornece as ferramentas para revelar o comportamento local de cada coeficiente (Strang, 2016) por meio de derivadas parciais, as quais avaliam as transformações lineares conduzidas em cada unidade da rede.

Sabendo que uma transformação linear básica é definida pela interação entre o vetor de entrada (x), a matriz de pesos sinápticos (W) e o vetor de viés ou *bias* (b), o vetor de saída de uma camada (y) pode ser formalizado pela equação matricial:

$$y = W \cdot x + b$$

A partir dessa base, a função de perda se consolida como uma função composta altamente complexa (Candido, 2023). A derivada parcial funciona como uma medida de sensibilidade local dentro da arquitetura do modelo, calculando a taxa de variação instantânea da perda total em resposta a uma modificação infinitesimal aplicada a um único peso de maneira isolada. Mantendo-se os demais parâmetros constantes, o sistema mensura com precisão a responsabilidade daquela conexão específica (Bishop, 2006), computando o impacto de cada peso (w_{ij}) individual na perda global (L) ao decompor o erro por meio do produto de derivadas parciais (Rumelhart; Hinton; Williams, 1986):

$$\frac{\partial L}{\partial w_{ij}} = \frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \cdot \frac{\partial z_j}{\partial w_{ij}}$$

Nessa expressão, (z_j) representa a combinação linear dos dados de entrada antes da ativação, dada por ($z_j = \Sigma w_{ij} \cdot x_i + b_j$) enquanto ($a_j$) corresponde ao valor de ativação resultante da aplicação da função não linear sobre essa combinação.

Sob a perspectiva geométrica, o cálculo de uma derivada parcial determina a inclinação da reta tangente à curva de erro naquele ponto do espaço hiperdimensional. Essa inclinação revela se um aumento no valor numérico de um peso específico fará a perda total subir ou descer. Caso a taxa de variação seja positiva, o algoritmo compreende que elevar aquele parâmetro ampliará o erro do modelo linguístico (Strang, 2016). Por outro lado, um resultado negativo indica que o incremento do peso contribuirá para aproximar a máquina da resposta correta esperada, permitindo mapear como as oscilações numéricas se propagam através das sinapses artificiais até alterarem a distribuição probabilística final (Jurafsky; Martin, 2023).

Para garantir a viabilidade desse diagnóstico matemático, toda a estrutura interna do modelo deve ser rigorosamente diferenciável. As funções de ativação empregadas para introduzir não linearidade na rede precisam apresentar derivadas bem definidas em todo o seu domínio operacional (Nogare, 2025). Modelos contemporâneos utilizam funções como a *Gaussian Error Linear Unit* (GELU), que emprega a função de distribuição acumulada da distribuição normal padrão (Phi(x)) multiplicada pela própria variável de entrada (x), sendo formalizada pela relação (Hendrycks; Gimpel, 2016):

$$\mathrm{GELU}(x) = x \cdot \varphi(x)$$

A função GELU atua na suavização dos valores negativos e evita descontinuidades abruptas, garantindo que o fluxo de informações sobre as taxas de variação permaneça estável durante o treinamento. A análise minuciosa de toda essa infraestrutura interna viabiliza a unificação das informações coletadas (Strang, 2016). Em vez de tratar as derivadas parciais como dados dispersos, a abordagem algébrica organiza esses valores em uma estrutura vetorial única e coordenada, consolidando o vetor gradiente.

2.10. Gradientes como Vetor de Direção do Erro

Ao reunir de forma ordenada todas as derivadas parciais calculadas para os parâmetros, o sistema consolida o vetor gradiente, representado por (∇L). O gradiente aponta a direção de maior crescimento do erro dentro do espaço de alta dimensionalidade. Conforme detalha Bishop (2006), essa ferramenta fornece a orientação geométrica necessária para guiar os algoritmos de ajuste de parâmetros, mapeando a topografia do erro para determinar a trajetória de otimização.

Para atenuar os erros de predição textual, o algoritmo conduz as modificações de seus pesos no sentido oposto ao indicado por esse vetor, movendo os parâmetros em direção ao ponto de mínimo da função de custo (Rumelhart; Hinton; Williams, 1986). Esse procedimento de ajuste iterativo é formalizado pela equação geral do gradiente descendente:

$$\theta_{t+1} = \theta_t - \alpha \cdot \nabla_{\theta_t} L(\theta_t)$$

Onde o conjunto de parâmetros atualizado (θ_{t+1}) é calculado subtraindo-se do conjunto de parâmetros no tempo atual (θ_t) o produto entre a taxa de aprendizagem (α), que controla a magnitude do passo de atualização, e o vetor gradiente da função de perda em relação a esses parâmetros ($\nabla_{\theta_t} L(\theta_t)$).

A precisão no cálculo desse vetor impede modificações estocásticas que poderiam desestabilizar os padrões já consolidados. Para tornar esse processo computacionalmente viável com volumes massivos de dados, adota-se o Gradiente Descendente Estocástico (SGD), que avalia o gradiente a partir de subconjuntos aleatórios de dados denominados *mini-batches*.

As variações locais medidas pelas derivadas parciais refletem a interdependência entre as diferentes camadas de processamento da rede neural. O gradiente captura essas reações em cadeia, garantindo que a rota de correção seja matematicamente consistente para todo o conjunto. O desafio seguinte consiste em transportar essa informação do gradiente ao longo de toda a extensão profunda da arquitetura de forma eficiente, aplicando o erro de maneira reversa entre as camadas por meio do algoritmo de retropropagação.

2.11. Retropropagação e a Aplicação da Regra da Cadeia

A distribuição sistemática das cores calculadas pelo gradiente ao longo de todas as camadas ocultas é viabilizada pelo algoritmo de retropropagação (*backpropagation*). Esse método baseia-se na aplicação sequencial da regra da cadeia do cálculo diferencial, permitindo o transporte da informação do erro de maneira reversa através do encadeamento de derivadas parciais (Rumelhart; Hinton; Williams, 1986). O fluxo analítico inicia-se na camada final de predição e caminha em direção aos componentes iniciais do sistema, garantindo que cada unidade de processamento receba o *feedback* adequado sobre sua contribuição para a perda (Bishop, 2006).

A aplicação da regra da cadeia resolve a dificuldade de calcular o impacto de parâmetros que não possuem ligação direta com a saída final do modelo. Por meio do produto sucessivo de derivadas locais, a retropropagação decomporá a influência de camadas profundas e intermediárias. Esse encadeamento matemático conecta operações complexas e distantes de forma unificada, cruzando a extensão da rede com rigor analítico (Strang, 2016).

A eficiência desse algoritmo viabilizou o treinamento de arquiteturas de Aprendizado Profundo (*Deep Learning*), caracterizadas pelo empilhamento de múltiplas camadas de neurônios artificiais. Esse arranjo permite que o modelo aprenda representações hierárquicas, nas quais as camadas iniciais isolam padrões elementares e as camadas subsequentes sintetizam conceitos abstratos e semânticos de alto nível (Lecun; Bengio; Hinton, 2015). Durante o percurso reverso, a informação do erro é refinada à medida que atravessa os arranjos matriciais de cada estágio de transformação, desfazendo as composições de funções operadas durante a etapa de processamento direto.

O cálculo reverso exige que as funções de ativação empregadas nos neurônios artificiais possuam derivadas bem definidas em todo o seu domínio operacional (Bishop, 2006). Se a derivada de uma função colapsar para zero em qualquer ponto desse refluxo, o aprendizado naquela região da arquitetura será interrompido permanentemente, motivo pelo qual se cuida da seleção de funções não lineares perfeitamente diferenciáveis (Nogare, 2025). A conclusão bem-sucedida desse fluxo de propagação reversa fornece os valores numéricos necessários para efetuar as modificações práticas nos parâmetros, direcionando o sistema para a fase de otimização global e consolidação do aprendizado.

2.12. Otimização e a Consolidação do Aprendizado

A atualização sistemática dos pesos operada a partir dos dados fornecidos pela retropropagação constitui a essência da otimização. Durante essa fase, o modelo busca encontrar a configuração ideal de parâmetros que apresente a menor perda possível diante dos exemplos do corpus. O treinamento é formalizado como um problema de Estimação por Máxima Verossimilhança (MLE), cujo objetivo consiste em encontrar o conjunto otimizado de parâmetros (θ^*) que maximiza o somatório do logaritmo da probabilidade condicional $P(t_i \mid \theta)$ de o modelo gerar os tokens textuais alvo (t_i) observados na base de treinamento (Bishop, 2006):

$$\theta^* = \arg\max_\theta \sum_{i=1}^{N} \log P(t_i|\theta)$$

Estabelecida a convergência paramétrica, o desempenho global do sistema sobre uma sequência textual é avaliado por meio da Perplexidade (PP), que quantifica o nível de incerteza do estimador (Jurafsky; Martin, 2023) através da equação:

$$PP(T) = \exp\left(-\frac{1}{N}\sum_{i=1}^{N} \log P(t_i|t_1, \ldots, t_{i-1})\right)$$

Nessa métrica, a perplexidade do texto (T) é dada pela exponencial do inverso negativo do número total de tokens na sequência (N) multiplicado pelo somatório do logaritmo da probabilidade condicional do token (t_i) dado o seu contexto histórico esquerdo ($t_1, t_2, t_3, \ldots, t_{i-1}$).

Modelos generativos modernos utilizam técnicas de otimização estocástica avançadas, como o algoritmo Adam (Kingma; Ba, 2015), que substitui o SGD puro ao introduzir o conceito de momento matemático por meio do cálculo de médias móveis dos gradientes anteriores. O Adam ajusta automaticamente a taxa de aprendizagem de forma individual para cada parâmetro, conferindo estabilidade ao processo e evitando que o sistema fique retido em mínimos locais superficiais.

Adicionalmente, aplicam-se técnicas de regularização como o *Dropout*, que desativa unidades neurais aleatoriamente, e o *Weight Decay*, que penaliza conexões complexas demais. Tais mecanismos funcionam como diretrizes probabilísticas prévias (conhecidas na estatística como *priors* bayesianas) com o objetivo de mitigar o *overfitting*, ou seja, o cenário em que o sistema apenas memoriza os dados de treinamento em vez de generalizar seus padrões. A partir de tal proteção, a consolidação definitiva do aprendizado estabiliza as matrizes de parâmetros

do modelo, dispensando a consulta direta aos documentos originais durante a inferência (Pimentel; Carvalho, 2025).

Com a arquitetura devidamente otimizada, o sistema torna-se apto a operar sobre sequências textuais. Contudo, para gerenciar contextos extensos sem perda de coesão, o processamento puramente linear dá lugar a mecanismos de ponderação dinâmica baseados em atenção (Vaswani *et al.*, 2017), os quais viabilizam o cálculo de relevância entre termos distantes, preservando o sentido global e o rigor estrutural do texto produzido.

2.13. Mecanismo de Atenção e a Ponderação Contextual de Vetores

A superação das limitações de memória e representação que afetavam os modelos de linguagem tradicionais ocorreu com a introdução do mecanismo de atenção (Vaswani *et al.*, 2017). Essa ferramenta opera como uma ponderação contextual que calcula o grau de relevância mútua entre todos os componentes de uma sequência de texto. As arquiteturas baseadas em Redes Neurais Recorrentes (RNNs) e redes de Longa Memória de Curto Prazo (LSTMs) processavam os tokens de forma estritamente linear e sequencial, mantendo uma variável de estado oculto para carregar e sintetizar o histórico (Hochreiter; Schmidhuber, 1997). Esse processamento linear e sequencial causava uma perda gradual nos sinais de aprendizado do algoritmo, o que comprometia a retenção de informações localizadas em posições distantes do texto.

O mecanismo de atenção elimina essa barreira de processamento linear ao permitir a modelagem de dependências diretas entre palavras distantes, processando todos os tokens simultaneamente (Vaswani *et al.*, 2017). Ao analisar um termo específico, o sistema avalia simultaneamente todas as outras palavras presentes no documento, estimando o impacto de cada uma delas (Análise Macro, 2024). Essa dinâmica transforma a leitura sequencial em um mapeamento global de conexões semânticas, permitindo ao modelo resgatar o contexto necessário para manter a coesão estrutural.

A classificação central dessa tecnologia é o conceito de Autoatenção (*Self-attention*), que executa o cálculo de relevância de forma direta sobre a própria sequência de entrada (Vaswani *et al.*, 2017). Em vez de depender de estados ocultos temporais, a rede neural mantém o acesso integral a todos os pontos do texto simultaneamente. Essa abordagem confere aos modelos baseados em *Transformers* a capacidade de mapear correlações de longo alcance, reduzindo significativamente o tempo de processamento através de computação massiva em paralelo.

Inerente a esse processo, a atenção funciona como uma ponderação que direciona os recursos computacionais para os fragmentos textuais mais informativos, aprendendo quais

associações entre palavras possuem maior valor estatístico para preservar a coesão da narrativa (Jurafsky; Martin, 2023). Essa flexibilidade geométrica permite que as representações vetoriais (*embeddings*) se tornem contextuais, ou seja, a representação de uma mesma palavra sofre alterações dinâmicas baseadas nas palavras vizinhas, resolvendo problemas de ambiguidade e polissemia (Devlin *et al.*, 2019). Para que esse cálculo ocorra de forma estruturada, o sistema projeta os dados em novos espaços lineares por meio de matrizes de disposição específicas.

2.14. Self-Attention Como Produto Escalar Matricial Em Larga Escala

O funcionamento prático do mecanismo de autoatenção baseia-se na geração de três projeções matriciais distintas a partir da matriz de entrada X. Essa engenharia matemática, introduzida por Vaswani et al. (2017), projeta os dados em matrizes denominadas Consulta (Q), Chave (K) e Valor (V). Para discriminar o escopo operacional e as atribuições analíticas dessas projeções lineares na dinâmica de mapeamento conceitual, a Tabela 1 detalha os respectivos comportamentos funcionais estabelecidos na arquitetura.

Tabela 1: Especificações operacionais e propriedades dos vetores Query, Key e Value no Transformer

Vetor (Inglês)	Termo Em Português	Propriedade Operacional No Transformer
Query (Q)	Consulta	Vetor que define o que um determinado *token* busca identificar nos demais elementos da sequência para mapear sua própria relevância contextual.
Key (K)	Chave	Vetor que atua como o rótulo de indexação de cada *token*, servindo de critério de correspondência para as consultas (*queries*) disparadas pelo sistema.
Value (V)	Valor	Vetor que armazena o conteúdo semântico real do *token*, o qual é extraído e ponderado matematicamente quando há afinidade entre a *Query* e a *Key*.

Fonte: Elaborado pelos autores, fundamentado em Vaswani et al. (2017).

Essas matrizes são obtidas por transformações lineares da entrada multiplicada por matrizes de pesos específicos. Desse modo, a matriz de Consulta é formalizada pela equação ($Q = X \cdot W^Q$), na qual a matriz de entrada X é multiplicada pelos pesos específicos de consulta (W^Q). De forma análoga, as projeções de Chave e Valor são estabelecidas pelas relações estruturais ($K = X \cdot W^K$) e ($V = X \cdot W^V$), em que (W^K) e (W^V) representam as matrizes de projeção de pesos a serem aprendidas durante o treinamento para as operações de Chave e Valor, respectivamente.

A operação matemática subsequente consiste em computar o produto escalar entre a Consulta e a Chave transposta para gerar uma matriz de afinidade que mensura o alinhamento

geométrico e a semelhança semântica entre todos os pares de palavras de forma simultânea. Considerando a dimensionalidade dos vetores de chave, representada por (n_k), como o determinante para o fator de escala dado por $(\sqrt{n_k})$ os escores normalizados dessa matriz de afinidade são submetidos à função Softmax para gerar uma distribuição probabilística de atenção, convertendo os valores brutos em pesos que somam exatamente 1 (Bishop, 2006; Jurafsky; Martin, 2023). Esses fatores são multiplicados pela matriz de Valor (V), consolidando a equação global (Vaswani *et al.*, 2017):

$$\text{Attention}(\mathrm{Q}, \mathrm{K}, \mathrm{V}) = \text{softmax}\left(\frac{\mathrm{Q} \cdot \mathrm{K}^{\mathrm{T}}}{\sqrt{\mathrm{n_k}}}\right)$$

Na prática, o modelo utiliza a Atenção Multi-Cabeça (*Multi-Head Attention*), que fragmenta o espaço de representação total em (h) subespaços lineares menores. Cada um desses subespaços constitui uma cabeça de atenção independente, configurando uma unidade analítica autônoma que permite ao sistema focar simultaneamente em diferentes tipos de relações sintáticas e semânticas ao executar essas operações em paralelo. As saídas individuais geradas por cada cabeça são posteriormente concatenadas e projetadas de volta ao espaço dimensional original, conforme estabelece a função:

$$\text{MultiHead}(\mathrm{Q}, \mathrm{K}, \mathrm{V}) = \text{Concat}(\text{head}_1, \ldots, \text{head}_\mathrm{h})$$

Nesse arranjo, os pesos da projeção linear final responsável por unificar as diferentes cabeças de atenção computadas em paralelo são denotados pela matriz W^O. Devido ao fato de o mecanismo de atenção puro tratar os tokens como um conjunto estocástico sem ordem intrínseca, faz-se necessária a introdução de uma Codificação Posicional, a exemplo dos *Rotary Position Embeddings* (RoPE), antes do processamento.

Tal integração garante que a informação sobre a ordem cronológica e a sequência das palavras seja preservada nas operações de produto escalar, permitindo resolver ambiguidades posicionais com alto nível de precisão (Su *et al.*, 2024). Essa multiplicação matricial em larga escala aproveita o potencial das unidades de processamento paralelo, servindo como base para a estruturação de uma rede profunda que unifica as transformações globais do sistema.

2.15. A Arquitetura Transformer como Composição Global de Funções

Ao conectar as etapas estudadas, a arquitetura *Transformer* se consolida como uma composição global de funções matemáticas que operam de maneira encadeada. Modelos de linguagem autoregressivos utilizam uma estrutura do tipo *decoder-only*, caracterizada pelo processamento da esquerda para a direita. Cada bloco dessa pilha executa transformações

sucessivas expressas pelas equações de conexão residual e normalização de camada (*Layer Norm*):

$$x = x + Multi\ Head\ Attention(LayerNorm(x))$$

$$x = x + FFN\big(LayerNorm(x)\big)$$

Nas quais x atua como a matriz de ativações intermediárias que flui de maneira aditiva pela rede, enquanto (FFN) define a operação da rede *Feed-Forward* interna que processa cada posição de forma isolada, funcionando como uma memória associativa que retém os fatos aprendidos (Geva *et al.*, 2021). A atenção causal implementada nesse bloco aplica uma máscara sobre as posições futuras, impedindo que o modelo antecipe os elementos subsequentes durante a predição.

Após o pré-treinamento estatístico, o modelo passa por uma fase de alinhamento com preferências por meio de técnicas como o Aprendizado por Reforço com Feedback Humano (RLHF) ou a Otimização Direta de Preferências (DPO). O desempenho final obtido por essa composição de funções segue Leis de Escala (*Scaling Laws*) empiricamente definidas. Utilizando constantes associadas à arquitetura (E, A, B) e expoentes de decaimento (α, β), a perda (L) decresce de forma previsível por meio de uma lei de potência em função do número de parâmetros (N) e do volume de tokens de treinamento (D) (Kaplan *et al.*, 2020; Hoffmann *et al.*, 2022):

$$L(N, D) \approx E + \frac{A}{N^{\alpha}} + \frac{B}{D^{\beta}}$$

Na fase de inferência, o vetor contextualizado final passa pela última camada linear para gerar os escores brutos (*logits*), que são convertidos em probabilidades pela função *Softmax*. Para modular a previsibilidade e mitigar repetições monótonas na escolha do próximo token, insere-se o parâmetro de Temperatura (T) para redistribuir os pesos probabilísticos de acordo com a relação de proporcionalidade previamente detalhada na Seção 2.7 (Ackley; Hinton; Sejnowski, 1985). Por meio desse ajuste, o hiperparâmetro (T) altera a magnitude dos *logits* antes da normalização exponencial, controlando diretamente a distribuição de probabilidade do elemento subsequente.

Quando (T) se aproxima de zero, o modelo converge para escolhas determinísticas, enquanto valores elevados ampliam a entropia da distribuição. Esse ajuste é refinado pelo método de Amostragem por Núcleo (*Nucleus Sampling* ou *Top-p*), que restringe o espaço amostral aos tokens cuja probabilidade acumulada atinja um limiar p predefinido, descartando candidatos aberrantes e preservando a coesão de textos extensos (Holtzman *et al.*, 2020).

A naturalidade observada no discurso gerado decorre desse controle estatístico das distribuições condicionais, manipulando as correlações numéricas extraídas das bases de dados (Pimentel; Carvalho, 2025). Falhas nessa predição iterativa, contudo, podem desencadear cascatas de erros conhecidas como alucinações, nas quais o modelo gera respostas factualmente incorretas com alto grau de confiança (Ji *et al.*, 2023). Esse fenômeno constitui uma limitação estrutural do modelo estocástico cuja tarefa primordial é computar plausibilidades estatísticas baseadas em ocorrências prévias.

2.16. Alucinações Estocásticas e os Limites da Predição Probabilística

A transição da amostragem probabilística para a geração autônoma de sequências textuais longas expõe o principal desafio estrutural das arquiteturas autorregressivas. Como o objetivo primordial do sistema é maximizar a plausibilidade linguística, o modelo frequentemente gera alucinações. Formalmente, a alucinação em inteligência artificial caracteriza-se como uma resposta emitida com elevado grau de confiança, mas sem fundamentação factual no corpus de treinamento (Gemini, 2026). Esse desvio ocorre porque a rede prioriza a coerência sintática e o encadeamento gramatical em detrimento da fidedignidade histórica ou científica das afirmações.

Esse fenômeno vincula-se à mecânica de predição iterativa token a token, na qual a probabilidade do próximo termo, expressa por $P(t_{n+1})$, é calculada de forma estritamente condicionada ao histórico de elementos anteriores, representados por $(t_1, t_2, \ldots, t_n)$. Sob essa dinâmica, se o modelo seleciona um *token* factualmente impreciso devido a flutuações na amostragem por núcleo ou pela influência de uma temperatura T elevada, esse elemento passa a integrar o contexto condicionante das iterações subsequentes. Configura-se uma reação em cadeia em que a estimativa inicial incorreta contamina os cálculos probabilísticos dos passos seguintes, gerando uma cascata de alucinações cumulativas à medida que a resposta se estende (Gemini, 2026).

A raiz dessa limitação reside no fato de que a tarefa fundamental de um modelo de linguagem consiste na previsão estatística da próxima palavra baseada em padrões lexicais abstratos, o que faz o sistema operar de forma alheia a uma ordem da verdade estável ou a um referencial ontológico externo (Neves, 2024). A convergência dos parâmetros obtida pela minimização da entropia cruzada assegura apenas que as sequências geradas sejam verossímeis dentro da distribuição geométrica do corpus humano, sem garantir precisão factual. O modelo avalia a continuidade do discurso por critérios de máxima verossimilhança e correlação

multidimensional, tornando o erro factual indistinguível de uma estrutura gramaticalmente perfeita.

Diante desse cenário, investigações matemáticas contemporâneas indicam que as alucinações não constituem falhas temporárias de programação ou defeitos de engenharia passíveis de correção por filtragens adicionais. Demonstra-se que a existência desses desvios é matematicamente insuperável e intrínseca à própria natureza da inferência probabilística em alta dimensionalidade (Gonzàlez, 2026). A mesma mecânica matemática que faculta à arquitetura Transformer sua flexibilidade e capacidade de generalização contextual impossibilita a erradicação completa do comportamento confabulatório.

REFERÊNCIAS

ACKLEY, D. H.; HINTON, G. E.; SEJNOWSKI, T. J. A learning algorithm for Boltzmann machines. **Cognitive Science**, v. 9, n. 1, p. 147–169, 1985. Disponível em: https://onlinelibrary.wiley.com/doi/abs/10.1207/s15516709cog0901_7. Acesso em: 30 mai. 2026.

ALURA. Transformers: funcionamento e benefícios. [S. l.]: **Alura**, 2023. Disponível em: https://www.alura.com.br/artigos/transformers. Acesso em: 14 jun. 2026.

ALVES, Erika. Um breve estudo dos Transformers. **Medium**, 2025. Disponível em: https://erika-gl-alves.medium.com/um-breve-estudo-dos-transformers-6abbf1b77512. Acesso em: 14 jun. 2026.

ANÁLISE MACRO. O que é e como funciona a Arquitetura Transformer. [S. l.]: **Análise Macro**, 2024. Disponível em: https://analisemacro.com.br. Acesso em: 18 mai. 2026.

A TERRA É REDONDA. Alucinação da Inteligência Artificial. **A Terra é Redonda**, 2026. Disponível em: https://aterraeredonda.com.br/alucinacao-da-inteligencia-artifici al/. Acesso em: 14 jun. 2026.

BENGIO, Yoshua; DUCHARME, Réjean; VINCENT, Pascal; JANVIN, Christian. A neural probabilistic language model. **Journal of Machine Learning Research**, v. 3, p. 1137–1155, 2003. Disponível em: https://www.jmlr.org/papers/v3/bengio03a.html. Acesso em: 30 mai. 2026.

BISHOP, C. M. Pattern Recognition and Machine Learning. New York: **Springer**, 2006. Disponível em: https://www.microsoft.com/enus/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf. Acesso em: 30 mai. 2026.

BRAINS - Brazilian AI Networks. Os Transformers Ilustrados. **Brains.dev**, 2023. Disponível em: https://brains.dev/2023/os-transformers-ilustrados/. Acesso em: 14 jun. 2026.

BROWN, Tom B.; MANN, Benjamin; RYDER, Nick; SUBBIAH, Melanie; KAPLAN, Jared; DHARIWAL, Prafulla; NEELAKANTAN, Arvind; SHYAM, Pranav; SASTRY, Girish; ASKELL, Amanda; AGARWAL, Sandhini; HERBERT-VOSS, Ariel; KRUEGER, Gretchen; HENIGHAN, Tom; CHILD, Rewon; RAMESH, Aditya; ZIEGLER, Daniel M.; WU, Jeffrey; WINTER, Clemens; HESSE, Christopher; CHEN, Mark; SIGLER, Eric; LITWIN, Mateusz; GRAY, Scott; CHESS, Benjamin; CLARK, Jack; BERNER, Christopher; MCCANDLISH, Sam; RADFORD, Alec; SUTSKEVER, Ilya; AMODEI, Dario. Language models are few-shot learners. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 33., 2020, Vancouver. **Anais [...]**. Red Hook: Curran Associates, 2020. p. 1877–1901. Disponível em: https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf. Acesso em: 30 mai. 2026.

CALLIOLI, Carlos Alberto; DOMINGUES, Hygino H.; COSTA, Roberto C. F. Álgebra linear e aplicações. 6. ed. São Paulo: **Atual**, 1990.

GONZÀLEZ, Guillem Alsina. Alucinações da IA são matematicamente insuperáveis. **SAPO.pt**, 2026. Disponível em: https://sapo.pt/artigo/alucinacoes-da-ia-sao-matematicamente-insuperaveis-6981a5594db45ec3bcd75c0e. Acesso em: 14 jun. 2026.

HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1997. Disponível em: https://dl.acm.org/doi/10.1162/neco.1997.9.8.1735. Acesso em: 30 maio 2026.

HOLTZMAN, Ari; BUYS, Jan; FORBES, Maxwell; CHOI, Yejin. The curious case of neural text degeneration. In: International Conference on Learning Representations, 2020, Addis Ababa. Anais [...]. [S. l.]: **OpenReview.net**, 2020. Disponível em: https://openreview.net/forum?id=rygGQyrFvH. Acesso em: 30 maio 2026.

JURAFSKY, D.; MARTIN, J. H. Speech and Language Processing. 3. ed. (draft). [S. l.]: **Prentice Hall**, 2023. Disponível em: https://web.stanford.edu/~jurafsky/slp3/. Acesso em: fev. 2025.

NOGARE, Diego. Função de Ativação - O cérebro matemático das Redes Neurais Artificiais. **Diego Nogare**, 2025. Disponível em: https://diegonogare.net/2025/12/funcao-de-ativacao-o-cerebro-matematico-das-redes-neural-artificiais/. Acesso em: 14 jun. 2026.

PIMENTEL, Mariano; CARVALHO, Felipe. IA generativa e educação: práticas e teorizações. Porto Alegre: **SBC**, 2025. Disponível online: https://iaeducacao.uniriotec.br. Acesso em: 17 mai. 2026.

SHANNON, Claude Elwood. A mathematical theory of communication. **The Bell System Technical Journal,** v. 27, n. 3, p. 379–423, 1948. Disponível em: https://ieeexplore.ieee.org/document/6773024. Acesso em: 30 mai 2026.

STRANG, Gilbert. Introduction to Linear Algebra. 5. ed. Wellesley: **Wellesley-Cambridge Press**, 2016. Disponível em: https://math.mit.edu/~gs/linearalgebra/. Acesso em: 01 jun. 2026.

VASWANI, Ashish; SHAZEER, Noam; PARMAR, Niki; USZKOREIT, Jakob; JONES, Llion; GOMEZ, Aidan N.; KAISER, Lukasz; POLOSUKHIN, Illia. Attention is all you need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 30., 2017,

Long Beach. **Anais [...]**. Red Hook: Curran Associates, 2017. p. 5998–6008. Disponível em: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html. Acesso em: 27 mai. 2026.

WIKIPEDIA. Alucinação (inteligência artificial). In: Wikipédia: a enciclopédia livre. [S. l.], 2026. Disponível em: https://pt.wikipedia.org/wiki/Alucina%C3%A7%C3%A3o_(intelig%C3%AAncia_artificial). Acesso em: 14 jun. 2026.

PARTE 2

INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO

CAPÍTULO 4

PERSPECTIVA HISTÓRICA DA TECNOLOGIA NA EDUCAÇÃO: DA INVENÇÃO DA ESCRITA À INTELIGÊNCIA ARTIFICIAL

A HISTORICAL PERSPECTIVE ON TECHNOLOGY IN EDUCATION: FROM THE INVENTION OF WRITING TO ARTIFICIAL INTELLIGENCE

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1244



Resumo: O texto investiga a relação intrínseca entre educação e tecnologia sob a perspectiva da *longue durée* de Fernand Braudel e da antropologia filosófica. Defende-se que a técnica não é um elemento externo ou recente, mas uma condição ontológica do humano; o ser humano se constitui através da exteriorização do pensamento em suportes exossomáticos, o que Bernard Stiegler denomina "memória terciária". A análise percorre cinco "ondas tecnológicas": a escrita, que externalizou a memória e inaugurou o pensamento linear ; a imprensa, que permitiu a reprodutibilidade técnica e a fundação da escola moderna industrial ; o audiovisual, que expandiu o alcance geográfico e sensorial do ensino ; a internet, que quebrou o monopólio da informação e introduziu a conectividade ubíqua ; e a inteligência artificial generativa, que externaliza o raciocínio discursivo e desafia a própria natureza do ato de conhecer. Conclui-se que, embora a tecnologia automatize tarefas de execução, o núcleo irredutível da educação permanece sendo o "trabalho da memória", a atribuição de sentido e o julgamento ético, processos que exigem a presença e o esforço consciente do sujeito humano.

Palavras-chave: Educação. Tecnologia. *Longue Durée*. Inteligência Artificial. Memória Terciária.

Abstract: This text investigates the intrinsic relationship between education and technology from the perspective of Fernand Braudel's *longue durée* and philosophical anthropology. It argues that technics is not an external or recent element but an ontological condition of being human; humanity constitutes itself through the exteriorization of thought onto exosomatic supports, a process Bernard Stiegler terms "tertiary memory". The analysis traces five "technological waves": writing, which exteriorized memory and inaugurated linear thinking ; the printing press, which enabled technical reproducibility and the foundation of the modern industrial school; audiovisuals, which expanded the geographical and sensory reach of teaching ; the internet, which broke the monopoly on information and introduced ubiquitous connectivity ; and generative artificial intelligence, which exteriorizes discursive reasoning and challenges the very nature of the act of knowing. It concludes that while technology automates execution tasks, the irreducible core of education remains the "work of memory," the attribution of meaning, and ethical judgment—processes that require the presence and conscious effort of the human subject.

Keywords: Education. Technology. *Longue Durée*. Artificial Intelligence. Tertiary Memory.

1 INTRODUÇÃO

A trajetória da educação confunde-se com a própria história da técnica, consolidando uma trama em que o ato de ensinar e os suportes de registro do saber constituem-se como faces de uma mesma moeda. Essa afirmação, aparentemente simples, encerra uma densidade histórica que desafia qualquer análise fundada no tempo curto dos eventos e das inovações pontuais. Para compreendê-la em sua radicalidade, é preciso adotar o que Fernand Braudel (Tomich, 2011) denominou *longue durée*.

Trata-se de uma perspectiva temporal que recusa a sedução do efêmero e volta o olhar para as estruturas quase imóveis que subjazem à superfície agitada dos acontecimentos. Sob essa lente, a história da educação revela não uma sucessão de rupturas tecnológicas, mas uma

permanência estrutural perturbadora: desde a primeira marca impressa numa parede de caverna até os algoritmos generativos da inteligência artificial contemporânea, o ato de externalizar o pensamento em um suporte material permanece como o gesto fundador e irredutível de toda pedagogia. Para Stiegler (1998), é precisamente essa exteriorização que constitui o humano enquanto tal: não há humanidade anterior à técnica, pois ambas emergem no mesmo movimento.

Como aponta Tomich (2011, p. 39) em sua leitura de Braudel, a *longue durée* não deve ser confundida com imobilidade absoluta. Trata-se, antes, de uma "relação histórica mais ou menos estável" que oferece o "terreno mais abrangente para a interpretação histórica", o pano de fundo contra o qual as variações cíclicas e os eventos de curto prazo ganham inteligibilidade. Aplicada ao campo da educação e da técnica, essa perspectiva revela que o artefato muda, como mudou da pedra ao papiro, do códice ao livro impresso, do quadro-negro ao *tablet*, mas a estrutura profunda permanece.

Ora, que estrutura permanece?

Permanece a necessidade humana de projetar a inteligência para além do corpo biológico, de romper as barreiras do esquecimento e do isolamento espacial por meio de suportes exossomáticos, definidos por Stiegler (1998, p. 177) como o prosseguimento da vida por meios não-vivos, mediante a inscrição da memória em matéria inorgânica. O que se altera no tempo médio das conjunturas é a forma do suporte; o que persiste na longa duração é a função estrutural que esse suporte desempenha na constituição mesma do saber e de sua transmissão.

Essa permanência estrutural, contudo, não existe no vácuo. Ela se inscreve em relações sociais determinadas e é precisamente aqui que a análise convoca o instrumental conceitual marxiano. Em Marx (1982), a distinção entre infraestrutura e superestrutura oferece uma chave interpretativa de grande potência. A infraestrutura compreende o conjunto das forças produtivas e das relações de produção; a superestrutura abarca as instituições jurídicas, políticas, culturais e educacionais que sobre ela se erguem.

A escola, enquanto instituição, pertence à superestrutura; os artefatos técnicos que suportam o ato pedagógico, enquanto instrumentos materiais inscritos nas relações de produção de uma época, participam da infraestrutura. No entanto, a tensão entre essas duas dimensões é mais complexa do que uma simples determinação mecânica da segunda pela primeira, já que os suportes técnicos da educação não apenas refletem a base material de cada sociedade, mas também a reproduzem, transmitindo, por meio dos conteúdos e das formas

pedagógicas que viabilizam, as condições de inteligibilidade do mundo que cada época considera legítimas.

A escrita cuneiforme, o *scriptorium* medieval, a imprensa de Gutenberg e a plataforma digital são, simultaneamente, produtos de estruturas sociais específicas e agentes ativos na perpetuação ou transformação dessas mesmas estruturas. A título de exemplo, Stiegler (1998) analisa precisamente a passagem da escrita cuneiforme para a escrita linear como uma mutação no regime de conservação da memória coletiva, demonstrando que cada novo suporte não apenas registra o saber, mas reorganiza as condições de sua produção e transmissão.

O cruzamento entre a *longue durée* braudeliana e a dialética marxiana infraestrutura/superestrutura permite, assim, uma abordagem que recusa tanto o determinismo tecnológico, ou seja, a ideia de que cada nova tecnologia revoluciona necessariamente as práticas educativas, quanto o idealismo pedagógico da ilusão de que a escola opera em uma esfera autônoma, imune às condicionantes materiais de cada época. O que se propõe aqui é, antes, uma história de ritmos descompassados.

Assim sendo, enquanto os artefatos técnicos inovam no tempo curto e médio, a estrutura profunda da relação educação-técnica persiste, e a instituição escolar, enraizada na longa duração, digere lentamente e frequentemente frustra as promessas de transformação que cada novo objeto tecnológico traz consigo.

Este capítulo se debruça sobre essa simbiose constitutiva entre educação e técnica, investigando como os objetos técnicos de cada época moldaram as práticas pedagógicas e de que maneira a mente humana se expandiu e se contraiu ao amparar-se em suportes exossomáticos historicamente situados, o que Stiegler (1998) conceitua como o processo pelo qual cada época técnica reconfigura as formas de retenção e transmissão da memória coletiva.

A análise percorre, pois, em perspectiva de longa duração, as grandes inflexões dessa relação, desde os primeiros sistemas de escrita até as tecnologias digitais contemporâneas, atentam tanto às permanências estruturais quanto às rupturas conjunturais, e sempre vigilante para a assimetria entre o tempo da inovação técnica e o tempo lento da mudança institucional e pedagógica.

Iniciaremos este percurso compreendendo a tecnologia não como um adendo instrumental, mas como uma interface de ampliação cognitiva. Ao longo da discussão, será demonstrado que cada inovação como, por exemplo, do estilete sumério que sulcou a argila aos algoritmos contemporâneos que processam a linguagem natural, desempenhou a função de estender o alcance dos sentidos e a capacidade de retenção da memória de trabalho, constituindo o que Stiegler (1998) denomina memória terciária, ou seja, a inscrição técnica que

libera o pensamento da dependência exclusiva da memória biológica. Tais tecnologias intelectuais, conforme proposto por Pierre Lévy (1993), não apenas facilitaram o acesso ao dado, mas reconfiguraram a própria ecologia cognitiva e as formas de comunicação entre as gerações.

A análise avançará, ao longo do capítulo, para a sistematização das grandes ondas tecnológicas, examinando as mutações educacionais provocadas pela escrita, pela reprodutibilidade técnica da imprensa e pela conectividade ubíqua da rede. Mais do que uma sucessão cronológica de inventos, essas ondas representam transformações estruturais na dinâmica da sala de aula e na própria epistemologia do papel docente, culminando no advento da cognição sintética. Nesse contexto, a inteligência artificial emerge como um novo interlocutor que, embora herde as promessas históricas de democratização e eficiência, impõe desafios inéditos de escala e personalização radical.

Por fim, o texto conduz a uma reflexão ética e antropológica sobre o que permanece essencialmente humano diante da crescente delegação de tarefas intelectuais às máquinas. Através do diálogo com o "trabalho da memória" e os dilemas da motivação na aprendizagem, exploraremos se a dependência desses auxiliares cognitivos resulta em uma redução da profundidade do pensar ou se libera o sujeito para níveis superiores de reflexão. Trata-se de uma análise crítica sobre o sentido do ato educativo em um cenário de automação técnica, onde o desafio primordial permanece a atribuição de significado e a construção da identidade frente ao conhecimento.

2 DA RELAÇÃO ONTOLÓGICA ENTRE EDUCAÇÃO E TECNOLOGIA

É recorrente, nos debates contemporâneos sobre educação, o uso da palavra tecnologia como sinônimo exclusivo de computadores, *smartphones* e internet. Professores são instados a incorporar tecnologia em suas aulas, como se ela fosse um elemento externo e recente que precisasse ser adicionado a uma prática pedagógica que, sem ela, existiria em estado puro. Essa concepção, no entanto, além de equivocada historicamente, empobrece a própria compreensão do que significa ensinar e aprender.

Tal prática esquece, pois, que a escrita cuneiforme gravada em placas de argila na Mesopotâmia, que os papiros egípcios com seus hieróglifos, que as tabuinhas de cera romanas nas quais os estudantes apagavam e reescreviam suas lições, que os pergaminhos medievais copiados laboriosamente por monges, que o mataborrão que absorvia o excesso de tinta das penas de ganso, que a máquina de escrever sobre a qual gerações de escritores e jornalistas

produziram seus textos, que o aparelho de celular por meio do qual se manda um meme[1] ou um emoji[2] em uma conversa, que todos esses artefatos eram, cada um em seu tempo, a mais avançada tecnologia disponível para registrar, transmitir e organizar o pensamento humano. Cada época teve a sua tecnologia educacional, e é precisamente aí que reside o equívoco central a ser desfeito. Trata-se de compreender que a relação entre educação e tecnologia não é acidental nem recente. Ela é ontológica.

Dizer que essa relação é ontológica significa afirmar que ela diz respeito ao próprio ser da educação humana, e não a um de seus atributos contingentes. Não existe educação humana sem algum suporte técnico mediando o processo de transmissão e construção do conhecimento. Mesmo nas sociedades de oralidade primária, anteriores à invenção da escrita, já havia tecnologias cognitivas sofisticadas: fórmulas mnemônicas, ritmos, cantos, narrativas épicas que funcionavam como dispositivos de armazenamento e recuperação de informações vitais para a sobrevivência e a coesão do grupo. Lévy (1993, p. 46) observa, a esse respeito, que "a humanidade construiu outros tempos, mais rápidos, mais violentos que os das plantas e animais" precisamente porque dispõe de "extraordinários instrumentos de memória e de propagação das representações". O ser humano sempre interveio na natureza com a mediação de artefatos, e esses artefatos sempre tiveram uma dimensão pedagógica, pois são transmitidos culturalmente de geração em geração. Técnica e educação nascem juntas na espécie humana.

A antropologia filosófica oferece uma chave interpretativa fundamental para compreender essa ligação constitutiva. Gehlen (1980), ao desenvolver sua análise do ser humano como animal biologicamente não especializado, mostrou que o homem, diferentemente de outros animais, não está adaptado naturalmente a nenhum nicho ecológico específico. Sua sobrevivência depende da capacidade de transformar o ambiente por meio de artefatos que substituem ou fortalecem seus órgãos deficientes. "A técnica é parte e parcela da própria essência do homem", sintetiza Gehlen (1980, p. 4).

Nessa perspectiva, fabricar ferramentas não é uma habilidade que o ser humano possui a mais, como um bônus evolutivo. Trata-se de uma condição de possibilidade de sua existência

[1] O termo meme possui origem no grego *mimema* (algo imitado), conceito adaptado para a biologia e sociologia para descrever uma unidade de transmissão cultural. Trata-se de uma imagem, vídeo ou frase de caráter humorístico, satírico ou irônico que se difunde de forma célere pelas redes sociais. Como artefato de interação, o meme carrega um sentido partilhado por uma comunidade, funcionando como um suporte de memória e crítica social cuja compreensão depende de códigos culturais comuns.

[2] Proveniente do japonês (*e* = imagem + *moji* = letra/caractere), o emoji é um ícone digital que sintetiza emoções ou conceitos. Embora se apresente como um pictograma universal, seu sentido é frequentemente ressignificado dentro de comunidades específicas; um único ícone pode adquirir valores simbólicos distintos dependendo do contexto da interação, atuando como um elemento de pontuação emocional e pragmática na escrita contemporânea.

como espécie. E se fabricar ferramentas é constitutivo do humano, aprender a fabricá-las e a usá-las é igualmente constitutivo, o que significa que a educação sempre esteve inscrita no coração da relação entre o ser humano e seus artefatos. A tecnologia, portanto, não é algo que foi adicionado de fora ao processo educativo, pois ela é a substância mesma de que esse processo é feito.

A história das tecnologias da escrita ilustra com nitidez essa tese. Quando, por volta de 3200 a.C., escribas sumérios passaram a registrar transações comerciais em tabuinhas de argila usando um estilete de bambu, estavam inaugurando o que Lévy (1993) chamaria de uma das grandes revoluções das "tecnologias intelectuais", aquelas que mediam a relação entre o ser humano e o conhecimento. A escrita cuneiforme não era apenas uma forma de contabilidade: era uma tecnologia cognitiva que transformava a própria estrutura do pensamento, permitindo arranjar informações em listas, tabulações e séries ordenadas que a memória oral jamais poderia sustentar com a mesma precisão. O autor é preciso a esse respeito:

> É um pouco como se a tabuinha de argila, o papiro, o pergaminho ou a fita magnética repetissem incansavelmente, mecanicamente, aquilo que confiamos a eles; sem tentar compreendê-lo, sem conectá-lo a outros elementos de informação, sem interpretá-lo. A escrita é uma forma de estender indefinidamente a memória de trabalho biológica. As tecnologias intelectuais ocupam o lugar de auxiliares cognitivos dos processos controlados, aqueles que envolvem a atenção consciente e dispõem de tão poucos recursos no sistema cognitivo humano (Lévy, 1993, p. 91-92).

Cada suporte de escrita ao longo da história corresponde, assim, a um conjunto de possibilidades cognitivas e pedagógicas distintas. As placas de argila eram pesadas e permanentes, portanto, adequadas para registros duradouros, mas pouco práticas para o ensino cotidiano. Os rolos de papiro permitiam textos mais longos e eram mais leves, mas frágeis e caros, o que tornava o acesso a eles um privilégio de classes. As tabuinhas de cera, ao contrário, eram reutilizáveis: o estudante romano gravava suas letras com o estilete pontiagudo e as apagava com a extremidade plana, num ciclo de aprendizagem que antecipava, em dois milênios, a lógica do caderno de exercícios. O pergaminho permitia dobrar as folhas e criar o códex, precursor do livro encadernado, tornando a navegação no texto mais intuitiva do que o rolo. A pena de ganso e o mata-borrão compunham um sistema tecnológico integrado de produção de escrita que exigia treino, destreza e uma pedagogia específica para seu domínio. A máquina de escrever mecanizou esse processo e alterou profundamente a relação do escritor com seu texto, ao separar o traço individual da caligrafia da letra impressa e uniforme. Em cada um desses momentos, havia uma tecnologia, havia uma pedagogia do seu uso e havia uma transformação nas formas de pensar e de aprender.

A invisibilidade dessas tecnologias passadas é ela mesma um fenômeno digno de reflexão. Quando não percebemos mais que a lousa, o pincel, o livro didático ou o caderno são tecnologias, é porque elas foram tão completamente absorvidas pelas práticas cotidianas que se tornaram transparentes. Cupani (2011, p. 13), filósofo da tecnologia, ao analisar a ubiquidade dos artefatos na vida humana, chama atenção para o fato de que "raramente participamos da criação de objetos tecnológicos", mas que "é quase impossível hoje em dia que não sejamos usuários deles".

Essa condição de usuários permanentes e inadvertidos é o que faz com que a tecnologia antiga desapareça de nosso campo de visão crítica. Ela se torna parte da paisagem, e paisagens não são questionadas, são apenas habitadas. O perigo pedagógico dessa invisibilidade é justamente o de tratar as novas tecnologias como ruptura absoluta, como se o passado educacional fosse um vazio técnico que apenas agora estivesse sendo preenchido, quando na verdade o que ocorre é uma transformação dentro de uma continuidade muito mais longa e profunda.

É nessa continuidade que reside a contribuição mais fértil de uma perspectiva histórico-filosófica para o debate contemporâneo sobre educação e tecnologia. Compreender que a sala de aula sempre foi um ambiente tecnológico não significa uma redução da importância das transformações digitais em curso, mas sim o reconhecimento da natureza técnica do humano. Como argumenta Stiegler (1998, p. 141), o ser humano só se constitui através da técnica, que é a sua própria exteriorização. Esta atua também como uma memória externa que permite a transmissão da experiência para além da finitude do corpo. Enquanto certas ferramentas ampliam a força ou o movimento, as tecnologias intelectuais funcionam como suportes de retenção que salvaguardam o patrimônio cognitivo da espécie.

Trata-se, portanto, de fornecer o horizonte conceitual a partir do qual as inovações podem ser avaliadas com mais rigor e serenidade, despindo a técnica de seu fetiche pela novidade. Se a tecnologia é o "horizonte de toda possibilidade futura" (Stiegler, 1998, p. 6), ela o é justamente por ser o depósito do passado, ou seja, uma exteriorização do saber que permite ao aluno pensar melhor e lembrar mais.

É necessário, contudo, operar uma distinção entre o que o senso comum denomina tecnologia velha e o fenômeno da tecnologia invisível. A tecnologia velha refere-se habitualmente a objetos que pertencem a uma mesma linhagem técnica, mas que foram superados em desempenho por modelos mais recentes, como ocorre com versões anteriores de *smartphones* em face das gerações atuais. A tecnologia invisível, por sua vez, é aquela que foi tão profundamente incorporada às práticas cotidianas que perdeu seu caráter de novidade para

se tornar infraestrutura. É o caso da lapiseira, da mesa e cadeira da sala de aula, da lousa ou do caderno. São todos artefatos que, em termos de representações sociais, não são mais vistos como tecnologia. Cupani (2011, p. 13) observa que essa onipresença dos artefatos os torna transparentes; eles deixam de ser percebidos como técnica para serem habitados como paisagem.

Essa invisibilidade é, na verdade, o estágio final da sedimentação cultural de uma ferramenta, constituindo o verdadeiro substrato ontológico da educação. A técnica não é apenas motor ou alavanca, tampouco se reduz ao último lançamento da indústria digital. Ela é a própria invenção do humano que, ao exteriorizar o conhecimento em objetos, permite que a educação ocorra como um processo de transmissão geracional contínuo. Stiegler (1998, p. 141) nomeia esse processo de retenção terciária, designando com isso a inscrição técnica que preserva a memória coletiva para além dos limites do corpo e da experiência individual. O erro pedagógico reside em ignorar as tecnologias invisíveis, pois é nelas que o pensamento se apoia para poder, então, dialogar com o que há de novo.

Desde o primeiro graveto riscado na argila, desde as primeiras pinturas nas paredes das cavernas de Lagoa Santa (MG) e São Raimundo Nonato (PI) que comunicavam algo a ser aprendido e preservado, o ser humano tem produzido extensões de sua mente. Sendo assim, "o espírito humano não é um centro organizador em torno do qual giram tecnologias intelectuais, como satélites a seu serviço. Em si, não é nada além de um agenciamento de satélites de todos os tamanhos" (Lévy, 1993, p. 106). A mente humana sempre foi, desde suas origens, uma mente tecnológica.

As funções que a tecnologia serve à cognição humana ultrapassam o simples registro e revelam-se em dimensões permanentes e dinâmicas. Sendo assim, escreve-se:

- para lembrar depois,
- para comunicar a outrem no presente e,
- para alcançar gerações futuras.

Contudo, a tecnologia atua de forma ainda mais profunda no auxílio aos processos controlados do intelecto. Escreve-se:

- para organizar o pensamento, transformando o fluxo mental em uma estrutura lógica e visível,
- para demonstrar, como o matemático que externaliza as etapas de um teorema para validar sua própria descoberta,
- para aprender, pois o ato de registrar é, em si, um método de internalização e síntese de novos conceitos.

Essas funções evidenciam que toda tecnologia educacional, das tabuinhas de cera às plataformas digitais, responde a um propósito cognitivo fundamental, qual seja, ampliar o que sozinhos não conseguimos reter, alcançar, organizar ou perpetuar. A tecnologia deixa de ser apenas um suporte passivo para se tornar um interlocutor que apoia o raciocínio, permitindo que a educação ocorra não apenas como transmissão, mas como uma constante assistência na produção do pensamento.

Os estudos históricos sobre a inserção de tecnologias na educação confirmam essa perspectiva de longa duração. Bruzzi (2016, p. 476) recorda que "cadernos, lápis, canetas, lousas, mimeógrafos, projetor de slides, ábacos, jogos, globos e mapas, laboratórios, cartazes, gravuras estão inseridos no espaço escolar desde o final do século XIX", compondo um repertório tecnológico que, à sua época, também foi objeto de resistências, debates e aprendizagem por parte de professores e estudantes. O padrão se repete, pois, cada nova tecnologia gera inicialmente estranhamento, depois adaptação, depois naturalização, e finalmente invisibilidade. O computador e a internet não escapam a esse ciclo; encontram-se ainda em suas fases intermediárias.

Reconhecer a natureza ontológica da relação entre educação e tecnologia tem, portanto, consequências práticas decisivas. A pergunta não pode ser "devemos ou não usar tecnologia na educação?", pois essa pergunta pressupõe equivocadamente que existe uma educação sem tecnologia à qual poderíamos retornar. A pergunta relevante é sempre outra: quais tecnologias usar, para quais fins pedagógicos, com quais critérios de seleção e avaliação? É essa pergunta que permite superar tanto o tecnoentusiasmo acrítico, que celebra toda novidade como progresso automático, quanto o tecnoceticismo reativo, que recusa toda inovação como ameaça. Gehlen (1980, p. 4) já alertava que as produções humanas "tomam o lugar e potencializam os órgãos de que o ser humano inicialmente dispõe", o que significa que elas ampliam possibilidades, mas também constrangem e redefinem práticas. Toda tecnologia educacional é, simultaneamente, uma abertura e um fechamento, pois, abre certas formas de aprender e torna outras menos acessíveis ou necessárias. Compreender isso é o primeiro passo para uma pedagogia verdadeiramente reflexiva em relação a seus próprios instrumentos.

3 ONDAS TECNOLÓGICAS E RECONFIGURAÇÃO DA EDUCAÇÃO

Se a parte anterior deste capítulo propôs reconhecer a relação entre educação e técnica como ontológica, constitutiva do ser humano como espécie, cabe agora percorrer, em detalhe analítico, as grandes inflexões históricas dessa relação. Trata-se de examinar as ondas que, na

formulação proposta, não se anulam, mas se ressignificam, sedimentando-se umas sobre as outras como estratos geológicos de uma mesma montanha de práticas pedagógicas. Para isso, é preciso recorrer à *longue durée* não apenas como moldura temporal, mas como instrumento interpretativo ativo.

Braudel (1969, p. 41) defende que a história de longa duração é a história das estruturas que resistem ao tempo, das permanências que governam o ritmo profundo das civilizações. Aplicada ao campo da transmissão do saber, essa perspectiva revela uma estrutura constante que atravessa milênios: a função docente como ato de mediação entre o saber acumulado e o aprendiz que ancora na ignorância. O que muda, em cada onda, é o suporte por meio do qual essa mediação se efetua e, consequentemente, a natureza do papel do mediador. O que permanece é a assimetria de saber entre quem ensina e quem aprende, bem como a necessidade social de reduzi-la de modo sistemático.

Pierre Lévy (1993, p. 46), em sua análise das tecnologias da inteligência, também percebe essa permanência profunda ao observar que "cada técnica de comunicação ou de tratamento da informação abre novas possibilidades de desenvolvimento cognitivo individual e social", mas o faz sempre sobre o substrato das tecnologias anteriores, que não desaparecem, mas mudam de função. A escrita não eliminou a oralidade; a imprensa não aboliu a escrita manuscrita; a internet não extinguiu o livro. O que cada onda faz é reorganizar a ecologia cognitiva disponível, criando novas hierarquias entre os suportes e novos regimes de atenção e memória.

Essa perspectiva é convergente com o que Stiegler (1998, p. 177) denomina de processo de gramatização: o movimento pelo qual as continuidades do fluxo vital, sejam a voz, o gesto, o pensamento, são discretizadas e inscritas em suportes externos, tornando-se reproduzíveis, transmissíveis e, portanto, ensináveis. Cada onda tecnológica na história da educação pode ser lida, nessa chave, como um novo momento do processo de gramatização, um novo patamar de exteriorização da memória coletiva, que inaugura uma nova organização do aprender.

3.1 Primeira onda: escrita, codificação e linearidade

A invenção da escrita é, sem exagero, a maior revolução pedagógica da história humana. Não apenas porque criou novos suportes para o conhecimento, mas porque transformou radicalmente a natureza do próprio conhecimento, o que Havelock (1988, p. 130) sintetizou ao mostrar que, antes da escrita, "os poetas eram chamados para dar continuidade

às funções didáticas de Homero", sendo a transmissão do saber inseparável da performance oral e do corpo do mestre.

Nas sociedades de oralidade primária, o conhecimento habitava, exclusivamente, a memória dos viventes. Lévy (1993, p. 50) descreve esse estado cognitivo com precisão:

> Numa sociedade oral primária, quase todo o edifício cultural está fundado sobre as lembranças dos indivíduos. A inteligência, nestas sociedades, encontra-se muitas vezes identificada com a memória, sobretudo com a auditiva. A escrita suméria, ainda muito próxima de suas origens orais, denota a sabedoria representando uma cabeça com grandes orelhas.

Sob esse regime, aprender significava escutar e repetir, e o educador era o guardião insubstituível de um saber que pereceria com ele. A memória biológica constituía o único suporte do patrimônio cultural, o que conferia ao mestre uma autoridade existencial, não apenas epistêmica: ele era, literalmente, o livro vivo da comunidade.

Com a escrita cuneiforme da Mesopotâmia e, posteriormente, com o alfabeto fonético grego, essa estrutura sofre uma transformação irreversível. O saber se separa do corpo do mestre e passa a existir de modo autônomo em um suporte externo. Platão, no Fedro, reconheceu esse deslocamento com a ambivalência característica dos pensadores que habitam as fronteiras de uma mudança civilizacional. Para Platão (*apud* Derrida, 2005, p. 21), a escrita era um *phármakon* ao mesmo tempo remédio e veneno: "remédio" porque liberta a memória do esforço de retenção, "veneno" porque substitui o saber vivo pela letra morta, entregando o conhecimento a quem não dispõe do mestre para interrogá-lo.

Essa ambivalência platônica é, em si mesma, um dado pedagógico de primeira ordem, pois registra, no próprio momento de sua emergência, a ansiedade que toda nova tecnologia cognitiva provoca nos guardiões do saber anterior. O mesmo padrão se repetirá, com variações, em cada onda subsequente. A escrita foi acusada de matar a memória viva; a imprensa, de democratizar perigosamente o acesso ao conhecimento; o rádio, de superficializar o pensamento; a internet, de fragmentar a atenção; e a inteligência artificial, de substituir o raciocínio. A história das ansiedades pedagógicas é, ela própria, uma das mais constantes estruturas de longa duração.

Do ponto de vista cognitivo, a escrita inaugurou o que Lévy (1993, p. 56) denomina de "pensamento linear": a possibilidade de dispor conceitos em sequência lógica verificável, de criar listas, tabelas e hierarquias que a memória oral não poderia sustentar sem distorção. O que antes era circularidade narrativa tornou-se progressão argumentativa. Essa transformação tem consequências pedagógicas profundas, uma vez que o aprendizado deixa de ser uma

imersão em narrativas que se repetem e passa a ser a navegação por um texto que avança e pode ser retomado, corrigido e acumulado. É o nascimento do que poderíamos chamar de aprendizado verificável: porque está escrito, pode ser confrontado, criticado e ampliado, o que é a estrutura básica de toda epistemologia científica.

A implicação pedagógica mais profunda da escrita, no entanto, é aquela que Lévy (1993, p. 56) expressa ao dizer que ela "é uma forma de estender indefinidamente a memória de trabalho biológica". Ao externalizar o pensamento em um suporte estável, o ser humano libera recursos cognitivos para operar em níveis de abstração mais elevados. Quem não precisa memorizar os fatos pode pensar sobre eles. A escrita não empobreceu o pensamento humano; reorganizou suas prioridades.

3.2 Segunda onda: a imprensa, reprodutibilidade e nascimento da escola moderna

Se a escrita tornou o conhecimento independente do corpo do mestre, a prensa de tipos móveis de Gutenberg, em meados do século XV, fez algo igualmente revolucionário, tornando-o independente da escassez. Antes dela, um manuscrito era um objeto único, produzido com enorme dispêndio de tempo e recurso humano. Depois dela, o mesmo texto podia ser reproduzido em centenas, depois milhares de cópias. O que isso significou para a educação é difícil de superestimar.

Burke (2002, p. 173) descreve a imprensa como a tecnologia que, "desde o século XVI", passou a ser "associada com o progresso do conhecimento". Mas o historiador é criterioso ao apontar que esse progresso trouxe consigo novos problemas, quais sejam, a explosão informacional, a dificuldade de verificar a qualidade do que se imprimia e a fragmentação do conhecimento em especialidades. Já em 1470, o humanista Niccolò Perotti propunha "um projeto defendendo a censura erudita" diante da multiplicação descontrolada dos textos impressos (Burke, 2002, p. 173). A tecnologia da reprodutibilidade criou, ao mesmo tempo, a possibilidade da escola moderna e a necessidade de um novo tipo de mediador que não fosse mais o guardião do livro único, mas o professor capaz de orientar a navegação num oceano de textos.

Rodrigues (2012, p. 197), ao analisar o impacto da tipografia sobre a escola moderna, observa que a imprensa "introduziu pela primeira vez a ideia de industrialização de uma mercadoria" no campo do saber, e que essa lógica industrial impregnou a própria arquitetura da sala de aula: as fileiras de carteiras, a divisão por séries, a uniformidade dos livros didáticos, a progressividade linear do currículo. "A linha reta", escreve o autor, "pode ser vista na sala de

aula ao observar-se as cadeiras dos alunos em relação à mesa do professor, as filas em ângulos retos e a própria arquitetura dos prédios escolares". A escola que conhecemos é, em grande medida, uma invenção tipográfica.

Essa associação entre tipografia e escola moderna tem implicações que vão além da organização física. Rodrigues (2012, p. 197) desenvolve o argumento de que o livro impresso impôs ao pensamento ocidental uma lógica de racionalidade linear, de descontextualização do saber e de impessoalidade do conhecimento, já que "a escola devia também sofrer suas transformações modernizantes. Não seria mais privilégio de classe, mas entraria para o esforço nacional pela instrução pública (homogeneizante), complementando sua função com a nobre tarefa de formar os cidadãos para o trabalho industrial".

Assim, ao mesmo tempo em que democratizava o acesso ao texto, a imprensa o uniformizava. Ao permitir que vários alunos tivessem o mesmo livro, criou a condição de possibilidade para o currículo comum, mas também para o pensamento padronizado. A pedagogia da imprensa é uma pedagogia da norma, pois há um texto correto, um uso correto da língua, uma sequência correta de aprendizagens. Burke (2002, p. 183) aponta que essa tendência levou, no limite, à especialização e à fragmentação do conhecimento, pois "a divisão do trabalho intelectual não se limitou a enciclopédias. Houve uma tendência geral para a especialização e fragmentação em detrimento do ideal geral do conhecimento".

Sob a perspectiva da *longue durée*, a imprensa inaugurou a fase em que a escola deixou de ser artesanal e passou a ser industrial, no sentido mais preciso, ela produz, em série, trabalhadores intelectuais com competências padronizadas. As consequências dessa virada ainda se fazem sentir nas formas dominantes de avaliação, nos currículos nacionais e na lógica do livro didático como espinha dorsal do ensino. A segunda onda tecnológica está na fundação da escola que ainda habitamos.

3.3 Terceira onda: audiovisual, ensino de massa e expansão geográfica

O rádio e a televisão representam a terceira grande onda tecnológica da educação, e sua especificidade histórica é a de ter dissociado a transmissão do saber da necessidade da leitura. Com eles, o conhecimento ganhou voz, imagem e presença sensorial, tornando-se acessível a populações que a barreira do letramento mantinha às margens dos sistemas educacionais formais.

No Brasil, essa promessa foi formulada com precisão por Fernando de Azevedo, para quem, como recorda a literatura da área, rádio e cinema "influenciam fortemente o sistema de

relações humanas" e constituem "poderosos instrumentos educativos e culturais, de informação, de propaganda e de ensino, que compõem fatores de primeira ordem de educação popular pelo seu poder extraordinário de sugestão, principalmente num país de numerosos iletrados" (Azevedo *apud* Gilioli, 2008, p. 44). Anísio Teixeira, por sua vez, defendia o rádio como instrumento de "democratização do acesso à informação, à cultura e ao conhecimento científico, independentemente da localização geográfica ou condição socioeconômica das pessoas" (Abrunhosa *apud* Klein *et al.*, 2020, p. 12).

O movimento escolanovista brasileiro dos anos 1930 abraçou com entusiasmo a promessa radiofônica. A Rádio Escola Municipal do Rio de Janeiro, as iniciativas de radioeducação nos estados e a articulação de educadores como Edgar Roquette-Pinto expressam o que se poderia chamar de tecnoentusiasmo progressista: a crença de que a tecnologia da comunicação de massa poderia ser o vetor de uma democratização cultural que as estruturas sociais e a insuficiência da rede escolar tornavam impossível apenas pelas vias tradicionais.

A televisão, ao acrescentar a imagem ao som, ampliou esse alcance sensorial. Sua chegada ao Brasil em 1950, pela TV Tupi de Assis Chateaubriand, inaugurou um novo regime de percepção coletiva. McLuhan (1964, p. 7), ao propor que "o meio é a mensagem", apontou com perspicácia que a televisão, como meio, "teve impacto profundo na cultura porque envolvia os espectadores de maneira sensorial e imersiva, alterando assim nossa experiência e percepção do mundo" (*apud* Friesen *et al.*, 2024, p. 120). A sala de aula não ficou imune a essa transformação, pois o audiovisual passou a ser integrado às práticas pedagógicas como suplemento, mas também como modelo implícito de comunicação, moldando as expectativas dos alunos em relação ao ritmo, ao estímulo sensorial e à forma da transmissão de conteúdos.

Do ponto de vista estrutural, a terceira onda representa um deslocamento fundamental: pela primeira vez, a transmissão do saber prescinde do encontro físico entre mestre e discípulo. O professor pode ser uma voz anônima num rádio ou um rosto na televisão, e o aprendiz pode estar a milhares de quilômetros. A educação de massa, no sentido mais radical da expressão, torna-se tecnicamente possível. Isso cria, simultaneamente, uma possibilidade de democratização sem precedentes e uma armadilha de passividade, pois o aluno-ouvinte ou o aluno-telespectador é, estruturalmente, um receptor. O modelo pedagógico que essa tecnologia sugere é bancário, para usar a expressão de Freire (1970), pois o saber flui em uma única direção, do emissor especializado para o receptor silencioso.

A lição estrutural da terceira onda, sob a lente da *longue durée*, é que a ampliação do alcance geográfico da educação não resolve, por si mesma, o problema da qualidade da relação

pedagógica. O rádio levou conhecimento ao sertão; a televisão democratizou o acesso à informação; ambos, no entanto, encontraram seus limites na incapacidade de responder às perguntas dos ouvintes, de adaptar-se aos ritmos individuais de aprendizagem e de estabelecer o diálogo como condição do ensino. A tecnologia resolveu o problema do alcance e criou o problema da interatividade.

3.4 Quarta onda: internet e quebra do monopólio da informação

A internet representa a quarta e, até o momento, mais radical transformação das condições de transmissão do saber. Seu impacto sobre a educação decorre de uma característica estrutural sem precedentes, já que pela primeira vez na história, qualquer indivíduo com acesso a uma conexão pode, simultaneamente, ser receptor e emissor de informação, aluno e professor, aprendiz e produtor de conhecimento.

Tim Berners-Lee desenvolveu, nos laboratórios do CERN, entre 1989 e 1991, o protocolo que daria origem à *World Wide Web*, concebida como uma "resposta à necessidade de compartilhar dados entre os colaboradores de projetos de pesquisa" por meio de uma "interface mais amigável" e baseada no conceito de hipertexto (Monteiro *apud* Klein *et al.*, 2020, p. 13). O que emergiu dessa inovação técnica ultrapassou em muito a intenção original, já que a *Web* se tornou o maior arquivo e o maior mercado de informação já construído pela humanidade, e sua disseminação foi tão acelerada que, como aponta Lévy (2023), "a proporção da população mundial que tinha acesso à internet era de cerca de 1% em 1990, 4% em 1999, 24% em 2009, 51% em 2018 e 65% em 2023" (*apud* Romanowski *et al.*, 2022, p. 87).

Para a educação, essa disponibilidade massiva de informação tem consequências estruturais profundas. O professor deixa de ser o guardião do acesso ao conhecimento, papel que a escassez do livro lhe havia conferido e que a imprensa já relativizara, mas não eliminara, e passa a ser, potencialmente, um curador, um orientador de percursos num ambiente em que a informação é abundante, mas a capacidade de avaliá-la e articulá-la é escassa. Como observa Kenski (1994 *apud* Klein *et al.*, 2020, p. 13), "a internet facilitou o acesso às informações", mas esse acesso facilitado trouxe consigo o problema da qualidade: um "congestionamento de pluralidade de teorias sobre determinado conteúdo" que, sem mediação pedagógica competente, produz um conhecimento "descartável que se apaga e ligeiramente é esquecido" (Klein *et al.*, 2020, p. 280).

Simão Neto (2002) descreveu a inserção das tecnologias digitais na escola em cinco ondas: desde os registros administrativos até a internet como vetor de informação e

comunicação. Na sua periodização, a chegada da internet representa a quinta onda da informática educacional, aquela em que o foco desloca-se do equipamento para a conectividade, do *hardware* para o fluxo (Simão Neto *apud* Romanowski; Bigan, 2022, p. 81). Essa ênfase na conectividade é o traço definitório da quarta onda nos termos aqui adotados: o que muda estruturalmente não é apenas o suporte, mas a topologia da relação pedagógica. A internet permite, em princípio, que o aprendizado se dê em redes horizontais, colaborativas e sem hierarquia fixa de autoridade epistêmica.

Essa potência democrática, no entanto, coexiste com tensões profundas. A desigualdade de acesso permanece como uma fratura estrutural: no Brasil, como documentam dados do IBGE e do CETIC (2019). Enquanto 76% dos estudantes de escolas urbanas tinham acesso diário à internet, nas escolas rurais mais da metade não dispunha dessa conectividade (Romanowski *et al.*, 2022, p. 87). A quarta onda, como as anteriores, distribuiu seus benefícios de modo assimétrico.

Além disso, a internet trouxe um desafio cognitivo específico que a pedagogia ainda não resolveu: a difusão de informações falsas, os mecanismos de câmara de eco das redes sociais e a lógica de engajamento dos algoritmos que priorizam o estímulo emocional sobre a verificação factual. Lévy (2023) alertou que somente "uma verdadeira educação para o pensamento crítico no novo ambiente de comunicação" poderá converter o potencial democrático da internet em cidadania renovada, o que exige "aprender a discernir as fontes de dados em termos de organização de categorias, narrativas e agendas dominantes" (*apud* Romanowski *et al.*, 2022, p. 82). A quarta onda ampliou o acesso e tornou mais urgente do que nunca a mediação pedagógica.

3.5 Quinta onda: a inteligência artificial

A inteligência artificial generativa representa a quinta onda e, em vários aspectos, a mais desconcertante de todas. As ondas anteriores transformaram os suportes do saber e as condições de acesso a ele, mas deixaram intacta a fronteira entre o humano que pensa e o artefato que registra, reproduz ou transmite. A IA generativa, ao ser capaz de produzir linguagem, raciocínio, argumentação e criação em nível que rivaliza, pelo menos na superfície, com a produção humana, embaralha essa fronteira de um modo historicamente inédito.

Do ponto de vista técnico, a IA generativa atual opera a partir de modelos de linguagem de grande escala (*Large Language Models*, LLMs) que, treinados sobre vastos *corpora* de texto humano, desenvolvem a capacidade de gerar respostas coerentes e contextualmente

pertinentes a qualquer prompt. O que emerge dessa arquitetura não é inteligência no sentido filosófico, mas o que Canavilhas e Essenfelder (2022 *apud* Friesen et al., 2024, p. 117) descrevem como um "guarda-chuva" de tecnologias que "se popularizou não só a partir do avanço tecnológico, mas também de um marketing agressivo", cujo objetivo declarado é o desenvolvimento de "agentes racionais", ou seja, "sistemas que resolvem problemas sem se preocupar em imitar o pensamento ou o comportamento humano" (Russel; Norvig *apud* Friesen *et al.*, 2024, p. 117).

Para a educação, essa capacidade cria possibilidades que as ondas anteriores não podiam oferecer. A IA pode personalizar o ensino "de acordo com as necessidades e preferências dos alunos", melhorar "a avaliação e o feedback" e "ampliar o acesso e a inclusão" (Unesco, 2019; Unicef, 2021 *apud* Friesen *et al.*, 2024, p. 120).

Mais do que isso, a IA constitui, pela primeira vez na história, um interlocutor disponível permanentemente, capaz de responder a perguntas, explicar conceitos de múltiplas maneiras, gerar exercícios adaptados e acompanhar o raciocínio do aprendiz em tempo real. Ela não resolve o problema da motivação, da disciplina ou do sentido, mas resolve o problema da disponibilidade do mestre. Tem-se, por meio dela, um mestre algorítmico que nunca dorme, nunca se cansa e nunca envergonha o aluno pela pergunta que considera óbvia.

A analogia com a escrita como *phármakon* platônico retorna com força neste ponto. A IA generativa é, simultaneamente, remédio e veneno para o processo educativo. Como remédio, oferece suporte cognitivo sem precedentes, liberando o aluno para tarefas de síntese e criação que as tarefas mecânicas de pesquisa e formatação anteriormente obstruíam. Como veneno, arrisca produzir o que Friesen *et al.* (2024, p. 122) descrevem como o paradoxo central da inteligência artificial aplicada à educação: "o objetivo é preparar os estudantes para passar em exames ou ajudá-los a se autorrealizarem, a atingirem seu máximo potencial?" A delegação do raciocínio à máquina pode liberar o humano para pensar melhor, ou pode simplesmente dispensá-lo de pensar.

Lévy (1993, p. 95) havia antecipado o contorno estrutural dessa questão ao observar que "é bem possível que novas tecnologias intelectuais, baseadas na informática, se estabilizem, e estas tornariam 'irracionais', ou ao menos muito grosseiros, raciocínios que utilizassem a lógica clássica e a teoria das probabilidades." Na quinta onda, o que está em jogo não é apenas o acesso ao saber ou a escala da sua transmissão, mas a própria natureza do ato de conhecer: se a cognição pode ser parcialmente terceirizada a sistemas algorítmicos, o que resta como tarefa especificamente humana no processo educativo?

Stiegler (1998, p. 177) fornece uma chave para essa questão ao afirmar que cada novo suporte técnico não apenas externaliza a memória, mas reconfigura as condições de sua produção. A IA não é apenas um novo arquivo ou um novo canal de transmissão: ela é um novo agente no processo de constituição do saber, um interlocutor que dialoga com o pensamento humano e, ao fazê-lo, o transforma. O desafio pedagógico da quinta onda é aprender a pensar *com* a IA, e não apenas aprender a usar a IA, o que é uma distinção de ordem epistemológica, não apenas instrumental.

3.6 O que as ondas têm em comum: externalização, democratização e deslocamento do educador

Percorridas as cinco ondas, é possível identificar uma gramática estrutural que as atravessa. Cada uma delas realiza, à sua maneira, três operações fundamentais. A primeira é a *externalização cognitiva*: transfere para fora da mente humana uma tarefa que antes era feita internamente, liberando recursos cognitivos para operações de ordem superior. A escrita externalizou a memória; a imprensa externalizou a cópia; o audiovisual externalizou a voz e a presença; a internet externalizou o arquivo; a IA externaliza o raciocínio discursivo. Cada externalização libera espaço mental e, ao fazê-lo, redefine o que se espera do aprendiz.

A segunda operação é a *democratização do acesso*: cada onda amplia, em alguma medida, o círculo dos que podem acessar o saber. A escrita permitiu a transmissão além do encontro presencial; a imprensa tornou o livro economicamente acessível às classes médias; o audiovisual alcançou os iletrados; a internet dissolveu fronteiras geográficas; a IA coloca, em princípio, um tutor sofisticado ao alcance de quem dispõe de um *smartphone*. Nenhuma dessas democratizações foi completa ou igualitária, mas todas representaram expansões reais do universo dos que podem aprender.

A terceira operação é o *deslocamento do papel do educador*, pois cada onda transforma o que se espera do professor. O educador da oralidade era a memória viva; o da escrita, o intérprete autorizado; o da imprensa, o guia num mundo de textos; o do audiovisual, o produtor e animador; o da internet, o curador e mediador crítico; o da IA, o formulador de perguntas que a máquina não pode fazer sozinha. Esse deslocamento progressivo não elimina o professor: elimina certos professores que se recusam a ser deslocados, e convoca outros que compreendem que seu valor não está naquilo que a máquina pode replicar, mas naquilo que apenas a presença humana, com sua dimensão afetiva, ética e existencial, pode oferecer.

Burke (2002, p. 183), em sua análise dos "problemas causados por Gutenberg", observou com agudeza que a imprensa tornou o ideal da *pansofia* cada vez mais irrealista: "o conhecimento universal já não está ao alcance do homem". A internet e a IA radicalizaram esse problema, pois em um ambiente de abundância informacional, a síntese substitui a enciclopédia como ideal pedagógico. O que se almeja não é mais saber tudo, mas saber como perguntar, como avaliar, como articular e como criar significado a partir do que está disponível. Essa é a tarefa que nenhuma onda tecnológica, por mais potente que seja, pode realizar por nós.

Cada onda tecnológica na história da educação foi acompanhada de uma pergunta que se repetiu, com variações lexicais, mas com estrutura invariante: *o que, desta vez, o ser humano perde ao delegar essa tarefa à máquina?* Platão perguntou o que se perderia ao confiar à escrita a tarefa da memória viva. Os humanistas do século XVI perguntaram o que se perderia ao multiplicar os textos além da capacidade de discernimento do leitor. Os escolanovistas dos anos 1930 perguntaram o que se perderia ao substituir o mestre humano pela voz do rádio. Os professores de hoje perguntam o que se perde quando o aluno entrega a redação ao ChatGPT.

Essa pergunta é incômoda porque não tem resposta simples, e é precisamente por isso que ela é pedagogicamente essencial. Ela obriga a distinguir entre a tarefa delegada e o aprendizado que essa tarefa propiciava. Vejamos um exemplo: a calculadora liberou os estudantes do cálculo manual, mas revelou que parte do aprendizado de matemática ocorria no processo de calcular, não no resultado. A IA libera os estudantes da busca, da síntese e da redação de primeiros rascunhos, mas coloca em suspeita o que exatamente aprendiam ao realizar essas tarefas. É possível que aprendessem muito; é possível que aprendessem pouco e que a verdadeira aprendizagem sempre tenha requerido o diálogo, a reflexão e a elaboração criativa que a IA não pode realizar por ninguém.

Sob a perspectiva da *longue durée*, o que essa pergunta revela é que o problema da educação nunca foi técnico. Ele é, e sempre foi, o problema do sentido: por que aprender? O que vale a pena saber? Como o saber se converte em sabedoria? As tecnologias transformam as condições em que essas perguntas são feitas, mas não as respondem. Cada onda amplia o repertório de meios disponíveis para ensinar e aprender; nenhuma resolve a questão do fim. E é essa questão, qual seja, o fim do aprender, que constitui o núcleo irredutível do trabalho pedagógico e a âncora de toda reflexão séria sobre educação e tecnologia.

Stiegler (1998, p. 6) descreveu a técnica como "o horizonte de toda possibilidade futura". Se ele tem razão, então cada nova onda tecnológica não é apenas uma mudança de suporte, antes, é uma ampliação do horizonte do possível. O que a educação precisa fazer, em

cada onda, é garantir que o aprendiz não apenas use esse horizonte mais amplo, mas o habite com consciência, autonomia e responsabilidade. Esse é o trabalho que nenhuma onda realiza sozinha, e que toda pedagogia digna do nome tem a obrigação de assumir.

4 TECNOLOGIA, MEMÓRIA E MOTIVAÇÃO: O QUE FICA PARA O HUMANO?

As ondas tecnológicas analisadas nas seções anteriores compartilham uma gramática comum. Cada uma delas externaliza uma função cognitiva que antes residia exclusivamente no interior do sujeito. A escrita externalizou a memória; a imprensa, a cópia; o audiovisual, a presença; a internet, o arquivo; a inteligência artificial, o raciocínio discursivo. Essa progressão levanta uma questão que vai além da instrumentalidade e toca o núcleo mesmo da filosofia da educação: o que resta ao humano quando a máquina assume, com crescente competência, as tarefas que definiam o aprendiz?

Para responder a essa pergunta com rigor, é preciso distinguir dois planos que o debate corrente frequentemente confunde. O primeiro é o plano da *execução*: realizar o cálculo, redigir o rascunho, recuperar a informação. O segundo é o plano da *constituição do sujeito*: o processo pelo qual, ao executar certas tarefas, o ser humano se forma, amadurece e constrói sua identidade intelectual e moral. A tecnologia pode assumir o primeiro plano com eficiência crescente. O segundo, pela sua própria natureza, permanece irredutível à delegação maquínica. É justamente aí que reside a tensão central deste debate.

4.1 O trabalho da memória e o risco da atrofia

Ecléa Bosi (1994, p. 55), ao investigar a memória de velhos trabalhadores urbanos, formulou uma distinção que transcende seu contexto original e ilumina o problema pedagógico contemporâneo: a memória não é um depósito estático de registros, mas um *trabalho*. Recordar é um ato de reconstituição ativa, carregado de afeto, de seleção e de sentido. "A memória", escreve a autora, "permite a relação do corpo presente com o passado e, ao mesmo tempo, interfere no processo atual das representações" (Bosi, 1994, p. 46). Lembrar, nessa perspectiva, não é recuperar um arquivo intacto: é refazer, com os materiais disponíveis no presente, uma experiência que o tempo foi sedimentando.

Essa compreensão tem consequências diretas para a discussão sobre tecnologia e aprendizagem. Quando delegamos integralmente à máquina a tarefa de lembrar coisas como datas, fórmulas, referências, estruturas argumentativas, não apenas liberamos a memória

biológica de um fardo, mas sim suprimimos o processo de reconstituição ativa que é, em si mesmo, formador. O esforço de recordar organiza o pensamento, estabelece conexões, revela lacunas e ativa o que Bosi (1994, p. 55) chama de "trabalho de reconstrução do passado", que é ao mesmo tempo trabalho de constituição do sujeito que recorda. Uma educação que elimina esse esforço não apenas simplifica o aprendizado, mas arrisca empobrecer o sujeito que aprende.

Stiegler (1998, p. 177), em perspectiva convergente, argumenta que a exteriorização técnica da memória, o que ele denomina retenção terciária, não é neutra, pois reconfigura as formas pelas quais a memória secundária (a memória individual, biológica) se constitui e opera. Cada novo suporte técnico não apenas guarda o que o sujeito poderia lembrar, mas altera as condições nas quais o sujeito se relaciona com o passado. Quando o suporte muda radicalmente como ocorre com a inteligência artificial generativa, capaz de sintetizar, conectar e articular com velocidade e abrangência sem precedentes, muda também o regime de constituição da memória individual, e com ele, potencialmente, o modo como o sujeito se compreende no tempo.

O risco que se configura não é o da ignorância enciclopédica, pois nenhuma época pôde, sozinha, dominar a totalidade do saber, mas o da *atrofia do processo*. Lévy (1993, p. 56) havia observado que as tecnologias intelectuais "ocupam o lugar de auxiliares cognitivos dos processos controlados, aqueles que envolvem a atenção consciente". Se essas tecnologias assumem não apenas o registro, mas a síntese e a articulação, suprimem precisamente os processos que requerem atenção consciente e que, ao serem exercidos, desenvolvem a capacidade de pensar. A questão pedagógica não é saber o que a máquina pode fazer, mas discernir quais processos devem permanecer com o humano porque é no seu exercício, e não em seu resultado, que reside o aprendizado essencial.

4.2 O dilema da motivação: para que aprender o que a máquina faz melhor?

O questionamento sobre a motivação para aprender diante da eficiência maquínica não é novo. Cada onda tecnológica o reedita, com atores e ferramentas diferentes, mas com estrutura invariante. Quando a calculadora se tornou acessível nas décadas de 1970 e 1980, professores de matemática em todo o mundo debateram se havia ainda sentido em ensinar a tabuada e o cálculo manual. Quando os processadores de texto substituíram a máquina de escrever, o debate se deslocou para a caligrafia e a ortografia. Quando os sistemas de navegação por GPS se popularizaram, a pergunta se voltou para a cartografia e a orientação espacial. Hoje,

com a inteligência artificial capaz de redigir textos, resolver equações diferenciais e analisar fontes históricas, o debate recomeça em escala ampliada.

Havelock (1988, p. 130) já havia assinalado que a escrita, ao libertar o conhecimento do corpo do mestre e da necessidade da memorização oral, criou exatamente essa crise de motivação, já que se o texto pode guardar, por que lembrar? A resposta histórica foi a gradual redefinição do que se esperava do aprendiz letrado. Não mais a memorização exaustiva dos textos sagrados ou das tradições épicas, mas a capacidade de ler criticamente, de confrontar fontes e de produzir argumentos originais. O que parecia uma perda ante o abandono da memorização revelou-se uma liberação para formas mais elevadas de elaboração intelectual.

O padrão se repete, mas a escala da perturbação varia. A calculadora liberou da aritmética mecânica; a IA generativa ameaça liberar da argumentação, da síntese e da criação. tarefas que os currículos sempre consideraram como as mais elevadas e, portanto, as mais propriamente educativas. Friesen *et al.* (2024, p. 122) formulam o paradoxo com precisão: "o objetivo é preparar os estudantes para passar em exames ou ajudá-los a se autorrealizarem, a atingirem seu máximo potencial?" Se o exame pode ser resolvido pela máquina, o exame deixa de ser um instrumento válido de avaliação; se a autorrealização é o objetivo, ela requer precisamente o esforço que a máquina suprime.

A resposta pedagógica não pode ser a negação da ferramenta. Platão temeu que a escrita destruísse a memória viva; os escolanovistas temeram que o rádio substituísse o diálogo; professores do século XX temeram que a calculadora embotasse o raciocínio matemático. Nenhum desses temores se concretizou na forma catastrófica prevista, embora todos identificassem tensões reais que exigiram respostas pedagógicas concretas. A resposta à quinta onda não será diferente. Não veremos a exclusão sistemática da IA do processo educativo, mas a redefinição do que, em presença dela, ainda precisa ser aprendido pelo próprio sujeito, em seu próprio corpo e tempo.

4.3 A redefinição pedagógica: do saber fazer ao saber avaliar

Diante da eficiência crescente dos sistemas algorítmicos, a questão pedagógica central se desloca. A pergunta deixa de ser "você sabe fazer?" e passa a ser "você sabe por que fazer, quando pedir à máquina que faça e como avaliar o que ela fez?" Essa transição não é uma capitulação diante da tecnologia. Antes, é o reconhecimento de que a finalidade da educação nunca foi a execução de tarefas que máquinas pudessem realizar, mas o desenvolvimento de capacidades que são irredutíveis à automação.

Burke (2002, p. 183) havia percebido que a imprensa, ao multiplicar os textos além da capacidade de qualquer leitor de os dominar, tornou o ideal da *pansofia* cada vez mais irrealista. O que emergiu em seu lugar foi a necessidade de um novo tipo de competência: não a enciclopédia, mas o discernimento; não a quantidade de saber retido, mas a capacidade de navegar, avaliar e articular o saber disponível. A internet radicalizou essa exigência; a IA a leva ao seu ponto mais agudo. Em um ambiente onde a informação é abundante e o argumento pode ser gerado automaticamente, a habilidade crítica se torna o recurso escasso e, portanto, o mais valioso.

Lévy (1993, p. 95) havia antecipado essa reconfiguração ao observar que novas tecnologias intelectuais tornam "irracionais" formas de raciocínio que, antes delas, eram as mais sofisticadas disponíveis. A lógica formal e o cálculo de probabilidades foram, durante séculos, o ápice do raciocínio disciplinado; os sistemas computacionais os tornaram rotineiros. Da mesma forma, a pesquisa bibliográfica exaustiva e a síntese de fontes primárias foram, durante décadas, habilidades nobres do trabalho acadêmico; a IA generativa as torna acessíveis a qualquer usuário em segundos. O que a educação precisa cultivar, portanto, não são as habilidades que a tecnologia pode replicar, mas aquelas que ela não pode, quais sejam, o julgamento ético, a curiosidade genuína, a capacidade de fazer perguntas que ainda não têm resposta e a responsabilidade pelo sentido do que se produz.

Gehlen (1980, p. 4) argumentou que o ser humano só sobrevive porque é capaz de substituir seus órgãos deficientes por artefatos. Mas a história da educação demonstra que essa substituição não é jamais completa, pois o artefato não pensa no lugar do humano. Ele, o artefato pensa *com* o humano, transformando o tipo de esforço que se requer dele. A calculadora não substituiu o matemático; redefiniu o que significa ser um matemático competente. A IA não substituirá o educador; redefinirá o que significa ser um aprendiz responsável e um professor indispensável. A tensão, em última análise, é sempre pedagógica: cabe à educação decidir o que, de cada onda, merece ser preservado como esforço humano irrenunciável.

Nessa perspectiva, a tecnologia retira o fardo da execução mecânica e, com isso, nos coloca diante de uma escolha que é, em si mesma, educativa: ou nos tornamos mais rasos, delegando à máquina não apenas a tarefa, mas o esforço de pensar que a tarefa suscitava; ou ficamos livres para pensar com mais profundidade, ocupando o espaço cognitivo liberado com formas de elaboração que nenhum algoritmo, por mais sofisticado, pode realizar no lugar do sujeito. Essa escolha não é técnica: é política, ética e, fundamentalmente, pedagógica.

5 CONCLUSÃO

A trajetória percorrida ao longo deste trabalho, da argila suméria aos modelos de linguagem de grande escala, confirma uma tese que a perspectiva da *longue durée* braudeliana permite enunciar com clareza, qual seja, a tecnologia sempre esteve na sala de aula. O que muda, de onda em onda, não é a presença da técnica no ato pedagógico, mas o grau de externalização cognitiva que ela propõe e, com isso, a natureza do esforço que resta ao sujeito que aprende.

A primeira onda, representada pela invenção da escrita, externalizou a memória e liberou o pensamento da dependência exclusiva do corpo do mestre. A segunda, a imprensa, externalizou a cópia e tornou possível a escola de massa, com seus currículos uniformes e seus livros didáticos. A terceira, o audiovisual, externalizou a presença e levou o conhecimento além das fronteiras geográficas que o analfabetismo e a distância impunham. A quarta representada pelo nascimento da internet, externalizou o arquivo e transformou o problema do acesso ao saber em problema da avaliação crítica de sua qualidade. A quinta, a inteligência artificial generativa, externaliza o raciocínio discursivo e coloca em suspeita o que, exatamente, ainda precisa ser aprendido pelo próprio sujeito.

Em cada onda, o padrão se reproduz, pois a nova tecnologia é recebida com uma combinação de entusiasmo e ansiedade; provoca resistências entre os guardiões do saber anterior; exige adaptações nos currículos, nas metodologias e no papel do educador; e, por fim, sedimenta-se nas práticas cotidianas até tornar-se invisível como tecnologia velha que já não é percebida como tecnologia, mas habitada como paisagem, nos termos de Cupani (2011, p. 13). Esse ciclo não é uma anomalia. Antes, é a estrutura profunda da relação entre educação e técnica, e reconhecê-la é o primeiro passo para abordá-la com serenidade e rigor.

A contribuição de Ecléa Bosi (1994) ilumina a dimensão que o debate tecnicista frequentemente obscurece, pois a memória não é apenas um repositório, mas um trabalho que forma o sujeito no próprio ato de lembrar. Quando a tecnologia suprime esse trabalho, corre-se o risco de empobrecer não o repertório de informações disponíveis, mas a capacidade de constituição identitária do aprendiz. A crise de motivação que cada onda suscita sintetizada na frase "para que aprender o que a máquina faz melhor?" só pode ser respondida a partir dessa perspectiva, pois aprendemos não apenas para saber, mas para nos tornar o que somos ao aprender.

O desafio da educação, portanto, não é resistir à externalização cognitiva que cada nova onda propõe, nem a aceitar passivamente como progresso inevitável. É, antes, cultivar o que nenhuma ferramenta pode fazer pelo aprendiz que é atribuir sentido ao que se sabe, construir uma identidade intelectual e moral em relação ao saber, e manter viva a curiosidade que é a mola propulsora de todo aprendizado genuíno. Como Stiegler (1998, p. 6) observou, a técnica é "o horizonte de toda possibilidade futura", mas horizonte não é destino. É o limite que convida ao movimento.

Historicamente, o conceito de aprender migrou da retenção para a conexão e, mais recentemente, para o diálogo — inclusive o diálogo com sistemas algorítmicos que falam a língua humana com fluência desconcertante. Cada onda não apaga a anterior. Antes, aproveita de cada etapa que a antecedeu o que propicia avanços. Hoje utilizamos a escrita, inaugurada há cinco milênios, para dialogar com modelos de linguagem treinados sobre bilhões de textos também escritos ao longo de milênios. A primeira onda ressoa na quinta; a argila suméria está na arquitetura de todo processador de linguagem natural. O que a educação precisa garantir é que essa ressignificação progressiva não resulte em um esquecimento da experiência vivida, o que Bosi (1994) chama de memória viva, e da capacidade de reflexão identitária que distingue o sujeito que aprende do sistema que processa.

A pergunta que retorna a cada ciclo, qual seja, "o que o ser humano perde ao delegar essa tarefa à máquina?", não tem resposta definitiva, e é precisamente por isso que ela é pedagogicamente essencial. Ela obriga professores, gestores e aprendizes a distinguir, em cada momento, entre a tarefa que pode ser delegada e o aprendizado que o esforço de a realizar propiciava. A resposta não é técnica, antes é ética, política e existencial. E é nessa ordem de questões que a educação encontra sua irredutibilidade.

Ferramentas ampliam o gesto. A argila ampliou a memória; o livro ampliou o alcance; a tela ampliou a presença; o algoritmo amplia o raciocínio. Mas o gesto que vale a pena ampliar que é a pergunta genuína, a curiosidade cultivada, o julgamento responsável, o cuidado com o outro que aprende, esse gesto nasce de uma decisão que nenhuma tecnologia pode tomar no lugar do educador. Educar ainda é, e sempre será, fundamentalmente, um ato humano.

REFERÊNCIAS

BOSI, Ecléa. **Memória e sociedade**: lembranças de velhos. 3. ed. São Paulo: Companhia das Letras, 1994.

BRUZZI, Demerval Guilarducci. Uso da tecnologia na educação, da história à realidade atual. **Polyphonía**, Goiânia, v. 27, n. 1, p. 475-487, jan./jun. 2016. Disponível em: https://doi.org/10.12957/teias.%y.53758. Acesso em: 10 fev. 2026.

BURKE, Peter. **Uma história social do conhecimento**: de Gutenberg a Diderot. Tradução de Plínio Dentzien. Rio de Janeiro: Jorge Zahar, 2002.

CUPANI, Alberto. **Filosofia da tecnologia**: um convite. Florianópolis: Editora da UFSC, 2011.

DERRIDA, Jacques. **A farmácia de Platão**. 3. ed. Tradução de Rogério Costa. São Paulo: Iluminuras, 2005.

FREIRE, Paulo. **Pedagogia do oprimido**. Rio de Janeiro: Paz e Terra, 1970.

FRIESEN, Norm; HOOD, Nina; WILLIAMSON, Ben. Artificial intelligence in education: ideologies, imaginaries and the ongoing reproduction of schooling. **Learning, Media and Technology**, v. 49, n. 1, p. 115-128, 2024.

GEHLEN, Arnold. **Man in the age of technology**. New York: Columbia University Press, 1980.

GILIOLI, Renato de Sousa Porto. **Educação e cultura no rádio brasileiro**: concepções de radioescola em Roquette-Pinto. 2008. Tese (Doutorado em Educação) — Universidade de São Paulo, São Paulo, 2008.

HAVELOCK, Eric A. **A musa aprende a escrever**: reflexões sobre a oralidade e a literacia da Antiguidade ao presente. Tradução de Maria Leonor Santa Bárbara. Lisboa: Gradiva, 1988.

KLEIN, Leandro Mattana; BERTOTTI, Rudimar Gomes; COSTA, Reginaldo Rodrigues da; MARTINS, Rosilda Baron (org.). **Educação, tecnologia e sociedade**: perspectivas e desafios. Curitiba: Appris, 2020.

LACERDA, Avâner Conceição. **A história da tecnologia na educação**: do quadro de giz à realidade virtual. 2001. Dissertação (Mestrado em Educação) — Universidade Federal de Santa Catarina, Florianópolis, 2001.

LÉVY, Pierre. **A esfera semântica**: computação, cognição, economia da informação. Tradução de Paulo Neves. São Paulo: Annablume, 2023.

LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática. Tradução de Carlos Irineu da Costa. Rio de Janeiro: Editora 34, 1993.

MARX, Karl. Prefácio à Contribuição à Crítica da Economia Política. In: MARX, Karl. **Para a crítica da economia política**. São Paulo: Abril Cultural, 1982 [1859].

McLUHAN, Marshall. **Understanding media**: the extensions of man. New York: McGraw-Hill, 1964.

PINTO, Álvaro Vieira. **O conceito de tecnologia**. São Paulo: Civilização Brasileira, 1972.

RODRIGUES, Doriedson do Socorro. Tecnologia e educação: reflexões marxianas sobre a dimensão ontológica e os determinantes históricos. **Trabalho & Educação**, Belo Horizonte, v. 21, n. 3, p. 193-209, set./dez. 2012.

ROMANOWSKI, Joana Paulin; DOMIT, Juliana; VELASQUE, Marcos Rodrigues; BIANCHESSI, Cleber. Possibilidades das novas tecnologias com a ruptura das metodologias de ensino tradicionais. In: ROMANOWSKI, Joana Paulin et al. (org.). **Educação e tecnologia**. Curitiba: UNINTER, 2021. p. 75-91.

SIMÃO NETO, Antônio. Cenários da educação a distância. **Boletim Técnico do SENAC**, Rio de Janeiro, v. 28, n. 3, p. 24-31, set./dez. 2002.

STIEGLER, Bernard. **Technics and Time, 1**: The Fault of Epimetheus. Tradução de Richard Beardsworth e George Collins. Stanford: Stanford University Press, 1998.

TOMICH, Dale. A Ordem do Tempo Histórico: a Longue Durée e a Micro-História. **Almanack**, Guarulhos, n. 02, p. 38-51, 2º semestre de 2011.

UNESCO. **Artificial intelligence in education**: challenges and opportunities for sustainable development. Paris: UNESCO, 2019.

UNICEF. **Reimagining education**: the role of AI and digital learning for achieving quality education. New York: UNICEF, 2021.

CAPÍTULO 5

INTELIGÊNCIA ARTIFICIAL NA ESCOLA: POSSIBILIDADES, TENSÕES E NOVOS PAPÉIS

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1249



Resumo: O presente texto examina a inserção da inteligência artificial (IA) na educação sob uma perspectiva crítica e antropológica e adota como eixo teórico central a noção de relação com o saber desenvolvida por Bernard Charlot. O estudo parte da contextualização histórica das promessas tecnológicas na escola e argumenta que a IA representa uma ruptura qualitativa em relação às tecnologias anteriores em razão de três características articuladas, a saber, a capacidade de personalização em escala, a operação sobre linguagem natural e a datificação da experiência de aprendizagem. A partir desse diagnóstico, o texto examina o funcionamento dos Sistemas Tutores Inteligentes e dos sistemas adaptativos de aprendizagem, de modo a problematizar os limites da eficiência algorítmica quando confrontada com a concepção charlotiana de saber como processo que exige atividade, sentido e prazer. Na sequência, analisa a reconfiguração do papel docente diante da tecnologia e recusa os dois extremos do debate público, representados pela substituição do professor pela máquina ou pela naturalização acrítica da ferramenta. A seção seguinte trata dos riscos e limites da IA na escola e aborda os vieses algorítmicos que reproduzem desigualdades de raça, gênero e classe. Também discute a privacidade dos dados dos estudantes; o risco de empobrecimento da experiência educativa pela substituição do erro produtivo, do diálogo e da curiosidade pela eficiência do feedback automático; e a crise de autoria e de integridade acadêmica provocada pela IA generativa. As considerações finais recusam falsos consensos e formulam perguntas abertas sobre que escola se quer construir com a tecnologia disponível e a serviço de quem ela efetivamente opera, mobilizando a provocação de Charlot em "Educação ou Barbárie" como horizonte crítico.

Palavras-chave: Inteligência artificial na educação. Relação com o saber. Viés algorítmico. Reconfiguração da docência. Formação humana.

Abstract: This paper examines the integration of artificial intelligence (AI) in education from a critical and anthropological perspective, drawing on Bernard Charlot's theoretical framework of the relationship with knowledge as its central axis. The study begins by situating the current technological moment within a broader historical trajectory of technological promises in schooling, arguing that AI represents a qualitative rupture from previous educational technologies due to three interconnected features: the capacity for large-scale personalization, the operation over natural language, and the datafication of the learning experience. Building on this diagnosis, the text analyses the functioning of Intelligent Tutoring Systems and adaptive learning platforms, questioning the limits of algorithmic efficiency when confronted with Charlot's conception of knowledge as a process that demands activity, meaning, and pleasure. The paper then addresses the reconfiguration of the teaching role in the face of AI, rejecting both extremes of the public debate: the replacement of teachers by machines and the uncritical naturalization of technology. A subsequent section examines the risks and limits of AI in schools, focusing on algorithmic biases that reproduce racial, gender, and class inequalities; student data privacy; the risk of impoverishing the educational experience by replacing productive error, dialogue, and curiosity with automated feedback efficiency; and the crisis of authorship and academic integrity provoked by generative AI. The concluding section refuses easy consensus and poses open questions about what kind of school we want to build with the AI at hand and whose interests it truly serves, drawing on Charlot's provocation in Education or Barbarism as a critical horizon.

Keywords: Artificial intelligence in education. Relationship with knowledge. Algorithmic bias. Reconfiguration of teaching. Human formation.

1 INTRODUÇÃO E CONTEXTUALIZAÇÃO HISTÓRICA

Que transformações produz uma tecnologia quando ela não apenas acrescenta recursos ao cotidiano, mas reconfigura as próprias condições de produção, transmissão e apropriação do saber? Essa questão, que poderia parecer especulativa em outras épocas, apresenta-se hoje como uma indagação urgente e concreta, especialmente no campo da educação. A escola, instituição historicamente constituída para mediar a relação das novas gerações com o patrimônio humano acumulado, enfrenta, no presente, uma alteração de escala e de natureza sem precedentes imediatos, qual seja, a emergência da inteligência artificial (IA) como agente ativo nos processos de ensino e aprendizagem.

Para compreender esse fenômeno em sua profundidade, é preciso situá-lo em uma trajetória histórica mais ampla, que tem seu ponto de inflexão relevante nos anos 1990, quando o acesso à rede mundial de computadores começou a se expandir de forma gradual e depois acelerada. A internet não apenas ampliou o volume de informações disponíveis como alterou estruturalmente a dinâmica da sala de aula, deslocando o professor da posição de detentor exclusivo do acesso ao conhecimento. Conforme observa Charlot (2013), o docente que organiza longas listas de conteúdos a serem memorizados encontra, nesse novo cenário, uma concorrência impossível de vencer:

> Nenhum professor pode, hoje em dia, rivalizar com o Google na coleta de documentos, gráficos, imagens, fotografias, vídeos etc. Mas ainda é necessário entrar com as palavras corretas no mecanismo de pesquisa, ser capaz de escolher entre os múltiplos links propostos e saber avaliar as informações que são apresentadas. (Charlot, 2013, p. 47)

Essa constatação aponta para uma distinção fundamental, que percorre toda a reflexão charlotiana que é a diferença entre informação e saber. Não se trata de uma oposição simples ou hierárquica, mas de uma distinção de natureza. Conforme Charlot (2013, p. 47), o saber é um conjunto de informações articuladas, hierarquizadas, ordenadas e às vezes sistematizadas, que permitem responder a uma questão, resolver um problema, entrar em um universo de significado e sentido. A informação, nessa perspectiva, é matéria-prima; o saber é resultado de um trabalho intelectual do sujeito sobre essa matéria. A internet e suas tecnologias derivadas intensificaram o fluxo da primeira, mas não produziram automaticamente o segundo.

Essa distinção não é apenas epistemológica. Ela é, antes de tudo, antropológica. Em sua obra fundante, Charlot (2000) define a relação com o saber como a relação com o mundo, com o outro e consigo mesmo de um sujeito confrontado com a necessidade de aprender.

Aprender, nessa perspectiva, não é acumular dados externos. Antes, trata-se de um processo pelo qual o sujeito humano se constitui, se humaniza e se singulariza ao apropriar-se do patrimônio produzido pela espécie ao longo de sua história. O sujeito que aprende é, indissociavelmente, humano, social e singular (Charlot, 2000, p. 33). Nenhuma tecnologia, por mais sofisticada que seja, pode realizar esse processo no lugar do aprendiz.

A chegada da internet às escolas revelou, portanto, uma tensão que não era apenas técnica ou infraestrutural. Era uma tensão pedagógica e antropológica que se interrogava como transformar o acesso ampliado à informação em efetiva construção de saberes. Essa questão permaneceu, em grande medida, sem resposta adequada. As experiências de inserção de computadores nas salas de aula, ao longo dos anos 1990 e 2000, repetiram um padrão que a história da educação já havia registrado com a televisão educativa, com o rádio escolar e com os laboratórios de línguas. As promessas de renovação pedagógica radical, formuladas com entusiasmo em cada novo ciclo tecnológico, tenderam a se dissolver diante da persistência das estruturas escolares, das condições concretas de trabalho docente e, sobretudo, da ausência de formação adequada.

1.1 A redução de ruídos históricos: promessas tecnológicas e os limites do passado

Como compreender essa recorrência? Por que as promessas de transformação educacional ancoradas em novas tecnologias se repetem, a cada geração, com a mesma intensidade retórica e os mesmos resultados limitados? Trata-se do que se pode denominar, provisoriamente, de ruído histórico caracterizado pela camada de expectativas infladas, de projetos interrompidos e de soluções técnicas aplicadas a problemas pedagógicos cujas raízes são de outra ordem.

A televisão educativa chegou ao Brasil nos anos 1960 e 1970 com a promessa de democratizar o acesso à educação formal em um território continental. Os resultados, amplamente documentados, ficaram aquém das expectativas, pois a transmissão unilateral de conteúdos não resolvia o problema central representado pela ausência de mobilização interna do aluno. Nos anos 1990, o computador pessoal e, depois, a internet foram recebidos como a solução definitiva para as deficiências do ensino tradicional. No início dos anos 2000, a lousa digital interativa ocupou o mesmo lugar discursivo. Em cada um desses ciclos, o entusiasmo inicial cedeu lugar a uma incorporação parcial e muitas vezes superficial das tecnologias à prática escolar, sem que houvesse uma transformação substantiva nas relações pedagógicas.

Charlot (2013) observa que a presença das tecnologias digitais de informação e comunicação (TDIC) nas escolas produz efeitos culturais e pedagógicos em parte contraditórios. De um lado, ampliam as possibilidades de acesso e diversificam as formas de exposição ao conhecimento. De outro, o uso predominante que os jovens fazem dessas tecnologias não é o estudo. Eles as utilizam para trocar mensagens, compartilhar imagens, buscar entretenimento. O método mais frequente de uso acadêmico das TDIC é, conforme o autor, o copiar-colar, que representa o grau quase zero de atividade de aprendizagem (Charlot, 2013, p. 44). A tecnologia, nesse caso, não apenas deixa de potencializar o aprendizado como pode, em certas condições, tornar-se um obstáculo à mobilização intelectual que o aprender genuíno exige.

Esse diagnóstico não é uma condenação das tecnologias. É, antes, um convite à precisão analítica. O que torna a inteligência artificial uma promessa distinta, no cenário contemporâneo, não é sua novidade em si, mas o conjunto de características que a diferenciam qualitativamente das tecnologias educacionais anteriores. Menciono, agora, três aspectos que merecem destaque.

O primeiro diz respeito à capacidade de personalização em escala. Diferentemente da televisão ou do computador conectado à internet, os sistemas de IA baseados em aprendizado de máquina podem adaptar conteúdos, ritmos e desafios em tempo real, de acordo com os padrões de resposta de cada estudante (Litwak; Gama, 2024 *apud* Goethel e Fonseca, 2025). Essa capacidade de ajuste individualizado, que antes exigia a presença constante e atenta de um tutor humano, torna-se tecnicamente viável em larga escala.

O segundo aspecto está relacionado à natureza dos dados que a IA mobiliza. As tecnologias anteriores operavam sobre conteúdos estruturados previamente. A IA generativa, em especial a partir de 2022, opera sobre volumes massivos de linguagem natural, produzindo respostas, textos, análises e materiais didáticos de forma fluida e contextualizada. Isso representa uma ruptura com o modelo de banco de informações estático, pois a IA passa a ser um interlocutor ativo no processo pedagógico. Conforme apontam Heggler, Szmoski e Miquelin (2025), os algoritmos de aprendizado de máquina são treinados com dados fornecidos pela população em geral, o que implica não apenas potencialidades, mas também riscos de reprodução de vieses e desigualdades sociais já presentes nesses dados.

O terceiro aspecto, e talvez o mais relevante do ponto de vista educacional, é a "datificação[3]" do processo de aprender. Pela primeira vez na história da educação escolarizada,

[3] O termo datificação refere-se ao processo sociotécnico de tradução de aspectos da vida social, bem como de comportamentos e interações humanas, em dados digitais que podem ser quantificados e processados por

é possível registrar, armazenar e analisar automaticamente uma quantidade enorme de dados sobre o comportamento cognitivo dos estudantes como, por exemplo, o tempo de resposta, os erros mais frequentes, os padrões de navegação, as estratégias de resolução de problemas. Esse processo, que a literatura internacional tem denominado de datificação do ensino, transforma a experiência escolar em fluxo contínuo de dados analisáveis. Isso abre possibilidades inéditas de diagnóstico pedagógico, mas também coloca questões éticas e políticas que a escola ainda não aprendeu a enfrentar.

1.2 O saber como processo antropológico e as tensões da era da ia

É precisamente nesse ponto que a fundamentação teórica de Bernard Charlot se revela não apenas pertinente, mas necessária. A perspectiva charlotiana recusa qualquer concepção de saber como objeto externo ao sujeito, como coisa que se transfere, se armazena ou se acessa. O saber, para Charlot (2000, p. 61), é sempre relação, esta com o mundo, com os outros e consigo mesmo. Aprender é, portanto, um ato antropológico antes de ser um ato informacional. Charlot (2000) situa essa afirmação na condição humana de nascimento inacabado, pois o filhote humano torna-se humano apenas ao se apropriar de uma parte do patrimônio que a espécie construiu ao longo de sua história, patrimônio que se apresenta sob a forma de saberes, instrumentos, práticas, sentimentos e formas de relação (Charlot, 2001, p. 42).

Essa estrutura antropológica do aprender tem uma consequência direta para a avaliação das tecnologias educacionais. Se aprender é uma mobilização interna do sujeito, um engajamento em atividade intelectual que exige sentido e implica prazer, então nenhuma tecnologia pode substituir esse processo. O que as tecnologias podem fazer é ampliar ou restringir as condições nas quais essa mobilização ocorre. Conforme sintetiza Charlot (2013, p. 74), a transformação da escola está estritamente relacionada com a atividade do aluno. Nenhum texto legal, nenhum computador em si mesmo conseguiu transformar a escola. A última instância é a atividade do aluno.

Esse princípio, formulado antes da emergência da IA generativa, adquire nova urgência no presente. Os sistemas de IA podem gerar respostas, resumos, exercícios e avaliações de forma autônoma. Podem produzir materiais aparentemente personalizados e oferecer

algoritmos. Segundo Van Dijck (2017, p. 33), essa lógica permite a conversão de ações cotidianas em metadados, o que possibilita a monitoração e a predição de comportamentos em larga escala, resultando em uma reestruturação das práticas institucionais e pedagógicas.

feedback imediato. O risco, já identificado por docentes em pesquisas recentes, é que os estudantes passem a confiar cegamente nas respostas geradas pela IA, sem verificar a veracidade das informações, e que a facilidade de acesso a respostas prontas comprometa o desenvolvimento da autonomia intelectual (Trindade, 2025, p. 8). Em termos charlotianos, o risco é que o sujeito delegue à máquina o processo de mobilização que somente ele deveria realizar.

A questão que se coloca, portanto, não é se a IA deve ou não estar presente na escola, pois essa presença já é uma realidade que se expande aceleradamente. A questão é em que condições a IA pode contribuir para aprofundar a relação com o saber, e em que condições ela tende a empobrecê-la, transformando o aprender em mero acesso a respostas prontas? Essa indagação orienta as análises que se desenvolvem nas seções seguintes, nas quais serão examinadas as possibilidades pedagógicas, as tensões estruturais e as implicações para os novos papéis do professor nesse cenário.

Conforme nos lembra Charlot (2022, p. 47), não é a tecnologia em si que determina os efeitos educacionais, mas a relação do homem com a tecnologia. Uma relação inculta ou mágica com a IA reproduzirá, em escala ampliada, os mesmos equívocos históricos que marcaram as ondas tecnológicas anteriores. Uma relação crítica, fundamentada e pedagógica, por outro lado, pode permitir que a escola ocupe, com humanidade, um mundo crescentemente tecnológico.

2 O NÚCLEO TÉCNICO E A PERSONALIZAÇÃO EM ESCALA

Que operações realiza um sistema de inteligência artificial quando aplicado à aprendizagem? E em que medida essas operações dialogam com o que Charlot (2000) compreende como a dimensão interna e subjetiva do ato de aprender? Responder a essas perguntas exige, inicialmente, uma aproximação ao funcionamento dos principais dispositivos de IA em contextos educativos, para depois problematizá-los a partir de uma perspectiva pedagógica mais ampla.

Os Sistemas Tutores Inteligentes (STI) constituem uma das categorias mais estudadas e desenvolvidas na intersecção entre IA e educação. Sua origem remonta aos anos 1970, com os primeiros experimentos em ensino assistido por computador, mas sua forma contemporânea incorporou, a partir dos anos 1990, modelos computacionais inspirados na psicologia cognitiva. Um STI opera, em linhas gerais, a partir de quatro componentes articulados entre si, que compreendem um modelo do domínio, que representa o conhecimento a ser ensinado; um modelo do aprendiz, que registra o estado atual de

compreensão do estudante; um modelo pedagógico, que determina como e quando intervir; e uma interface, que faz a mediação de interação (Trindade, 2025). A integração desses componentes permite que o sistema diagnostique lacunas de aprendizagem e ofereça intervenções adaptadas ao perfil do estudante, sem que o professor precise conduzir cada ajuste manualmente.

Os sistemas adaptativos de aprendizagem representam uma evolução e uma expansão dessa lógica. Ao incorporar algoritmos de aprendizado de máquina, esses sistemas não apenas respondem a um modelo pedagógico pré-definido, mas atualizam continuamente o perfil do estudante com base nos dados gerados pela própria interação. Plataformas desse tipo coletam informações sobre o tempo de resposta em cada questão, os padrões de erro, as trilhas de conteúdo percorridas, as sequências de acesso e a frequência de retorno a determinados materiais. A partir dessas informações, os algoritmos inferem estados cognitivos e emocionais do estudante, como dificuldade, desengajamento ou domínio de um conceito, e ajustam o percurso de aprendizagem em tempo real (Heggler; Szmoski; Miquelin, 2025). Conforme descrevem Dakakni e Safa (2023 *apud* Heggler; Szmoski; Miquelin, 2025, p. 7), a personalização pode contribuir para a identificação dos pontos fortes e fracos dos alunos, adaptando-se ao que o estudante já sabe e às suas novas necessidades de aprendizagem.

Essa capacidade adaptativa é apresentada, nos discursos sobre IA educacional, como uma resposta à heterogeneidade das turmas. As salas de aula brasileiras, marcadas por diferenças de ritmo, de repertório cultural e de trajetória escolar, representam um desafio permanente para a organização do ensino. A lógica do ensino simultâneo, que pressupõe um estudante médio e aplica o mesmo ritmo a todos, produz historicamente exclusões que a escola ainda não conseguiu resolver. A possibilidade de que um sistema computacional acompanhe individualmente o percurso de cada estudante, sinalizando dificuldades antes que se acumulem e sugerindo caminhos alternativos de acesso ao conteúdo, configura uma promessa que merece ser levada a sério.

2.1 A Personalização e seus Limites: o saber entre o algoritmo e o sujeito

Entretanto, uma análise mais cuidadosa revela que a personalização algorítmica opera sobre uma concepção de aprendizagem que é, em aspectos fundamentais, distinta daquela que Charlot (2000; 2013) articula. Para os sistemas adaptativos, aprender significa, em última instância, adequar respostas a padrões esperados e percorrer trilhas de conteúdo com taxas crescentes de acerto. O progresso é medido por métricas de desempenho compostas por

percentual de acertos, tempo médio de resposta, nível de complexidade alcançado. Essa mensuração é, certamente, útil como indicador parcial. Mas ela não captura o que Charlot (2013, p. 47) identifica como o núcleo do processo educativo que é o saber como um conjunto de informações articuladas, hierarquizadas e ordenadas que permitem entrar em um universo de significado e sentido.

O saber, em Charlot (2000, p. 61), não é um objeto separado do sujeito que pode ser transferido de um repositório a outro. Ele é sempre resultado de uma atividade intelectual do próprio aprendiz, uma atividade que exige mobilização interna, que produz sentido e que, quando encontra seu objeto verdadeiro, gera prazer. A equação pedagógica fundamental que Charlot sintetiza como atividade, sentido e prazer não pode ser decomposta em variáveis mensuráveis sem perda essencial de seu conteúdo (Charlot, 2013, p. 75). Um estudante pode percorrer uma trilha adaptativa do início ao fim, registrar cem por cento de acertos e não ter aprendido nada no sentido pleno do termo, pois não encontrou razão interna para se mobilizar intelectualmente.

Conforme Goethel e Fonseca (2025), a personalização algorítmica, ao determinar caminhos ideais e adaptar o aprendizado conforme padrões preditivos, corre o risco de limitar a autonomia do estudante na escolha de motivos e objetivos da atividade. Essa dinâmica pode impedir a busca exploratória e a superação de desafios que dão origem ao sentido pessoal, transformando a resolução de problemas em uma sequência de operações reativas focadas na validação algorítmica. A atividade educativa, nesse cenário, deixa de ser orientada pela necessidade de compreender e passa a ser orientada pela busca de acertos, o que corresponde à substituição da atividade pela ação sem motivo interno.

Não se trata, portanto, de uma oposição simples entre tecnologia e humanismo. Trata-se de uma distinção entre dois modos de conceber o que é aprender. Se o objetivo educacional for aumentar desempenhos em tarefas específicas e mensuráveis, os sistemas adaptativos podem ser eficazes. Se o objetivo for a formação de um sujeito capaz de dar sentido ao mundo, de relacionar saberes, de formular perguntas próprias e de construir uma relação ativa com o conhecimento, então a eficiência algorítmica, por si só, é insuficiente. Essa distinção não invalida os sistemas adaptativos. Pelo contrário, ela os situa em seu lugar adequado que é o de ser instrumentos que podem apoiar determinadas dimensões do processo educativo, mas que não podem substituir a mediação pedagógica humana.

2.2 Vieses Algorítmicos e o Problema da Equidade

Uma segunda ordem de problemas emerge quando se examina não a eficácia, mas a equidade dos sistemas baseados em IA. Heggler, Szmoski e Miquelin (2025) demonstram, em revisão sistemática que os vieses algorítmicos na educação podem comprometer tanto a equidade quanto a eficácia dos processos educativos. Os autores identificam que esses vieses não surgem apenas durante o treinamento e o desenvolvimento da programação; podem ocorrer também fora dessas etapas, pois os modelos algorítmicos são treinados para resolver determinados problemas do mundo real, mas podem apresentar distorções quando seus resultados interagem com os fenômenos sociais (Heggler; Szmoski; Miquelin, 2025, p. 9).

O mecanismo central do problema é que os algoritmos de aprendizado de máquina aprendem a partir de dados históricos. Se esses dados refletem padrões de desigualdade, discriminação ou exclusão já presentes na sociedade, o algoritmo tende a reproduzir e, em alguns casos, ampliar essas desigualdades. Um caso emblemático, documentado por Baker e Hawn (2022 *apud* Heggler; Szmoski; Miquelin, 2025), ocorreu no Reino Unido em 2020, quando um sistema automatizado de classificação de notas atribuiu resultados mais baixos a estudantes de escolas públicas e resultados superiores, além das previsões dos professores, a estudantes de escolas privadas. O algoritmo havia sido treinado com dados que codificavam desigualdades estruturais preexistentes, e sua aplicação as reproduziu em larga escala.

No campo da educação brasileira, o risco é particularmente relevante. As desigualdades de acesso, de infraestrutura, de repertório cultural e de trajetória escolar que caracterizam o sistema educacional do país produzem dados que, quando usados para treinar algoritmos sem as devidas correções, podem resultar em sistemas que recomendam trilhas mais pobres para estudantes que já enfrentam desvantagens. Conforme Williams (2024 *apud* Heggler; Szmoski; Miquelin, 2025, p. 4), se historicamente as alunas foram menos propensas a se inscrever em cursos de computação, um sistema de recomendação pode continuar sugerindo cursos de humanidades para estudantes do sexo feminino com base nas tendências históricas, reforçando a segregação que se pretendia superar.

Esses problemas não são inevitáveis, mas também não são solucionáveis apenas por ajustes técnicos. Eles exigem uma reflexão política sobre quais dados são coletados, por quem, com que finalidade e sob qual forma de controle institucional será aplicada sobre eles. Conforme apontam Ito-Jaeger *et al.* (2023 *apud* Heggler; Szmoski; Miquelin, 2025), questões relativas à proteção de dados e ao preconceito algorítmico são complexas demais para serem resolvidas apenas pelos desenvolvedores. Nesse sentido, as partes interessadas, incluindo

professores, gestores, estudantes e famílias, devem estar ativamente envolvidas nas discussões sobre sua implementação.

2.3 Usos instrumentais e usos sistêmicos: uma distinção necessária

Para além dos sistemas adaptativos e dos algoritmos de personalização, é preciso distinguir dois planos em que a IA opera nas instituições educacionais. O primeiro plano é o instrumental, que corresponde ao uso de ferramentas de IA como recursos de apoio ao trabalho pedagógico cotidiano. Nesse plano, o professor utiliza plataformas de geração de materiais didáticos, de correção automatizada de textos, de sugestão de atividades ou de síntese de conteúdo. Docentes pode relatar, por exemplo, o uso do ChatGPT para gerar ideias de atividades e exemplos contextualizados, ou para revisar textos e adaptar linguagens de acordo com o perfil da turma (Trindade, 2025, p. 8). Nesse uso instrumental, a IA atua como ferramenta auxiliar, e o professor mantém o controle sobre as decisões pedagógicas.

O segundo plano é o sistêmico, e é nele que as implicações se tornam mais profundas e mais controversas. O uso sistêmico da IA ocorre quando os algoritmos passam a operar sobre decisões curriculares, avaliatórias e de gestão escolar, isto é, quando a IA deixa de ser uma ferramenta nas mãos do professor e passa a ser uma infraestrutura que organiza o próprio funcionamento da escola. Algoritmos de análise preditiva utilizados para classificar estudantes em grupos de risco, sistemas de recomendação curricular que determinam quais conteúdos cada turma deve priorizar e plataformas que geram relatórios de desempenho usados para avaliação docente são exemplos de uso sistêmico. Eles colocam em jogo dimensões de poder, de autonomia e de responsabilidade que vão muito além do debate técnico.

A distinção entre usos instrumentais e sistêmicos é relevante porque os riscos associados a cada um são de natureza diferente. O uso instrumental pode ser monitorado, avaliado e corrigido pelo professor a partir de sua experiência e de seu julgamento pedagógico. O uso sistêmico, por sua vez, tende a se tornar invisível e difícil de questionar, pois os algoritmos que tomam ou informam decisões estruturais raramente são transparentes em seus critérios. Conforme Tang e Su (2024 *apud* Heggler; Szmoski; Miquelin, 2025, p. 31), as decisões educacionais tomadas por modelos preditivos de IA levam a preconceitos e discriminação, mas não estão claros quais recursos e atributos algorítmicos seriam necessários para reduzir esse preconceito. A opacidade dos algoritmos, nesse cenário, é ela mesma um problema político.

No Brasil, o contexto torna essas questões ainda mais urgentes. A disparidade no acesso a recursos tecnológicos, a insuficiência de formação docente específica e a infraestrutura

desigual entre instituições públicas e privadas constituem barreiras que tornam a implementação da IA na educação um fenômeno socialmente assimétrico. A ausência dessas condições não elimina a chegada da IA às escolas; ela apenas assegura que essa chegada seja desigual, beneficiando mais intensamente os contextos que já dispõem de melhores condições (Paixão *et al.*, 2025 *apud* Moura; Carvalho, 2025).

Retomando a perspectiva de Charlot (2022), é possível afirmar que o que está em jogo não é a tecnologia em si, mas a relação que a escola e os sujeitos que nela habitam estabelecem com ela. Uma relação inculta ou mágica com a IA, que a receba como solução autossuficiente para os problemas educacionais ou que a implante sem reflexão crítica sobre seus efeitos, reproduzirá os mesmos padrões de fracasso que marcaram as ondas tecnológicas anteriores. Uma relação crítica, pedagogicamente orientada e politicamente responsável, pode, ao contrário, fazer da IA um instrumento que amplia as condições em que a relação com o saber pode se desenvolver de forma mais plena e mais justa.

3 A RECONFIGURAÇÃO DA DOCÊNCIA DIANTE DA INTELIGÊNCIA ARTIFICIAL

Que professor a era da inteligência artificial requer? E o que, nessa figura, permanece insubstituível? Essas perguntas percorrem com crescente insistência os debates pedagógicos contemporâneos, e as respostas que circulam tendem, com frequência preocupante, a dois extremos igualmente imprecisos. De um lado, o alarmismo que anuncia a extinção da profissão docente ante a eficiência dos algoritmos. De outro, o otimismo ingênuo que descarta qualquer risco, afirmando que o professor apenas ganhará tempo e recursos com a automação de tarefas. Nenhuma dessas posições resiste a uma análise mais cuidadosa da natureza do ato educativo.

A posição mais fundamentada é também a mais exigente, afirmando que o professor não será substituído pela máquina, mas tendo sua função se transformada de modo substantivo. Essa transformação não é meramente instrumental, como se bastasse aprender a operar novas ferramentas. Ela implica uma revisão profunda do que o professor faz, de como faz e, sobretudo, do que ninguém mais pode fazer em seu lugar. Compreender essa transformação exige, antes de qualquer coisa, examinar o que a IA efetivamente não consegue realizar no interior do processo pedagógico.

3.1 O que a inteligência artificial não faz: saber, sentido e a presença do outro

Charlot (2013) já havia identificado, antes mesmo da emergência da IA generativa, a crise do que chamou de professor de informações. Esse profissional, que organiza listas de conteúdos a serem memorizados, foi tornado obsoleto não pela IA, mas pela própria internet. A novidade da IA é que ela avança sobre um território que a internet ainda não havia ocupado de forma tão direta, realizando tarefas tais como a produção de textos coerentes, a elaboração de explicações aparentemente personalizadas, a sugestão de atividades e até a simulação de diálogo pedagógico. Isso poderia sugerir que o professor de saber, aquele que Charlot (2013, p. 47) define como capaz de gerir a informação de forma eficaz, de construir e transmitir sentido, de incentivar nos jovens uma mobilização para aprender que se estenda além da simples coleta de informações, também estaria ameaçado. Essa inferência, entretanto, é equivocada.

A razão é de ordem antropológica. Conforme Charlot (2000, p. 53), o sujeito só pode aprender pela mediação do outro, seja de forma direta ou indireta. Essa afirmação não é uma preferência pedagógica; é uma constatação sobre a estrutura do aprender humano. O saber, como se viu, não é um objeto que se transmite de um depósito a outro. Ele é resultado de uma atividade do sujeito sobre o mundo, atividade que é sempre socialmente situada e historicamente mediada. O outro, que aparece nessa mediação, não é apenas uma fonte de informação. É uma presença que interpela, que provoca, que frustra quando necessário e que reconhece o aprendiz como sujeito capaz. Charlot (2000) descreve essa dimensão ao articular a relação com o saber como também uma relação com os outros e consigo mesmo. Aprender é entrar em um campo de sentido habitado por outras vozes, outros olhares, outras expectativas.

Um sistema de IA, por mais sofisticado que seja seu modelo de linguagem, não é um outro nesse sentido. Ele não tem uma história, não possui uma relação com o saber que lhe seja própria, não é afetado pelo que o estudante aprende ou deixa de aprender. O docente ouvido por Trindade (2025, p. 8) sintetizou essa percepção de forma precisa ao afirmar que a IA não ensina valores, não compreende emoções e não substitui a mediação humana. O participante seguinte complementou que a tecnologia pode ampliar a atuação docente, mas o professor continua sendo o responsável por dar sentido ao processo educativo (Trindade, 2025, p. 8). Essas narrativas docentes expressam, de forma intuitiva, uma compreensão que a teoria charlotiana sustenta com rigor, isto é, o saber exige um outro humano, não porque a transmissão de conteúdo o requeira, mas porque a construção de sentido é um processo intersubjetivo que não se realiza sem a presença, mesmo que mediada, de outra consciência.

3.2 Do transmissor ao curador e interlocutor crítico

Se a função de transmissor de conteúdo foi progressivamente esvaziada pela internet e agora pela IA, o que ocupa seu lugar? A literatura recente tem proposto a metáfora do professor curador como uma das figuras emergentes da docência contemporânea. A curadoria, no campo da informação e da cultura, designa a prática de selecionar, organizar, contextualizar e apresentar materiais de forma que adquiram coerência e sentido para um determinado público. Transposta para a prática pedagógica, essa metáfora aponta para um professor que não mais detém o monopólio do acesso ao conhecimento, mas que exerce um papel insubstituível na qualificação do que circula, na organização das trajetórias de aprendizagem e na problematização das fontes.

Trindade (2025) observa que o uso da IA tem promovido uma redefinição do papel docente, deslocando o professor de uma posição centralizadora para a de mediador do conhecimento e curador de informações. Esse deslocamento exige novas competências de natureza técnica, pedagógica e ética, de modo que o educador seja capaz de aliar o domínio da tecnologia à sensibilidade humana, orientando os estudantes na construção crítica e responsável do saber. Essa combinação de dimensões não é acessória; ela é o núcleo do que a docência contemporânea precisa ser.

Uma das possibilidades mais interessantes nesse novo papel diz respeito ao uso pedagógico das próprias limitações da IA. Os modelos de linguagem de grande porte produzem, com regularidade conhecida, o fenômeno denominado alucinação que é a geração de informações factualmente incorretas, mas formalmente coerentes e fluentes. Uma IA pode afirmar com aparente segurança datas erradas, atribuir citações inexistentes a autores reais, ou construir argumentos logicamente válidos a partir de premissas falsas. Esse fenômeno, que representa um risco quando o estudante usa a IA sem criticidade, transforma-se em oportunidade pedagógica quando o professor o converte em material de análise. Propor que os estudantes verifiquem as afirmações de um sistema de IA, identifiquem os erros, pesquisem as fontes corretas e reflitam sobre os mecanismos que tornam a tecnologia convincente mesmo quando erra é uma sequência didática que mobiliza habilidades de leitura crítica, pesquisa documental e raciocínio epistêmico que dificilmente seriam ativadas por uma aula expositiva convencional.

Nesse sentido, o professor se converte também em formulador de perguntas que a IA não responde bem. Os sistemas de inteligência artificial têm dificuldades estruturais com perguntas abertas, que exijam tomada de posição ética situada, que demandem compreensão

de contextos locais e particulares, ou que requeiram a articulação entre saberes de campos distintos sem uma resposta canônica disponível nos dados de treinamento. Um professor que orienta seus alunos a fazerem à tecnologia exatamente esse tipo de pergunta, e que depois conduz a análise coletiva das respostas obtidas, está exercendo uma função que nenhum sistema automatizado pode cumprir que é a de ser mediador de uma experiência de descoberta sobre os próprios limites do conhecimento disponível.

3.3 A mediação pedagógica como necessidade antropológica

A fundamentação charlotiana permite ir além das metáforas para compreender por que a mediação pedagógica humana é uma necessidade estrutural do processo educativo, e não apenas uma preferência culturalmente condicionada. Charlot (2000, p. 53) afirma que o sujeito não tem uma relação com o saber, ele é relação com o saber. Isso significa que aprender não é adicionar informação a um receptáculo neutro, mas transformar-se a si mesmo pela apropriação do mundo. Essa transformação é sempre singular, pois cada sujeito chega ao saber a partir de sua história, de seus desejos, de suas identificações e de suas resistências. Nenhum algoritmo, por mais refinado que seja seu modelo do aprendiz, consegue captar essa singularidade em sua profundidade, pois ela não é redutível aos padrões de resposta que os dados de interação registram.

O professor, nessa perspectiva, não é apenas alguém que organiza conteúdo ou que monitora percursos. Ele é um mediador que reconhece o aprendiz como sujeito, que o interpela a partir de sua história particular, convidando-o a entrar em um campo de sentido que vai além do que o estudante já conhece. Charlot (2013) descreve o professor como um trabalhador da contradição, cuja função é manter um mínimo de coerência em uma sociedade atravessada por tensões entre o ideal da formação humana e a lógica do desempenho e da concorrência (Charlot, 2013, p. 59). Essa função, por sua própria natureza, é ética e política antes de ser técnica. Ela não pode ser delegada a um sistema que opera sem valores, sem projeto de sociedade e sem responsabilidade pelo destino de sujeitos concretos.

Essa dimensão ética da mediação pedagógica é também o que distingue o professor do tutor inteligente quando o processo educativo envolve a formação de sujeitos para além do domínio de competências técnicas. Goethel e Fonseca (2025) observa que a IA, ao deslocar o centro da mediação para sistemas controlados por interesses de mercado, corre o risco de institucionalizar uma nova forma de colonização da subjetividade, o que os autores denominam colonialismo algorítmico. O antídoto a esse risco não é a recusa da tecnologia,

mas a reafirmação da intencionalidade pedagógica como princípio orientador de toda a prática docente. O algoritmo pode ser uma ferramenta útil nesse processo, desde que não substitua o vínculo humano que sustenta o processo educativo (Goethel e Fonseca, 2025, p. 16).

3.4 A formação docente para além do operacional

Como preparar professores para essa nova configuração da docência? Essa é uma das questões mais urgentes e mais mal respondidas no debate atual sobre IA e educação. A tendência dominante nas iniciativas de formação existentes aponta para um treinamento operacional que abrange desde a utilização do ChatGPT e a elaboração de prompts eficazes até a integração de plataformas adaptativas ao planejamento de aula, o que se mostra necessário, porém insuficiente para garantir a eficácia da tecnologia na educação.

Moura e Carvalho (2025) identificam que muitos professores ainda não tiveram acesso à capacitação adequada para lidar com as novas tecnologias e, especialmente, com os recursos baseados em IA, o que pode gerar resistência, insegurança e uso inadequado desses dispositivos no processo pedagógico. A insuficiência de formação, entretanto, não se resolve apenas com cursos de capacitação técnica. Ela se resolve, principalmente, com uma formação que desenvolva no professor a capacidade de compreender o que a IA é e como ela funciona, de analisar criticamente os efeitos pedagógicos de seu uso e de tomar decisões fundamentadas sobre quando, como e para quê utilizá-la.

Paixão *et al.* (2025 *apud* Moura; Carvalho, 2025) sublinham que a formação docente, regulamentada por diretrizes nacionais, assume papel central nesse processo, pois é por meio dela que se constrói a competência profissional necessária para integrar tecnologias de forma crítica e significativa. Não se trata, portanto, de uma formação orientada apenas para a eficiência operacional, mas de uma formação que inclua dimensões epistemológicas, éticas e políticas do uso da IA na educação. O professor que compreende como um algoritmo de recomendação opera, quais dados ele coleta e quais padrões ele tende a reproduzir, está em condições muito melhores de exercer o julgamento pedagógico necessário para decidir se e como utilizar esse recurso com seus alunos.

Essa formação mais ampla implica, ainda, desenvolver nos professores a capacidade de problematizar a própria IA diante dos alunos, tornando o uso das ferramentas digitais uma oportunidade de letramento crítico. Conforme aponta Goethel e Fonseca (2025), o letramento crítico em IA não se limita ao domínio operacional das ferramentas, mas abrange a capacidade de compreender, analisar criticamente e usar a IA de forma ética e responsável. Para que o

professor possa cultivar esse letramento nos estudantes, ele precisa, antes, tê-lo desenvolvido em si mesmo. Isso demanda tempo, apoio institucional e, sobretudo, uma política de formação continuada que não trate a IA como mais uma ferramenta a ser incorporada ao repertório docente, mas como um fenômeno que reconfigura as condições estruturais em que o ensino e a aprendizagem ocorrem.

Essa reconfiguração das condições estruturais é também o que torna insuficiente qualquer resposta puramente individual ao desafio da IA na educação. O professor que enfrenta sozinho, sem apoio técnico, sem tempo para formação e sem diálogo com seus pares, a tarefa de compreender e integrar criticamente tais ferramentas tecnológicas, está condenado a estratégias de sobrevivência que, como observou Charlot (2013, p. 49), tendem a esvaziar o sentido das inovações ao reinterpretá-las na lógica do mínimo esforço possível. A transformação da docência no contexto da IA é, portanto, também uma questão de política educacional, de gestão escolar e de cultura institucional. O futuro da educação digital depende não apenas da disponibilidade de ferramentas tecnológicas, mas principalmente da valorização do trabalho docente, do diálogo com a comunidade escolar e da construção coletiva de práticas pedagógicas transformadoras e humanizadas (Moura; Carvalho, 2025, p. 13).

4 LIMITES, ÉTICA E RISCOS DA INTELIGÊNCIA ARTIFICIAL NA ESCOLA

Compreender o que a inteligência artificial pode oferecer à educação é indispensável, mas igualmente necessário é mapear o que ela ameaça. Os riscos associados à incorporação crescente da IA nos ambientes escolares não são especulativos nem futuros, pois já se manifestam em experiências documentadas, em dados de pesquisa e nas narrativas dos próprios professores e estudantes. Examiná-los com rigor não é uma postura reacionária; é uma condição para que o uso da tecnologia seja, de fato, educativo.

4.1 O sentido pessoal delegado ao algoritmo

Um dos riscos mais silenciosos e mais profundos identificados pela pesquisa recente diz respeito ao que Goethel e Fonseca (2025) denominam delegação do sentido pessoal ao algoritmo. As autoras, fundamentadas na Teoria da Atividade de Leontiev, argumentam que a atividade de estudo dotada de sentido é aquela em que o sujeito reconhece sua ação como significativa em sua trajetória de vida, em suas necessidades e projetos. Quando essa dimensão é substituída por um sistema de instruções automáticas ou pela delegação de tarefas à IA,

perde-se a mediação cultural necessária para a constituição da subjetividade (Goethel; Fonseca, 2025, p. 14).

O mecanismo descrito pelas autoras manifesta-se quando, ao delegar à tecnologia etapas cruciais da atividade de estudo, como a organização de um argumento, a síntese de fontes, a formulação de hipóteses ou a elaboração de uma redação, o estudante mantém formalmente a tarefa, mas perde o processo que lhe conferia sentido. A ação se converte em operação, e o sujeito, em executor. Goethel e Fonseca (2025, p. 19) sintetizam essa transformação com precisão, afirmando que a delegação da organização da aprendizagem ou de etapas cruciais das tarefas ao algoritmo pode transformar a atividade do estudante de uma busca motivada por propósito em uma sequência de ações operacionais focadas na validação algorítmica.

Evidências neurocientíficas corroboram essa análise pedagógica. O estudo de Kosmyna *et al.* (2025, *apud* Goethel; Fonseca, 2025) revelou que participantes que utilizaram modelos de linguagem de grande porte para redigir textos e depois foram solicitados a escrever sem auxílio tecnológico apresentaram conectividade neural mais fraca e menor engajamento das redes cerebrais associadas ao planejamento e ao raciocínio crítico, em comparação com aqueles que sempre escreveram de forma autônoma. No mesmo estudo, oitenta e três por cento dos participantes do grupo que utilizou IA falharam em fornecer uma citação correta de seus próprios ensaios na sessão posterior, e a percepção de propriedade do texto foi significativamente menor para esse grupo. O conhecimento simplesmente não foi internalizado, e a memória, fragilizada.

Essa constatação dialoga diretamente com a perspectiva charlotiana. Charlot (2013, p. 75) propõe que aprender é atividade, sentido e prazer, portanto, uma equação que não se decompõe sem perda essencial. O erro produtivo, a resistência do objeto ao pensamento, a busca frustrada e a descoberta subsequente são elementos constitutivos do processo de aprender. Quando a IA elimina o esforço antes que ele produza sentido, elimina também a possibilidade de que o aprendiz se reconheça como sujeito capaz. A eficiência técnica, nesse caso, opera contra a formação.

4.2 Vieses algorítmicos e a reprodução de desigualdades de raça, gênero e classe

A falta de neutralidade nos sistemas algorítmicos resulta na manutenção de desigualdades de raça, gênero e classe, uma vez que as ferramentas operam sobre padrões históricos de exclusão. Heggler, Szmoski e Miquelin (2025, p. 9) argumentam que os vieses

técnicos influenciam a avaliação acadêmica e as recomendações personalizadas, o que pode "comprometer a equidade e a eficácia dos processos educativos". Na prática, essas disparidades manifestam-se em falhas de monitoramento remoto, na segregação de gênero em trilhas de aprendizagem e na redução do suporte institucional destinado a estudantes vulnerabilizados, consolidando assimetrias em vez de superá-las (Heggler, Szmoski, Miquelin, 2025).

Charlot (2022) já havia alertado para a ilusão de neutralidade dos algoritmos ao analisar o fenômeno do Big Data, já que os dados nunca são neutros, pois incorporam, em sua própria coleta e categorização, toda uma série de escolhas que codificam hierarquias, normas e preconceitos sociais (Charlot, 2022, p. 147). No campo educacional, essa codificação tem consequências diretas sobre quais estudantes são identificados como potencialmente capazes, quais são apontados como de risco, e quais trajetórias lhes são recomendadas. A aparência de objetividade do algoritmo não apenas encobre essas escolhas, antes, as legitima.

4.3 Privacidade de dados e aprofundamento da desigualdade estrutural

A implantação de sistemas de IA nas escolas pressupõe a coleta massiva de dados sobre os estudantes, quais sejam, padrões de acesso, tempos de resposta, sequências de erro, comportamentos de navegação, resultados em avaliações e, em alguns sistemas mais invasivos, até dados biométricos e de reconhecimento facial. Esse processo de datificação da educação, transforma a experiência escolar em fonte de extração de valor para plataformas e empresas que desenvolvem os sistemas.

Charlot (2022, p. 144) descreve esse fenômeno em termos que se aplicam diretamente ao contexto educacional. Segundo o autor, por sermos constantemente registrados em bases de dados interconectadas, acabamos por ser socialmente definidos por um perfil digital de dados que não solicitamos e que é muito difícil controlar. Nas escolas, essa dinâmica é agravada pelo fato de que os dados são coletados de crianças e adolescentes, sujeitos em formação que raramente têm consciência de que sua experiência de aprendizagem está sendo transformada em insumo para o treinamento e o aprimoramento de sistemas comerciais.

No Brasil, a desigualdade de infraestrutura entre redes públicas e privadas produz um paradoxo particularmente grave. Os estudantes com melhores condições socioeconômicas, matriculados em escolas com maior capacidade de investimento, terão acesso a sistemas de IA mais sofisticados e mais bem integrados ao currículo. Os estudantes das redes públicas, especialmente nas periferias urbanas e nas regiões interioranas, enfrentarão a chegada fragmentada e desigual das mesmas tecnologias. Conforme sintetiza o levantamento de Paixão

et al. (2025, *apud* Moura; Carvalho, 2025), a disparidade no acesso a recursos tecnológicos, a insuficiência de formação docente e a infraestrutura desigual entre instituições públicas e privadas constituem barreiras que tornam a implementação da IA um fenômeno socialmente assimétrico. A ausência dessas condições não impede a chegada da tecnologia às escolas; ela apenas assegura que essa chegada aprofunde as desigualdades já existentes.

Paixão *et al.* (2025, *apud* Moura; Carvalho, 2025) apontam, ainda, que escolas com maior infraestrutura conseguem utilizar melhor as ferramentas de IA, enquanto outras permanecem com métodos tradicionais, aprofundando o fosso entre diferentes redes de ensino e contextos sociais. A questão da privacidade tampouco se apresenta de forma igualitária. Como registram professores ouvidos em pesquisa recente, há desconfiança crescente quanto ao modo como as plataformas armazenam e utilizam as informações dos estudantes, e a falta de transparência sobre esses processos compromete a confiança nas ferramentas e expõe os mais vulneráveis a riscos que os mais protegidos conseguem, ao menos em parte, evitar (Moura; Carvalho, 2025, p. 11).

4.4 O empobrecimento da experiência educativa

Além dos riscos que afetam diretamente grupos específicos, há um risco de alcance mais geral que consiste no empobrecimento da experiência educativa quando a eficiência técnica se torna o valor dominante na organização do ensino. A escola, como instituição, sempre foi um lugar de tensão produtiva. Essa condição manifesta-se entre o que o estudante já sabe e o que ainda não sabe, entre o conforto do familiar e o desconforto do novo, entre a velocidade do desejo e a lentidão necessária à compreensão profunda. Essas tensões não são disfunções a serem eliminadas, mas representam, em grande medida, o próprio motor do aprendizado.

Goethel e Fonseca (2025, p. 9) observam que a correção automatizada e o feedback instantâneo, embora otimizem o tempo, podem desviar o foco do estudante da compreensão profunda e da reflexão sobre o erro para uma simples busca pela validação e correção superficial. O erro deixa de ser uma oportunidade de reconstrução do raciocínio e passa a ser um dado a ser corrigido pelo sistema. O diálogo, que em uma relação pedagógica genuína produz sentido ao revelar perspectivas inesperadas e ao confrontar interpretações distintas, é substituído por uma interação padronizada orientada para o desempenho mensurável.

A curiosidade, essa disposição que Charlot (2013, p. 75) identifica como condição para que o aprender seja fonte de prazer, é particularmente vulnerável a esses processos. Um sistema

adaptativo que antecipa o próximo passo antes que o estudante o procure, que sugere a resposta antes que a pergunta se forme, reduz o espaço em que a curiosidade pode se desenvolver. Conforme a análise de Goethel e Fonseca (2025), a personalização algorítmica, ao determinar caminhos ideais e adaptar o aprendizado conforme padrões preditivos, pode impedir a busca exploratória e a superação de desafios que dão origem ao sentido pessoal. A experiência de não saber, de buscar e de encontrar, que está no núcleo do que significa aprender, corre o risco de ser substituída pela experiência de receber.

4.5 A crise de autoria e a necessidade de repensar a avaliação

A disseminação das ferramentas de IA generativa entre estudantes colocou em crise um dos pilares tradicionais da avaliação escolar, a saber, a suposição de que o texto, o argumento e a solução apresentados pelo estudante são produto de seu próprio trabalho intelectual. Essa suposição, que sustentava a validade das provas dissertativas, das redações e dos trabalhos acadêmicos, perdeu sua solidez a partir do momento em que qualquer estudante com acesso à internet pode solicitar a um modelo de linguagem que elabore um texto coerente e aparentemente original sobre qualquer tema.

A questão não é apenas ética, embora o seja também. É, antes de tudo, pedagógica. Williams (2024, *apud* Heggler; Szmoski; Miquelin, 2025) inclui a autoria entre as implicações éticas centrais do uso de plataformas como o ChatGPT na educação, ao lado da privacidade dos dados e do viés algorítmico. Mas as consequências para o processo formativo vão além da desonestidade acadêmica. Goethel e Fonseca (2025, p. 9) demonstram que a geração de rascunhos por IA fragiliza a memória, reduz a percepção de autoria e compromete o engajamento ativo na síntese crítica das informações e na formulação de argumentos próprios. O estudante que entrega um texto gerado por IA não apenas viola uma norma institucional; ele priva a si mesmo do processo que tornaria aquele texto algo seu.

Essa realidade exige uma revisão das formas de avaliar, e não apenas uma intensificação dos mecanismos de detecção. Avaliar, na perspectiva charlotiana, é verificar não apenas se o resultado esperado foi alcançado, mas se o processo que conduz a esse resultado envolveu a mobilização intelectual do sujeito. Isso aponta para modalidades avaliativas que privilegiam o processo sobre o produto tais como portfólios que documentam o percurso de aprendizagem, avaliações orais que exigem elaboração em tempo real, propostas que demandam a articulação com a realidade concreta do estudante, e tarefas cujo contexto específico não pode ser replicado por um modelo genérico de linguagem. A avaliação, nesse sentido, precisa ser

redesenhada não para vencer a tecnologia em seu próprio terreno, o da geração de texto, mas para habitar um terreno que a IA não ocupa que é o da singularidade do sujeito, de sua história, de suas questões e de seu modo próprio de se relacionar com o saber.

Essa reconfiguração avaliativa é também uma resposta ao risco identificado por Goethel e Fonseca (2025) de que a padronização produzida pelos modelos de linguagem comprometa a diversidade e a originalidade do pensamento. Quando os ensaios gerados por IA apresentam baixo desvio estatístico entre si, isso não é apenas um dado técnico; é um sintoma de que a mediação algorítmica tende a aproximar os produtos cognitivos dos estudantes de um denominador comum, apagando as singularidades que a educação deveria, ao contrário, cultivar.

5 CONSIDERAÇÕES FINAIS

Este texto não chegou a um consenso. Isso não é uma falha; é a única conclusão intelectualmente honesta possível diante de um fenômeno que ainda está em curso, cujas contradições são reais e cujas consequências dependem, em grande medida, de escolhas que ainda estão sendo feitas. A inteligência artificial na escola é, ao mesmo tempo, uma possibilidade genuína de ampliar o alcance e a qualidade do ensino e um risco concreto de empobrecer a experiência formativa, aprofundar desigualdades e substituir a relação com o saber por uma relação com métricas de desempenho. Essas duas dimensões coexistem, e reconhecê-las sem hierarquizá-las apressadamente é o ponto de partida de qualquer reflexão que mereça o nome de pedagógica.

Há, no entanto, uma tensão central que nenhuma síntese conciliatória consegue dissolver entre a promessa de personalização e o risco de desumanização. Os sistemas adaptativos prometem ajustar o ensino a cada estudante, reconhecendo sua singularidade cognitiva e respondendo às suas necessidades em tempo real. Mas a personalização algorítmica opera sobre dados, padrões e probabilidades, e a singularidade que ela reconhece é a do perfil de desempenho, não a do sujeito. Charlot (2000, p. 33) ensinou que o sujeito é indissociavelmente humano, social e singular, e que a relação com o saber é sempre também uma relação com o mundo, com os outros e consigo mesmo. Nenhum algoritmo, por mais refinado que seja seu modelo do aprendiz, consegue captar essa dimensão existencial do aprender. O que se personaliza, nos sistemas de IA educacional, é o percurso técnico; o que permanece intocado, e que precisaria ser o centro de qualquer projeto educativo, é a questão do sentido.

Goethel e Fonseca (2025) demonstraram com precisão o paradoxo segundo o qual a IA promete otimizar a aquisição de conhecimento e pode, simultaneamente, esvaziar o sentido pessoal do estudo, elemento fundamental para a formação da consciência. Quando a delegação ao algoritmo substitui o engajamento do sujeito, a atividade de aprender se converte em uma sequência de operações reativas. O estudante produz resultados sem ter construído o processo que os tornaria seus. A eficiência técnica, aqui, opera contra a formação.

A essa tensão interna soma-se uma tensão externa de natureza política. Charlot (2022) expõe a questão com a clareza que a urgência histórica exige, argumentando que, no cenário descrito em "Educação ou Barbárie", hoje é necessário escolher (Charlot, 2022, p. 153). O autor não faz referência direta à inteligência artificial nessa formulação, mas a indagação se aplica com plena pertinência ao cenário atual. A barbárie, para Charlot, não é apenas violência explícita; é toda situação em que se nega a humanidade do outro, em que se trata o ser humano como objeto, dado, perfil, variável. A escola que adota a IA sem reflexão crítica, que deixa os algoritmos decidirem quais estudantes têm potencial e quais são de risco, que substitui o diálogo pedagógico pela eficiência do feedback automático, está, nesse sentido, em risco de barbárie. Não porque a tecnologia seja maléfica, mas porque a relação inculta ou mágica com ela tende a produzir exatamente o que a educação deveria combater, que consiste na redução do sujeito à sua utilidade mensurável.

A questão formulada Charlot (2022) e adaptada ao presente consiste em identificar a serviço de quem opera a IA nas escolas. A resposta não é simples, e é justamente por isso que ela não pode ser evitada. Os sistemas de IA educacional são, em sua maioria, desenvolvidos por empresas privadas com interesses comerciais. Os dados que alimentam esses sistemas são coletados dos estudantes sem que eles ou suas famílias tenham plena consciência das condições de uso. As decisões tomadas ou informadas por algoritmos afetam trajetórias acadêmicas e profissionais. Em um contexto de profunda desigualdade social como o brasileiro, onde a infraestrutura tecnológica das escolas públicas é cronicamente insuficiente, a chegada da IA tende a beneficiar mais os que já estão em situação de vantagem. Sem políticas públicas robustas, sem regulação transparente, sem formação docente crítica e sem participação das comunidades escolares nas decisões sobre quais sistemas adotar, tal tecnologia não democratiza o ensino. Pelo contrário, ela o estratifica.

O que resta, então, como horizonte? Não a recusa da tecnologia, que seria tão ingênua quanto seu acolhimento acrítico. O que resta é a reafirmação da educação como processo de humanização, no sentido exato que Charlot (2022, p. 150) lhe confere que é o de ocupar o mundo com humanidade. Para isso, a IA pode ser, em certas condições, um recurso. Mas as

condições importam tanto quanto o recurso. Uma tecnologia utilizada por um professor que compreende seus mecanismos, que a emprega com intencionalidade pedagógica, que mantém o diálogo como centro do processo e que forma os estudantes para questionar as respostas que ela produz é uma IA a serviço da educação. Um sistema algorítmico implantado sem reflexão, que substitui o professor, que classifica estudantes segundo critérios opacos, quiçá preconceituosos, e que transforma a escola em plataforma de coleta de dados é uma ferramenta a serviço de outros interesses.

E é precisamente essa distinção que exige que as perguntas permaneçam abertas. Que reestruturação pedagógica é necessária para que a tecnologia sirva à emancipação do sujeito e não apenas ao controle métrico? Como a escola pode incorporar a IA sem renunciar ao erro produtivo, do diálogo genuíno e da curiosidade como disposição formativa? Como avaliar o que um estudante aprendeu, em uma era em que a IA pode gerar qualquer produto cognitivo solicitável, sem reduzir a avaliação à detecção do engano? Como garantir que a personalização algorítmica não se torne uma forma sofisticada de predeterminar trajetórias e confirmar desigualdades? Como formar professores que não sejam apenas usuários de ferramentas, mas críticos das condições em que essas ferramentas operam?

Nenhuma dessas perguntas tem resposta técnica. Todas elas são perguntas políticas, éticas e pedagógicas. E é justamente por isso que elas precisam ser respondidas nas escolas, nas salas de professores, nos conselhos de gestão, nas políticas públicas e nos espaços de formação docente. A IA está chegando às escolas. A questão não é se ela chegará, mas que escola a receberá, e que projeto de formação humana orientará seu uso. Essa escolha, como Charlot (2022) nos lembra, é nossa.

Referências

BAKER, Ryan; HAWN, Aaron. Algorithmic bias in education. **International Journal of Artificial Intelligence in Education**, [s.l.], v. 32, n. 4, p. 1052-1092, 2022. Disponível em: https://link.springer.com/article/10.1007/s40593-021-00285-9. Acessoe em: 9 abr. 2026.

BIRD, Kelli A.; CASTLEMAN, Benjamin L.; SONG, Yifeng. Are algorithms biased in education? Exploring racial bias in predicting community college student success. **Journal of Policy Analysis and Management**, [s. l.], v. 43, n. 1, p. 1-28, 2024. Disponível em: https://doi.org/10.1002/pam.22569. Acesso em: 10 abr. 2026.

CHARLOT, Bernard. **Da relação com o saber**: elementos para uma teoria. Tradução de Bruno Magne. Porto Alegre: Artes Médicas Sul, 2000.

CHARLOT, Bernard. **Relação com o saber, formação dos professores e globalização**: questões para a educação hoje. Porto Alegre: Artes Médicas Sul, 2005.

CHARLOT, Bernard. **Da relação com o saber às práticas educativas**. São Paulo: Cortez, 2013.

CHARLOT, Bernard. **Educação ou barbárie**: uma escolha para a sociedade contemporânea. Tradução de Sandra Pina. São Paulo: Cortez, 2022.

DAKAKNI, Dina; SAFA, Nour. Artificial intelligence in the L2 classroom: implications and challenges on ethics and equity in higher education: A 21st century Pandora's Box. **Computers and Education: Artificial Intelligence**, [s. l.], v. 5, p. 1-10, 2023. Disponível em: https://doi.org/10.1016/j.caeai.2023.100179. Acesso em: 9 abr. 2026.

DAVIS, Jenny L.; WILLIAMS, Andrew; YANG, Michael W. Algorithmic reparation. **Big Data & Society**, [s. l.], v. 8, n. 2, p. 1-12, 2021. Disponível em: https://doi.org/10.1177/20539517211044808. Acesso em: 9 abr. 2026.

GASKINS, Nettrice. Interrogating Algorithmic bias: from speculative fiction to liberatory design. **TechTrends**, [s. l.], v. 67, n. 3, p. 417-425, 2023. Disponível em: https://doi.org/10.1007/s11528-022-00783-0. Acesso em: 9 abr. 2026.

GOETHEL, Elisiane Spencer Quevedo; FONSECA, Débora Cristina. A inteligência artificial na educação e a formação da consciência: uma análise leontieviana sobre o risco do sentido pessoal delegado ao algoritmo. **Obutchénie: Revista de Didática e Psicologia Pedagógica**, Uberlândia, v. 9, e2025-20, 2025. Disponível em: http://doi.org/10.14393/OBv9.e2025-20. Acesso em: 9 abr. 2026.

HEGGLER, João Marcos; SZMOSKI, Romeu Miqueias; MIQUELIN, Awdry Feisser. As dualidades entre o uso da inteligência artificial na educação e os riscos de vieses algorítmicos. **Educação & Sociedade**, Campinas, v. 46, e289323, 2025. Disponível em: https://doi.org/10.1590/ES.289323. Acesso em: 9 abr. 2026.

ITO-JAEGER, Stefanie et al. TrustScapes: a visualisation tool to capture stakeholders' concerns and recommendations about data protection, algorithmic bias, and online safety. **International Journal of Qualitative Methods**, [s. l.], v. 22, p. 1-10, 2023. Disponível em: https://doi.org/10.1177/16094069231186965. Acesso em: 9 abr. 2026.

JOHNSON, Gabbrielle M. Algorithmic bias: on the implicit biases of social technology. **Synthese**, [s. l.], v. 198, n. 10, p. 9941-9961, 2020. Disponível em: https://doi.org/10.1007/s11229-020-02696-y. Acesso em: 9 abr. 2026.

LEONTIEV, Alexei Nikolaevich. **Actividad, conciencia, personalidad**. Buenos Aires: Ciencias del Hombre, 1978.

MOURA, Maria José de Freitas de Oliveira; CARVALHO, Francisco Daniel Sousa. Cultura digital e inteligência artificial na prática docente. **Revista Caderno Pedagógico**, Curitiba, v. 22, n. 9, p. 1-16, 2025. Disponível em: https://doi.org/10.54033/cadpedv22n9-001. Acesso em: 9 abr. 2026.

PAIXÃO, Rodrigo et al. Inteligência artificial e educação: desafios e perspectivas para a educação básica brasileira. **Lumen et Virtus**, São José dos Pinhais, v. 16, n. 55, p. 1-15, 2025. Disponível em: https://doi.org/10.35699/lumenetvirtus.v16i55. Acesso em: 9 abr. 2026.

TANG, Li; SU, Man. Ethical implications and principles of using artificial intelligence models in the classroom: A systematic literature review. **Universidad: Revista de la Unión de Universidades**, [s. l.], v. 8, n. 5, p. 25-36, 2024. Disponível em: https://doi.org/10.9781/ijimai.2024.02.010. Acesso em: 9 abr. 2026.

TRINDADE, Lenilsa dos Santos. Inteligência artificial no ensino: contribuições, desafios e implicações para o papel docente. **Revista Caderno Pedagógico**, Curitiba, v. 22, n. 13, p. 1-15, 2025. Disponível em: https://doi.org/10.54033/cadpedv22n13-001. Acesso em: 9 abr. 2026.

VAN DIJCK, José. **A sociedade da plataforma**: conectividade em um mundo datificado. Tradução: Bruno Magne. Rio de Janeiro: Editora FGV, 2017.

WILLIAMS, Tiffany R. The ethical implications of using generative chatbots in higher education. **Frontiers in Education**, [s. l.], v. 9, p. 1-7, 2024. Disponível em: https://doi.org/10.3389/feduc.2023.1331607. Acesso em: 9 abr. 2026.

YODER-HIMES, David R. et al. Racial, skin tone, and sex disparities in automated proctoring software. **Frontiers in Education**, [s. l.], v. 7, p. 1-16, 2022. Disponível em: https://doi.org/10.3389/feduc.2022.881449. Acesso em: 9 abr. 2026.

ZAJKO, Michael. Conservative AI and social inequality: conceptualizing alternatives to bias through social theory. **AI & Society**, [s. l.], v. 36, n. 3, p. 1047-1056, 2021. Disponível em: https://doi.org/10.1007/s00146-021-01153-9. Acesso em: 9 abr. 2026.

PARTE 3

Educação Inclusiva na Perspectiva da Neurodiversidade

CAPÍTULO 6

FUNDAMENTOS DA NEURODIVERSIDADE E O PARADIGMA DA INCLUSÃO

FOUNDATIONS OF NEURODIVERSITY AND THE INCLUSION PARADIGM

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1245



Resumo: O presente capítulo analisa a evolução do conceito de educação inclusiva no Brasil, contrastando a ênfase histórica na acessibilidade física com a necessidade emergente de uma acessibilidade pedagógica voltada à diversidade cognitiva. Investiga-se a origem do conceito de neurodiversidade, cunhado por Judy Singer, destacando sua natureza política e sua distinção fundamental em relação à neurodivergência. A partir das contribuições de Georges Canguilhem sobre o normal e o patológico, o texto discute como o modelo social da deficiência pode ser aplicado ao campo da cognição para remover barreiras escolares. O mapeamento das condições contempladas — TDAH, Transtorno Afetivo Bipolar, Transtorno Opositor Desafiador e transtornos alimentares — fundamenta a defesa de uma abordagem baseada em forças, que reconheça a variação neurológica como uma vantagem evolutiva e um direito humano.

Palavras-chave: Neurodiversidade. Educação Inclusiva. Acessibilidade Pedagógica. Neurodivergência. Modelo Social da Deficiência.

Abstract: This chapter analyzes the evolution of the concept of inclusive education in Brazil, contrasting the historical emphasis on physical accessibility with the emerging need for pedagogical accessibility focused on cognitive diversity. It investigates the origin of the neurodiversity concept, coined by Judy Singer, highlighting its political nature and its fundamental distinction from neurodivergence. Based on Georges Canguilhem's contributions regarding the normal and the pathological, the text discusses how the social model of disability can be applied to the field of cognition to remove school barriers. The mapping of the conditions covered — ADHD, Bipolar Disorder, Oppositional Defiant Disorder, and eating disorders — supports the advocacy for a strength-based approach that recognizes neurological variation as an evolutionary advantage and a human right.

Keywords: Neurodiversity. Inclusive Education. Pedagogical Accessibility. Neurodivergence. Social Model of Disability.

1 INTRODUÇÃO

Que imagem evoca a palavra *inclusão* no imaginário do educador brasileiro? Para uma grande parte dos profissionais que percorrem os corredores de nossas escolas, a resposta ainda se materializa em objetos concretos, palpáveis e mensuráveis tais como a rampa de acesso, o piso tátil, a sinalização em braile ou a vaga reservada no estacionamento. Há algo reconfortante nessa concretude. Ela permite que o cumprimento legal de uma obrigação seja fotografado, inspecionado por uma comissão de avaliação e registrado em formulários oficiais. Como se diz no popular: "Fica bem na foto!" A inclusão, nessa acepção predominante, converte-se em obra de engenharia civil.

Essa redução, todavia, não é gratuita nem fortuita. Ela resulta de um longo percurso histórico em que o movimento pela inclusão precisou lutar, palmo a palmo, contra a segregação institucionalizada de pessoas com deficiências físicas e sensoriais. A batalha pela rampa foi real, foi necessária e ainda está longe de ser vencida por completo no território brasileiro.

Entretanto, ao consolidar-se como símbolo máximo da acessibilidade, a rampa produziu um efeito colateral silencioso, já que tornou invisível tudo aquilo que não pode ser resolvido com concreto e argamassa.

É precisamente nessa invisibilidade que reside a questão central deste capítulo. Como compreender a inclusão quando a barreira não está no degrau, mas na estrutura da aula expositiva? Quando o impedimento não é o corredor estreito, mas o tempo fixo da prova? Quando o obstáculo não está na porta da escola, mas na rigidez de um currículo pensado para um único tipo de intelecto? Para responder a essas perguntas, é necessário percorrer o trajeto histórico do conceito de educação inclusiva no Brasil, examinar seus limites na prática escolar e preparar o terreno para a entrada de um conceito que tem reorganizado o campo, qual seja a neurodiversidade.

2 O conceito de educação inclusiva e seus limites na prática escolar

A educação inclusiva não nasceu como uma política. Nasceu como uma insurgência. Antes que qualquer decreto ou lei a formalizasse, ela existia como demanda de famílias que recusavam ver seus filhos confinados em instituições especializadas, separados da vida ordinária de seus bairros, de suas comunidades, de seus coetâneos. Nesse sentido, compreender a inclusão exige que nos aproximemos de sua história não apenas como sequência de marcos jurídicos, mas como processo social carregado de tensões, avanços, recuos e disputas de sentido.

O ponto de inflexão internacional mais amplamente citado é a Declaração de Salamanca, adotada pela Conferência Mundial sobre Necessidades Educacionais Especiais, organizada pela Unesco em 1994, na Espanha. Esse documento estabeleceu o princípio de que toda criança tem o direito fundamental à educação em escolas regulares, independentemente de suas condições físicas, intelectuais, sociais, emocionais ou linguísticas (Unesco, 1994). A Declaração de Salamanca representou, naquele momento, uma virada paradigmática, pois deslocou o foco da adaptação do indivíduo para a adaptação da escola. O problema deixava de ser o aluno "diferente" e passava a ser a instituição que não sabia lidar com a diferença.

No Brasil, esse princípio encontrou eco gradual e desigual. A Constituição Federal de 1988 já assegurava o atendimento educacional especializado às pessoas com deficiência, preferencialmente na rede regular de ensino. A Lei de Diretrizes e Bases da Educação Nacional, Lei nº 9394/1996, reafirmou esse direito e estabeleceu a obrigação dos sistemas de ensino de assegurar currículos, métodos e recursos específicos para atender às necessidades dos

estudantes. Contudo, a distância entre o texto legal e a realidade escolar revelou-se, desde o início, considerável.

A Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva, lançada em 2008, procurou conferir maior sistematicidade ao processo, definindo o público-alvo da educação especial e orientando os sistemas educacionais a estruturar o Atendimento Educacional Especializado como suporte complementar à escolarização regular. Esse marco foi significativo porque explicitou, em termos de política pública, que a educação especial não substitui a educação regular, mas a complementa. A lógica da substituição, que por décadas havia justificado a existência das classes e escolas especiais como destino do aluno com deficiência, começava a ser formalmente contestada.

A Lei Brasileira de Inclusão da Pessoa com Deficiência, a Lei n.º 13.146, de 6 de julho de 2015, aprofundou esse caminho ao incorporar a perspectiva dos direitos humanos e o modelo social da deficiência. Conforme estabelece o artigo 27 desse diploma legal, "a educação constitui direito da pessoa com deficiência, assegurados sistema educacional inclusivo em todos os níveis e aprendizado ao longo de toda a vida" (Brasil, 2015). A definição de deficiência adotada pela Lei é especialmente relevante para os propósitos deste texto, pois compreende o impedimento não como atributo isolado do indivíduo, mas como resultado da interação entre as características da pessoa e as barreiras impostas pelo ambiente social. Onde não há barreira, há participação. Onde a escola remove a barreira, a deficiência perde sua força incapacitante.

Essa formulação jurídica é, em si mesma, uma mudança de paradigma. Ela transfere parte da responsabilidade pela exclusão do corpo do aluno para a estrutura da instituição. No entanto, a adoção formal desse princípio não se traduziu, de modo automático, em transformações pedagógicas nas salas de aula brasileiras. O que se observou, ao longo das décadas seguintes à Declaração de Salamanca e dos marcos legislativos nacionais, foi um processo que Mantoan (2011) denomina de inclusão marginal caracterizado pela presença física do aluno com deficiência na escola regular sem que essa escola tivesse se reorganizado para acolhê-lo de fato.

> Na área da Educação Inclusiva, historicamente, as políticas públicas vieram como marcos fundamentais em busca da garantia de direitos como acesso, permanência e qualidade no contexto educacional para as pessoas com deficiência, desde o provimento de serviços especializados até a proposta e efetividade de práticas educacionais e educativas na perspectiva inclusiva. A sala de aula da educação básica se configurou como um vasto campo de práticas e estudos avaliativos que mapearam o desenvolvimento do conhecimento proposto pela área da educação especial na escola comum não só acolhendo, mas, principalmente, assumindo o seu caráter educativo. (Zanata, 2011, p. 6)

Esse diagnóstico aponta para um dos nós centrais da inclusão na prática escolar brasileira que diz respeito à distinção entre acessibilidade física e acessibilidade pedagógica. A primeira diz sobre as condições materiais e arquitetônicas que permitem ao estudante com deficiência ocupar o espaço escolar. A segunda trata das condições didáticas, curriculares e relacionais que permitem a esse estudante aprender e desenvolver-se nesse espaço. No percurso histórico da inclusão no Brasil, o investimento público e a atenção normativa concentraram-se, de maneira desproporcional, na acessibilidade física, enquanto a acessibilidade pedagógica permaneceu como uma zona de tensão mal resolvida.

Não se trata de minimizar a importância das rampas, do braille ou da Língua Brasileira de Sinais. Essas conquistas representam lutas históricas de movimentos sociais organizados e garantem condições elementares de existência para estudantes com deficiências físicas e sensoriais. O problema surge quando a acessibilidade física é tratada como o horizonte da inclusão, e não como seu ponto de partida. Quando a escola ergue a rampa, muitas vezes até de forma improvisada e matematicamente inconsistente, e considera sua obrigação cumprida, ela revela uma concepção estreita de deficiência e uma incapacidade de perceber as barreiras que não se medem em centímetros.

A resistência à adaptação curricular é, talvez, a manifestação mais visível dessa limitação. Ao longo de décadas de implementação das políticas de inclusão, gestores escolares, docentes e sistemas de ensino demonstraram, com frequência, uma disposição muito maior para reformar o espaço físico do que para reformar as práticas pedagógicas. Modificar o currículo, flexibilizar o tempo das avaliações, diversificar as estratégias de ensino, oferecer múltiplos modos de representação do conhecimento são ações que exigem uma revisão profunda da identidade profissional do professor e do projeto político-pedagógico da escola. Elas ameaçam certezas consolidadas sobre o que é ensinar e o que é aprender.

A Lei de Diretrizes e Bases já determinava, em seu artigo 59, que os sistemas de ensino deveriam assegurar "currículos, métodos, recursos e organização específicos para atender às suas necessidades" (Brasil, 1996). Essa determinação, reiterada pela Lei Brasileira de Inclusão e pela Política Nacional de Educação Especial, encontrou na prática escolar uma resistência que não é apenas cultural ou atitudinal. Ela é também estrutural, pois professores sem formação específica, turmas superlotadas, ausência de materiais adequados e inexistência de apoio técnico-pedagógico sistemático compõem um quadro que inviabiliza, com frequência, mesmo as boas intenções.

Observa-se, portanto, que a trajetória da inclusão no Brasil pode ser lida como a história de um conceito que se expandiu juridicamente, mas cuja expansão encontrou enormes

dificuldades de materialização no interior das instituições escolares. A cada novo marco legal, o horizonte da inclusão se alargou. A cada novo alargamento, as demandas sobre a escola tornaram-se mais complexas e mais exigentes de transformação real. E foi exatamente nesse contexto de tensão entre a letra da lei e a realidade da sala de aula que um novo conceito começou a ganhar relevância, trazendo consigo perguntas que o paradigma dominante não estava equipado para responder.

Como compreender, nesse cenário, as necessidades de estudantes cujas barreiras não são físicas nem sensoriais? Que lugar ocupa, no arcabouço da educação inclusiva, o adolescente cujo cérebro processa a informação de maneira radicalmente diferente da maioria, mas sem apresentar nenhuma marca visível que o identifique, de forma rotulada, como "deficiente"? Essas perguntas introduzem, de maneira natural, o tema central que orientará os próximos desenvolvimentos deste texto que são o conceito de neurodiversidade e suas implicações para a prática educacional.

A neurodiversidade não chegou à educação como uma política de governo. Chegou como chegam, geralmente, os conceitos que transformam campos consolidados, pela força de experiências que não cabem nas categorias disponíveis. Para compreender o que esse conceito representa e por que ele interpela tão diretamente o paradigma da inclusão, é preciso examinar sua origem, sua trajetória e os debates que ele suscita. É esse o percurso que este capítulo se propõe a trilhar nas seções seguintes.

3 A NEURODIVERSIDADE COMO CATEGORIA: ORIGEM, DEFINIÇÃO E ALCANCE

As perguntas que encerraram a seção anterior sobre como a escola deve responder às barreiras invisíveis não encontraram, durante décadas, uma linguagem conceitual adequada nos debates sobre educação inclusiva. Isso ocorreu porque o paradigma predominante confundia a inclusão com a educação especial, focando em uma acessibilidade concreta e mensurável. Esse modelo era sólido para descrever a acessibilidade física e relativamente desenvolvido para tratar das deficiências sensoriais e intelectuais, mas permanecia em silêncio diante de mentes que funcionavam de maneira radicalmente diferente da maioria sem exibir nenhum sinal exterior reconhecível.

Essa lacuna revela uma contradição interna no próprio movimento inclusivo. Ao priorizar a "batalha pela rampa" ou a garantia do "intérprete de libras", muitos defensores da educação especial acabaram por restringir o direito à escola ao fornecimento de suportes

externos e recursos técnicos visíveis. Ao focar naquilo que pode ser fotografado, inspecionado ou contratado, o movimento acabou por ignorar as demandas da neurodiversidade, tornando invisível tudo o que não se resolve com reformas estruturais ou auxílios técnicos protocolares.

É nesse cenário de invisibilidade forçada que o conceito de neurodiversidade germina e "bate à porta", exigindo o reconhecimento de que a educação inclusiva é um horizonte muito mais amplo que a educação especial. Ele interpela a escola a olhar para o que não se mede com trena nem se resolve apenas com a presença de um profissional de apoio, afirmando que essas mentes não apenas existem, mas possuem direitos pedagógicos que o paradigma tradicional, focado no visível e no palpável, falhou em assegurar.

Isso posto, a palavra neurodiversidade foi cunhada pela socióloga australiana Judy Singer em sua tese de graduação em honras, apresentada à Universidade de Tecnologia de Sydney em 1998, sob o título *"Odd People In: The Birth of Community Amongst People on the Autistic Spectrum"*. Singer, que se identifica como portadora da síndrome de Asperger, e cuja mãe e filha também se situavam no espectro autista, partia de uma posição simultaneamente biográfica e intelectual, vivenciava, ela mesma e de múltiplos lugares de fala, a questão que investigava. A motivação inicial da tese não era, portanto, abstrata. Era a busca por uma categoria que fizesse sentido para vidas que os sistemas classificatórios disponíveis descreviam apenas pelo que lhes faltava, nunca pelo que possuíam.

Conforme relata Ortega (2008), o ponto de partida de Singer era a percepção de que uma nova categoria de diferença humana havia emergido na consciência coletiva. Para a socióloga, as variações neurológicas representavam um fenômeno que as divisões tradicionais de etnia, classe, gênero e mesmo de deficiência não conseguiam mais explicar de forma isolada. O que Singer propunha era o estabelecimento de um novo eixo de análise fundamentado nas diferenças entre os tipos de mente, recusando a redução dessas vidas a meras listas de sintomas ou falhas. A motivação da tese não era uma abstração teórica, mas a busca por uma categoria que conferisse sentido e dignidade a existências que os sistemas de classificação vigentes descreviam apenas pelo que lhes faltava.

O aparecimento do conceito não foi casual. Singer identificou, em seu texto, um conjunto de condições históricas que tornaram possível a emergência desse novo paradigma. A influência do feminismo forneceu às mães de crianças autistas a autoconfiança necessária para questionar o modelo psicanalítico dominante, que as responsabilizava pelo transtorno dos filhos. A ascensão de grupos de apoio aos pacientes e a subsequente redução da autoridade médica, viabilizadas em grande medida pela disseminação da internet, permitiram que indivíduos até então isolados se organizassem em comunidades, trocassem informações e

desenvolvessem uma identidade coletiva (Ortega, 2008). O que havia sido vivido como estigma individual começou a ser ressignificado como pertencimento a uma minoria neurológica.

A internet, nesse processo, cumpriu um papel que Singer descreveu como o de "dispositivo protético que une autistas isolados e socialmente desabilitados em um organismo social coletivo capaz de ter uma voz pública" (Singer, 2017, p. 14). Essa imagem é iluminadora, pois enquanto os ambientes convencionais acentuavam as dificuldades de comunicação presencial, o meio digital permitia que pessoas cujo estilo cognitivo as isolava encontrassem, enfim, uma forma de expressão coletiva. Houve, portanto, uma oposição funcional, a tecnologia atuava como um nivelador de barreiras relacionais, antecipando, no plano comunicacional, o que a educação inclusiva ainda lutava para fazer no plano pedagógico.

Ao mesmo tempo em que Singer desenvolvia sua tese na Austrália, o jornalista norte-americano Blume publicava, em 1998, um artigo no *The Atlantic* em que utilizava o mesmo termo, propondo a ideia de pluralismo neurológico (Blume, 1998). Ambos chegaram ao conceito por caminhos distintos e sem conhecimento mútuo. A convergência, por si mesma, revela algo sobre o momento histórico, qual seja, a percepção de que havia uma demanda epistêmica latente por uma categoria que nomeasse a variação neurológica como fenômeno de interesse coletivo, e não apenas como problema clínico individual. Conforme observa Rosa e Bucco (2025), o termo passou a ser adotado como conceito central para compreender a variabilidade natural da cognição humana, com sua origem diretamente ligada às intensas batalhas da comunidade autista por direitos e contra a discriminação.

Essa gênese já anuncia algo fundamental sobre o estatuto epistemológico do conceito de neurodiversidade que não nasceu nos laboratórios de neurociência. Antes, emergiu nos fóruns de internet, nas associações de famílias, nas narrativas autobiográficas de pessoas que se recusavam a ser definidas apenas por aquilo que não conseguiam fazer segundo os padrões dominantes. Trata-se, portanto, de uma categoria que tem, em sua constituição, uma dimensão política inseparável da dimensão descritiva. Ignorar essa dupla natureza é empobrecer o conceito.

3.1 Neurodiversidade e neurodivergência: uma distinção necessária

O amadurecimento do campo produziu, ao longo das décadas seguintes, uma distinção conceitual de grande importância prática que consiste na diferença entre neurodiversidade e neurodivergência. Essa distinção não é meramente terminológica. Ela implica uma diferença

de escala e de referência que tem consequências diretas para a forma como os conceitos podem ser operacionalizados na pesquisa educacional e na formulação de políticas.

Neurodiversidade é uma propriedade da espécie, não dos indivíduos. Assim como se fala em biodiversidade para descrever a variedade de formas de vida em um ecossistema, fala-se em neurodiversidade para descrever a variedade de tipos de funcionamento neurológico que coexistem na espécie humana. Nesse sentido, todos os seres humanos integram a neurodiversidade, pois o cérebro que processa informação de maneira convencional e o cérebro que processa de maneira incomum são igualmente expressões da variação neurológica da espécie. Afirmar que "todos somos neurodiversos" não é um equívoco, já que é, precisamente, o ponto central da proposta de Singer. A neurodiversidade é a condição de uma espécie que evoluiu produzindo cérebros diferentes, não porque algo deu errado, mas porque a variação é, ela própria, uma vantagem evolutiva.

Neurodivergência, por sua vez, é uma condição individual. Uma pessoa é neurodivergente quando seu funcionamento neurológico diverge de maneira significativa das normas estatísticas e sociais prevalentes em seu contexto. O termo foi cunhado por Kassiane Asasumasu, ativista autista norte-americana, para designar qualquer forma de funcionamento neurológico que difira do padrão neurotípico, seja por condições como o Transtorno do Espectro Autista, o Transtorno de Déficit de Atenção e Hiperatividade, a dislexia, a discalculia, a síndrome de Tourette, entre outras. Uma pessoa neurodivergente não tem menos neurodiversidade que uma pessoa neurotípica; ela tem um tipo específico de funcionamento neurológico que difere da maioria.

Essa distinção esclarece um equívoco frequente nos debates educacionais que é o de tratar neurodiversidade como sinônimo de deficiência ou de dificuldade. A neurodiversidade não é uma dificuldade. É um fato biológico sobre a espécie. O que pode gerar dificuldades é a interação entre um funcionamento neurológico divergente e um ambiente escolar construído com base no pressuposto de que todos os cérebros aprendem da mesma maneira. A dificuldade, nesse caso, não reside no cérebro do estudante; reside no desajuste entre a arquitetura do processo educativo e a realidade neurológica desse estudante. Trata-se, exatamente, do deslocamento que o modelo social da deficiência operou em relação ao modelo biomédico, aplicado agora ao campo da cognição.

Essa convergência entre o paradigma da neurodiversidade e o modelo social da deficiência não é fortuita. Como apontam Rosa e Bucco (2025), a neurodiversidade está intrinsecamente ligada aos desdobramentos do modelo social de deficiência, reivindicado em contraposição ao modelo médico predominante à época de sua formulação. No modelo

biomédico, a origem da limitação está no corpo ou na mente do indivíduo, e a solução é terapêutica, baseada em corrigir, compensar, reabilitar. No modelo social, a origem da limitação está nas barreiras construídas pela sociedade, e a solução é política e pedagógica, qual seja, remover as barreiras, redesenhar os ambientes, transformar as práticas.

Para a educação, essa mudança de perspectiva tem implicações profundas. Se o problema está no estudante, a tarefa da escola é normalizá-lo, aproximá-lo do padrão esperado, suprindo suas lacunas. Se o problema está na escola, a tarefa é reorganizar seus processos de tal forma que diferentes tipos de mente possam neles participar e progredir. A segunda perspectiva é, evidentemente, mais exigente e mais perturbadora das práticas consolidadas. Ela questiona não apenas o que se ensina, mas como se ensina, como se avalia, como se organiza o tempo, o espaço e as relações no interior da instituição escolar.

3.2 O alcance do conceito: da política de identidade à categoria analítica

Desde sua formulação original, o conceito de neurodiversidade acumulou usos, interpretações e disputas que importa examinar com atenção crítica, porque deles depende a consistência de seu emprego como ferramenta analítica no campo educacional.

Em sua dimensão política, a neurodiversidade funciona como uma categoria de identidade coletiva, análoga às que sustentaram outros movimentos de direitos civis. Os indivíduos autistas que se organizaram nos fóruns da internet nos anos 1990 utilizaram o conceito para reivindicar, como lembra Ortega (2008), não a cura de sua condição, mas o reconhecimento de sua diferença como legítima e merecedora de direitos. O Dia do Orgulho Autista, celebrado em 18 de junho desde 2005 com o lema "Aceitação, não cura", expressa com nitidez essa dimensão de recusa à medicalização e afirmação do direito de existir como se é, sem que a existência seja condicionada à adequação a um padrão neurológico hegemônico.

Essa politização do conceito tem, contudo, seus riscos teóricos, que a própria Singer reconheceu em trabalhos subsequentes. O perigo, identificado por Ortega (2008) com base na crítica de Agnes Heller aos movimentos biopolíticos, é o da homogeneização interna. Ao se construir uma identidade coletiva fundada em predicados neurológicos, o movimento pode reproduzir, dentro de si, a lógica de exclusão que combate externamente, apagando a singularidade de cada pessoa em nome de uma identidade de grupo. Uma pessoa autista com necessidades de suporte intensivo vive realidades muito distintas de uma pessoa autista de alto funcionamento e equacioná-las sob uma mesma categoria política pode obscurecer diferenças cruciais para a elaboração de respostas educativas e sociais adequadas.

Em sua dimensão analítica, por sua vez, a neurodiversidade opera como conceito que convida a uma reclassificação epistemológica das diferenças neurológicas. Aqui, a contribuição de Georges Canguilhem torna-se indispensável. Em sua obra "O normal e o patológico", Canguilhem (2009) desenvolve uma distinção conceitual de grande alcance discutindo a diferença entre anomalia e anormalidade. A anomalia é um fato descritivo; designa um desvio da média estatística, uma variação individual que impede dois seres de se substituírem completamente um ao outro. A anormalidade, por sua vez, é um julgamento de valor que implica uma condição que restringe a vida, gera sofrimento e reduz a capacidade do ser vivo de instituir novas normas para si mesmo.

> A anomalia é a consequência de variação individual que impede dois seres de poderem se substituir um ao outro de modo completo. Ilustra, na ordem biológica, o princípio leibnitziano dos indiscerníveis. No entanto, diversidade não é doença. O anormal não é o patológico. Patológico implica *páthos*, sentimento direto e concreto de sofrimento e de impotência, sentimento de vida contrariada. (Canguilhem, 2009, p. 44).

A importância dessa distinção para o debate sobre neurodiversidade é imediata. O que o modelo biomédico tem feito, historicamente, é tratar a anomalia neurológica como se fosse necessariamente patologia, convertendo o fato descritivo do desvio estatístico em julgamento negativo sobre o valor de uma vida. Canguilhem (2009) nos lembra, porém, que o patológico não é a ausência de norma biológica, mas uma norma diferente, comparativamente repelida pelo ambiente. O cérebro que não aprende pela exposição prolongada a textos escritos não é, em si mesmo, um cérebro doente. É um cérebro cujas normas vitais são contrárias às do ambiente escolar construído para uma maioria de outro tipo. Remova-se a barreira ambiental, reorganize-se o processo de ensino, e aquele mesmo cérebro pode revelar capacidades que o ambiente anterior suprimia.

Essa articulação entre Canguilhem (2009) e Singer (2017) permite compreender a neurodiversidade não apenas como uma reinvindicação política, mas como uma consequência lógica de uma filosofia da biologia que recusa a naturalização da norma estatística como norma de valor. O neurodivergente não é, por definição, inferior ao neurotípico. É, apenas, diferente. E essa diferença pode ser fonte de dificuldade ou de potência, dependendo, em larga medida, de como o ambiente escolar responde a ela.

Thomas Armstrong (2011), em *"The Power of Neurodiversity"*, aprofunda essa perspectiva ao propor que as condições neurológicas frequentemente enquadradas como déficits carregam, ao mesmo tempo, capacidades que o modelo do déficit sistematicamente ignora. O que a escola chama de déficit de atenção em um estudante com TDAH pode ser reinterpretado como uma

forma de atenção difusa e distribuída que foi evolutivamente vantajosa para ambientes que exigem vigilância ampla. O que se nomeia como dificuldade de leitura em um estudante com dislexia pode coexistir com habilidades viso-espaciais e de pensamento tridimensional que os testes padronizados de alfabetização simplesmente não capturam. Armstrong não propõe, com isso, uma romantização das dificuldades. Propõe que o olhar pedagógico seja, simultaneamente, sensível às barreiras e atento às potências.

O alcance do conceito de neurodiversidade, portanto, não se limita a nomear um conjunto de condições clínicas. Ele implica uma mudança no modo como a escola concebe sua tarefa fundamental que não é uniformizar mentes, mas criar condições para que intelectos diferentes possam aprender e desenvolver-se cada uma segundo suas próprias normas vitais, para usar a linguagem de Canguilhem (2009). Essa reorientação exige, antes de tudo, que a escola reconheça que a norma pedagógica dominante não é universal; é, como toda norma, uma construção histórica que expressa as necessidades e os valores de determinados grupos em determinados contextos.

Essa consciência crítica da norma pedagógica é o passo que distingue uma inclusão superficial de uma inclusão genuína. A inclusão superficial acomoda o estudante neurodivergente no espaço físico da escola, mas mantém intocados os processos que o excluem. A inclusão genuína parte do reconhecimento de que o problema não está no estudante; está em uma escola que, ao construir seus processos com base em um único tipo de funcionamento neurológico, produz exclusão como efeito estrutural, não como desvio.

Compreender essa distinção é o que permite, finalmente, articular o debate sobre neurodiversidade com o paradigma da inclusão examinado na seção anterior. Os limites da inclusão tradicional, expressos na preponderância da acessibilidade física sobre a acessibilidade pedagógica, decorrem, em última instância, de uma concepção estreita de diferença. Enquanto esta última era concebida como deficiência física ou sensorial, a resposta inclusiva assumiu a forma de adaptação arquitetônica. Quando o campo da diferença se expande para abarcar a variação neurológica, torna-se evidente que nenhuma reforma de alvenaria é suficiente. O que se exige, agora, é uma reforma das práticas, das concepções e das relações que constituem o cotidiano pedagógico.

É para esse horizonte que os textos seguintes se voltam, com o propósito de examinar como a educação brasileira tem respondido a esse desafio e quais instrumentos teóricos e práticos o paradigma da neurodiversidade coloca à disposição dos educadores que buscam, genuinamente, construir escolas em que diferentes tipos de mente possam não apenas frequentar, mas, trabalhar, aprender e florescer.

4 Delimitação e justificativa do recorte: por que neurodiversidade?

Toda escolha conceitual é também uma escolha política. Quando uma pesquisa opta por empregar o termo neurodiversidade em lugar de categorias como "distúrbio de aprendizagem", "transtorno mental" ou "déficit cognitivo", essa preferência não é meramente terminológica. Ela implica um posicionamento sobre onde localizar a origem das dificuldades vividas pelo estudante, sobre quem carrega a responsabilidade pela exclusão e sobre quais transformações são necessárias para que a escola cumpra seu propósito social. Faz-se necessário, portanto, explicitar as razões que justificam esse recorte conceitual e suas implicações para a pesquisa educacional, compreendendo que este conceito não é estático, tendo evoluído de uma proposta seminal para uma perspectiva contemporânea mais ampla.

A primeira justificativa, que fundamenta e atravessa todas as demais, reside no reconhecimento do lugar de fala (Ribeiro, 2017) de quem escreve este texto. Esta pesquisa é conduzida por uma autoria neurodivergente, o que confere ao trabalho um caráter não apenas analítico, mas também reivindicatório. Trata-se da afirmação do direito de existir e de ser aceito na integridade de um funcionamento neurológico que, embora divirja da norma, não deve ser reduzido à patologia. Ao assumir essa perspectiva, o texto expõe uma lacuna profunda nas políticas de inclusão ao evidenciar a invisibilidade dos profissionais neurodivergentes. Se a neurodiversidade é frequentemente negada no corpo discente, ela é quase absoluta e violentamente silenciada no corpo docente. Não há censos escolares que a identifique, nem debates que reconheçam a existência de professores autistas, bipolares ou com outras variações cognitivas. A escola inclusiva que se discute é, paradoxalmente, uma instituição onde profissionais presumidamente típicos tentam ensinar a uma diversidade cada vez mais ampla de alunos, muitas vezes em um descompasso de reciprocidade empática, por não compreenderem as lógicas e os processos de análise de mundo que fogem ao padrão. Incluir a neurodiversidade como recorte é, portanto, trazer à luz esses profissionais invisíveis e reconhecer que a pluralidade, que abrange também o cego, o surdo e o cadeirante, deve incluir também a diversidade da arquitetura cerebral em todos os níveis da hierarquia escolar.

A segunda justificativa é de ordem epistemológica. Os modelos puramente médicos das condições de saúde que, neste trabalho, são compreendidas sob o prisma da neurodiversidade[4] descrevem tais variações como desvios patológicos de uma norma biológica

[4] Neste trabalho, o conceito de neurodiversidade é compreendido de forma ampla, abrangendo: 1. Transtorno do Espectro Autista (TEA); 2. Transtorno de Déficit de Atenção e Hiperatividade (TDAH); 3. Transtornos Específicos de Aprendizagem (Dislexia, Discalculia, Disgrafia); 4. Transtorno do Desenvolvimento de Linguagem (TDL); 5. Transtorno do Desenvolvimento da Coordenação (Dispraxia/TDC); 6. Transtorno do

presumidamente universal, que demandam intervenção corretiva centrada no indivíduo. Historicamente, a proposta seminal de neurodiversidade, cunhada por Judy Singer no final da década de 1990, surgiu para descrever variações inatas no neurodesenvolvimento, como o Autismo e o TDAH, sob a ótica do modelo social da deficiência (Singer, 2017). Entretanto, este trabalho filia-se à proposta mais ampla de neurodiversidade, defendida por autores como Nick Walker e Robert Chapman, que expandem o conceito para além das condições inatas, abrangendo o que se denomina "neurodivergência clínica" (Walker, 2021).

Nesta proposta ampliada, condições como o Transtorno Afetivo Bipolar, a Esquizofrenia e os Transtornos de Ansiedade são integradas ao paradigma, pois, embora possuam dimensões clínicas que exigem cuidado especializado, elas também resultam em funcionamentos cerebrais que divergem da norma estatística ou neurotípica (Chapman, 2023). O modelo médico tem uma utilidade clínica inegável, pois orienta diagnósticos, indica terapêuticas e permite que pessoas em sofrimento real recebam cuidados especializados. No entanto, quando transportado integralmente para o ambiente escolar sem a mediação de uma perspectiva educacional crítica, ele produz efeitos problemáticos. A escola que recebe o estudante apenas sob a ótica do diagnóstico clínico tende a enxergá-lo como objeto de tratamento, e não como sujeito de aprendizagem. Ela espera que o aluno se adapte ao seu funcionamento padrão depois de suficientemente medicado ou terapeuticamente assistido, em vez de reorganizar seus próprios processos para acolher a diversidade cognitiva presente em suas salas.

A terceira justificativa é política. Conforme argumentam Rosa e Bucco (2025), o termo neurodiversidade está intrinsecamente ligado às lutas de movimentos sociais que reivindicaram o reconhecimento das diferenças neurológicas como dimensão legítima da variação humana, e não como falha a ser corrigida. Adotar esse enquadramento na pesquisa educacional significa filiar-se a uma tradição que situa a questão da diferença no campo dos direitos e não apenas no campo da terapêutica.

A transição da proposta seminal para a proposta mais ampla permite que a escola reconheça que o estudante com TDAH não precisa, primariamente, de correção. Pelo contrário, precisa de condições pedagógicas que reconheçam seu modo específico de processar, atender e aprender. Da mesma forma, afirma-se que o estudante com transtorno

Processamento Sensorial (TPS); 7. Síndrome de Tourette; 8. Altas Habilidades/Superdotação; 9. Transtorno Afetivo Bipolar (TAB); 10. Esquizofrenia; 11. Depressão Maior; 12. Transtornos de Ansiedade Severa e Síndrome do Pânico; 13. Transtorno de Personalidade Borderline (TPB); 14. Transtorno Obsessivo Compulsivo (TOC); 15. Transtornos Alimentares (Anorexia e Bulimia); 16. Transtorno Opositor Desafiador (TOD); 17. Sinestesia; 18. Afantasia; 19. Outros Transtornos de Tique; 20. Epilepsia.

bipolar não é um paciente que ocupa um banco escolar enquanto aguarda estabilização clínica. É, no entanto, um sujeito de direito à educação, cujas oscilações de humor e energia têm implicações pedagógicas que a escola precisa aprender a manejar. No cenário brasileiro, reforça-se a necessidade de uma visão interseccional dessa política, onde a neurodivergência é lida junto a marcadores de raça e classe, ampliando ainda mais o escopo da acessibilidade institucional.

Há, nessa opção conceitual, uma consequência direta para o processo de desestigmatização. O estigma, conforme elaborado por Erving Goffman (1988) em sua análise clássica, é um atributo que descredita profundamente o indivíduo, reduzindo-o de uma "pessoa completa e comum" a uma "pessoa estragada e reduzida". No ambiente escolar, os diagnósticos de transtornos mentais ou de aprendizagem frequentemente operam como estigmas, não pela má-fé dos educadores, mas pela lógica do modelo que os orienta. O estudante que recebe um laudo clínico passa a ser percebido, com frequência, por aquilo que não consegue fazer, pelo que lhe falta em relação à norma esperada. Suas dificuldades tornam-se mais visíveis do que suas potências; seu diagnóstico passa a preceder sua subjetividade. Damian Milton (2012) contribui para esta discussão ao propor a teoria do "problema da dupla empatia", sugerindo que a falha de comunicação e compreensão não reside apenas no sujeito neurodivergente, mas na falta de disposição do ambiente e das pessoas neurotípicas em compreender uma lógica de funcionamento distinta da sua.

O conceito de neurodiversidade opera um deslocamento nessa lógica. Ao situar a variação neurológica no campo da diferença humana legítima, e não no campo da patologia, ele cria condições para que o olhar do educador se reorganize. Em lugar de perguntar "o que esse aluno não consegue fazer que os outros conseguem?", a perspectiva da neurodiversidade convida a perguntar "de que maneira específica esse aluno processa, aprende e se relaciona com o conhecimento, e quais condições pedagógicas permitiriam que essa maneira fosse produtiva?". Essa inversão não é trivial. Ela muda o ponto de partida da relação pedagógica, deslocando-o da falta para a potência.

Thomas Armstrong (2011) sistematizou essa perspectiva ao descrever a abordagem baseada em forças, em contraposição ao modelo baseado em déficits. Para Armstrong, a escola tradicional construiu uma cultura da deficiência que obscurece o potencial de estudantes cujos funcionamentos cognitivos diferem da norma. O que o sistema classifica como Déficit de Atenção pode ser reinterpretado, em uma perspectiva evolutiva, como uma forma de atenção distribuída e hipervigilante, vantajosa em contextos que exigem monitoramento amplo do ambiente. O que se nomeia como rigidez cognitiva em contextos de autismo pode ser a outra

face de uma capacidade de foco e sistematização que encontra, em nichos específicos como o campo da computação, da matemática e do desenvolvimento tecnológico, condições para um desempenho excepcional. O que os transtornos de humor trazem de sofrimento e instabilidade pode coexistir com uma sensibilidade afetiva e uma capacidade empática que enriquecem, quando bem acolhidas, o próprio processo de aprendizagem.

Essa afirmação não tem o propósito de romantizar o sofrimento dos estudantes neurodivergentes. O transtorno bipolar gera crises que incapacitam. O TDAH produz dificuldades reais de organização e persistência que se traduzem em desempenho escolar comprometido. Os transtornos alimentares podem colocar em risco a vida. Reconhecer as potências associadas a esses funcionamentos não é minimizar as dificuldades; é recusar a redução do estudante às suas dificuldades, insistindo que a escola deve ser capaz de enxergar o ser humano inteiro que está diante de si. Na proposta mais ampla aqui adotada, entende-se que o tratamento clínico de condições como a bipolaridade é uma ferramenta de garantia de qualidade de vida, sobrevida e manutenção da agência do sujeito, sem que isso anule sua natureza intrínseca como ser neurodivergente.

A quarta justificativa é de ordem metodológica. O recorte pela neurodiversidade permite agregar, sob um mesmo enquadramento analítico, condições que os manuais diagnósticos classificam em categorias distintas, mas que compartilham, do ponto de vista educacional, um conjunto de desafios comuns, quais sejam, a invisibilidade da diferença, a resistência institucional à adaptação, a tendência à culpabilização do estudante e a insuficiência dos modelos pedagógicos tradicionais para acolhê-las. Essa transversalidade analítica é, ela mesma, uma contribuição do paradigma da neurodiversidade ao campo educacional, pois permite que se construam respostas pedagógicas que não dependam exclusivamente do diagnóstico específico de cada aluno, mas que partem do reconhecimento mais amplo de que a variação cognitiva é um fato constitutivo de qualquer grupo humano.

Por fim, a quinta justificativa é de natureza ética e se articula com os princípios da educação inclusiva examinados nas seções anteriores. Se o princípio fundador da inclusão é o de que a escola deve se reorganizar para acolher todos os estudantes, independentemente de suas características, então o paradigma da neurodiversidade representa a extensão mais consequente desse princípio ao campo da cognição. Ele interpela a escola não apenas quanto à sua arquitetura física, mas quanto à sua arquitetura pedagógica, curricular e relacional. Ele exige que a instituição se pergunte, de forma sistemática e honesta, se seus processos foram desenhados para um único tipo de mente ou para a pluralidade de mentes que, de fato, compõe sua comunidade.

Essa interpelação não é confortável, mas é necessária. É essa a razão pela qual este trabalho adota o recorte da neurodiversidade: não como modismo conceitual, mas como ferramenta crítica que permite formular perguntas que o paradigma médico, por si só, não autoriza. As seções que se seguem procuram explorar as implicações concretas desse recorte, examinando as condições específicas contempladas e o modo como se manifestam no ambiente escolar.

5 Mapeamento das condições contempladas

Que sujeitos habitam o campo coberto pelo conceito de neurodiversidade tal como o empregamos neste trabalho? Esta pergunta exige uma resposta cuidadosa, pois a amplitude do conceito é, simultaneamente, uma de suas forças e um de seus riscos. Conforme detalhado anteriormente, a proposta mais ampla de neurodiversidade abarca uma vasta gama de variações, desde as inatas até as clínicas e sensoriais. Contudo, para fins de aprofundamento analítico e considerando as urgências do contexto acadêmico e as potencialidades da Inteligência Artificial Generativa que serão exploradas adiante, este trabalho opera um mapeamento focado em oito grupos de condições.

Esta seleção representa os principais desafios de estigma e invisibilidade na escola contemporânea, servindo como base para discutir como as tecnologias emergentes podem atuar como próteses cognitivas e mediadoras de inclusão.

O transtorno de déficit de atenção e hiperatividade

O Transtorno de Déficit de Atenção e Hiperatividade é, entre as condições aqui contempladas, a que reúne o maior volume de literatura educacional. Caracteriza-se por padrões persistentes de desatenção, hiperatividade e impulsividade que interferem no funcionamento em múltiplos contextos. No ambiente escolar, manifesta-se tipicamente como dificuldade de sustentar a atenção em tarefas prolongadas e agitação motora. O modelo do déficit tende a negligenciar que esses estudantes frequentemente demonstram hiperfoco em atividades de interesse genuíno. Esta assimetria indica que o problema não reside na ausência de capacidade atentiva, mas na incompatibilidade entre as demandas de um ambiente rígido e o modo específico de funcionamento desse cérebro.

Como propõe Armstrong (2011), o padrão de atenção desses estudantes assemelha-se ao de um caçador, sendo distribuído e sensível a variações, o que se torna desadaptado apenas

quando confrontado com a exigência escolar de imobilidade e foco prolongado em estímulos únicos. Armstrong (2011, p. 34) argumenta que o TDAH pode ser visto como um conjunto de traços que foram, no passado da espécie, evolutivamente vantajosos, incluindo a busca por novidades, a criatividade e a capacidade de reagir rapidamente a estímulos ambientais. Ele afirma que o que chamamos de patologia é, muitas vezes, uma atenção errante que, em contextos adequados, permite conexões originais que mentes focadas jamais fariam. Nesse sentido, tanto o TDAH quanto o transtorno bipolar representam a expressão de cérebros que possuíam alta funcionalidade em um ambiente natural que já não existe. O processo civilizatório criou um ambiente humano artificial onde esse tipo de configuração cognitiva não é mais bem-vinda, transformando o que era uma vantagem adaptativa em um transtorno passível de medicalização por não se ajustar à rigidez produtivista da vida moderna.

O transtorno afetivo bipolar

O Transtorno Afetivo Bipolar apresenta ao ambiente escolar desafios que permanecem, em grande medida, invisíveis. Caracteriza-se pela alternância entre episódios de humor elevado ou irritável e episódios de humor deprimido[5]. Nesta pesquisa, defende-se que a bipolaridade é parte da natureza intrínseca do sujeito e que o tratamento clínico não visa a cura, mas a garantia de qualidade de vida. Tratar a bipolaridade é um ato de preservação da agência, permitindo ampliar a sobrevida, reduzir riscos de suicídio e mitigar o declínio cognitivo precoce.

A abordagem da neurodiversidade convida a escola a reconhecer a realidade neurológica das oscilações de humor, manejando-as pedagogicamente sem permitir que o diagnóstico consuma a subjetividade do aluno. Para Armstrong (2011, p. 104), a bipolaridade traz consigo o dom do humor, sugerindo que as fases em que este último se encontra elevado

[5] No uso cotidiano, chamamos de "humor" uma disposição passageira, algo que muda com uma boa notícia ou com um contratempo. No Transtorno Afetivo Bipolar, o termo adquire outro sentido, referindo-se ao estado de fundo da mente, persistente e relativamente independente dos acontecimentos externos. Esse estado oscila entre dois polos. No primeiro deles, a depressão, a pessoa não experimenta apenas tristeza, mas o apagamento da vontade, a sensação de que pensar exige esforço deliberado, de que o futuro simplesmente não e que não se resolve com uma boa notícia ou com uma conquista do dia. No polo oposto, denominado mania ou hipomania conforme sua intensidade, a mente acelera, surgem projetos grandiosos, a necessidade de sono diminui drasticamente e a pessoa pode tomar decisões graves sem perceber os riscos. É preciso desfazer um equívoco frequente que é o de conceber a bipolaridade como representada nos memes compartilhados nas redes sociais, em que comportamentos oscilam de um polo ao seu oposto em minutos. Isso não acontece. As fases do TAB possuem uma estabilidade neuroquímica própria e podem durar dias, semanas ou meses. É precisamente essa persistência que torna o transtorno tão grave, já que os medicamentos levam dias para produzir efeito, deixando a pessoa vulnerável ao estado em que se encontra por um período que não pode ser abreviado pela força da vontade ou pela mudança das circunstâncias.

podem ser motores de uma criatividade e produtividade prodigiosas. Ele observa que muitos dos maiores artistas e pensadores possuíam essa configuração neuronal, utilizando a intensidade emocional como matéria-prima inovadora. Entre um bipolar que trata e um autista que não necessita de medicação, a semelhança reside na vulnerabilidade compartilhada diante de uma estrutura social que não foi desenhada para seus ritmos. Essa incompatibilidade é agravada pelo que Byung-Chul Han (2015) denomina sociedade do desempenho, onde a pressão por produtividade ininterrupta e o excesso de positividade exaurem o cérebro neurodivergente. Na visão de Han (2015), a era contemporânea é marcada por doenças neuronais que são subprodutos da nossa própria forma de viver, evidenciando que ritmos emocionais singulares são transformados em patologia por um sistema que não tolera a inconstância nem o repouso.

O transtorno opositor desafiador

O Transtorno Opositor Desafiador é frequentemente lido pela escola como um problema de caráter, e não como uma condição neurológica. Caracteriza-se por humor irritável e comportamento desafiador sistemático diante de figuras de autoridade. No entanto, por trás da aparente recalcitrância, frequentemente residem uma inteligência crítica aguçada e uma sensibilidade elevada à injustiça percebida. A falha de entendimento entre o estudante com TOD e a instituição é um exemplo do problema da dupla empatia de Milton (2012), onde o ambiente típico falha em compreender o imperativo ético ou a necessidade de agência do estudante.

Armstrong (2011) sugere que o comportamento desafiador pode ser uma manifestação de uma necessidade profunda de integridade e autonomia. Embora o autor não dedique um capítulo isolado ao TOD, ele discute extensivamente como temperamentos fortes e mentes independentes são frequentemente patologizados por um sistema que valoriza a docilidade. A resposta pedagógica eficaz não é a imposição de autoridade, mas a construção de relações que ofereçam escolha e controle, reconhecendo o potencial de liderança e o senso crítico desses sujeitos, que se recusam a aceitar normas arbitrárias sem fundamentação lógica.

Os transtornos alimentares

A inclusão dos transtornos alimentares, como a anorexia e a bulimia, neste mapeamento justifica-se por sua dimensão cognitiva e relacional. A obsessão pelo controle e

os padrões de pensamento perfeccionista que caracterizam essas condições permeiam o desempenho acadêmico e são frequentemente alimentados pela cultura escolar de comparação. Reconhecer a neurobiologia desses transtornos é um passo necessário para romper estigmas. Situá-los no campo da neurodiversidade significa admitir que eles implicam modos específicos de processar informações sobre performance e controle, que têm repercussões diretas na vivência das demandas escolares.

Armstrong (2011, p. 120) toca na questão do perfeccionismo e da rigidez ao discutir como certas configurações cerebrais buscam a ordem e a simetria de forma intensa. No caso dos transtornos alimentares, essa força de sistematização e controle volta-se contra o próprio corpo. A escola deve reconhecer que o esforço hercúleo de autodisciplina demonstrado por esses estudantes é uma potência cognitiva que está sendo mal direcionada. Transformar esse ambiente exige uma redução das pressões de performance padronizada, permitindo que a necessidade de excelência do estudante encontre caminhos criativos e saudáveis de expressão.

Essa dinâmica é amplificada pelo que Byung-Chul Han (2015) descreve como a transição da sociedade disciplinar para a sociedade do desempenho, agora viciada na efemeridade das imagens digitais de plataformas como o Instagram e o TikTok. Vivemos em um tempo em que o corpo perfeito e a vida idealizada são compartilhados e impostos diariamente na ordem social como metas inalcançáveis ou produtos a serem comprados. A sociedade do desempenho alimenta e estimula, ininterruptamente, dentre outros, os transtornos alimentares ao transformar a imagem no critério supremo de valor. Nesse contexto, a escola pode auxiliar estudantes com tais questões se atuar como um espaço de resistência e desconexão. Ao privilegiar a arte, as disciplinas científicas e o mergulho em saberes sólidos, a instituição de ensino oferece aos alunos algo concreto e profundo em que focar, servindo como uma âncora intelectual que impede o sujeito de sucumbir diante da dispersão das telas. Trata-se de oferecer um horizonte de sentido que permita ao estudante desligar-se da tirania das redes sociais e construir uma subjetividade que não dependa da validação visual constante e padronizada do algoritmo (Han, 2015).

Transtornos específicos de aprendizagem: dislexia, discalculia e disgrafia

Os transtornos específicos de aprendizagem constituem um dos pilares da proposta seminal de neurodiversidade, mas sua análise em uma perspectiva ampla exige reconhecer que eles não são falhas de inteligência, mas variações no processamento de símbolos, sons e números. No curso de Matemática, a discalculia, por exemplo, é frequentemente confundida

com a incapacidade intelectual, quando na verdade representa uma divergência na percepção de grandezas e na automação de processos numéricos.

Armstrong (2011, p. 79) é enfático ao afirmar que a dislexia e as dificuldades correlatas trazem vantagens significativas no processamento visual-espacial. Ele argumenta que esses estudantes possuem uma mente tridimensional, capaz de visualizar objetos de múltiplos ângulos e de perceber padrões globais que mentes focadas no processamento linear de texto não conseguem detectar. Para o autor, o dom da dislexia é uma habilidade de síntese e de pensamento holístico. Na matemática e na engenharia, essa capacidade de visualização geométrica é uma força poderosa que a escola frequentemente anula ao exigir apenas a manipulação mecânica de símbolos e fórmulas.

Esquizofrenia e o Espectro Psicótico

A esquizofrenia e os fenômenos associados, como as paranoias, representam uma das fronteiras mais complexas da neurodiversidade clínica. Incluí-las significa reconhecer que o processamento da realidade e a percepção sensorial podem divergir drasticamente do padrão neurotípico sem que isso anule o direito do sujeito à permanência escolar. O estigma da periculosidade e a patologização extrema frequentemente empurram esses indivíduos para fora da escola.

Armstrong (2011, p. 165) discute como o espectro psicótico está intimamente ligado ao pensamento divergente e à capacidade de fazer associações remotas entre ideias. Ele sugere que a mente esquizofrênica habita um mundo de significados expandidos, onde as fronteiras entre conceitos são fluidas, o que é a base de muitos processos de inovação radical. A abordagem da neurodiversidade defende que, com o suporte clínico e institucional adequado, as variações de pensamento podem ser integradas ao ambiente pedagógico, permitindo que o estudante seja compreendido como uma mente que habita uma realidade cognitiva singular e rica em metáforas e conexões inesperadas.

Fenômenos ansiosos e depressivos: fobia social, pânico e depressão maior

Os fenômenos ansiosos, incluindo a fobia social e o transtorno do pânico, bem como a depressão maior, alteram de forma profunda e persistente a disponibilidade cognitiva e a interação social. O estudante com fobia social não experimenta apenas uma timidez acentuada, mas um processamento hipervigilante do julgamento alheio que pode paralisar sua participação

acadêmica. Reconhecer essas condições como parte da neurodiversidade é admitir que a arquitetura emocional do cérebro é indissociável de sua arquitetura cognitiva.

Armstrong (2011) associa a ansiedade e a depressão a uma sensibilidade elevada ao ambiente e a uma profundidade de processamento afetivo. Ele argumenta que indivíduos com tendência à depressão frequentemente possuem uma visão mais realista e menos enviesada da realidade, o que ele chama de realismo depressivo. Essa sensibilidade, embora gere sofrimento em ambientes hostis ou competitivos, é uma força de empatia e de análise crítica profunda. A escola deve atuar na redução dos fatores de estresse, permitindo que essa profundidade emocional se torne um recurso para a compreensão humana e para o engajamento ético com o mundo.

Neste cenário, a formação docente fundamentada no paradigma da neurodiversidade revela-se um instrumento de justiça curricular. O olhar atento do professor, muitas vezes sensibilizado por sua própria vivência neurodivergente, permite identificar que o que parece desinteresse ou incapacidade é, na verdade, uma barreira funcional específica. Um exemplo reside na fobia social: para este estudante, a obrigatoriedade de um seminário ou de apresentações públicas não constitui um desafio pedagógico ordinário, mas uma fonte de paralisia. Nestes casos, a avaliação deve ser obrigatoriamente adaptada para não se transformar em um instrumento de opressão. Condições como a depressão, o pânico e a fobia social exigem cuidados metodológicos específicos, sob o risco de que o ambiente escolar se torne tão hostil que force o estudante ao abandono ou à permanência sob sofrimento psíquico severo. Escrever sobre estas realidades é, portanto, um ato de fornecer subsídios para que outros educadores possam enxergar seus alunos na plenitude de suas existências e construir avaliações que respeitem a arquitetura emocional de cada mente.

Altas Habilidades e Superdotação

As altas habilidades representam, talvez, o maior paradoxo da inclusão escolar. Por apresentarem um desempenho intelectual superior, esses estudantes são frequentemente esquecidos pelas políticas de suporte. Constituem o paradoxo da invisibilidade no âmbito da neurodiversidade: sua alta performance camufla um rastro de incompreensão, isolamento social e danos emocionais severos gerados pela assincronia entre o desenvolvimento intelectual e o emocional.

Armstrong (2011) enfatiza que esses estudantes frequentemente enfrentam o que se denomina assincronia, um descompasso entre o desenvolvimento cognitivo acelerado e a maturidade emocional ou social que segue o ritmo cronológico. Esta disparidade gera um sentimento de inadequação profunda, pois o sujeito compreende conceitos complexos mas carece de ferramentas emocionais para lidar com a frustração ou com a exclusão social. O tédio na escola não é uma simples falta de ocupação, mas uma forma de erosão psíquica intensa. Diante de um currículo repetitivo e lento, mentes superdotadas podem desenvolver comportamentos disruptivos que mimetizam o TDAH ou o TOD, como forma de reagir a um ambiente que anula sua sede de saber.

É imperativo considerar que as altas habilidades frequentemente se sobrepõem ao Transtorno do Espectro Autista (TEA), gerando o que a literatura denomina Dupla Excepcionalidade. Nestes casos, a comorbidade não é uma exceção, mas uma configuração que desafia o olhar clínico e pedagógico tradicional. Como destaca Armstrong (2011), muitos estudantes superdotados são erroneamente diagnosticados como autistas devido à sua intensidade e ao seu isolamento social, que decorre mais do tédio perante conteúdos simplistas do que de uma barreira na comunicação social. Simultaneamente, há sujeitos autistas com alto desempenho intelectual cujas potências são ignoradas pelo foco exclusivo no déficit. A escola, ao não compreender essa interseção, produz uma redução das possibilidades de desenvolvimento desse aluno, tratando a inteligência superior e a neurodivergência como campos mutuamente excludentes, quando na verdade eles frequentemente habitam a mesma arquitetura cerebral.

Han (2015) observa que, na sociedade do desempenho, o superdotado é frequentemente capturado pela armadilha do perfeccionismo absoluto, onde o erro não é visto como parte do aprendizado, mas como uma falha ontológica catastrófica. A pressão constante por resultados excepcionais transforma o talento em um fardo de vigilância ininterrupta sobre si mesmo, levando à redução da saúde mental e ao isolamento, uma vez que a dificuldade em encontrar pares intelectuais reforça a percepção de ser um estrangeiro em seu próprio meio. Para esses alunos, a escola deve ser um lugar que desafie sua capacidade de síntese e que acolha suas paixões intensas por temas específicos, garantindo que o potencial do indivíduo não se transforme em frustração ou em um senso de desajuste permanente perante a cultura da mediocridade institucional.

6 A comorbidade como regra, não como exceção

Que configurações assumem, na prática, as condições neurodivergentes quando se apresentam a um professor? E como compreender o fato de que, na maior parte dos casos, elas não chegam sozinhas?

Um aspecto transversal a todas as condições contempladas neste mapeamento que merece atenção especial é a comorbidade. Bukstein (2001) define-a como a presença de mais de um transtorno em uma mesma pessoa em um período específico, podendo ser examinada tanto pela perspectiva atual quanto pela perspectiva ao longo da vida. Na prática clínica e educacional, as condições neurodivergentes raramente se apresentam de forma isolada. O estudante com TDAH frequentemente apresenta também transtorno de ansiedade. O estudante com transtorno bipolar pode ter diagnóstico concomitante de TDAH ou de transtorno alimentar. O estudante com TOD frequentemente coexiste com TDAH, e ambos podem coexistir com dificuldades de aprendizagem específicas. Conforme apontam Rosa e Bucco (2025), não existe um consenso na literatura sobre os grupos que compõem o campo das neurominorias, mas há ampla evidência de que essas condições tendem a aparecer em configurações complexas e sobrepostas.

Essa sobreposição não é marginal, mas a norma. Sartorius (2013) afirma que a ocorrência simultânea de transtornos mentais é a regra, e não a exceção, e que todo paciente deve ser abordado com essa premissa em mente. Conforme demonstram AL-Asadi, Klein e Meyer (2015), em estudo com 12.665 participantes avaliados por meio de ferramenta diagnóstica on-line, praticamente todos os transtornos psicológicos investigados apresentaram co-ocorrências significativas entre si, tanto para homens quanto para mulheres.

Os autores identificaram relações consistentes entre a presença simultânea de transtornos e taxas mais elevadas de ideação suicida, maior gravidade dos sintomas, pior qualidade de vida e menor suporte social. Além disso, observaram que a comorbidade díade, isto é, a presença concomitante de dois transtornos, aumentava os riscos psicossociais de forma mais acentuada do que a presença isolada de qualquer um dos dois transtornos considerados separadamente (Al-Asadi; Klein; Meyer, 2015).

Para Sartorius (2013), é necessário aceitar o fato de que a comorbidade de diversas doenças, e em particular a ocorrência simultânea de transtornos mentais e físicos, é a regra e não a exceção, e que todo paciente e, acrescentamos, todo estudante neurodivergente, deve ser abordado com essa premissa em mente. No contexto educacional, isso assume um significado específico e urgente. A superdotação, por exemplo, compõe com certa frequência perfis

duplamente excepcionais, nos quais convive com TDAH, transtorno do espectro autista ou transtornos de aprendizagem específicos.

Conforme observam Arky e Webb (2025), crianças que são ao mesmo tempo superdotadas e apresentam desafios de aprendizagem podem ser difíceis de compreender, já que seus pontos fortes podem encobrir suas dificuldades, ou suas dificuldades podem mascarar a superdotação, de modo que, em alguns casos, nenhum dos dois aspectos é reconhecido. Esse fenômeno, conhecido na literatura como dupla excecionalidade, ilustra com precisão o problema da comorbidade invisível como, por exemplo, o laudo de TDAH que, às vezes, não diz que há superdotação junto. O diagnóstico de bipolaridade, às vezes, não revela o transtorno de ansiedade que o acompanha. O professor, então, acredita compreender seu aluno, mas sua compreensão é apenas parcial.

Há ainda uma dimensão epistemológica nessa questão que merece atenção. A comorbidade psiquiátrica carrega, implicitamente, uma suposição de independência entre os transtornos, o que frequentemente não corresponde à realidade clínica (Allsopp; Read; Corcoran, 2023). As fronteiras diagnósticas são, muitas vezes, construções que não capturam a totalidade da vivência subjetiva. O próprio grupo-tarefa responsável pela elaboração do DSM-5 reconheceu que as fronteiras entre as categorias de transtornos são mais fluidas do que se acreditava anteriormente (APA, 2013). Essa fluidez não invalida o diagnóstico clínico, mas o relativiza e é exatamente essa relativização que o professor precisa incorporar à sua prática.

Essa realidade tem uma implicação pedagógica fundamental, pois o diagnóstico clínico, quando existe, raramente oferece ao professor uma descrição completa do estudante que tem diante de si. Um laudo de TDAH não informa sobre a ansiedade que acompanha a dificuldade de organização. Um diagnóstico de transtorno bipolar não descreve as estratégias de enfrentamento que o estudante desenvolveu ao longo de anos de convivência com sua condição. Um relatório médico, por mais preciso que seja, captura a condição; não captura o sujeito.

É precisamente aí que o paradigma da neurodiversidade oferece sua contribuição mais relevante para a prática pedagógica, pois convoca o educador a conhecer o estudante para além do diagnóstico. Não a ignorar o diagnóstico, que pode ser uma informação preciosa para orientar adaptações, mas a não se deixar reduzir por ele. A pergunta que a perspectiva da neurodiversidade coloca ao professor não é "qual é o diagnóstico desse aluno?", mas "quem é esse aluno, como ele aprende, o que o mobiliza, o que o paralisa, e o que eu preciso ajustar na minha prática para que ele possa aprender?".

7 Considerações (in)conclusas: o caminho para a ação pedagógica

As discussões apresentadas ao longo deste capítulo buscaram não apenas definir a neurodiversidade, mas situá-la como um campo de forças políticas, éticas e pedagógicas que exige uma revisão profunda dos marcos educacionais brasileiros. Ao partirmos do lugar de fala e transitarmos da proposta seminal para uma perspectiva ampliada, reconhecemos que a inclusão cognitiva é mais do que a simples aceitação de laudos clínicos; é o reconhecimento da variação humana como uma vantagem evolutiva e um direito à dignidade.

A pergunta que encerrou a seção anterior — "quem é esse aluno?" — é simples em sua formulação, mas extremamente exigente na prática. Ela serve como o fio condutor que este capítulo deixa aberto para as reflexões subsequentes. A escola brasileira avançou, nos últimos trinta anos, em sua capacidade de nomear e identificar as diferenças que habitam suas salas. O desafio que permanece, e que os capítulos seguintes procurarão examinar, é o de transformar esse reconhecimento em uma reengenharia pedagógica efetiva.

Esta conclusão permanece aberta e inconclusa pois a teoria, por si só, não remove as barreiras da sala de aula. É necessário agora investigar como o ambiente escolar, pressionado pela patologização e cercado por marcos legais, pode encontrar nas tecnologias emergentes, especificamente na Inteligência Artificial Generativa, as ferramentas necessárias para dar visibilidade a esses sujeitos e personalizar o cuidado humano sem substituí-lo. O que foi aqui fundamentado como direito deve, a partir de agora, ser explorado como prática e possibilidade técnica.

Referências

AL-ASADI, Ali M.; KLEIN, Britt; MEYER, Denny. Multiple comorbidities of 21 psychological disorders and relationships with psychosocial variables: a study of the online assessment and diagnostic system within a web-based population. **Journal of Medical Internet Research**, v. 17, n. 3, e55, 2015. Disponível em: https://www.jmir.org/2015/3/e55/. Acesso em: 10 abr. 2026.

ALLSOPP, Kelly; READ, John; CORCORAN, Rhiannon. Psychiatric comorbidity: a concept in need of a theory. **Psychological Medicine**, v. 53, n. 14, p. 6513–6521, 2023. Disponível em: https://www.google.com/search?q=https://doi.org/10.1017/S003329172300052X. Acesso em: 10 abr. 2026.

AMERICAN PSYCHIATRIC ASSOCIATION. **Diagnostic and statistical manual of mental disorders**: DSM-5. 5. ed. Arlington: American Psychiatric Association, 2013.

ARKY, Beth; WEBB, James T. **Twice-exceptional kids**: both gifted and challenged. Child Mind Institute, 2025. Disponível em: https://childmind.org/article/twice-exceptional-kids-both-gifted-and-challenged/. Acesso em: 10 abr. 2026.

ARMSTRONG, Thomas. **The power of neurodiversity**: unleashing the advantages of your differently wired brain. Philadelphia: Da Capo Lifelong Books, 2011.

ARMSTRONG, Thomas. **Neurodiversity in the classroom**: strength-based strategies to help students with special needs succeed in school and life. Alexandria: ASCD, 2012.

BLUME, Harvey. On the neurological underpinnings of geekdom. **The Atlantic**, [s. l.], v. 282, n. 3, p. 11-12, set. 1998. Disponível em: https://www.theatlantic.com/magazine/archive/1998/09/neurodiversity/305909/. Acesso em: 9 abr. 2026.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 9 abr. 2026.

BRASIL. **Decreto nº 6.949, de 25 de agosto de 2009**. Promulga a Convenção Internacional sobre os Direitos das Pessoas com Deficiência e seu Protocolo Facultativo, assinados em Nova York, em 30 de março de 2007. Brasília, DF: Presidência da República, 2009. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2009/decreto/d6949.htm. Acesso em: 9 abr. 2026.

BRASIL. **Decreto nº 12.686, de 20 de outubro de 2025**. Institui a Política Nacional de Educação Especial Inclusiva e a Rede Nacional de Educação Especial Inclusiva. Brasília, DF: Presidência da República, 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/d12686.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996**. Estabelece as diretrizes e bases da educação nacional. Brasília, DF: Presidência da República, 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, DF: Presidência da República, 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 14.254, de 30 de novembro de 2021**. Dispõe sobre o acompanhamento integral para educandos com dislexia ou Transtorno do Déficit de Atenção com Hiperatividade (TDAH) ou outro transtorno de aprendizagem. Brasília, DF: Presidência da República, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14254.htm. Acesso em: 9 abr. 2026.

BRASIL. Ministério da Educação. **Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva**. Brasília, DF: MEC, SEESP, 2008. Disponível em: http://portal.mec.gov.br/arquivos/pdf/politicaeducespecial.pdf. Acesso em: 9 abr. 2026.

BUKSTEIN, Oscar Gary. Comorbidity. In: **International Encyclopedia of the Social & Behavioral Sciences**. Oxford: Elsevier, 2001. p. 2351-2354.

CANGUILHEM, Georges. **O normal e o patológico**. Tradução de Maria Thereza Redig de Carvalho Barrocas. 6. ed. Rio de Janeiro: Forense Universitária, 2009.

CHAPMAN, Robert. **Empire of normality**: neurodiversity and capitalism. London: Pluto Press, 2023.

CONNOR, David J.; VALLE, Jan. Estudos da Deficiência na Educação, passado, presente e futuro? **Educação & Realidade**, Porto Alegre, v. 49, e141799, 2024. Disponível em: https://www.scielo.br/j/edreal/a/yK8hJ7k5T9VzX9S9v9X9S9v/. Acesso em: 9 abr. 2026.

FONSECA, Kátia de Abreu; CRISTOVAM, Maria Osvalda de Castro Feitosa; CAPELLINI, Vera Lucia Messias Fialho; MENEZES, Vinícius Iuri de (Org.). **Práticas inclusivas**: antigas questões, novas possibilidades. São Paulo: Summus, 2011.

GOFFMAN, Erving. **Estigma**, notas sobre a manipulação da identidade deteriorada. Tradução de Márcia Bandeira de Mello Leite Nunes. 4. ed. Rio de Janeiro: LTC, 1988.

HAN, Byung-Chul. **Sociedade do cansaço**. Petrópolis: Vozes, 2015.

MANTOAN, Maria Teresa Eglér. **A educação especial no Brasil**, da exclusão à inclusão escolar. Campinas: Unicamp, 2011.

MILTON, Damian. On the ontological status of autism: the 'double empathy problem'. **Disability & Society**, London, v. 27, n. 6, p. 883-887, 2012. Disponível em: https://www.tandfonline.com/doi/full/10.1080/09687599.2012.710008. Acesso em: 10 abr. 2026.

ORTEGA, Francisco. O sujeito cerebral e o movimento da neurodiversidade. **Mana**, Estudos de Antropologia Social, Rio de Janeiro, v. 14, n. 2, p. 477-509, 2008. Disponível em: https://www.scielo.br/j/mana/a/7k9n8vH8W6bX8fR3mC9p4mF/. Acesso em: 9 abr. 2026.

RIBEIRO, Djamila. **O que é lugar de fala?** Belo Horizonte: Letramento, 2017.

ROSA, Rodrigo Assunção; BUCCO, Eduardo Antony. Neurodiversidade, expandindo as fronteiras da diversidade, equidade e inclusão nas organizações. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 23, n. 1, e2023-0196, 2025. Disponível em: https://www.scielo.br/j/cebape/a/K8w3L9fD8jS6G5H4V2bT1qR/. Acesso em: 9 abr. 2026.

SARTORIUS, Norman. Comorbidity of mental and physical diseases: a main challenge for medicine of the 21st century. **Shanghai Archives of Psychiatry**, v. 25, n. 2, p. 68–69, 2013. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC4054544/. Acesso em: 10 abr. 2026.

SINGER, Judy. **NeuroDiversity**, the birth of an idea. Lexington: Amazon Digital Services, 2017.

UNESCO. **Declaração de Salamanca e linha de ação sobre necessidades educativas especiais**. Adotada pela Conferência Mundial sobre Necessidades Educacionais Especiais, Acesso e Qualidade, Salamanca, Espanha, 7-10 de junho de 1994. Brasília, DF: CORDE, 1994. Disponível em: http://portal.mec.gov.br/seesp/arquivos/pdf/salamanca.pdf. Acesso em: 9 abr. 2026.

WALKER, Nick. **Neuroqueer heresies**: notes on the neurodiversity paradigm, autistic empowerment, and postnormal possibilities. Fort Worth: Autonomous Press, 2021.

ZANATA, Eliana Marques. Prefácio. *In*: FONSECA, Kátia de Abreu; CRISTOVAM, Maria Osvalda de Castro Feitosa; CAPELLINI, Vera Lucia Messias Fialho; MENEZES, Vinícius Iuri de (Org.). **Práticas inclusivas**: antigas questões, novas possibilidades. Campina Grande: Quipá Editora, 2022.

CAPÍTULO 7

O AMBIENTE ESCOLAR, PATOLOGIZAÇÃO E O MARCO LEGAL

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1246



Resumo: O presente capítulo examina as formas pelas quais o ambiente escolar brasileiro produz, de modo estrutural, a invisibilização e a patologização do estudante neurodivergente. Partindo da análise da escola como instituição historicamente constituída para um sujeito médio, discute-se como a norma neurotípica converte a diferença neurológica em desvio moral e comportamental, desencadeando o que se denomina cadeia da patologização. Investiga-se o mascaramento (*masking*) como estratégia de sobrevivência imposta ao estudante neurodivergente, articulando esse fenômeno às noções de injustiça epistêmica, de lugar de fala e de epistemicídio cognitivo. Examina-se, ainda, a exigência escolar de autocontrole como uma forma de violência institucional sobre sujeitos cujas funções executivas e regulação emocional divergem do pressuposto neurotípico. No plano normativo, analisa-se a arquitetura do desamparo produzida pelo ordenamento jurídico brasileiro, identificando as lacunas deixadas pela Lei Brasileira de Inclusão, pela Lei de Diretrizes e Bases, pela Lei da Reforma Psiquiátrica e pelo Decreto n.º 12.686/2025, lacunas que mantêm determinadas condições neurodivergentes num interregno entre os regimes da educação regular e da educação especial. A neurodiversidade é então situada como pauta de direitos e reconhecimento, em diálogo com os movimentos feminista, negro e LGBTQIAP+, e fundamentada no conceito de justiça cognitiva. Por fim, o capítulo denuncia a ausência do professor neurodivergente tanto nas pesquisas educacionais quanto nos quadros docentes das escolas, propondo que a agenda de investigação sobre neurodiversidade incorpore esse sujeito não apenas como agente de inclusão, mas como sujeito de direitos cujas condições de trabalho precisam ser examinadas e reorganizadas.

Palavras-chave: Masking Neurodivergente. Patologização Escolar. Desamparo Normativo. Justiça Cognitiva. Professor Neurodivergente.

Abstract: This chapter examines the ways in which the Brazilian school environment structurally produces the invisibilization and pathologization of neurodivergent students. Starting from an analysis of the school as an institution historically constituted around an average subject, it discusses how the neurotypical norm converts neurological difference into moral and behavioral deviance, triggering what is here called the pathologization chain. Masking is investigated as a survival strategy imposed on neurodivergent students, articulated with the notions of epistemic injustice, place of speech, and cognitive epistemicide. The school's demand for self-control is also examined as a form of institutional violence against subjects whose executive functions and emotional regulation diverge from the neurotypical assumption. At the normative level, the chapter analyzes the architecture of abandonment produced by the Brazilian legal framework, identifying the gaps left by the Brazilian Inclusion Law, the Law of Directives and Bases of National Education, the Psychiatric Reform Law, and Decree No. 12,686/2025 — gaps that leave certain neurodivergent conditions in an interregnum between regular and special education regimes. Neurodiversity is then situated as a rights and recognition agenda, in dialogue with the feminist, Black, and LGBTQIAP+ movements, and grounded in the concept of cognitive justice. Finally, the chapter denounces the absence of neurodivergent teachers both in educational research and in school faculties, proposing that the neurodiversity research agenda incorporate this subject not merely as an agent of inclusion, but as a subject of rights whose working conditions need to be examined and reorganized.

Keywords: Neurodivergent Masking. School Pathologization. Normative Abandonment. Cognitive Justice. Neurodivergent Teacher.

1 INTRODUÇÃO

Há uma cena que se repete, com variações, nos relatos que educadores fazem de seus primeiros encontros com estudantes neurodivergentes. O estudante que não para quieto, que responde antes que a pergunta termine, que esquece o material, que olha pela janela durante a explicação mais importante da aula. O estudante que chora sem motivo aparente numa segunda-feira, que recusa a atividade sem justificativa, que reage com uma explosão desproporcional à menor frustração. O estudante que passa a semana inteira em silêncio e na semana seguinte parece ter energia demais para o espaço que ocupa. O que a escola faz, em geral, com esses estudantes? Encaminha. Encaminha para a coordenação, encaminha para a psicologia, encaminha para a família, encaminha para o médico e, até em situações que beiram à comédia ou a tragédia, en

caminha para o padre, o pai-de-santo ou pastor. E quando o médico devolve um laudo, a escola arquiva o papel e acredita ter cumprido sua parte.

Es

sa narrativa, que se poderia nomear como a cadeia da patologização, revela um modo de funcionamento institucional que este capítulo se propõe a examinar criticamente. Não se trata de atribuir má-fé aos educadores que operam dentro dessa lógica, visto que muitos o fazem com genuíno cuidado e com os recursos conceituais e estruturais disponíveis. Trata-se, antes, de compreender como a escola, enquanto instituição, foi historicamente constituída de tal modo que sua resposta natural à diferença neurológica é a medicalização, e não a reorganização pedagógica. E de examinar as consequências dessa resposta para os sujeitos que nela se encontram.

O capítulo organiza-se em duas seções principais. A primeira analisa o ambiente escolar como sistema que produz, de modo estrutural, a invisibilização e a patologização do estudante neurodivergente. A segunda seção situa a neurodiversidade no campo dos direitos e do reconhecimento, estabelecendo paralelos com outros movimentos sociais que precederam e iluminam a luta das neurominorias, e fundamentando, no conceito de justiça cognitiva, a reivindicação pelo pleno estatuto de sujeito de direito desses estudantes.

2 O AMBIENTE ESCOLAR OPRESSIVO: INVISIBILIDADE, PATOLOGIZAÇÃO E MEDICALIZAÇÃO

Que escola é essa que invisibiliza o estudante neurodivergente? Não é uma escola maligna, nem necessariamente negligente. É, em geral, uma escola ocupada. Ocupada com conteúdos a cumprir, com turmas numerosas, com avaliações externas que cobram resultados padronizados e com professores que raramente tiveram, em sua formação inicial, qualquer contato substantivo com as especificidades das condições neurodivergentes. Essa escola não invisibiliza por omissão deliberada; invisibiliza por *design*. Seu desenho original pressupõe um estudante médio, normal, como filosofa Canguilhem (2009) e tudo aquilo que diverge desse pressuposto entra num registro de problema a ser resolvido, não de diferença a ser acolhida.

Canguilhem (2009) nos oferece uma chave para compreender esse mecanismo que cria a confusão entre a norma estatística, a qual descreve a frequência, e a norma de valor, que prescreve um ideal. Quando a escola toma o funcionamento neurológico da maioria como padrão de referência e o converte em expectativa normativa, ela transforma automaticamente toda divergência em desvio, portanto, "anormal". Por exemplo, estudante com TDAH não consegue corresponder à expectativa de atenção sustentada e silenciosa; o estudante bipolar oscila onde se espera constância; o estudante com TOD contesta onde se espera deferência. Em todos esses casos, o ambiente escolar lê a diferença como falha moral, comportamental ou cognitiva, antes de lê-la como variação neurológica que demanda resposta pedagógica diferenciada.

Essa leitura equivocada tem uma trajetória conhecida. Conforme descreve Ortega (2008), a biologização dos transtornos mentais produziu um duplo movimento. De um lado, desculpabilizou famílias e pacientes ao transferir a origem dos conflitos de dinâmicas relacionais ou falhas comportamentais para o campo da neurobiologia.

Assim, o que outrora era rotulado como "falha de criação" ou "comportamento inadequado" por permissividade familiar, passa a ser lido estritamente como um problema de saúde mental a ser remediado. De outro, abriu caminho para a medicalização de modos de ser e reagir ao mundo, fossem eles variações legítimas da personalidade ou mesmo comportamentos considerados ilegítimos e cruéis.

Atos que antes evocavam o rigor educativo da família ou a sanção moral da comunidade, como a agressividade ou a falta de empatia, são agora frequentemente absorvidos pelo diagnóstico psiquiátrico, substituindo o juízo ético pelo relatório clínico. A escola, ao receber esses diagnósticos sem uma mediação pedagógica crítica, limita-se a arquivar o laudo,

acreditando que a gestão do comportamento do aluno é agora uma tarefa exclusiva da medicina e da psicologia.

O resultado mais visível desse processo é o que se pode denominar a sequência do laudo. Nela, o comportamento atípico é observado pelo professor, comunicado à família e encaminhado ao especialista, que emite um diagnóstico. Tal diagnóstico é devolvido à escola, que registra o aluno numa categoria especial de necessidade, orienta os pais a manter o acompanhamento clínico e, frequentemente, sugere medicação como pré-condição para que o estudante seja "gerenciável" em sala. Nessa sequência, a escola raramente se pergunta o que pode mudar em suas próprias práticas. A pergunta que se faz é: o que precisa ser ajustado no estudante para que ele caiba na escola que temos?

Essa inversão da responsabilidade é, em si mesma, uma forma de violência institucional. Nesse sentido, a escola que exige que um aluno em uso de antipsicóticos sedativos, que causam sonolência prolongada como efeito colateral, chegue, e chegue com "pleno funcionamento", às sete horas da manhã, sob pena de reprovação por falta, ou diminuição de suas notas por falta de "participação", não está exercendo autoridade pedagógica. Antes, está exercendo violência sobre um corpo cujos limites ela se recusa a reconhecer. Do mesmo modo, a escola que exige apresentações orais de um aluno com fobia social severa, invocando a premissa da "preparação para o mercado de trabalho", não está preparando esse aluno para o mundo; está impondo ao seu sofrimento uma justificativa que ignora completamente sua realidade neurológica.

2.1 A invisibilidade estrutural e o mascaramento forçado

Há uma dimensão da opressão escolar sobre o estudante neurodivergente que é particularmente insidiosa porque não assume a forma de proibição ou de punição explícita. Ela assume a forma de expectativa implícita. O estudante aprende, desde cedo, que para ser aceito e, fundamentalmente, para ser ouvido, precisa parecer neurotípico. Esse processo de simulação da normalidade é definido por Devon Price (2022) como "*masking*" ou, em bom português, "mascaramento".

O *masking* pode ser definido um esforço contínuo e frequentemente exaustivo de suprimir traços naturais, imitar padrões de interação social convencionais e ocultar sintomas para evitar o estigma. Para Price (2022), o *masking* não é uma escolha estética ou uma simples adaptação social, mas uma estratégia de sobrevivência imposta por um ambiente que pune a divergência com a exclusão ou a patologização. Segundo o autor, "o *masking* é um trabalho de

representação constante e exaustivo que exige que uma pessoa autista suprima as suas formas naturais de ser a fim de sobreviver num mundo que não foi construído para ela" (Price, 2022, p. 42, tradução nossa).

O mascaramento impõe custos elevados em múltiplas esferas. Do ponto de vista cognitivo, ele exige uma demanda ininterrupta de recursos atencionais e funções executivas que deveriam estar voltados para a aprendizagem. O estudante que monitora cada gesto, tom de voz e expressão facial para se ajustar ao padrão esperado dispõe de uma capacidade reduzida para processar os conteúdos pedagógicos. Trata-se de uma forma de exaustão invisível que pode culminar no que a literatura descreve como esgotamento autista ou *burnout* neurodivergente, um estado de colapso físico e mental que a escola raramente associa à pressão pela conformidade.

Sob a ótica da filosofia social, o mascaramento revela o que Miranda Fricker (2007) denomina injustiça epistêmica, especificamente em sua modalidade de injustiça testemunhal. A injustiça testemunhal ocorre quando um falante recebe um nível reduzido de credibilidade devido a um preconceito do ouvinte em relação à sua identidade. Fricker (2007, p. 1, tradução nossa) explica que "a injustiça testemunhal ocorre quando o preconceito faz com que um ouvinte atribua um nível reduzido de credibilidade à palavra de um falante".

Na estrutura escolar, o estudante, e também outros membros da comunidade escolar, que exibe comportamentos atípicos é frequentemente desacreditado ou lido como alguém cuja percepção da realidade é falha ou puramente emocional. Diante desse cenário, o *masking* opera como uma tentativa de restaurar essa credibilidade perdida. Quando o estudante simula a neurotipicidade, sua voz ressoa com mais força na comunidade escolar, não porque suas ideias mudaram, mas porque ele removeu o "ruído", a "sombra" ou a "sujeira" da divergência que disparava o preconceito do interlocutor.

Contudo, esse ganho de voz é uma armadilha ética profunda. Fricker (2007) argumenta que ser privado de ser reconhecido como um sujeito capaz de transmitir conhecimento é uma ofensa à dignidade humana. Ao forçar o estudante a fingir ser o que não é para que sua palavra seja validada, a escola pratica uma violência que condiciona a cidadania epistêmica à anulação do ser. A redução da injustiça testemunhal por meio do mascaramento é, portanto, uma prova da existência de uma injustiça epistêmica estrutural, a qual afirma que o neurodivergente só é aceito como par se abdicar de sua verdade neurológica.

A aplicação do conceito de lugar de fala de Djamila Ribeiro (2017) à neurodiversidade permite compreender que essa pressão pelo mascaramento não é um evento isolado, mas um reflexo da estrutura de poder que define quem é o sujeito universal. Ribeiro argumenta que "o

lugar de fala não é um 'sou eu que falo', mas uma posição social que autoriza ou desautoriza o discurso" (Ribeiro, 2017, p. 58). No ambiente escolar, a norma neurotípica ocupa, junto a outras tantas, o lugar do "sujeito universal", aquele que não precisa se explicar nem se mascarar porque sua arquitetura cerebral é a métrica da normalidade.

O sujeito neurodivergente, ao contrário, é empurrado para a subalternidade, onde sua fala é constantemente mediada pelo crivo do diagnóstico médico ou pela suspeita de incapacidade. Essa subalternidade é consolidada por uma ordem social que, ao adotar o "normal estatístico" de Georges Canguilhem (2009) como parâmetro de validade, impõe aos que divergem uma redução de valor humano, enquadrando suas existências sob o signo da loucura.

Ao traçar um paralelo com os movimentos feministas e negros discutidos por Ribeiro, percebe-se que a neurodiversidade enfrenta o que a autora chama de "interdição do dizer". Quando pessoas neurodivergentes, dentre elas um estudante, tenta expressar suas necessidades ou sua forma singular de aprender, elas é frequentemente e sistematicamente silenciadas por uma estrutura que rotula sua subjetividade como sintoma de um transtorno. Como assevera Ribeiro (2017, p. 64), "o silenciamento é a própria anulação da subjetividade, é a interdição do ser".

O mascaramento forçado é a materialização desse silenciamento no interior da subjetividade do estudante. Ele finge ser neurotípico para ocupar um lugar de fala que, na verdade, não é seu, alienando-se de sua própria identidade para obter uma validação que a instituição nega à sua natureza original. Portanto, romper com a necessidade do *masking* exige que a escola reconheça a neurodivergência como um local social legítimo, retirando-a da margem da patologia e devolvendo-lhe a autoridade de narrar a própria existência.

2.2 A redução da autonomia: a norma do autocontrole

Nenhuma outra dimensão da opressão escolar sobre o estudante neurodivergente é tão frequentemente confundida com disciplina legítima quanto a exigência de autocontrole. A escola valoriza o autocontrole como virtude pedagógica fundamental. O aluno que sabe se sentar, calar, esperar, organizar e persistir é o aluno que corresponde ao ideal institucional. Essa exigência remete ao que Michel Foucault (1999) descreve como a produção de corpos dóceis, um mecanismo disciplinar que visa tornar o indivíduo útil e submisso através de uma vigilância constante e de uma ortopedia do comportamento. Na escola, a disciplina não busca apenas a ordem, mas a fabricação de um sujeito cuja anatomia política seja inteiramente

gerenciável pela instituição. Esse ideal, porém, pressupõe um aparato neurológico específico que consiste em funções executivas robustas, regulação emocional estável e um sistema de resposta à recompensa que tolera gratificações diferidas.

Pensemos em um primeiro exemplo ilustrativo, cujas lógicas de barreira se aplicam a diversas outras configurações neurodivergentes: o caso do estudante com TDAH. Para esse sujeito, esse pressuposto é problemático em sua base. As funções executivas, que incluem planejamento, inibição de respostas impulsivas, memória de trabalho e flexibilidade cognitiva, são exatamente as funções cujo desenvolvimento é mais afetado na condição. Exigir do estudante com TDAH o mesmo nível de autocontrole que se exige de um estudante neurotípico é, do ponto de vista neurológico, equivalente a exigir de alguém com mobilidade reduzida que suba uma escada sem rampa disponível. A dificuldade não é ausência de esforço ou de vontade; é a inadequação do ambiente às capacidades reais do sujeito. Quando a escola ignora essa assimetria, ela deixa de exercer pedagogia e passa a exercer o que Foucault (1999, p. 118) identifica como uma disciplina que "dissocia o poder do corpo", transformando a resistência neurológica em uma falha moral passível de correção.

Como um segundo exemplo, que nos permite observar como a autonomia é afetada não apenas pela falha na execução, mas pela oscilação drástica do tônus vital, consideremos o estudante com transtorno bipolar. Nesse caso, a questão é qualitativamente diferente e igualmente grave. A autonomia, no contexto do transtorno bipolar, não é um atributo estável que o sujeito pode mobilizar regularmente. Em fases hipomaníacas ou maníacas, a sensação de autonomia e capacidade pode estar inflada além do que o desempenho real justifica, levando o estudante a tomar decisões impulsivas, comprometer-se com projetos que não conseguirá concluir e demonstrar uma confiança que desconcerta professores e colegas. Em fases depressivas, a autonomia colapsa, de modo que levantar da cama, ir à escola ou abrir o caderno podem ser tarefas que demandam um esforço que o organismo simplesmente não tem como sustentar.

A escola que avalia o estudante neurodivergente, seja ele com tdah ou bipolaridade, pela mesma régua com que avalia os demais está, inadvertidamente, penalizando-o por sua condição. O conceito de falta, de prazo e de desempenho mínimo são instrumentos pedagógicos concebidos para um funcionamento neurológico regular. Aplicados sem mediação a um estudante cujo funcionamento oscila de modo significativo ao longo do tempo, esses instrumentos não medem aprendizagem, mas a estabilidade neurológica, que não é produto do esforço escolar e não deveria ser critério de avaliação.

Compreender essa dimensão da opressão escolar não implica propor a abolição das exigências pedagógicas. Implica reconhecer que as exigências precisam ser calibradas às capacidades reais dos estudantes, e que uma escola que não consegue fazer essa calibração não está ensinando a todos: está ensinando aos que já chegam equipados para corresponder ao seu modelo. Os demais, a escola apenas tolera ou exclui, conforme a paciência institucional disponível no momento.

3 NEURODIVERSIDADE COMO PAUTA DE DIREITOS E RECONHECIMENTO

Como a luta pela neurodiversidade se situa no campo mais amplo das lutas por reconhecimento? Essa pergunta não é retórica. Ela importa porque a resposta determina, em larga medida, que tipo de estratégia, seja ela pedagógica, jurídica ou política, se mostra adequada para fazer avançar os direitos das pessoas neurodivergentes e dentre elas, os estudantes com essa condição. Se a questão é apenas técnica, basta produzir melhores diagnósticos e melhores protocolos de atendimento. Se a questão é política, é necessário compreender a neurodivergência como uma dimensão da diversidade humana que articula demandas de justiça que o campo técnico, por si só, não pode satisfazer.

Connor e Valle (2024), ao historiar o campo dos Estudos da Deficiência na Educação, reconstroem com precisão a genealogia política dessa pauta. O movimento pelos direitos das pessoas com deficiência não emergiu no vácuo, mas na esteira dos movimentos pelos direitos civis dos negros norte-americanos, que demonstraram, na prática, que um país resistente à mudança podia ser confrontado e transformado por grupos minoritários organizados. Seguindo essa trilha, o movimento feminista, o movimento *gay* de libertação e, posteriormente, o movimento pelos direitos das pessoas com deficiência foram construindo, cada um a seu tempo, a compreensão de que a exclusão social não é um fado natural, mas um produto histórico que pode e deve ser contestado.

O paralelo entre esses movimentos e a emergência da neurodiversidade como pauta de direitos não é meramente analógico. É estrutural. Em todos esses casos, o ponto de partida é o mesmo, revelando um grupo humano que o poder dominante define pela negativa, por aquilo que lhe falta em relação à norma eleita como universal. As mulheres, definidas pela ausência de razão plena que o pensamento filosófico ocidental lhes atribuiu por séculos. As pessoas negras, definidas pela ausência de civilidade que o projeto colonial precisava postular para se legitimar. As pessoas LGBTQIAP+, definidas pela ausência de normalidade e

moralidade sexual que a heterossexualidade compulsória impõe como parâmetro. E os sujeitos neurodivergentes, definidos pela ausência de funcionamento cognitivo que a escola-padrão toma como dado universal.

Em todos esses casos, a insurgência começa pelo mesmo gesto da recusa de ser definido pela falta. A neurodiversidade, como movimento, realiza esse gesto ao propor que as diferenças neurológicas não são déficits a serem corrigidos, mas variações a serem reconhecidas. Singer (2017) articula essa proposta com clareza ao situar os sujeitos neurodivergentes como minorias neurológicas que lutam por direitos civis e reconhecimento, em um movimento análogo às lutas de raça, gênero e sexualidade. Essa analogia não nivela as especificidades de cada luta, que são distintas em suas histórias e em suas demandas concretas. Ela identifica, porém, a lógica comum que as anima: o questionamento de uma norma que se apresenta como universal e é, na realidade, particular.

3.1 O *masking* e o armário: opressões análogas

O conceito de mascaramento, examinado na seção anterior, ganha um significado político mais preciso quando posto em diálogo com experiências análogas de grupos que também foram compelidos a dissimular sua identidade para sobreviver em ambientes hostis. Na experiência LGBTQIAP+, esse processo é descrito pela metáfora do armário como o espaço simbólico no qual a identidade é mantida oculta para evitar perseguição, exclusão ou violência. No contexto histórico do racismo, o processo análogo foi descrito como branqueamento comportamental, a pressão exercida sobre pessoas negras para que adotassem modos de falar, de se vestir e de se portar que as aproximassem do padrão branco dominante.

O mascaramento neurodivergente opera segundo a mesma lógica, ainda que com suas especificidades. O estudante com paranoia que mobiliza um esforço exaustivo para ocultar a desconfiança sistemática, tentando mimetizar uma naturalidade que não sente para evitar o julgamento alheio. O estudante com depressão que, mesmo diante de um profundo esvaziamento de sentido, performa uma funcionalidade mecânica e um entusiasmo artificial para não transparecer sua letargia. O estudante com discalculia que desenvolve estratégias elaboradas para esconder sua dificuldade com o raciocínio numérico, simulando uma compreensão imediata por receio de que sua limitação cognitiva seja confundida com falta de empenho. Em todos esses casos, o que se pede ao sujeito é que suprima sua experiência real para produzir uma aparência de conformidade com a norma dominante.

Conforme analisa Ribeiro (2017) ao discutir o lugar de fala, o silenciamento dos grupos subalternizados não é apenas uma questão de quem tem direito a falar, mas de quem tem direito a existir plenamente no espaço público. O lugar de fala, entendido como o posicionamento social a partir do qual o sujeito enuncia sua experiência, é sistematicamente negado aos grupos que o poder dominante define pela negativa. O estudante neurodivergente que é forçado a mascarar sua experiência no ambiente escolar está sendo privado de seu lugar de fala, no sentido mais substantivo do termo, pois está sendo privado da possibilidade de existir no espaço pedagógico a partir de sua perspectiva genuína.

3.2 Justiça cognitiva e o estatuto de sujeito de direito

Como fundamentar, em termos filosóficos e jurídicos, o reconhecimento do estudante neurodivergente como sujeito pleno de direitos? O conceito de justiça cognitiva oferece uma base especialmente fecunda para essa fundamentação. Cunhado no campo dos estudos pós-coloniais, em particular por Boaventura de Sousa Santos (2009), o conceito diz respeito ao reconhecimento de que existem múltiplas formas de conhecer o mundo, e que a hierarquização dessas formas segundo os critérios da racionalidade científica ocidental produz uma epistemicídio, conceituado como o assassinato de saberes e de formas de experiência que não correspondem ao modelo dominante.

Transportado para o campo educacional e para o debate sobre neurodiversidade, o conceito de justiça cognitiva ilumina uma dimensão da exclusão que os marcos jurídicos convencionais raramente capturam que é a dimensão epistêmica. O estudante neurodivergente não é apenas prejudicado em sua trajetória escolar pela ausência de adaptações razoáveis. Ele é prejudicado, antes disso, pela desconsideração de seu modo específico de processar, de aprender e de se relacionar com o conhecimento como uma forma legítima de cognição. A escola que só reconhece como válido o caminho de aprendizagem convencional não apenas exclui esse estudante, como também desautoriza sua experiência cognitiva como inferior ou deficiente.

Esse desautorizar tem consequências que se estendem muito além do desempenho acadêmico. Ele produz o que a literatura tem chamado de injustiça epistêmica, conceito desenvolvido por Fricker (2007), já mencionada, para descrever o dano que se faz a um sujeito especificamente em sua capacidade de conhecer e de ser reconhecido como produtor de conhecimento. O estudante que aprende, ao longo de anos escolares, que seu modo de processar é um defeito e não uma diferença desenvolve uma relação comprometida com o

próprio pensamento, pois desconfia de suas intuições, subestima suas capacidades, internaliza a avaliação negativa que o ambiente lhe devolve sistematicamente.

A reivindicação pela neurodiversidade como pauta de direitos inclui, portanto, uma dimensão epistêmica que vai além das adaptações curriculares e dos laudos médicos. Inclui o reconhecimento de que diferentes modos de processar informação, de sustentar atenção, de regular emoções e de se relacionar com o conhecimento são formas legítimas de cognição, e não desvios de um padrão único que a escola tem o dever de corrigir. Essa reivindicação se apoia, juridicamente, em um conjunto normativo, desde a Lei Brasileira de Inclusão até a Convenção sobre os Direitos das Pessoas com Deficiência, que consagram o modelo social e retiram do indivíduo a responsabilidade pela exclusão.

No plano normativo, a Lei Brasileira de Inclusão (Lei n.º 13.146/2015) já consagra, em seu artigo 27, que a educação constitui direito da pessoa com deficiência, assegurado o sistema educacional inclusivo em todos os níveis. A definição de deficiência adotada por esse diploma, que compreende o impedimento não como atributo isolado, mas como resultado da interação com as barreiras do ambiente, é a tradução jurídica precisa do que o paradigma da neurodiversidade afirma no plano conceitual, uma vez que a dificuldade não está apenas no sujeito, mas está no encontro entre o sujeito e um ambiente que não foi concebido para ele.

Reconhecer o estudante neurodivergente como sujeito pleno de direito significa, na prática, reconhecer que ele não precisa se tornar neurotípico para merecer educação de qualidade. Significa reconhecer que a escola tem a obrigação, não apenas a benevolência, de reorganizar seus processos para que esse estudante possa aprender. E significa reconhecer que os movimentos que historicamente lutaram pelo reconhecimento de outros grupos subalternizados deixaram uma herança política que a luta pela neurodiversidade pode e deve reivindicar, a qual demonstra que a normalidade não é universal, e que a diferença não é uma tragédia a superar, mas uma dimensão da condição humana que merece, plenamente, existir.

Nas seções que se seguem, o texto se volta para o exame específico do marco legal brasileiro e das práticas pedagógicas que esse reconhecimento demanda. A pergunta que orienta esse percurso não é mais se o estudante neurodivergente tem direito a uma escola que o acolha. A resposta a essa pergunta já está dada, jurídica e filosoficamente. A pergunta que permanece aberta é: como essa escola, concretamente, se constitui?

4 O marco legal da inclusão e suas lacunas

Iniciemos com uma necessária digressão, uma imagem que se revela quando nos permitimos enxergar aquilo que a estrutura escolar frequentemente ignora. Ao observar o estudante com transtorno bipolar, com transtorno opositor desafiador, com fobia social ou com transtorno alimentar dentro da escola brasileira, surge a figura de alguém que estagnou no meio da ponte. Esse sujeito não se encontra mais do lado de cá, onde os sistemas de exclusão histórica simplesmente não o reconheciam como detentor de qualquer política educacional específica. Entretanto, ele também não alcançou o outro lado, onde uma escola genuinamente inclusiva o aguardaria com práticas reorganizadas, professores formados e recursos adequados. Ele permanece no meio, exposto e sem um apoio institucional definido. Encontra-se em uma zona normativa de desamparo que o ordenamento jurídico brasileiro, tomado em sua extensão e complexidade, ainda não resolveu de maneira satisfatória.

Nomear esse desamparo com precisão é, antes de qualquer proposta de superação, o gesto intelectual mais necessário. Porque o desamparo desses estudantes não resulta de ausência de legislação; resulta de uma estrutura normativa que avançou na direção correta em vários momentos, mas que produziu, no processo, zonas de opacidade que as instituições escolares exploram para justificar a inação. O que falta não é o texto da lei; é a articulação entre os textos disponíveis e a clareza sobre quem tem a obrigação de agir.

4.1 O que a Lei Brasileira de Inclusão garante e o que silencia

A Lei Brasileira de Inclusão da Pessoa com Deficiência, Lei n.º 13.146, promulgada em 6 de julho de 2015, representa um avanço normativo inegável na história dos direitos educacionais brasileiros. Em seu artigo 2.º, ela consagra uma definição de deficiência que abandona o modelo biomédico puro e adota o modelo biopsicossocial, compreendendo a pessoa com deficiência como aquela que possui impedimentos de longo prazo de natureza física, mental, intelectual ou sensorial, os quais, em interação com uma ou mais barreiras, podem ter obstruída sua participação plena e efetiva na sociedade em igualdade de condições com as demais pessoas (Brasil, 2015).

Essa formulação é filosoficamente densa e revela que a deficiência possui uma natureza estritamente dialógica. Ela não deve ser entendida como um atributo isolado ou uma característica intrínseca ao corpo ou à mente, mas sim como o resultado do encontro entre as particularidades do sujeito e as barreiras impostas pelo ambiente. Nesse sentido, a deficiência nasce da relação, pois um impedimento só se torna uma deficiência quando o meio é incapaz de acolher a diversidade do funcionamento humano. Assim, a dificuldade não reside na

constituição do indivíduo, mas na rigidez de um espaço que não foi concebido para a sua forma de existir.

Para os estudantes neurodivergentes contemplados neste trabalho, essa definição abre uma janela jurídica relevante. O transtorno bipolar, a discalculia, os transtornos de caráter psicóticos, a fobia social, os transtornos alimentares e o transtorno opositor desafiador, dentre outros, são, clinicamente, impedimentos que podem ser de longo prazo e de natureza mental. Quando esses impedimentos interagem com as barreiras escolares, como a rigidez de horários incompatível com o ciclo medicamentoso, a obrigatoriedade de apresentações orais para um estudante com fobia social grave, ou a inflexibilidade de prazos para um estudante em fase depressiva, configura-se, em tese, a situação descrita pelo artigo 2.º da LBI. Esses estudantes seriam, nessa leitura, pessoas com deficiência para fins legais, com todos os direitos que daí decorrem.

Entretanto, entre essa leitura extensiva da LBI e a prática das instituições escolares, abre-se um abismo. Esse abismo tem um nome técnico que é o Público-Alvo da Educação Especial, o PAEE, definido pelo Decreto n.º 12.686, de 20 de outubro de 2025, que instituiu a Política Nacional de Educação Especial Inclusiva. O decreto delimita o público que acessa os recursos específicos da educação especial, como o Atendimento Educacional Especializado e o profissional de apoio escolar, a três grupos: pessoas com deficiência, pessoas com Transtorno do Espectro Autista e pessoas com altas habilidades ou superdotação.

Essa delimitação, necessária do ponto de vista do financiamento e da gestão dos recursos do Fundeb, produz um efeito colateral grave, pois coloca em evidência que os estudantes com transtornos de humor, transtornos funcionais e transtornos de conduta não integram, por força de decreto, o público prioritário dos recursos especializados.

O resultado é um paradoxo jurídico de consequências humanas sérias. O estudante com transtorno bipolar, com fobia social ou com transtorno opositor desafiador integra, em tese, o público da educação inclusiva, um direito amplo garantido pela LDB a todos os estudantes. Mas frequentemente não acessa os recursos da educação especial, que são os instrumentos concretos pelos quais a inclusão se materializa em sala de aula.

A escola tem o dever genérico de incluí-lo, mas argumenta, com algum respaldo normativo, que não dispõe dos recursos específicos necessários para fazê-lo. O estudante fica entre dois regimes jurídicos: não é suficientemente "normal" para que a escola regular o acolha sem adaptações, e não é suficientemente "deficiente" para que o sistema de educação especial lhe destine recursos. Fica no meio da ponte. Exposto. Desamparado.

4.2 A Lei da Reforma Psiquiátrica e o vácuo educacional

A Lei n.º 10.216, de 6 de abril de 2001, conhecida como a Lei da Reforma Psiquiátrica, estabelece os direitos das pessoas acometidas de transtorno mental e redireciona o modelo assistencial em saúde mental para serviços comunitários e de base territorial. Em seu artigo 2.º, parágrafo único, ela assegura ao portador de transtorno mental, entre outros direitos, o de ser tratado com humanidade e respeito e no interesse exclusivo de beneficiar sua saúde, visando alcançar sua recuperação pela inserção na família, no trabalho e na comunidade (Brasil, 2001). A Lei aponta, portanto, para um projeto de inserção social que inclui, implicitamente, a inserção educacional.

O problema é que a Lei n.º 10.216 é uma lei de saúde, não uma lei de educação. Ela regula os cuidados no sistema de saúde mental, não a organização dos sistemas de ensino. O estudante com transtorno bipolar tem seus direitos de saúde protegidos por essa legislação, mas quando entra na escola, essa proteção não o segue de modo automático. A articulação entre o sistema de saúde mental e o sistema educacional, prevista como princípio na Lei n.º 10.216 e reiterada pela intersetorialidade do Decreto n.º 12.686 de 2025, permanece, na prática, muito mais declarada do que realizada. A escola e a clínica continuam, em grande medida, como territórios separados que se comunicam por laudos e encaminhamentos, sem uma coordenação sistemática que transforme o cuidado clínico em suporte pedagógico efetivo.

4.3 O avanço pontual da Lei n.º 14.254 e seus limites

Em novembro de 2021, a Lei n.º 14.254 trouxe um avanço normativo específico que merece exame cuidadoso, pois garantiu o acompanhamento integral para educandos com dislexia, TDAH ou outro transtorno de aprendizagem, determinando que as redes de ensino garantissem, entre outras providências, a elaboração de um plano individualizado para esses estudantes. Esse diploma criou, pela primeira vez, uma obrigação legal de suporte pedagógico explicitamente vinculada a condições neurodivergentes que não se enquadram nos critérios tradicionais de deficiência e que anteriormente dependiam de interpretações extensivas da LDB ou da LBI para terem seus direitos reconhecidos.

A Lei n.º 14.254 tem relevância simbólica e prática considerável, pois reconhece que há condições neurodivergentes que demandam resposta pedagógica específica mesmo quando não configuram deficiência nos termos estritos da legislação vigente. Ao fazê-lo, ela amplia o campo da obrigação estatal e reduz o espaço da discricionariedade institucional. No entanto,

os limites da norma são também expressivos. Ela cobre explicitamente a dislexia, o TDAH e os "outros transtornos de aprendizagem", expressão cujo alcance é objeto de disputa interpretativa. Ela não menciona os transtornos de humor, os transtornos de conduta, a fobia social ou os transtornos alimentares. Esses estudantes continuam, mesmo após a Lei n.º 14.254, numa zona de incerteza normativa que a escola frequentemente traduz em inação.

4.4 A arquitetura do desamparo: quem são os sujeitos entre os regimes

Quem são, concretamente, os estudantes que habitam essa zona de desamparo? São jovens que chegam à escola com diagnósticos clínicos estabelecidos ou em processo de estabelecimento, mas cujas condições não produzem marcas físicas imediatamente visíveis. São estudantes que, em fases de descompensação, apresentam comportamentos que a escola interpreta como indisciplina, falta de vontade ou manipulação, e que em fases de compensação parecem tão funcionais quanto qualquer outro colega, o que reforça a desconfiança institucional de que suas dificuldades são voluntárias. São jovens que frequentemente não têm diagnóstico formal porque o acesso ao serviço de saúde mental no Brasil é desigual e o processo diagnóstico pode levar anos, e que, sem o laudo que documentaria sua condição, ficam inteiramente desprovidos de qualquer mecanismo de proteção escolar.

São também estudantes que, quando têm diagnóstico, deparam-se com uma escola que não sabe o que fazer com aquele papel. O laudo de transtorno bipolar não tem um protocolo escolar correspondente da forma como o laudo de deficiência auditiva tem. O diagnóstico de fobia social não aciona um processo de elaboração de Plano de Atendimento Educacional Especializado como o diagnóstico de deficiência intelectual aciona. O resultado é que o laudo vai para a ficha do estudante e lá permanece, sem que o professor regente receba qualquer orientação sobre como ajustar suas práticas, sem que a coordenação desenvolva qualquer protocolo de suporte, sem que a família receba qualquer informação sobre os direitos pedagógicos que aquele diagnóstico poderia amparar.

Esses estudantes são, para usar uma figura que Guimarães Rosa teria reconhecido, os que vivem no meio do caminho, nem bem no começo nem chegados ao fim. A escola padrão não os alcança plenamente porque foi construída para o estudante médio que eles não são. A educação especial não os alcança plenamente porque seus critérios de acesso foram desenhados para condições que eles não têm, ao menos não formalmente. Ficam num interregno que o sistema produz e que o sistema nega: produz porque sua própria arquitetura

normativa gera essas zonas de opacidade; nega porque nenhuma instituição quer reconhecer formalmente que deixou alguém cair entre as malhas da rede.

4.5 Caminhos para o aperfeiçoamento das políticas públicas

Nomear o problema com precisão é condição necessária para propor sua superação. Os estudantes em situação de desamparo normativo não são vítimas de má vontade dos sistemas, mas de uma arquitetura de políticas públicas que não acompanhou a expansão do conhecimento sobre neurodiversidade e sobre o que é necessário para que diferentes tipos de mente aprendam. Há caminhos concretos para corrigir essas lacunas, e eles se organizam em três planos, quais sejam, o normativo, o institucional e o formativo.

No plano normativo, a primeira medida necessária é a revisão do conceito operacional de Público-Alvo da Educação Especial para incluir, de modo explícito, os estudantes com transtornos de humor de longo prazo, com transtornos de conduta que impactem significativamente a trajetória escolar, com fobia social clinicamente estabelecida e com transtornos alimentares que afetem o funcionamento acadêmico. Essa inclusão não precisa ser feita por lei ordinária; pode ser operacionalizada por resolução do Conselho Nacional de Educação, que tem competência para normatizar as diretrizes da educação especial. O precedente já existe na forma como a Lei n.º 14.254 ampliou o escopo das obrigações pedagógicas para além dos critérios tradicionais de deficiência.

A segunda medida normativa necessária é a regulamentação do artigo 28, inciso III, da LBI, que prevê a elaboração de projetos pedagógicos que institucionalizem o atendimento educacional especializado e os serviços de adaptação razoável. Essa regulamentação precisa ser suficientemente específica para cobrir as condições neurodivergentes que hoje ficam no limbo, estabelecendo protocolos claros para a elaboração de planos individualizados, para a flexibilização de prazos e formatos de avaliação e para a articulação entre escola e serviços de saúde mental. O Decreto n.º 12.686 de 2025 avança nessa direção ao prever o Plano de Atendimento Educacional Especializado e o Plano Educacional Individualizado, mas sua aplicação precisa ser expressamente estendida aos estudantes em situação de desamparo normativo.

No plano institucional, a medida mais urgente é a criação de protocolos escolares de acolhimento para estudantes com diagnósticos de saúde mental, que orientem as equipes pedagógicas sobre como proceder ao receber um laudo clínico que não se enquadra nos

critérios tradicionais do PAEE. Esses protocolos precisam responder a perguntas práticas, tais como:

- quem na escola é responsável por elaborar o plano individualizado?
- que adaptações são obrigatórias independentemente de qualquer recurso adicional?
- como a escola se articula com os serviços de saúde que acompanham o estudante?
- quais informações a família tem direito de receber sobre as adaptações disponíveis?

A ausência de respostas institucionalizadas a essas perguntas é o mecanismo mais eficiente de perpetuação do desamparo.

Uma segunda medida institucional indispensável é a criação de espaços de articulação sistemática entre os sistemas de educação e de saúde mental nos municípios e estados, nos moldes do que o Decreto n.º 12.686 prevê em seu artigo 3.º, inciso VIII, ao estabelecer a articulação intersetorial como diretriz da política. Essa articulação precisa superar o formato atual de encaminhamento unidirecional, no qual a escola manda o aluno para o serviço de saúde e aguarda passivamente um laudo. Ela precisa assumir a forma de um trabalho conjunto, no qual o profissional de saúde que acompanha o estudante informa o professor sobre as implicações pedagógicas da condição, e o professor informa o profissional de saúde sobre as barreiras concretas que o ambiente escolar impõe ao estudante.

No plano formativo, a lacuna mais grave é a formação inicial e continuada dos professores. Como seria possível esperar que um professor faça adaptações pedagógicas para um estudante com transtorno bipolar, afantasia ou epilepsia se sua formação nunca o apresentou ao que esse transtorno implica pedagogicamente? Como esperar que um coordenador pedagógico elabore um protocolo de acolhimento para um estudante com fobia social se nunca discutiu, em sua trajetória formativa, a diferença entre timidez, ansiedade e fobia social clinicamente estabelecida? A inclusão efetiva de estudantes neurodivergentes demanda que o conhecimento sobre neurodiversidade entre nos currículos dos cursos de formação de professores, não como disciplina optativa ou conteúdo tangencial, mas como eixo estruturante da formação pedagógica.

A quarta medida proposta, de natureza transversal aos três planos, é a participação dos próprios estudantes neurodivergentes e de suas famílias na construção das políticas educacionais que lhes dizem respeito. O lema que animou a Convenção sobre os Direitos das

Pessoas com Deficiência, "nada sobre nós, sem nós"[6], precisa ser operacionalizado também no campo das neurodivergências que hoje habitam o limbo normativo. As políticas públicas que buscam atender esses estudantes têm sido, historicamente, construídas por especialistas clínicos e gestores que raramente são eles próprios neurodivergentes. Trazer para o centro das decisões as vozes e as experiências de quem vive nos corpos e nas mentes que as políticas buscam alcançar é, ao mesmo tempo, uma exigência de justiça epistêmica e uma condição para que as políticas sejam, de fato, eficazes.

O desamparo dos estudantes que vivem entre os regimes normativos não é inevitável. É uma escolha coletiva que a sociedade brasileira tem feito, por ação ou omissão, ao não exigir que seu ordenamento jurídico e suas práticas institucionais deem conta da complexidade real dos seres humanos que passam pelas escolas. Reverter essa escolha é possível e necessário. Os instrumentos jurídicos, em grande medida, já existem; precisam ser completados, articulados e efetivados. A vontade política de fazê-lo é o que permanece em disputa.

5 O PROFESSOR NEURODIVERGENTE: O SILÊNCIO DA LITERATURA E O VAZIO DAS POLÍTICAS

Há uma ausência que percorre toda a literatura sobre educação inclusiva e que raramente é nomeada com a clareza que merece, qual seja, a ausência do professor. Não do professor como agente de inclusão, papel para o qual existe vasta produção acadêmica, mas do professor como sujeito a ser incluído no ambiente educacional. A pesquisa educacional brasileira construiu um campo robusto sobre como o docente deve acolher o estudante neurodivergente, mas não construiu, até o momento, quase nada sobre o que acontece quando o próprio docente é neurodivergente. Esse silêncio é, em si mesmo, um dado.

A hipótese que este pesquisador propõe, ainda na qualidade de hipótese a ser investigada empiricamente, é que os profissionais da educação neurodivergentes estão sub-representados nos quadros das escolas brasileiras, e que essa sub-representação é invisível porque os mecanismos de exclusão que a produzem raramente chegam a ser nomeados como tal. Para compreender por que essa hipótese merece investigação rigorosa, é preciso examinar dois conjuntos de evidências: os dados sobre a participação das pessoas com deficiência no

[6] O lema "Nada sobre nós, sem nós" surgiu como um grito de resistência do movimento de pessoas com deficiência, consolidando-se nos anos 1990, embora suas raízes remontem a conversas dos anos 1980 na África do Sul. Popularizou-se internacionalmente em 1998 com o livro homônimo de David Werner e a Declaração de Madri em 2002.

mercado de trabalho brasileiro e o que a literatura sobre o *masking* nos ambientes profissionais pode sugerir sobre a experiência de docentes neurodivergentes.

5.1 A neurodiversidade e o mercado de trabalho: o que os dados revelam

Imagine um professor de língua portuguesa, muito inteligente, que descreve de forma bastante vívida a cena de um jantar, e ao se referir aos "pratos" servidos na bela noite, escreve sistematicamente "patos". Imagine que esse professor troque letras com sons semelhantes (ex: P/B, T/D, F/V) e aglutine palavras escrevendo "acasa" em vez de "a casa". Esse professor, que respira o mundo das letras e dos grandes autores, ironicamente tem uma neurodivergência conhecida como disortografia[7]. Seria ele facilmente contratado seja em âmbito privado, seja para lecionar na escola pública?

Que tipo de inclusão é possível quando os próprios agentes chamados a produzi-la não reconhecem, em si mesmos, a condição que pretendem acolher? A escola brasileira construiu, nas últimas décadas, um discurso robusto sobre educação inclusiva, sustentado por legislação avançada, por documentos orientadores e por formações continuadas que multiplicaram o vocabulário da diferença nos corredores institucionais. Contudo, esse discurso encontra um limite que raramente é nomeado com a clareza que merece que é a quase ausência de professores neurodivergentes nos quadros docentes das escolas, seja porque nunca chegaram a esses espaços, seja porque, ao chegarem, precisaram aprender a fazer suas identidades desaparecerem.

Os dados da Pesquisa Nacional por Amostra de Domicílios Contínua de 2022, divulgada pelo Instituto Brasileiro de Geografia e Estatística, permitem dimensionar a extensão estrutural desse problema. Conforme registrado pelo IBGE (2023), apenas 29,2% das pessoas com deficiência em idade de trabalhar estavam ocupadas no mercado de trabalho, em contraste com 60,7% das pessoas sem deficiência. Dos 17,5 milhões de pessoas com deficiência em idade laboral, apenas 5,1 milhões integravam a força de trabalho, e mais da metade delas, 55%, trabalhava na informalidade, com renda média 30,8% inferior à das pessoas sem deficiência. Esses números constroem um cenário de exclusão que se aprofunda quando se examina o percurso de chegada ao exercício da docência.

[7] A disortografia é um transtorno específico de aprendizagem, de base neurobiológica, caracterizado por dificuldades persistentes na escrita correta, ortografia e gramática, independentemente do nível intelectual. Afeta a precisão ortográfica, estruturação de frases e uso de regras gramaticais, sendo comum a troca, omissão ou inversão de letras e sílabas.

O percurso de um professor até a sala de aula é longo e seletivo. Exige aprovação em processos seletivos cujos formatos privilegiam competências de desempenho social; exige sobrevivência em ambientes universitários que, conforme apontam estudos recentes, representam espaços de alto sofrimento para estudantes neurodivergentes, marcados por isolamento, inadaptação sensorial e evasão (Rosa; Bucco, 2025).

Mas o dado que mais ilumina o problema está em outro recorte do IBGE (2023), o qual afirma que mesmo entre as pessoas com deficiência que concluíram o ensino superior, apenas 51,2% estavam ocupadas no mercado de trabalho, enquanto 80,8% das pessoas sem deficiência com a mesma formação estavam ocupadas. A escolaridade superior, portanto, não equaliza as oportunidades. O que esse dado revela é que a barreira de acesso ao trabalho não é de qualificação. É de preconceito, de ambiente inadequado e de uma cultura institucional que não sabe ou não quer lidar com a diferença neurológica ou funcional.

Os poucos professores neurodivergentes que atravessam esse funil e chegam às escolas encontram, na maioria das vezes, uma segunda exigência que é a de se tornarem invisíveis. O *masking*, a camuflagem social pela qual o indivíduo neurodivergente suprime, modifica ou simula comportamentos para parecer neurotípico, opera na escola não como uma escolha, mas como uma condição de permanência (Rosa; Bucco, 2025). Aquele professor que fala alto por intensidade afetiva é rotulado de mal-educado ou agressivo. Aquele que se organiza de forma diferente é visto como desorganizado. Aquele que precisa de previsibilidade e estrutura é considerado inflexível. Aquele que tem dificuldade com as convenções tácitas das relações institucionais é descrito como problemático, difícil, inadequado ao trabalho em equipe. Conforme ilustrado em registro autoetnográfico analisado à luz de Canguilhem (2009), o conflito entre a norma vital do sujeito neurodivergente e a norma institucional da escola produz, sistematicamente, a leitura patologizante do professor como funcionário problema.

A consequência direta desse ambiente é a invisibilização deliberada. Professores neurodivergentes que não foram diagnosticados, e muitos não foram, especialmente aqueles de gerações anteriores ao avanço dos critérios diagnósticos, constroem carreiras inteiras sob o custo psíquico brutal do mascaramento contínuo. Os que têm diagnóstico frequentemente optam por não o revelar, por saber que a revelação pode significar a mudança de como são percebidos por gestores, colegas e famílias de estudantes. A pesquisa *Neurodiversidade no Mercado de Trabalho*, realizada em 2024 pela Consultoria Maya em parceria com a Universidade Corporativa Korú, registrou que quase a metade dos 12 mil profissionais e estudantes entrevistados nunca trabalhou diretamente com pessoas neurodivergentes, o que sugere não a ausência dessas pessoas nos ambientes de trabalho, mas sua invisibilidade sistemática

(Consultoria Maya; Korú, 2024). Na escola, essa invisibilidade é ainda mais densa porque a própria instituição não a reconhece como um problema; trata-a como normalidade.

O caso do autismo permite tornar ainda mais concreto o argumento. Segundo dados da Organização Mundial da Saúde, estima-se que cerca de 85% das pessoas no espectro autista estejam desempregadas no Brasil que é um país que, conforme estimativas, conta com aproximadamente 2 milhões de pessoas autistas (Instituto Inclusão Brasil, 2023). A neuropsicóloga Joana Portolese, da Faculdade de Medicina da USP, registra que apenas entre 10% e 20% das pessoas autistas consideradas leves chegam ao mercado de trabalho ou à universidade (Instituto Inclusão Brasil, 2023). Esse funil de exclusão opera muito antes da escola, já no processo seletivo de ingresso na universidade, nos ambientes universitários que não foram desenhados para acolher modos distintos de processar, comunicar e existir. Quando se pensa especificamente na docência, o funil se estreita ainda mais, pois a profissão exige, cotidianamente, um desempenho de habilidades sociais que os processos seletivos avaliam e que os ambientes de trabalho pressupõem como naturais.

É nesse ponto que a questão deixa de ser apenas de justiça trabalhista e se torna uma questão pedagógica de primeira ordem. O conceito do *Problema da Dupla Empatia*, formulado pelo sociólogo autista Damian Milton (2012) e sistematizado em estudos recentes sobre neurodiversidade e fenomenologia, propõe que as dificuldades de comunicação entre pessoas neurodivergentes e neurotípicas não são uma falha unilateral do sujeito neurodivergente, mas um desencontro mútuo de estilos cognitivos, normas e disposições vivenciais. A escola, construída quase integralmente por e para perfis neurotípicos, perpetua esse desencontro ao ofertar uma inclusão gerida apenas por aqueles que nunca precisaram ser incluídos. Uma escola de professores exclusivamente neurotípicos lidando com alunos neurodivergentes é uma escola em que a dupla empatia jamais pode ser plenamente ativada porque falta, do lado institucional, a experiência vivida da diferença.

Retomemos, pois, o princípio, já mencionado anteriormente, que sintetiza esse problema, um dos mais antigos do movimento pelos direitos das pessoas com deficiência: "nada sobre nós, sem nós. Aplicado à escola, ele significa que uma educação inclusiva genuína não pode ser concebida, implementada e avaliada apenas por sujeitos neurotípicos que estudaram a neurodivergência como objeto exterior. Ela precisa ser co-construída com e por pessoas que conhecem, por dentro, o que significa aprender, relacionar-se e trabalhar de modo diferente do que a instituição presume como natural. Conforme nos mostram Rosa e Bucco (2025), a ausência das neurominorias como grupo representado nas pesquisas e nas práticas de diversidade, equidade e inclusão nas organizações é, simultaneamente, um problema teórico e

um problema político. Na escola, essa ausência é ainda mais grave porque é nela que se formam as representações sobre o que é a neurodivergência. Temos, nesse caso, pois, representações sendo construídas, cotidianamente, sem a voz dos que deveriam ser seus sujeitos.

A dimensão desse abismo pode ser aferida pelos dados sobre o ambiente corporativo, que guarda analogias com a escola. A pesquisa da Consultoria Maya e Korú (2024) revelou que 86,4% dos profissionais entrevistados nunca participaram de treinamentos ou programas relacionados à neurodiversidade no ambiente de trabalho, e que apenas 22% das empresas possuem políticas formais de apoio ao tema (Metadados, 2025). Se esses números já são expressivos no setor privado, é razoável supor que sejam ainda mais severos no serviço público educacional, onde as culturas institucionais tendem a ser mais rígidas e as pressões por conformidade mais veladas, mas não menos eficazes. A escola que declara incluir alunos neurodivergentes sem formar, acolher e reter professores neurodivergentes produz um discurso de inclusão que, na prática, reproduz a mesma lógica que pretende superar que é a de que a diferença deve ser administrada por quem nunca a viveu.

Essa sub-representação é tanto mais difícil de enfrentar quanto mais difícil de documentar. As condições neurodivergentes, especialmente aquelas que não produzem marcas físicas visíveis, raramente aparecem nas declarações funcionais dos servidores públicos ou nos registros das instituições de ensino privado. O professor que foi diagnosticado com TDAH, com dislexia, com transtorno do espectro autista ou com transtorno bipolar carrega, em silêncio, uma condição que a escola não reconhece como parte de sua composição humana.

A exclusão, neste contexto, não é apenas um fenômeno estatisticamente mensurável; é também uma representação que os atores sociais constroem e reproduzem e que justifica, ainda que silenciosamente, práticas institucionais de não acolhimento. Uma escola mais inclusiva para os alunos neurodivergentes talvez precise começar por se perguntar, com honestidade, quantos de seus professores, aqueles que foram chamados de difíceis, de problemáticos, de desajustados, eram, na verdade, neurodivergentes que nunca tiveram o direito de sê-lo.

5.2 O *masking* no ambiente profissional docente

A literatura internacional sobre neurodiversidade tem desenvolvido, nas últimas décadas, um campo de investigação específico sobre o *masking* em ambientes profissionais. Pearson e Rose (2021), em análise conceitual do *masking* autista, definem o fenômeno como a supressão consciente ou inconsciente de respostas naturais neurodivergentes e a adoção de

alternativas em domínios tão variados quanto a interação social, a experiência sensorial, a cognição, o movimento e o comportamento. Pesquisadores como Hull *et al.* (2017) e Pryke-Hobbes *et al.* (2023) têm documentado como indivíduos neurodivergentes, especialmente pessoas autistas e com TDAH, desenvolvem estratégias sofisticadas de camuflagem de seus funcionamentos atípicos para corresponder às expectativas sociais e profissionais dos ambientes em que se inserem.

O que a pesquisa tem revelado, de modo cada vez mais consistente, é que o *masking* não é uma escolha em sentido pleno. Conforme argumentam Pearson e Rose (2021), trata-se de uma resposta compreensível ao estigma e à narrativa de déficit que historicamente cercou a neurodivergência. Mascarar-se é, em termos sociológicos, uma tentativa de "passar por normal", no sentido que Goffman (1963) conferiu ao conceito, qual seja, ocultar um atributo socialmente desacreditado para evitar a rejeição e as consequências que derivariam de sua visibilidade.

No contexto profissional, esse processo adquire uma dimensão ainda mais coercitiva porque, conforme registrado por Pryke-Hobbes *et al.* (2023), a manutenção do emprego é frequentemente percebida pelos próprios trabalhadores neurodivergentes como dependente de sua capacidade de mascarar-se com eficiência. Um dos participantes do estudo expressou com precisão o peso dessa lógica, ao avaliar que os riscos de ser diferente no trabalho podem efetivamente mudar a vida de alguém, pois estão ligados à renda.

A pesquisa identificou que o *masking* no ambiente de trabalho foi amplamente empregado como resposta adaptativa a uma série de desafios sociais do contexto profissional, funcionando habitualmente como estratégia de proteção contra ameaças de resultados negativos tanto sociais quanto trabalhistas. Esse caráter de proteção, contudo, tem um custo que a pesquisa documenta com crescente precisão. Participantes descreveram exaustão extrema como uma consequência danosa do *masking*, relatando que manter a máscara por longos períodos com pouco ou nenhum descanso comprometia a qualidade de seu próprio trabalho. Um dos participantes do estudo sintetizou o paradoxo, ao afirmar que o *masking* sacrifica as próprias capacidades, pois reduz o que se consegue ouvir, faz perder coisas, consome mais energia e impede o uso da mente nas formas em que se sabe que ela funciona muito bem.

A pesquisa da Homewood Health (2026), dedicada ao custo energético do *masking* em ambientes corporativos, articula a dimensão neurobiológica desse processo. Quando uma pessoa neurodivergente enfrenta pressão constante para mascarar-se, os sistemas de estresse do cérebro, particularmente a amígdala e a rede de saliência, entram em sobrecarga, enquanto

o córtex pré-frontal, responsável pelas funções executivas como planejamento, priorização e memória de trabalho, começa a se comprometer. O resultado é que a energia cognitiva que deveria estar disponível para o trabalho em si é consumida pelo trabalho de parecer que se está trabalhando normalmente. O *masking*, nesse sentido, é frequentemente confundido com profissionalismo aprendido, quando é, na verdade, uma estratégia de sobrevivência que muitas pessoas desenvolvem ainda na infância como forma de evitar críticas ou rejeição.

O *burnout* neurodivergente, que é uma consequência frequentemente relatada do *masking* prolongado, difere qualitativamente do esgotamento típico pelo trabalho. Conforme descrito pela Homewood Health (2026), trata-se de uma forma de exaustão marcada por névoa mental, baixa motivação, irritabilidade e incapacidade de iniciar ou concluir tarefas, que não desaparece após um fim de semana de descanso e frequentemente requer descanso profundo e, crucialmente, redução das demandas de *masking* para que haja recuperação. O Neurodiversity Hub (s.d.) oferece uma definição que ilumina a dimensão existencial do fenômeno, ao afirmar que o *burnout* neurodivergente cresce com anos de fingir ser algo que não se é, para caber num mundo que não foi desenhado para pessoas como você.

No ambiente docente, as condições para o *masking* são particularmente intensas. O professor está permanentemente em exposição diante dos alunos, dos colegas, dos gestores e das famílias. Sua autoridade profissional depende, em larga medida, de uma aparência de consistência, equilíbrio e domínio que pode ser extraordinariamente custosa para quem tem um funcionamento neurológico que produz variações significativas de energia, foco ou regulação emocional. Imagine uma professora com TDAH precisa manter a atenção em trinta alunos da educação infantil simultaneamente enquanto gerencia sua própria dificuldade de foco. Tente idelizar um professor bipolar super irritado em sua fase mista que precisa sustentar, por horas, uma presença agradável, educada e compassiva que a regência de turma exige. Elaboremos a visão de uma professora autista que enfrenta, todos os dias, os estímulos sensoriais do ambiente escolar e as demandas de interação social contínua que a função requer.

A pesquisa de Pryke-Hobbes *et al.* (2023) encontrou um dado que ilumina diretamente essa questão, pois muitos participantes do estudo expressaram que o *masking* era essencial para cumprir as responsabilidades do seu cargo, como ser educado, escutar e ceder a vez, visto que a habilidade de se comunicar adequadamente com outros no ambiente de trabalho era vital para o sucesso de suas funções. No caso do professor, a demanda comunicativa é ainda mais intensa do que na maioria das profissões, pois envolve não apenas a interação com pares adultos, mas a gestão simultânea de dezenas de sujeitos em desenvolvimento, com

necessidades, ritmos e demandas distintas. Para um docente neurodivergente, esse conjunto de exigências pode representar uma carga de *masking* que se acumula ao longo de décadas.

Há um elemento adicional que agrava o problema e que a pesquisa começa a documentar. O *masking* no ambiente de trabalho criou barreiras para que os participantes buscassem e recebessem apoio dos empregadores, pois a projeção de uma falsa impressão de que estavam lidando bem com as demandas deixou outros membros da equipe ignorantes de seus desafios cotidianos. Para os professores neurodivergentes, essa dinâmica é amplificada pelo fato de que a profissão é culturalmente associada a um ideal de doação e de gestão das dificuldades alheias, o que torna ainda mais difícil revelar as próprias. O professor que revela dificuldades de regulação emocional, de organização ou de gestão sensorial arrisca ver sua competência profissional questionada precisamente nas dimensões que sua função exige que ele domine.

O *masking* tem sido sugerido como relacionado a vários resultados negativos para pessoas autistas, incluindo diagnóstico tardio ou ausente, problemas de saúde mental, *burnout* e suicidalidade. Pearson e Rose (2021) situam essa cadeia de consequências em sua dimensão social, argumentando que o *masking* não é, em essência, um problema individual, mas uma resposta a um ambiente que estigmatiza determinadas formas de existir. Quando a escola, como instituição, não reconhece a neurodivergência como uma condição legítima entre seus professores, ela produz as condições para que o *masking* seja a única estratégia disponível para esses docentes sobreviverem profissionalmente. E ao produzir essas condições, ela também produz, silenciosamente, o esgotamento, o adoecimento e a exclusão de profissionais que, em outros contextos, poderiam ser precisamente aqueles com maior capacidade de reconhecer e acolher a diferença nos alunos.

Docentes neurodivergentes existem. Estão nas escolas brasileiras. Mas são quase inteiramente invisíveis na literatura educacional, que os menciona, quando muito, como casos pontuais de superação individual, nunca como um grupo cujas condições de trabalho precisam ser sistematicamente examinadas e reorganizadas. A legislação trabalhista, nesse ponto, é mais avançada do que a prática institucional, pois a Lei Brasileira de Inclusão, em seu artigo 37, garante ao trabalhador com deficiência o direito à colocação competitiva com acesso a recursos de tecnologia assistiva e adaptação razoável no ambiente de trabalho (Brasil, 2015). A Lei de Cotas, por sua vez, a Lei n.º 8.213 de 1991, determina que empresas com mais de cem empregados reservem entre 2% e 5% de suas vagas para pessoas com deficiência. Essas disposições, em tese, cobrem os neurodivergentes que se enquadram nos critérios de

deficiência da LBI. Na prática, a ausência de dados sobre professores neurodivergentes impede até mesmo que se avalie o cumprimento dessas obrigações no setor educacional.

5.3 Um vazio que não é neutro

A ausência de pesquisas sobre o professor neurodivergente não é um acidente da produção acadêmica. É o resultado de escolhas epistemológicas que determinam quais sujeitos merecem ser investigados e quais devem permanecer na invisibilidade. A educação inclusiva, como campo de estudos, construiu-se sobre uma distinção implícita entre os que incluem e os que são incluídos, entre os que têm e os que não têm deficiência, entre os profissionais e os usuários dos sistemas de ensino. Essa distinção, que pode ter sido analiticamente útil num primeiro momento, tornou-se uma limitação quando se percebe que os próprios agentes do sistema educacional podem ser membros dos grupos que o sistema exclui.

O estudante neurodivergente que experimenta a opressão escolar descrita nas seções anteriores se tornará, alguns anos depois, um candidato ao magistério, à coordenação pedagógica, à gestão escolar. Se os sistemas de seleção e de trabalho docente forem tão pouco preparados para acolher a diversidade neurológica dos profissionais quanto são para acolher a dos alunos, o ciclo da exclusão se perpetuará. O professor neurodivergente que não encontra condições adequadas de trabalho abandona a docência, ou nunca a alcança, ou a exerce no regime de masking permanente que, como a literatura documenta, conduz inevitavelmente ao esgotamento. Em qualquer dessas hipóteses, perde-se um profissional que poderia ter, para os estudantes neurodivergentes sob sua responsabilidade, algo que nenhuma metodologia pode oferecer que é a experiência vivida e a compreensão empaticamente fundamentada de um funcionamento semelhante.

Esta seção propõe, portanto, que a agenda de pesquisa sobre neurodiversidade na educação brasileira incorpore, com urgência, o professor neurodivergente como objeto de investigação. Não apenas como agente de inclusão, mas como sujeito de direitos cujas condições de trabalho precisam ser examinadas, documentadas e, onde necessário, reformadas. A inclusão que se pretende construir não pode ignorar que as escolas são habitadas não apenas por alunos, mas por adultos profissionais cuja diversidade neurológica é parte da ecologia institucional, e que merece, tanto quanto a dos alunos, reconhecimento e proteção.

6 Conclusão: da inserção à transformação

Ao percorrer as seções deste capítulo, um argumento central foi se consolidando por acumulação de evidências e de análises. Esse argumento pode ser formulado de maneira simples, ainda que suas implicações sejam complexas: a inclusão genuína não é um estado que se alcança inserindo estudantes diferentes numa escola que permanece idêntica a si mesma. É um processo que exige a transformação da escola, de seus processos, de suas concepções e de suas relações, de modo que ela se torne capaz de acolher a variação humana em toda a sua extensão.

Demonstrou-se, anteriormente, que a história da educação inclusiva no Brasil foi marcada por um movimento que, embora tenha avançado significativamente em termos jurídicos, permaneceu incompleto em termos pedagógicos. A acessibilidade física foi priorizada sobre a acessibilidade pedagógica; a inserção foi confundida com inclusão; e o surgimento do conceito de neurodiversidade abriu uma fronteira que o paradigma dominante ainda não estava equipado para enfrentar.

Já este capítulo aprofundou essa análise, mostrando que o ambiente escolar, tal como está constituído, produz ativamente a invisibilização e a patologização do estudante neurodivergente, e que o ordenamento jurídico brasileiro, embora tenha avançado em instrumentos normativos importantes, deixou em desamparo um conjunto significativo de estudantes cujas condições não se encaixam nos critérios tradicionais de deficiência.

A passagem do modelo médico para o modelo social da deficiência, que a LBI consagrou juridicamente em 2015, é o eixo teórico em torno do qual esse percurso se organizou. No modelo médico, a dificuldade reside no sujeito e a solução é terapêutica com o intuito de corrigir, compensar, normalizar. No modelo social, a dificuldade reside na interação entre o sujeito e as barreiras do ambiente, e a solução é política e pedagógica, tratando-se de remover as barreiras, reorganizar os processos, transformar as práticas. Essa passagem não é apenas conceitual. Ela tem consequências concretas para o modo como as escolas organizam seu tempo, seu espaço, seus instrumentos de avaliação, sua formação docente e suas relações com os estudantes e suas famílias.

O que as seções precedentes tornaram evidente é que essa passagem está incompleta. A escola brasileira recebeu o texto do modelo social nas leis e nos decretos, mas não interiorizou suas implicações nas práticas cotidianas. Continua a perguntar o que há de errado com o estudante quando deveria perguntar o que há de inadequado em seus próprios processos. Continua a esperar que o aluno se adapte quando deveria adaptar seus instrumentos a ele. E continua a ignorar a existência de profissionais neurodivergentes em seus próprios

quadros, perpetuando uma lógica que separa, de modo arbitrário, os que incluem e os que são incluídos.

O horizonte propositivo que este capítulo aponta não é utópico; é exigente. Ele demanda que a política educacional complete o movimento que iniciou, estendendo as proteções normativas aos estudantes que hoje habitam o limbo entre os regimes jurídicos. Demanda que a formação docente incorpore a neurodiversidade como eixo estruturante, não como conteúdo periférico. Demanda que a articulação intersetorial entre saúde mental e educação deixe de ser princípio declarado para tornar-se prática sistemática. E demanda, sobretudo, que a pesquisa educacional amplie seu olhar para incluir os profissionais da educação entre os sujeitos cujas condições de diversidade merecem investigação e proteção.

A escola que emerge desse horizonte não é apenas uma escola que acolhe alunos diferentes. É uma escola que reconhece a variação humana como constitutiva de qualquer comunidade de aprendizagem, sejam estudantes, docentes ou demais profissionais da educação, e que organiza seus processos com base nesse reconhecimento. Não a escola que tolera a diferença; a escola que a necessita, que encontra nela sua própria riqueza e sua própria razão de existir. Os capítulos seguintes examinarão como esse horizonte pode ser traduzido em práticas pedagógicas concretas, situadas no cotidiano das salas de aula e nas culturas institucionais das escolas brasileiras.

REFERÊNCIAS

ARMSTRONG, Thomas. **The power of neurodiversity**: unleashing the advantages of your differently wired brain. Philadelphia: Da Capo Lifelong Books, 2011.

ATTWOOD, Tony. **The Complete Guide to Asperger's Syndrome**. London: Jessica Kingsley Publishers, 2007.

BENI, Priscila Ferreira; BAGATINI, Daniela Duarte da Silva; BARONE, Dante Augusto de Couto; CARDOSO, Denise Andrade; GOMES, Fabiana Baptista Maurer; AMARAL, Liz; LOPES, Maura Corcini; JOVER, Renato Schneider Rivero. **Neurodiversidade e os traços neurodivergentes não evidentes**: guias de boas práticas institucionais e um modelo para inclusão no ensino superior e no ambiente de trabalho. SciELO Preprints, São Paulo, 2 jun. 2025. Disponível em: https://doi.org/10.1590/scielo.preprints.beni2025. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 8.213, de 24 de julho de 1991**. Dispõe sobre os Planos de Benefícios da Previdência Social e dá outras providências. [Lei de Cotas]. Brasília, DF: Presidência da República, 1991. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8213cons.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996**. Estabelece as diretrizes e bases da educação nacional. Brasília, DF: Presidência da República, 1996. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 10.216, de 6 de abril de 2001**. Dispõe sobre a proteção e os direitos das pessoas portadoras de transtornos mentais e redireciona o modelo assistencial em saúde mental. Brasília, DF: Presidência da República, 2001. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l10216.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Brasília, DF: Presidência da República, 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 9 abr. 2026.

BRASIL. **Lei nº 14.254, de 30 de novembro de 2021**. Dispõe sobre o acompanhamento integral para educandos com dislexia ou Transtorno do Déficit de Atenção com Hiperatividade (TDAH) ou outro transtorno de aprendizagem. Brasília, DF: Presidência da República, 2021. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/l14254.htm. Acesso em: 9 abr. 2026.

BRASIL. **Decreto nº 12.686, de 20 de outubro de 2025**. Institui a Política Nacional de Educação Especial Inclusiva e a Rede Nacional de Educação Especial Inclusiva. Brasília, DF: Presidência da República, 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/d12686.htm. Acesso em: 9 abr. 2026.

BRASIL. **Decreto nº 12.773, de 8 de dezembro de 2025**. Altera o Decreto nº 12.686, de 20 de outubro de 2025. Brasília, DF: Presidência da República, 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/d12773.htm. Acesso em: 9 abr. 2026.

CANGUILHEM, Georges. **O normal e o patológico**. Tradução de Maria Thereza Redig de Carvalho Barrocas. 6. ed. Rio de Janeiro: Forense Universitária, 2009.
CONNOR, David J.; VALLE, Jan. Estudos da Deficiência na Educação: passado, presente e futuro? **Educação & Realidade**, Porto Alegre, v. 49, e141799, 2024. Disponível em: https://doi.org/10.1590/2175-6236141799vs01. Acesso em: 9 abr. 2026.

FRICKER, Miranda. **Epistemic injustice**: power and the ethics of knowing. Oxford: Oxford University Press, 2007.

GOFFMAN, Erving. **Stigma**: Notes on the Management of Spoiled Identity. London: Penguin Books, 1963.

HOMEWOOD HEALTH. **Neurodiversity in Focus**: The Energy Drain of Masking. Life Lines, mar. 2026. Disponível em: https://homeweb.ca. Acesso em: 11 abr. 2026.

HULL, Laura; PETRIDES, K. V.; ALLISON, Carrie; SMITH, Paula; BARON-COHEN, Simon; LAI, Meng-Chuan; MANDY, William. "Putting on my best normal": social camouflaging in adults with autism spectrum conditions. **Journal of Autism and Developmental Disorders**, v. 47, n. 8, p. 2519-2534, 2017. Disponível em: https://doi.org/10.1007/s10803-017-3166-5. Acesso em: 11 abr. 2026.

IBGE. **Pesquisa Nacional por Amostra de Domicílios Contínua** — Pessoas com Deficiência 2022. Rio de Janeiro: IBGE, 2023. Disponível em: https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/37317-pessoas-com-deficiencia-tem-menor-acesso-a-educacao-ao-trabalho-e-a-renda. Acesso em: 9 abr. 2026.

INSTITUTO INCLUSÃO BRASIL. **Neurodiversidade e inclusão no mercado de trabalho**. São Vicente: Instituto Inclusão Brasil, 2023. Disponível em: https://institutoinclusaobrasil.com.br. Acesso em: 11 abr. 2026.

METADADOS. **Neurodiversidade**: a importância da inclusão no ambiente de trabalho. *Blog Metadados*, 20 maio 2025. Disponível em: https://www.metadados.com.br. Acesso em: 11 abr. 2026.

MILTON, Damian. On the ontological status of autism: the 'double empathy problem'. **Disability & Society**, v. 27, n. 6, p. 883-887, 2012. Disponível em: https://doi.org/10.1080/09687599.2012.710008. Acesso em: 11 abr. 2026.

NEURODIVERSITY HUB. **Masking**. [s.l.: s.n., s.d.]. Disponível em: https://www.neurodiversityhub.com. Acesso em: 11 abr. 2026.

ORTEGA, Francisco. O sujeito cerebral e o movimento da neurodiversidade. **Mana**: Estudos de Antropologia Social, Rio de Janeiro, v. 14, n. 2, p. 477-509, 2008. Disponível em: https://doi.org/10.1590/S0104-93132008000200005. Acesso em: 9 abr. 2026.

PEARSON, Amy; ROSE, Kieran. A Conceptual Analysis of Autistic Masking: Understanding the Narrative of Stigma and the Illusion of Choice. Autism in Adulthood, v. 3, n. 1, p. 52-60, 2021. Disponível em: https://doi.org/10.1089/aut.2020.0043. Acesso em: 11 abr. 2026.

PRICE, Devon. **Unmasking autism**: discovering the new faces of neurodiversity. New York: Harmony Books, 2022.

PRYKE-HOBBES, Amber; DAVIES, Jade; HEASMAN, Brett; LIVESEY, Adam; WALKER, Amy; PELLICANO, Elizabeth; REMINGTON, Anna. The workplace masking experiences of autistic, non-autistic neurodivergent and neurotypical adults in the UK. **PLoS ONE**, v. 18, n. 9, e0290001, 2023. Disponível em: https://doi.org/10.1371/journal.pone.0290001. Acesso em: 11 abr. 2026.

RIBEIRO, Djamila. **O que é lugar de fala?** Belo Horizonte: Letramento, 2017. (Feminismos Plurais).

ROSA, Rodrigo Assunção; BUCCO, Eduardo Antony. Neurodiversidade: expandindo as fronteiras da diversidade, equidade e inclusão nas organizações. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 23, n. 1, e2023-0196, 2025. Disponível em: https://doi.org/10.1590/1679-395120230196. Acesso em: 9 abr. 2026.

SANTOS, Boaventura de Sousa. **Uma epistemologia do Sul**: a reinvenção do conhecimento e a emancipação social. São Paulo: Cortez; CLACSO, 2009.

SINGER, Judy. **NeuroDiversity**: the birth of an idea. Lexington: Amazon Digital Services, 2017.

UNESCO. **Declaração de Salamanca e linha de ação sobre necessidades educativas especiais**. Brasília, DF: CORDE, 1994. Disponível em: http://portal.mec.gov.br/seesp/arquivos/pdf/salamanca.pdf. Acesso em: 9 abr. 2026.

CAPÍTULO 8

USO DA INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA

USE OF ARTIFICIAL INTELLIGENCE AS AN INCLUSIVE EDUCATION TOOL

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1247



Resumo: O presente capítulo investiga as potencialidades da inteligência artificial (IA) como ferramenta de promoção da educação inclusiva, propondo um deslocamento das representações sociais que oscilam entre a visão da tecnologia como ameaça ou como panaceia. Através do conceito de "terceira via", fundamentado na teoria das representações sociais, argumenta-se que a IA deve ser compreendida como um instrumento de equidade. O texto desenvolve a tese da IA como uma "prótese comunicacional", uma tecnologia assistiva que não substitui a capacidade intelectual do estudante neurodivergente, mas compensa as barreiras funcionais entre a intenção e a expressão. São analisadas as convergências estruturais entre a arquitetura da IA generativa e as necessidades específicas de sujeitos com TDAH, dislexia, TDL e autismo. Por fim, o capítulo apresenta um inventário de ferramentas e discute evidências empíricas preliminares, defendendo que o uso ético e intencional da IA na escola representa uma escolha política pelo direito à plena expressão e à participação acadêmica.

Palavras-chave: Inteligência Artificial na Educação. Neurodiversidade. Tecnologia Assistiva. Inclusão Escolar. Prótese Comunicacional.

Abstract: This chapter investigates the potential of artificial intelligence (AI) as a tool for promoting inclusive education, proposing a shift in social representations that fluctuate between viewing technology as a threat or as a panacea. Through the concept of a "third way," grounded in the theory of social representations, it is argued that AI should be understood as an instrument of equity. The text develops the thesis of AI as a "communicational prosthesis," an assistive technology that does not replace the neurodivergent student's intellectual capacity but compensates for the functional barriers between intention and expression. It analyzes the structural convergences between the architecture of generative AI and the specific needs of individuals with ADHD, dyslexia, DLD, and autism. Finally, the chapter presents an inventory of tools and discusses preliminary empirical evidence, advocating that the ethical and intentional use of AI in schools represents a political choice for the right to full expression and academic participation.

Keywords: A rtificial Intelligence in Education. Neurodiversity. Assistive Technology. School Inclusion. Communicational Prosthesis

1 INTRODUÇÃO

Imagine uma professora do quinto ano que, ao preparar sua aula de português, decide experimentar um recurso de inteligência artificial para ajudar um aluno com dislexia a acompanhar a leitura coletiva do texto. No dia seguinte, ela descobre que outro estudante, sem qualquer dificuldade de aprendizagem diagnóstica, utilizou o mesmo recurso para escrever inteiramente a sua redação. Dois usos, dois efeitos, uma mesma ferramenta. Diante dessa cena, que poderá se tornar corriqueira nas escolas brasileiras contemporâneas, a questão que se impõe não é se a inteligência artificial representa uma ameaça ou uma solução para a educação, mas como compreender a complexidade das relações que ela estabelece com os sujeitos que habitam a escola.

Nos últimos anos, o debate sobre a presença da inteligência artificial nos ambientes educacionais tem se organizado, predominantemente, em torno de duas representações sociais que disputam o senso comum pedagógico. A primeira é a representação da ameaça. Nela, a IA é percebida como instrumento de cola, de trapaça, de substituição do esforço cognitivo e, em última instância, de esvaziamento do papel do professor. A segunda é a representação da panaceia, na qual a IA é concebida como tecnologia redentora, capaz de resolver as históricas deficiências do sistema educacional, personalizar o aprendizado de cada estudante e democratizar o acesso ao conhecimento de modo quase automático.

Costa Neto (2026), em pesquisa empírica sobre as representações sociais da mídia jornalística a respeito da integração da IA à educação, demonstra que essas duas imagens convivem no imaginário coletivo de forma paradoxal, frequentemente coexistindo no discurso de um mesmo veículo, de uma mesma instituição ou de uma mesma matéria. Esse achado, produzido a partir da análise de corpus jornalístico nacional, evidencia a tensão não resolvida que atravessa o debate público sobre a IA na escola brasileira.

Nenhuma dessas representações, tomada isoladamente, é capaz de dar conta da complexidade do fenômeno. A representação da ameaça tende a universalizar o uso mais raso e irresponsável da tecnologia, como se toda a inteligência artificial fosse, por definição, uma fraude em potencial. A representação da panaceia comete o erro oposto ao suprimir as dimensões éticas, políticas e pedagógicas do processo de inclusão, reduzindo-o a um problema de ordem técnica que bastaria uma ferramenta suficientemente sofisticada para resolver. Ambas, paradoxalmente, conduzem ao mesmo lugar que é o da inação. Quem vê apenas a ameaça recusa a ferramenta. Quem vê apenas a solução delega à máquina o que cabe ao coletivo humano decidir. Em ambos os casos, o estudante neurodivergente, sujeito central deste livro, permanece à margem da decisão.

Para compreender como essas representações se formam e se consolidam, é oportuno recorrer à Teoria das Representações Sociais de Serge Moscovici. Como nos mostra o autor (Moscovici, 2007, p. 54), a finalidade primária de toda representação é tornar familiar o que é estranho. Ao se deparar com um objeto novo e perturbador como a inteligência artificial, o pensamento coletivo mobiliza dois processos indissociáveis, quais sejam, a ancoragem e a objetivação.

A ancoragem, conforme explica Jodelet (1993, p. 17), enraíza a representação e seu objeto para lhes dar sentido e coerência em relação aos valores sociais já consolidados. Conforme demonstra Costa Neto (2026) em sua análise do *corpus* jornalístico brasileiro, a IA é ancorada na figura do professor substituto ou do copista eletrônico, categorias antigas

convocadas para domesticar um elemento recém-chegado. A objetivação, por sua vez, transforma o conceito abstrato em imagem concreta, como a do robô (tutor eletrônico) que ensina, o algoritmo que avalia, a plataforma que personaliza, produzindo uma representação que simplifica, mas que também distorce, ao apagar as tensões éticas e pedagógicas que a tecnologia carrega consigo.

No caso da educação inclusiva, essa distorção tem consequências particularmente graves. Ao ancorar a IA nos extremos, ora como vilã, ora como redentora, a sociedade perde de vista a sua potencialidade mais relevante para os sujeitos neurodivergentes que é a de funcionar como instrumento de equidade. Esta é a terceira via que este capítulo se propõe a explorar. Não se trata de uma posição ingênua ou conciliatória entre os dois extremos, mas de uma perspectiva que parte das condições concretas de existência dos estudantes com autismo, TDAH, dislexia, altas habilidades e outras formas de neurodivergência, para perguntar de que maneira a inteligência artificial pode, quando orientada por uma prática pedagógica crítica e inclusiva, contribuir para reduzir as barreiras que o ambiente escolar impõe a esses sujeitos.

Como bem argumenta Linhares (2025), a IA deve ser compreendida como apoio complementar, jamais como substituta da interação humana e do julgamento pedagógico. Essa premissa é o eixo sobre o qual este capítulo se estrutura. A tecnologia, por mais avançada que seja, precisa estar a serviço da humanização do ensino, conforme sintetiza Moura, Lima e Santos (2024, p. 6) ao afirmar que a mediação docente, quando apoiada por tecnologias inteligentes, deve preservar o caráter humano da educação, unindo sensibilidade e técnica para que o processo de aprendizagem se torne inclusivo e transformador. Essa articulação, entre o técnico e o humano, entre o algoritmo e a escuta, entre a plataforma e o vínculo, é o que distingue o uso instrumental da IA do seu uso verdadeiramente inclusivo.

Os capítulos anteriores desta obra percorreram um longo e necessário caminho. Discutimos os fundamentos sobre a inteligência artificial, o conceito de neurodiversidade, tal como cunhado por Judy Singer ao final do século XX, e revisamos o paradigma da inclusão no Brasil, com suas conquistas legislativas e suas lacunas persistentes. Analisamos o ambiente escolar como espaço frequentemente opressor para estudantes neurodivergentes, examinamos o fenômeno do mascaramento e suas consequências para a saúde mental dos sujeitos, e apontamos a invisibilidade do professor neurodivergente nos estudos educacionais. Cada um desses temas desenhou um panorama de urgências, de silêncios e de potencialidades não realizadas.

É nesse horizonte que a inteligência artificial emerge não como resposta pronta, mas como questão aberta. Que usos da IA produzem equidade e quais produzem novas formas de

exclusão? Como diferenciar o emprego genuinamente inclusivo das ferramentas digitais inteligentes de seu uso performático, que apenas simula modernidade sem alterar as relações de poder que estruturam a escola? De que modo o professor, ele mesmo frequentemente invisibilizado e precarizado, pode ser parceiro nessa transformação, e não vítima de um processo de automação que o esvazie de sentido? Estas são as questões que orientam a reflexão que se inicia aqui.

Para respondê-las, este capítulo está organizado em torno de seis eixos temáticos. Inicialmente, situa a IA em seu contexto histórico e conceitual, articulando as noções de tecnologia assistiva e inteligência artificial. Em seguida, examina as potencialidades das ferramentas de IA para estudantes neurodivergentes, discutindo personalização, acessibilidade e mediação pedagógica. O terceiro eixo aborda as dimensões éticas do uso de IA na educação inclusiva, com atenção especial aos riscos de vieses algorítmicos e vigilância. O quarto eixo trata da formação docente como condição de possibilidade para a inclusão mediada por tecnologia. O quinto examina experiências e iniciativas concretas, nacionais e internacionais, que iluminam o que já existe de promissor. E o sexto, por fim, retoma o problema das representações sociais para propor uma reorientação do olhar coletivo sobre a IA, da suspeição ou da euforia para a responsabilidade pedagógica.

A inteligência artificial, neste contexto, não é apenas uma ferramenta. Ela é também uma representação dos atores sociais, um campo de disputas sobre o que significa aprender, quem tem direito a uma educação de qualidade e que tipo de escola queremos construir. Essa compreensão que articula o fenômeno técnico ao fenômeno social é precisamente o que Costa Neto (2026) demonstra ao analisar como o discurso midiático constrói sentidos sobre a IA na educação que moldam o imaginário coletivo e influenciam políticas pedagógicas. Compreender a IA assim, como fenômeno técnico e social ao mesmo tempo, é o primeiro passo para que ela possa, de fato, servir àqueles que a escola historicamente tem colocado à margem.

2 CONVERGÊNCIAS ENTRE IA GENERATIVA E NECESSIDADES DE ALUNOS NEURODIVERSOS

Que tipo de tecnologia é, afinal, a inteligência artificial generativa? E o que ela tem a ver, de maneira estrutural e não apenas circunstancial, com as necessidades educacionais dos estudantes neurodivergentes? Essas perguntas exigem que nos aproximemos da natureza técnica da IA generativa com a mesma seriedade com que nos aproximamos das

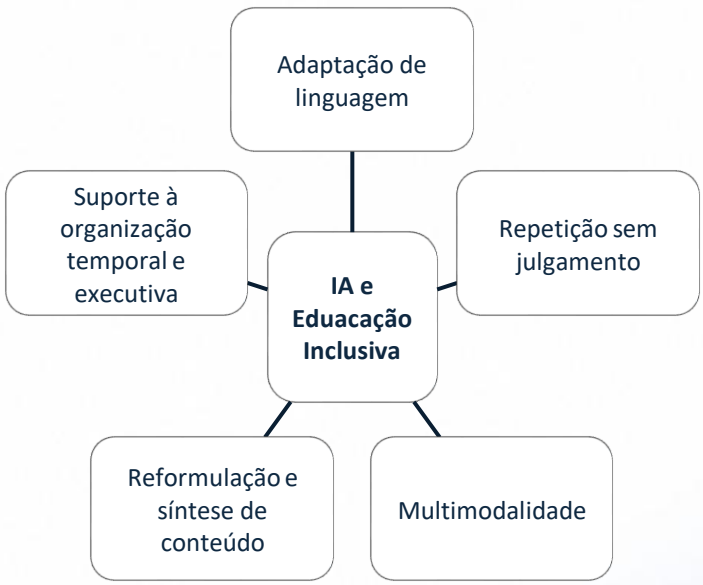
especificidades cognitivas de cada condição neurodivergente, pois é justamente nessa dupla aproximação que as convergências mais relevantes se revelam.

A inteligência artificial generativa, como vimos em capítulos anteriores, é uma categoria de sistema computacional capaz de produzir, a partir de grandes volumes de dados de treinamento, conteúdos novos e contextualmente coerentes. Textos, imagens, áudios, resumos e traduções são gerados por modelos que aprenderam padrões estatísticos a partir de corpus extensos da produção humana.

Diferentemente dos sistemas de IA baseados em regras fixas, os modelos generativos operam com flexibilidade, adaptando respostas ao contexto imediato da interação. Costa Neto (2026) observa que essa flexibilidade é precisamente o que torna a IA generativa um objeto socialmente difícil de estabilizar, já que, por ser ao mesmo tempo genérica e personalizável, ela desafia as categorias disponíveis no senso comum e exige dos processos de ancoragem um esforço redobrado de classificação. Essa instabilidade representacional, porém, é também o que torna a IA generativa tão relevante para contextos de alta heterogeneidade cognitiva, como é o caso das salas de aula inclusivas.

As capacidades estruturais dos modelos de inteligência artificial generativa apresentam pontos de contato diretos com as necessidades recorrentes de estudantes neurodivergentes. Essa interseção não é meramente acidental, mas decorre da flexibilidade da ferramenta em processar e reformatar informações de maneira personalizada. Essas convergências organizam-se em torno de cinco dimensões fundamentais, detalhadas na Figura 1, as quais encontram ressonância em perfis cognitivos específicos e fundamentam a potencialidade inclusiva da tecnologia em contextos pedagógicos orientados.

Figura 1. Possibilidades de uso de IA na Educação Inclusiva.



Fonte: produzido pelo autor.

A primeira e talvez mais imediata convergência diz respeito à adaptação de linguagem. Um modelo generativo pode reformular um enunciado complexo em linguagem mais simples, transformar um texto escrito em narrativa oral e vice-versa, converter uma instrução abstrata em passos concretos e sequenciados, ou ainda apresentar o mesmo conteúdo em níveis distintos de complexidade sem que o estudante precise solicitar auxílio diretamente ao professor, o que, para muitos alunos neurodivergentes, representa uma barreira em si.

Para estudantes com dislexia, essa capacidade de reformulação é especialmente relevante. Conforme aponta Nascimento (2025), a aprendizagem de alunos com dislexia e TDAH beneficia-se significativamente de recursos que transformem o modo de apresentação dos conteúdos, reduzindo a carga cognitiva imposta pela decodificação da escrita convencional. A IA generativa, ao oferecer versões reformuladas, simplificadas ou multimodais de um mesmo texto, atua exatamente nessa direção.

A segunda convergência estrutural refere-se à ausência de julgamento nas interações com a IA. Esse aspecto, frequentemente subestimado nos debates tecnológicos, é de grande relevância pedagógica para estudantes que carregam histórias de fracasso escolar, de vergonha e de exposição constante às correções públicas em sala de aula. Para alunos com fobia social ou com transtorno de ansiedade generalizada, condições frequentemente comórbidas com outras formas de neurodivergência, a possibilidade de tirar dúvidas, errar, reformular e tentar novamente sem o peso do olhar alheio representa uma abertura genuína.

O sistema de IA não suspira, não demonstra impaciência, não registra o erro no livro de chamada e não o comenta com outros estudantes. Como observa Santos (2024, p. 2), "a inteligência artificial representa uma nova forma de interação com o conhecimento, permitindo que pessoas com TDAH encontrem caminhos personalizados de aprendizagem por meio de sistemas inteligentes que se ajustam às suas demandas cognitivas". Essa ausência de julgamento facilita a retomada, o erro como parte do processo e a persistência sem punição.

A terceira convergência relaciona-se com a multimodalidade[8]. Estudantes do espectro autista frequentemente apresentam perfis sensoriais e cognitivos que privilegiam determinados canais de processamento em detrimento de outros. Para alguns, textos densos são incompreensíveis, mas diagramas visuais ou narrações em áudio podem ser altamente eficazes.

[8] Utilização de dois ou mais sistemas de signos, como texto, imagem, som e recursos espaciais, para a construção de sentidos em um mesmo evento comunicativo.

Para outros, a construção de mapas conceituais ou a organização visual da informação facilita a compreensão que a linearidade escrita não consegue promover.

A IA generativa, integrada a interfaces multimodais, pode gerar representações diversas de um mesmo conteúdo, adaptando o formato à demanda do usuário. Narciso *et al.* (2024, p. 406) apontam que "o objetivo principal é explorar como o uso de tecnologias digitais pode influenciar positivamente a comunicação, a socialização e o desenvolvimento cognitivo de indivíduos autistas", e os recursos de adaptação multimodal aparecem como um dos mecanismos mais promissores nesse campo.

A quarta convergência, não menos importante, é a capacidade de suporte às funções executivas. Estudantes com TDAH enfrentam dificuldades recorrentes de planejamento, priorização, início de tarefas, gestão do tempo e regulação da atenção. Essas dificuldades não derivam de falta de capacidade intelectual, mas de variações no funcionamento dos circuitos pré-frontais que regulam o comportamento dirigido a objetivos.

A IA generativa pode funcionar como andaime cognitivo externo ao organizar a tarefa em etapas menores, sugerir sequências de ação, oferecer lembretes e sintetizar o que já foi feito e o que ainda está pendente. Conforme observam Alves, Lima e Alves (2023, p. 400), "a utilização de jogos educativos digitais para crianças com TDAH estimula a atenção contínua e melhora a capacidade de concentração, contribuindo para o fortalecimento da memória de curto prazo". Essa lógica de suporte pode ser ampliada para o uso de ferramentas generativas que constroem, junto com o estudante, o itinerário da tarefa, reduzindo o impacto das dificuldades executivas.

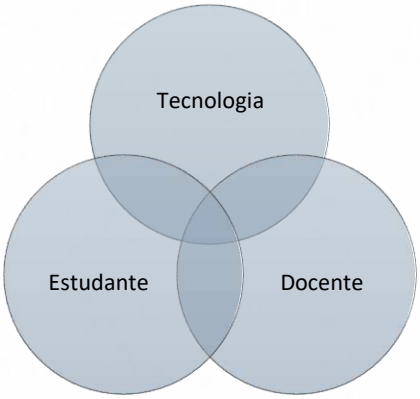
A quinta convergência que merece atenção é aquela que envolve estudantes com transtorno bipolar ou com outras condições que geram alta variabilidade no desempenho ao longo do tempo. Para esses sujeitos, o modelo escolar que exige ritmo constante, produção homogênea e avaliações em datas fixas é particularmente hostil. Um dia de hiperfoco pode ser seguido de dias de profunda dificuldade de concentração, sem que isso indique falta de comprometimento ou de capacidade. A IA generativa, ao ser disponibilizada de modo assíncrono e flexível, permite que o estudante retome o conteúdo no momento em que sua capacidade de processamento é maior, sem depender de condições presenciais e temporalmente rígidas. A personalização do ritmo, nesse caso, não é apenas uma conveniência pedagógica, mas uma condição de acesso ao conhecimento.

É importante sublinhar, como fazem Santos *et al.* (2025) e Linhares (2025), que essas convergências não operam de forma automática ou garantida. A flexibilidade da IA encontra a variabilidade da cognição neurodivergente, mas esse encontro só se torna pedagogicamente

produtivo quando mediado por um professor que conhece os sujeitos, que escolhe as ferramentas com intencionalidade e que acompanha o processo com sensibilidade.

A IA, por si só, não sabe que determinado aluno tem dislexia severa e que a reformulação em texto ainda é insuficiente para seu acesso. Não sabe que outro estudante está em um episódio depressivo e que qualquer tarefa nova será esmagadora. Não sabe que uma criança autista específica tem aversão a determinados estímulos sonoros produzidos pela síntese de voz. Quem sabe, ou pode saber, é o professor, e por isso a relação entre IA e inclusão é sempre triangular, visto que tecnologia, docente e estudante compõem um sistema pedagógico integrado, e o enfraquecimento de qualquer um desses vértices compromete o conjunto.

Figura 2 - Relação de mediação entre a IA e os sujeitos da educação



Fonte: produzido pelo autor.

Compreender essas convergências como estruturais significa reconhecer que a IA generativa não foi desenhada para estudantes neurodivergentes, mas que suas características fundamentais tais como flexibilidade, adaptabilidade, assincronicidade, ausência de julgamento, multimodalidade e capacidade de reformulação respondem, de modo não acidental, a necessidades que a escola convencional historicamente ignorou. Como ressaltam Santos *et al.* (2025, p. 4120), "a aprendizagem desses alunos demanda atenção diferenciada, considerando aspectos cognitivos, emocionais e comportamentais", e a IA generativa, quando bem orientada, tem condições de oferecer exatamente essa atenção diferenciada em escala, sem substituir a escola, mas ampliando as suas possibilidades inclusivas.

A questão que se abre, portanto, não é se a IA generativa pode contribuir para a educação de estudantes neurodivergentes, já que as convergências estruturais indicam que pode, mas em que condições, com quais salvaguardas éticas e sob qual orientação pedagógica esse potencial se realiza. É para esse horizonte que as próximas seções deste capítulo se dirigem.

3 A IA COMO PRÓTESE COMUNICACIONAL

Há uma pergunta que a escola raramente formula com clareza, mas que atravessa silenciosamente a trajetória de cada estudante neurodivergente: o que acontece quando o pensamento existe, mas não encontra passagem? Quando a ideia está lá, viva e organizada na mente do aluno, porém o canal que a escola exige para que ela chegue ao papel tal como a ortografia convencional, a linearidade sintática, o registro acadêmico, funciona como uma barreira intransponível? Esse é o problema que esta seção se propõe a nomear e a examinar, e é aqui que a metáfora da prótese comunicacional encontra sua justificativa mais precisa.

Comecemos por um exemplo simples, mas revelador. Os óculos não pensam. Não substituem o olhar, não interpretam o que é visto, não decidem o que merece atenção. Eles apenas removem uma barreira física entre a intenção de ver e o mundo que se oferece para ser visto. Quem usa óculos não se torna mais ou menos inteligente por causa deles. Torna-se, simplesmente, capaz de exercer uma capacidade que já possuía, mas que estava obstruída por uma limitação corporal. O mesmo princípio pode ser aplicado a uma série de instrumentos que a humanidade desenvolveu para ampliar a expressão humana tais como o teclado para quem não escreve à mão, o sintetizador de voz para quem não articula sons, a língua de sinais para quem não ouve, a bengala para quem não equilibra o passo. Todos esses instrumentos são próteses, não no sentido depreciativo do termo, mas no sentido rigoroso de dispositivos que preenchem a lacuna entre a intenção do sujeito e o mundo que a recebe.

É nesse sentido que propomos compreender a inteligência artificial como prótese comunicacional para estudantes neurodivergentes. A IA, quando utilizada com essa intencionalidade pedagógica, não substitui o pensamento do estudante. Não pensa por ele, não aprende em seu lugar, não lhe concede um saber que não possui. O que ela pode fazer e que a torna pedagogicamente relevante no contexto da inclusão é remover a barreira entre o pensamento que existe e a forma que a escola exige para reconhecê-lo. Conforme aponta Linhares (2025), as tecnologias assistivas baseadas em IA podem ampliar a autonomia dos estudantes com deficiência, não por suprimir a diferença cognitiva, mas por oferecer canais alternativos de expressão e acesso ao conhecimento.

Esse argumento encontra ressonância profunda no modelo social da deficiência, discutido com maior extensão nos capítulos anteriores desta obra. Ao deslocar o problema da deficiência do sujeito para o ambiente, o modelo social propõe que o que incapacita não é a condição em si, mas a ausência de condições que permitam ao sujeito funcionar. Conforme

analisam Lustosa e Figueiredo (2021, *apud* Andréia *et al.*, 2025), na perspectiva da inclusão, o que realmente dificulta o processo de aprendizagem não é a deficiência em si, mas os obstáculos existentes no ambiente educacional que impedem o pleno desenvolvimento dos estudantes. Transpondo essa lógica para o campo comunicacional, o problema do estudante neurodivergente não é que ele não tenha o que dizer. O problema é que o canal que a escola oferece e frequentemente exige como único não é o canal pelo qual o seu pensamento trafega com fluência.

Para tornar esse argumento concreto, considere três casos que habitam, de forma numerosa e invisível, as salas de aula brasileiras. O primeiro é o aluno com disortografia que tem a ideia, mas trava na ortografia. Esse estudante pode ter lido o texto com atenção, compreendido suas camadas de sentido, articulado internamente uma resposta sofisticada e original. Mas, quando sua mão toca o papel, a grafia vacila, as letras se invertem, a palavra certa foge e a palavra errada ocupa seu lugar. A nota que recebe não reflete o que sabe. Ao contrário, reflete o atrito entre o seu pensamento e o canal que a escola privilegia.

Softwares baseados em reconhecimento de voz, ditado inteligente ou ferramentas de revisão ortográfica mediadas por IA podem, nesse caso, funcionar como óculos para a escrita. Conforme observa Linhares (2025), algoritmos de IA podem auxiliar no reconhecimento de padrões e na reorganização de estímulos textuais para estudantes com transtornos específicos de aprendizagem como a disgrafia, dislexia ou disortografia, ampliando o acesso ao currículo sem suprimir a diferença cognitiva.

O segundo caso é o aluno com TDAH que organiza mentalmente, mas não consegue linearizar no papel. Esse estudante pode ter percorrido mentalmente toda a estrutura de um texto, identificado os argumentos principais, antecipado as objeções e previsto a conclusão. Mas, quando é chegada a hora de colocar essa estrutura na página, a atenção já migrou, a sequência se perdeu, e o que emerge é um conjunto de fragmentos que não comunica a arquitetura que existia na mente.

A escola avalia o produto escrito e conclui, equivocadamente, que o aluno não sabe. A IA generativa pode ser, nesse caso, o andaime que sustenta a externalização do que já existe internamente, pois pode receber do aluno os fragmentos, devolvê-los organizados, perguntar se a sequência faz sentido, sugerir conexões. Não é a IA que pensa a estrutura. Não! É o aluno que a confirma, ajusta e aprova. Como nos lembra Vygotsky (1997, p. 68), o uso de instrumentos e signos transforma a estrutura psicológica das funções mentais humanas, reorganizando a maneira como o indivíduo aprende, pensa e interage com o mundo, e é

exatamente nessa perspectiva que a IA pode operar como ferramenta mediadora do desenvolvimento.

O terceiro caso é o aluno com autismo que compreende profundamente, mas não encontra o registro acadêmico esperado. Esse estudante pode ter uma capacidade analítica extraordinária, dominando os conteúdos com precisão e originalidade. Mas sua forma de expressar esse domínio frequentemente diverge dos padrões que a escrita acadêmica exige. Ele pode ser excessivamente técnico onde se espera narrativa, excessivamente literal onde se espera metáfora, excessivamente direto onde se espera contextualização.

A escola, ao avaliar apenas o produto, tende a subnotar o que foi compreendido porque não reconhece a forma em que essa compreensão chegou. A IA pode funcionar, aqui, como mediador de registro. O aluno apresenta sua análise no formato em que a produz, e a IA auxilia na transposição para o registro esperado, sem alterar o conteúdo, apenas oferecendo o canal. Essa função é precisamente o que Narciso *et al.* (2024, p. 406) descrevem ao afirmarem que as tecnologias digitais podem influenciar positivamente a comunicação e o desenvolvimento cognitivo de indivíduos autistas, não para torná-los idênticos aos colegas neurotípicos, mas para ampliar os canais pelos quais sua expressão chega ao mundo.

Nos três casos, o que está em jogo não é a ausência de conteúdo, de compreensão ou de inteligência. O que está em jogo é a desigualdade comunicacional, ou seja, a distância entre o que o sujeito sabe e o que os canais disponíveis permitem que ele expresse. A escola convencional, ao privilegiar um único canal único tal como a escrita ortograficamente correta, linear, no registro acadêmico padrão como legítimo, produz uma forma de exclusão que, magicamente ou misteriosamente, não é visível como exclusão, porque aparece disfarçada de avaliação.

O estudante que não atinge os critérios não é reconhecido como alguém cujo canal de comunicação foi obstruído, mas como alguém que simplesmente não sabe ou não aprende. Essa é uma das formas mais perversas de barreira pedagógica, porque não se apresenta como barreira, mas como critério. Costa Neto (2026), ao analisar como a mídia representa a IA na educação, demonstra que essa mesma lógica opera na esfera pública, visto que o discurso que naturaliza a tecnologia como solução esconde as disputas sobre que tipo de aprendizagem, para quem e por quais canais a escola deve reconhecer.

É aqui que o conceito de acessibilidade pedagógica, discutido no Capítulo 6 desta obra, adquire sua dimensão mais radical. A acessibilidade pedagógica não se limita a rampas, fontes ampliadas ou tempo estendido para as provas. Ela implica questionar os próprios canais pelos quais o conhecimento é expresso e avaliado. Implica reconhecer que a escrita convencional, a

linearidade discursiva e o registro acadêmico são convenções culturais e históricas, não evidências naturais de que alguém sabe ou não sabe.

Segundo Bersch (2017, *apud* Linhares, 2025), a tecnologia assistiva é um conjunto de recursos e serviços que visam proporcionar ou ampliar habilidades funcionais de pessoas com deficiência, promovendo autonomia e inclusão. A prótese comunicacional é, portanto, uma forma de tecnologia assistiva que não trata o déficit, mas equipa o sujeito para se mover em um ambiente que, sem ela, seria inacessível.

A metáfora da prótese tem, naturalmente, seus limites, e é importante reconhecê-los. Uma prótese mal ajustada pode machucar mais do que ajudar. Uma prótese oferecida sem formação pode gerar dependência ao invés de autonomia. Uma prótese imposta sem que o sujeito a queira pode ser invasiva e redutora. Os mesmos riscos se aplicam ao uso da IA como prótese comunicacional.

O emprego irrefletido da ferramenta, sem acompanhamento pedagógico, sem atenção às especificidades do sujeito e sem um projeto de progressiva autonomia, pode gerar o efeito inverso, ou seja, ao invés de remover a barreira, pode simplesmente delegar ao algoritmo o que deveria continuar sendo do estudante. Como alertam Dwork e Mulligan (2013, *apud* Linhares, 2025), os sistemas automatizados tendem a reproduzir padrões discriminatórios existentes nos dados que os alimentam, e essa advertência se aplica também ao risco de que a IA, ao traduzir a expressão neurodivergente para o registro normativo, apague justamente o que havia de singular e valioso nessa expressão.

O ponto central, portanto, não é que a IA seja sempre uma prótese comunicacional eficaz, mas que ela pode sê-lo, quando orientada por uma intenção pedagógica clara e por uma escuta genuína das necessidades do sujeito. Essa possibilidade, por si só, já representa um deslocamento significativo em relação à lógica escolar que, ao longo de décadas, tratou a diferença comunicacional como déficit a ser corrigido, e não como especificidade a ser acolhida. Reconhecer a IA como instrumento capaz de remover barreiras sem apagar diferenças é reconhecer, ao mesmo tempo, que o que falta nas trajetórias de muitos estudantes neurodivergentes não é capacidade. É a possibilidade de completa expressão.

4 FERRAMENTAS DE IA EXISTENTES E CASOS DE USO DOCUMENTADOS NA INCLUSÃO

O argumento conceitual da prótese comunicacional, desenvolvido na seção anterior, encontra sua sustentação mais concreta na existência de um ecossistema crescente de

ferramentas baseadas em inteligência artificial que já estão sendo usadas, com maior ou menor sistematização, por estudantes neurodivergentes ao redor do mundo. Falar dessas ferramentas sem cair no catálogo acrítico exige um cuidado metodológico importante: o que se busca aqui não é simplesmente listar aplicativos e seus recursos, mas compreender de que maneira cada tipo de ferramenta responde a um tipo específico de barreira comunicacional, cognitiva, sensorial ou emocional. Organizada em quatro subseções, esta análise percorre o espectro amplo da neurodivergência tal como adotado neste livro, da dislexia ao transtorno bipolar, do autismo ao transtorno de processamento sensorial, da ansiedade severa ao transtorno obsessivo-compulsivo, e procura mostrar que as convergências entre as capacidades dessas ferramentas e as necessidades dos sujeitos não são circunstanciais, mas estruturais.

4.1 Ferramentas de linguagem e comunicação

A mais imediata das barreiras que os estudantes neurodivergentes enfrentam no ambiente escolar é, com frequência, a barreira da linguagem, entendida aqui não como vocabulário ou gramática, mas como o conjunto de operações que a escola exige para que o pensamento chegue à forma reconhecida como legítima. Escrever corretamente, ler com fluência, formular perguntas no registro adequado, interpretar enunciados ambíguos, identificar a informação relevante em um texto denso, cada uma dessas operações constitui um ponto de atrito específico para sujeitos com dislexia, disortografia, disgrafia, transtorno do desenvolvimento da linguagem (TDL) ou autismo. As ferramentas de IA focadas em linguagem e comunicação incidem exatamente sobre esses atritos.

Entre as mais documentadas está a categoria dos leitores de tela com IA, representados por *softwares* como o Speechify, que converte textos escritos em áudio natural, e o NaturalReader, que oferece funcionalidade semelhante com ajuste de velocidade e destacamento sincrônico das palavras. Para estudantes com dislexia, a separação entre o ato de decodificar e o ato de compreender é fundamental, dado que quando a decodificação é mediada pela voz sintetizada, a atenção pode se dedicar inteiramente ao sentido.

Conforme observam Costa Júnior *et al.* (2024, *apud* Andréia *et al.*, 2025), os leitores de tela com IA contribuem para o aumento no engajamento e a melhoria do desempenho acadêmico de estudantes com dificuldades visuais e cognitivas, e o mesmo princípio se aplica a estudantes com dislexia severa, para quem a leitura visual é um processo exaustivo e muitas vezes doloroso.

Na direção inversa, os sistemas de reconhecimento de voz e ditado inteligente funcionam como escrita mediada pela fala. O Dragon NaturallySpeaking é uma das plataformas mais utilizadas nesse campo, com capacidade de transcrever com alta precisão o que o usuário fala, permitindo que estudantes com disgrafia, com dificuldades motoras finas ou com dislexia severa produzam textos sem que a ortografia e a caligrafia se tornem obstáculos.

O Co:Writer, desenvolvido especificamente como tecnologia assistiva, combina ditado com sugestão preditiva de palavras e frases, e acrescenta feedback sonoro à medida que o usuário escreve, o que é especialmente útil para estudantes com disortografia que precisam ouvir o que estão compondo para avaliar se soa correto antes de ver o resultado na tela.

Para além do ditado simples, modelos de linguagem como o ChatGPT e o Gemini têm sido utilizados por estudantes e professores para reformular enunciados complexos em linguagem mais acessível, traduzir instruções ambíguas em passos concretos e oferecer explicações em múltiplos níveis de complexidade, o que atende diretamente ao princípio do Desenho Universal para Aprendizagem (DUA), que prevê múltiplos meios de representação do conhecimento.

Para estudantes com autismo, as ferramentas de comunicação assumem uma dimensão adicional de fazer a mediação do registro social. Aplicativos como o proloquo2go e outros baseados em Comunicação Aumentativa e Alternativa (CAA) há muito servem a crianças autistas não verbais ou minimamente verbais, mas a incorporação de IA generativa nesse campo está expandindo as possibilidades.

O Sachs Center, por exemplo, lançou em 2024 uma ferramenta gratuita que auxilia adultos no espectro autista a interpretar expressões idiomáticas, metáforas e linguagem indireta, respondendo em tempo real com explicações sobre o sentido convencional de frases que frequentemente confundem usuários autistas que processam a linguagem de forma mais literal.

Em contextos escolares, estudantes com transtorno do desenvolvimento da linguagem (TDL) se beneficiam de ferramentas que transformam textos orais em transcrições organizadas, como o Otter.ai, que grava e transcreve aulas com marcação de tempo, permitindo que o aluno revise o conteúdo no ritmo que precisa, identifique os pontos que não compreendeu e retome sem depender da memória auditiva imediata. Como aponta a literatura revisada por Linhares (2025), *softwares* de leitura automática, reconhecimento de voz e tradutores têm potencializado a acessibilidade curricular para estudantes com transtornos

específicos de aprendizagem, ampliando o conjunto de canais pelos quais o conhecimento pode ser recebido e expresso.

Estudantes com sinestesia, condição em que os sentidos se cruzam de formas não convencionais, frequentemente relatam que a forma escrita padrão dos textos escolares entra em conflito com o modo como processam as informações. Ferramentas que permitem alterar cores de fundo, tipografias, espaçamentos e contraste, como o Read&Write for Google, que integra leitura em voz alta com ajuste tipográfico e ferramentas de grifamento visual, podem criar interfaces que reduzem o ruído sensorial da leitura e tornam o processo menos perturbador.

O mesmo se aplica a estudantes com transtorno do processamento sensorial (TPS), para quem a leitura de textos densos pode desencadear sobrecarga. Recursos de simplificação textual mediados por IA, que condensam textos longos em versões mais curtas e estruturalmente mais claras, podem ser a diferença entre um conteúdo acessível e um inacessível.

4.2 Ferramentas de organização, planejamento e regulação

Se a seção anterior tratou das barreiras da expressão e da recepção, esta se volta às barreiras da organização, que afetam de modo especialmente incisivo os estudantes com TDAH, com transtorno de desenvolvimento da coordenação (TDC) em sua dimensão executiva, e com transtorno afetivo bipolar. Para esses sujeitos, o problema não é necessariamente o conteúdo, tampouco o canal comunicacional, mas a capacidade de iniciar, sequenciar, priorizar e sustentar a execução de uma tarefa até sua conclusão. A escola, ao exigir autonomia executiva sem oferecê-la, trata como falha moral o que é, na verdade, uma diferença neurológica no funcionamento dos circuitos pré-frontais.

O Goblin Tools é uma das ferramentas mais citadas por comunidades neurodivergentes *online* como transformador de suas rotinas escolares. Desenvolvido especificamente para pessoas com dificuldades de funcionamento executivo, o aplicativo oferece um conjunto de recursos que incluem o Magic ToDo, que decompõe uma tarefa complexa em etapas menores e ajusta o nível de granularidade conforme a necessidade do usuário, e o Compiler, que transforma um fluxo desordenado de pensamentos em uma lista estruturada de ações. A ferramenta também inclui o Judge, que analisa o tom emocional de mensagens recebidas, recurso especialmente útil para estudantes com sensibilidade à rejeição, condição frequente em pessoas com TDAH, que podem interpretar e-mails neutros como

agressivos. Como descreve ADDitude Magazine (2026), ao cobrir ferramentas de IA para estudantes neurodivergentes, o Goblin Tools pode apoiar o funcionamento executivo, a escrita e a comunicação de modo acessível e não punitivo.

Para o planejamento do tempo, ferramentas como o Motion e o MyStudyLife utilizam IA para organizar automaticamente a agenda do estudante, considerando prazos, prioridades e padrões de energia ao longo do dia. Para um estudante com TDAH que tem dificuldade crônica para iniciar tarefas e estimar o tempo necessário para completá-las, a externalização dessa gestão para um sistema inteligente não representa uma abdição de autonomia, mas uma estratégia de andaime cognitivo que permite que a energia mental disponível seja investida no aprendizado efetivo, e não no esforço estéril de gerir a própria agenda.

Como demonstram Alves, Lima e Alves (2023, p. 400), tecnologias digitais que apoiam a atenção contínua e a memória de trabalho contribuem para o desenvolvimento das funções executivas em estudantes com TDAH, e as ferramentas de organização baseadas em IA estendem esse princípio para além do conteúdo curricular, alcançando a gestão do próprio processo de aprendizagem.

Para estudantes com transtorno afetivo bipolar, a variabilidade do desempenho ao longo do tempo é uma das características mais difíceis de conciliar com a rigidez dos calendários escolares. A capacidade de acessar o conteúdo de forma assíncrona, de retomar tarefas em momentos de maior estabilidade e de usar a IA como escada organizacional nos dias de menor capacidade executiva representa um suporte que os sistemas convencionais raramente oferecem.

Ferramentas de *mind mapping* com IA, como o GitMind, que organiza fluxos de pensamento e cria mapas conceituais a partir de insumos livres, permitem que o estudante capte ideias no momento em que surgem, mesmo que fragmentadas, e organize-as posteriormente em estruturas coerentes, sem que a descontinuidade temporal do funcionamento bipolar resulte em perda irreparável de progresso.

Estudantes com transtorno obsessivo-compulsivo (TOC) frequentemente enfrentam dificuldades específicas com a entrega de trabalhos, já que o ciclo de revisão compulsiva, a incapacidade de reconhecer a tarefa como concluída e a ansiedade associada ao julgamento externo podem tornar cada atividade escolar uma fonte de sofrimento desproporcional ao seu conteúdo. Ferramentas de IA que definem um número máximo de revisões, que confirmam quando uma tarefa atinge os critérios estabelecidos ou que oferecem feedback estruturado e limitado em vez de aberto e ilimitado podem funcionar como contenção externa ao ciclo

obsessivo, não substituindo o tratamento terapêutico, mas reduzindo o atrito escola-TOC em contextos específicos.

4.3 Ferramentas de acessibilidade sensorial e motora com suporte de IA

A neurodivergência não é apenas cognitiva. Muitas das condições discutidas neste livro envolvem diferenças sensoriais e motoras que tornam o ambiente escolar fisicamente hostil muito antes de qualquer questão de aprendizagem propriamente dita. O transtorno do processamento sensorial (TPS), frequentemente associado ao autismo, mas também presente em sujeitos com TDC, TDAH severo e Síndrome de Tourette, pode fazer com que a sala de aula convencional, com seus sons sobrepostos, suas luzes fluorescentes, seus aromas e suas demandas posturais, seja um ambiente de sobrecarga permanente.

Para sujeitos com epilepsia, certas características visuais dos ambientes digitais, como flashes e padrões intermitentes, podem representar riscos concretos. As ferramentas de IA voltadas para acessibilidade sensorial e motora buscam remodelar os ambientes de interação para que esses sujeitos possam participar com menor custo energético e com maior segurança.

O reconhecimento de voz, já mencionado na subseção anterior, tem um papel central também aqui, mas em uma dimensão diferente. Trata-se não apenas de um canal alternativo de escrita, mas de interface de controle do ambiente digital. Para estudantes com dispraxia (TDC) que têm dificuldade com o uso preciso de mouse e teclado, a possibilidade de navegar por sistemas educacionais usando comandos de voz representa uma mudança de paradigma na acessibilidade motora.

Da mesma forma, estudantes com Síndrome de Tourette que apresentam tiques motores envolvendo as mãos podem beneficiar-se de interfaces por voz ou por gestos que não dependem do controle motor fino. Ferramentas como o Dragon NaturallySpeaking e os recursos de controle por voz integrados aos sistemas operacionais Windows, Mac e Chrome OS representam o estado atual mais desenvolvido desse campo.

A legendagem em tempo real é outra categoria de ferramentas com impacto documentado. Aplicativos como o Google Live Transcribe e o Otter.ai foram originalmente desenvolvidos para usuários surdos ou com deficiência auditiva, mas seu uso por estudantes com transtorno do processamento auditivo, frequentemente comórbido com outros perfis neurodivergentes, tem se revelado igualmente transformador.

Para esses estudantes, a sala de aula convencional, onde o professor fala sem pausa e o aluno precisa registrar simultaneamente, é um ambiente particularmente hostil. A transcrição

automática em tempo real oferece uma segunda via de acesso à informação oral, desacoplando a recepção da anotação e aliviando uma sobrecarga que para muitos é invisível, mas constante.

Para estudantes com afantasia, condição em que a capacidade de visualizar mentalmente imagens está ausente ou significativamente reduzida, os recursos de descrição automática de imagens mediados por IA representam uma prótese pedagógica relevante. Ferramentas como o Microsoft Seeing AI e o Be My Eyes, originalmente desenvolvidas para usuários com deficiência visual, descrevem imagens, gráficos, diagramas e fotografias em linguagem verbal, criando um acesso textual ao conteúdo visual que, para pessoas com afantasia, pode ser pedagogicamente mais eficaz do que a imagem em si. Esse é um exemplo particularmente interessante de como ferramentas desenvolvidas para uma condição específica migram, de modo orgânico, para o uso por sujeitos com perfis cognitivos distintos, mas com necessidades de acessibilidade relacionadas.

Para estudantes com epilepsia, as plataformas educacionais digitais representam um campo de risco se não forem projetadas com consciência das fotossensibilidades que algumas formas de epilepsia envolvem. A IA pode ser empregada aqui de forma preventiva, ao analisar o conteúdo audiovisual e identificar e filtrar elementos potencialmente gatilhantes antes que o estudante os visualize.

Embora essa aplicação específica ainda esteja em estágio de desenvolvimento e não seja generalizada nas plataformas educacionais brasileiras, o princípio já orienta diretrizes de acessibilidade como as WCAG (*Web Content Accessibility Guidelines*), que estabelecem parâmetros para o controle de animações e flashes em ambientes digitais. A integração da IA a esses sistemas de filtragem proativa representa um avanço possível e desejável.

4.4 Ferramentas de apoio emocional e autorregulação

A quarta categoria de ferramentas é também a mais delicada, tanto pelas possibilidades que abre quanto pelos riscos que carrega. Estudantes com fobia social, transtornos de ansiedade severa e síndrome do pânico, transtorno borderline de personalidade (TPB), depressão maior ou transtornos alimentares enfrentam dificuldades que não são primariamente cognitivas nem comunicacionais, mas que afetam profundamente sua capacidade de estar na escola, de participar, de se arriscar e de se expor ao julgamento que o ambiente escolar, por natureza, implica. As ferramentas de IA voltadas para apoio emocional e autorregulação buscam oferecer um espaço de suporte que a escola raramente tem condições de oferecer de forma suficiente.

O Woebot é a ferramenta mais estudada nesse campo. Desenvolvido em 2017 pela psicóloga clínica Alison Darcy, com base em técnicas de Terapia Cognitivo-Comportamental (TCC), o *chatbot* realiza check-ins diários de humor, auxilia os usuários a identificar pensamentos disfuncionais e os guia por exercícios de reformulação cognitiva. Uma pesquisa randomizada controlada publicada no JMIR Mental Health (Fitzpatrick *et al.*, 2017) demonstrou que o uso consistente do Woebot reduziu significativamente sintomas de ansiedade e depressão em jovens adultos. A ferramenta não pretende substituir a psicoterapia, e seu próprio design comunica isso ao usuário, mas pode funcionar como suporte entre sessões, especialmente em contextos em que o acesso à saúde mental é limitado ou inexistente, como ocorre na maior parte das escolas públicas brasileiras.

O Wysa é outra plataforma de apoio emocional com base em evidências clínicas, disponível em 30 países e com mais de 15 publicações científicas revisadas por pares. Seu funcionamento integra técnicas de TCC, terapia dialética comportamental (TDC), meditação e *motivational interviewing* em conversas conduzidas por um avatar de pinguim que combina leveza de interface com seriedade clínica.

Para estudantes com transtornos de ansiedade que experimentam crises de pânico ou dificuldades de regulação emocional antes de provas ou apresentações, o Wysa pode oferecer exercícios de respiração guiada, técnicas de *grounding* e psicoeducação acessível, funcionando como um recurso de autorregulação disponível 24 horas, sem filas e sem julgamento. A revisão sistemática de Moylan e Doherty (2025), publicada no Journal of Medical Internet Research, que analisou 160 estudos sobre *chatbots* de saúde mental entre 2020 e 2024, confirma o crescimento desse campo. Conforme a pesquisa, os sistemas baseados em modelos de linguagem de grande porte (LLMs) saltaram de 0% dos estudos em 2020 para 45% dos novos estudos em 2024, o que indica aceleração significativa do interesse clínico e científico.

Para estudantes com autismo que frequentemente enfrentam dificuldades de regulação emocional e de interpretação das próprias emoções e das emoções alheias, ferramentas como o EMooly, um aplicativo baseado em IA generativa e realidade aumentada desenvolvido por pesquisadores para apoiar a aprendizagem socioemocional de crianças autistas com o envolvimento ativo dos cuidadores, representam uma fronteira emergente e promissora (Lyu et al., 2024, apud Frontiers in Psychiatry, 2025).

Para estudantes com síndrome de Tourette, para quem a presença dos tiques em ambientes de alta pressão social pode gerar vergonha e isolamento, ferramentas que permitem a interação mediada por texto, sem câmera e sem a exposição do corpo no espaço social digital, podem reduzir o custo emocional da participação.

Estudantes com transtornos alimentares como anorexia e bulimia frequentemente desenvolvem relações complexas com o ambiente escolar, onde comentários sobre corpo, alimentação e aparência são comuns e onde as refeições coletivas podem ser fontes de sofrimento intenso. Ferramentas de IA de apoio emocional não tratam esses transtornos, mas podem oferecer um espaço de expressão, de identificação de gatilhos e de orientação para busca de ajuda.

Para estudantes com transtorno de personalidade borderline, cuja característica central é a instabilidade emocional intensa, a regularidade de uma ferramenta que responde sempre da mesma forma, sem mudança de humor, sem abandono e sem julgamento, pode ter um efeito estabilizador que não substitui a psicoterapia, mas a complementa no cotidiano escolar.

É aqui, no entanto, que a nota crítica se impõe com mais urgência. Os limites do uso de IA para apoio emocional são reais, documentados e precisam ser explicitados com clareza. A revisão sistemática de Moylan e Doherty (2025) aponta que apenas 16% dos estudos sobre *chatbots* de LLMs passaram por testes de eficácia clínica, o que significa que a vasta maioria das ferramentas disponíveis carece de validação terapêutica rigorosa.

O mesmo estudo destaca a tensão entre os sistemas baseados em regras como o Woebot, que priorizam segurança por meio de diálogos estruturados, e os sistemas LLM como o Replika, que correm o risco de gerar respostas não validadas devido à dependência de dados não curados. Em dezembro de 2024, a *American Psychological Association* (APA) formalizou uma queixa à *Federal Trade Commission* dos Estados Unidos contra um *chatbot* de IA generativa acusado de causar danos a crianças, episódio que ilustra concretamente os riscos de uso irresponsável nesse campo.

Há relatos documentados de que *chatbots* de simulação social podem aumentar a ansiedade em usuários autistas ao reproduzir interações que se assemelham, mas não reproduzem fielmente, as nuances da interação humana, gerando respostas desconcertantes justamente por serem quase, mas não completamente, humanas (MadeCurious, 2025). A dependência é outro risco, pois diferentemente das tecnologias assistivas tradicionais, que ampliam capacidades existentes, ferramentas de apoio emocional baseadas em IA podem, se mal orientadas, substituir relações humanas em vez de sustentá-las, aprofundando o isolamento que se propõem a mitigar.

Diante desses riscos, a posição que este livro defende é clara: as ferramentas de apoio emocional baseadas em IA têm potencial real de ampliar o suporte disponível a estudantes neurodivergentes em contextos escolares onde esse suporte é cronicamente insuficiente, mas devem ser utilizadas como complemento explicitamente reconhecido como tal, nunca como

substituto do vínculo humano, do acompanhamento clínico ou da responsabilidade institucional da escola com o bem-estar de seus estudantes. A inteligência artificial, nesse campo, só é inclusiva quando opera dentro de uma rede de cuidado que a precede e que a envolve. Quando essa rede não existe, a ferramenta não tem como preencher o vazio. Antes, apenas o torna mais tecnológico.

5 EVIDÊNCIAS EMPÍRICAS: O QUE AS PESQUISAS DIZEM SOBRE EFICÁCIA E IMPACTO

Toda vez que uma nova tecnologia entra no debate educacional com promessas de transformação, surge a necessidade imperiosa de distinguir o entusiasmo da evidência. O campo da inteligência artificial aplicada à educação inclusiva não é exceção a esse imperativo.

A quantidade de publicações sobre o tema cresceu de forma notável, sobretudo após 2022: conforme apurado por Liu *et al.* (2024, *apud* Hamilton *et al.*, 2025), o número de estudos sobre IA generativa na educação aumentou aproximadamente 175% entre 2022 e 2023 isoladamente. Esse crescimento acelerado, no entanto, não corresponde necessariamente a um aprofundamento da qualidade metodológica ou da robustez das evidências. Compreender o que as pesquisas efetivamente demonstram, e o que ainda não sabemos com rigor suficiente, é uma responsabilidade epistemológica que esta seção assume de frente.

A revisão de escopo conduzida por Hamilton et al. (2025), publicada na *ScienceDirect*, que analisou publicações entre 2020 e julho de 2024, identificou 21 fontes relevantes sobre o uso de IA generativa com estudantes neurodivergentes em contextos escolares de educação básica. Os resultados apontam aplicações predominantes em três áreas: aprendizagem personalizada, suporte administrativo a educadores e desenvolvimento de planos educacionais individualizados. Nos estudos sobre autismo, TDAH e dislexia, o uso de ferramentas de IA generativa foi associado a um aumento de engajamento, maior prazer na aprendizagem e possíveis melhorias nos resultados escolares. A revisão, porém, é enfática ao sinalizar que a maior parte dos estudos examinados são exploratórios ou conceituais, e não empíricos controlados, o que limita a generalização das conclusões.

Uma revisão sistemática mais rigorosa, publicada no PMC com base em critérios PRISMA e o framework PICOS, selecionou onze estudos experimentais publicados entre 2022 e 2025, abrangendo 3.033 participantes com transtornos específicos de aprendizagem, predominantemente dislexia (Andréia *et al.*, 2025, referenciando revisão internacional). Os resultados gerais indicam efeitos positivos no desempenho em leitura, escrita e compreensão

de texto, especialmente quando os sistemas adaptativos foram implementados com consistência e acompanhamento docente.

O risco de viés nos estudos revisados, avaliado pelas ferramentas ROBINS-I e JBI, foi classificado como moderado a alto em boa parte das pesquisas, o que exige cautela na interpretação dos dados. O próprio perfil dos estudos incluídos revela a assimetria do campo, pois intervenções com dislexia concentram a maioria das evidências, enquanto outras condições do espectro neurodivergente permanecem sub-representadas.

Essa assimetria é confirmada pela revisão sistemática de Shahini *et al.* (2024), que analisou 84 estudos sobre tecnologias assistivas baseadas em IA para crianças com transtornos do neurodesenvolvimento publicados entre 2018 e 2024. Dos estudos revisados, 56% focavam no TEA, 36% no TDAH e apenas 8% na dislexia, o que inverte a ênfase da revisão anterior e evidencia que a distribuição das pesquisas varia significativamente conforme a base de dados consultada e os critérios de inclusão adotados.

Os estudos sobre autismo, em sua maioria, reportaram resultados encorajadores no desenvolvimento de habilidades sociais e comunicacionais mediadas por robótica e interfaces de IA, com tecnologias baseadas em visão computacional atingindo até 88% de precisão no reconhecimento de emoções em crianças autistas em ambientes controlados. Os estudos sobre TDAH mostraram benefícios no engajamento e na manutenção da atenção por plataformas gamificadas e sistemas de tutoria inteligente. Em ambos os casos, o desempenho observado em ambientes de laboratório ou em contextos altamente controlados contrasta com a escassez de evidências sobre eficácia em sala de aula regular, o que permanece uma lacuna crítica.

A revisão de Liao *et al.* (2025), publicada nos *Proceedings da Association for Information Science and Technology*, centrou-se especificamente no uso de IA generativa por estudantes neurodivergentes no ensino superior, sintetizando 21 estudos publicados entre 2022 e 2025. Seus achados indicam que ferramentas como ChatGPT foram associadas a melhoras em fluência de escrita, gestão do tempo, autoeficácia e engajamento acadêmico.

Os estudantes com TDAH relataram benefícios especialmente na decomposição de tarefas e na organização de fluxos de trabalho. Já os estudantes com dislexia reportaram ganhos na produção de texto quando podiam usar a IA para reformular rascunhos iniciais em linguagem mais fluente. A revisão, porém, aponta persistência de desafios relacionados à personalização inadequada das ferramentas às necessidades individuais, à acessibilidade desigual das plataformas e às tensões em torno do uso ético na avaliação acadêmica.

O estudo de Chen e Zhu (2023, *apud* ScienceDirect, 2025) apresenta uma perspectiva que merece atenção especial por sua contracorrente. A investigação sugere que o uso de IA

generativa pode, em alguns casos, prejudicar estudantes com TDAH ao simplificar excessivamente o processamento de informações, potencialmente reduzindo a duração da atenção em vez de fortalecê-la. Embora esse achado não tenha sido replicado de forma sistemática, ele aponta para a necessidade de investigar os efeitos de longo prazo e os riscos de uso irrefletido, especialmente em populações com perfis cognitivos que podem ser afetados de modo contraintuitivo por interfaces projetadas para a facilidade.

No que diz respeito ao contexto brasileiro e latino-americano, o panorama das evidências é ainda mais incipiente. A produção científica nacional sobre o uso de IA com estudantes neurodivergentes em contextos escolares concentra-se majoritariamente em revisões bibliográficas e reflexões teóricas, com poucos estudos de campo que incluam mensuração de resultados de aprendizagem.

Santos *et al.* (2025), em revisão bibliográfica brasileira sobre personalização do aprendizado com IA para neurodivergentes, confirmam que a maioria dos estudos nacionais disponíveis até 2025 é de natureza qualitativa e exploratória, com amostras pequenas e sem grupos de controle, o que reflete tanto o estado nascente da pesquisa empírica na área quanto as limitações de infraestrutura e acesso tecnológico que caracterizam grande parte do sistema escolar público brasileiro

Melo Júnior *et al.* (2024) são igualmente explícitos ao reconhecer que a ausência de estudos longitudinais representa uma lacuna significativa que compromete a avaliação real dos impactos das intervenções tecnológicas na aprendizagem de longo prazo. Essa lacuna é parcialmente iluminada, no campo das representações, pela pesquisa de Costa Neto (2026), que demonstra como o discurso jornalístico nacional opera para naturalizar certas imagens da IA na educação antes mesmo que haja evidência empírica suficiente para sustentá-las, produzindo no senso comum uma convicção que os estudos científicos ainda estão longe de confirmar.

A revisão de Dumitru *et al.* (2026), que sintetiza 27 pesquisas publicadas entre 2020 e 2024 sobre IA no ensino superior para estudantes com deficiência, identifica cinco temas dominantes: aprendizagem personalizada, benefícios das tecnologias assistivas baseadas em IA, desafios de adoção, barreiras institucionais e boas práticas de implementação. O estudo aponta que as deficiências sub-representadas na literatura são justamente aquelas que concentram os maiores desafios de inclusão, como os transtornos de personalidade, os transtornos de humor severos e as condições sensoriais complexas. Essa lacuna não é neutra, pois significa que as populações mais marginalizadas são também aquelas sobre as quais menos

se pesquisa, perpetuando um ciclo de invisibilidade acadêmica que o campo precisa urgentemente romper.

Uma síntese honesta do estado atual das evidências permite afirmar que:

- há resultados positivos consistentes e documentados no uso de IA para melhorar o acesso ao currículo e a expressão acadêmica de estudantes com dislexia e TEA, em especial no que se refere à comunicação aumentativa e alternativa e à leitura mediada por síntese de voz;
- há evidências preliminares promissoras sobre os benefícios para TDAH em engajamento e gestão executiva, mas com qualidade metodológica variável, e, por fim;
- há escassez quase total de estudos longitudinais sobre qualquer condição neurodivergente, o que impede conclusões sobre sustentabilidade e transferência dos ganhos ao longo do tempo.

A pesquisa de campo em contextos de baixa renda, em países do Sul Global e em língua portuguesa é rara, o que compromete a validade ecológica das conclusões para o contexto escolar brasileiro. Como aponta a revisão de Linhares (2025), futuras pesquisas precisam ampliar investigações empíricas, incluindo experimentações em sala de aula e estudos de usabilidade com estudantes com deficiência, para que a construção de um uso humanizado, democrático e inclusivo das tecnologias inteligentes tenha sustentação empírica suficiente.

O que as pesquisas dizem, portanto, é simultaneamente encorajador e incompleto. Há promessas sólidas e evidências preliminares que justificam o investimento em pesquisa e em experimentação pedagógica. Mas não há ainda a base robusta de conhecimento empírico que permitiria afirmar, com segurança científica, que a IA transforma sistematicamente os resultados de aprendizagem de estudantes neurodivergentes em contextos escolares reais e diversos. Essa honestidade epistemológica não é derrota. Antes, é a condição de possibilidade de uma agenda de pesquisa séria, que traduza o entusiasmo legítimo do campo em evidências que sirvam, de fato, àqueles que mais precisam.

6 CONCLUSÃO

No início deste capítulo, evocamos a cena de uma professora que descobre, no mesmo dia, dois usos radicalmente distintos da inteligência artificial: um que abre possibilidade para um aluno com dislexia e outro que fecha o caminho de aprendizagem de um estudante que delegou à máquina o que deveria ser seu. Essa cena não foi escolhida à toa. Ela materializa, em um único episódio escolar, a tensão central que atravessa toda a discussão sobre IA e educação,

pois a tecnologia, em si, não tem direção. Quem a direciona são as escolhas pedagógicas, institucionais, políticas e éticas que a rodeiam.

O percurso realizado ao longo das seis seções deste capítulo procurou oferecer instrumentos para que as escolhas pedagógicas sejam mais conscientes e fundamentadas. Inicialmente, partimos do mapeamento das representações sociais dominantes sobre a inteligência artificial na educação, as quais foram identificadas empiricamente por Costa Neto (2026) em sua análise do *corpus* jornalístico brasileiro. A partir desse cenário, propusemos uma terceira via que não se caracteriza como mera conciliação, mas como um deslocamento teórico ao tomar a tecnologia como um instrumento de equidade.

Em continuidade a essa reflexão, identificamos as convergências estruturais entre as funcionalidades da IA generativa e as demandas de estudantes neurodivergentes. Demonstrou-se que essa relação não é acidental, pois está inscrita na própria arquitetura de flexibilidade e adaptabilidade de tais ferramentas. Nesse sentido, o argumento da prótese comunicacional permitiu nomear com precisão o problema que a tecnologia pode mitigar, qual seja, não a suposta falta de conteúdo, mas a barreira persistente entre o pensamento do estudante e a via pela qual a instituição escolar exige a sua expressão.

Por fim, o inventário das ferramentas existentes comprovou que esse potencial já possui aplicações concretas em plataformas e sistemas disponíveis para uso imediato. Contudo, o exame das evidências empíricas situou o debate no estado atual do conhecimento científico, revelando que, embora existam promessas fundamentadas e dados preliminares encorajadores, o campo ainda carece de estudos longitudinais robustos. Tal lacuna é necessária para que a discussão avance do território da esperança para o território da certeza metodológica.

A tese central que emerge desse percurso pode ser enunciada com clareza. A inteligência artificial, quando usada com intencionalidade pedagógica, com mediação docente qualificada e com atenção às especificidades de cada sujeito neurodivergente, tem potencial real de funcionar como instrumento de equidade educacional. Ela pode reduzir a desigualdade comunicacional sem apagar a diferença cognitiva. Pode oferecer canais alternativos de expressão sem homogeneizar os modos de pensar. Pode ampliar o acesso ao currículo sem substituir o vínculo humano que dá sentido ao ato de aprender. Mas nada disso ocorre de forma automática, e nada disso ocorre sem que a escola, o professor e as políticas públicas assumam responsabilidades que nenhuma tecnologia pode assumir em seu lugar.

É preciso reconhecer os limites do que foi aqui discutido. Este capítulo focou predominantemente nas potencialidades e nas condições de uso inclusivo da IA, reservando para a nota crítica da Seção 4 e para a honestidade epistemológica da Seção 5 os alertas sobre

riscos e lacunas. Esse foco foi intencional: o livro já havia, nos capítulos anteriores, mapeado o ambiente escolar como espaço frequentemente opressor para estudantes neurodivergentes, e este capítulo precisava abrir uma perspectiva de possibilidade sem ser ingênuo. Mas a perspectiva de possibilidade não esgota o debate. Os riscos permanecem reais e exigem tratamento específico.

É por isso que o Capítulo 9 se dedicará às dimensões que este capítulo intencionalmente não aprofundou: os riscos éticos do uso de IA na educação de populações vulneráveis, os vieses algorítmicos que podem reproduzir e ampliar desigualdades em vez de reduzi-las, as implicações políticas da privatização das ferramentas e das plataformas que mediam o acesso ao conhecimento, e os desafios da regulação em um campo que evolui mais rapidamente do que as instituições conseguem acompanhar. Esses não são temas periféricos. São o reverso necessário da moeda que este capítulo examinou pelo lado do brilho.

Guimarães Rosa, em Grande Sertão: Veredas, escreveu que o real não está na saída nem na chegada, mas na travessia. A inteligência artificial, para os estudantes neurodivergentes, não é destino nem ponto de partida. É travessia. E como toda travessia, seu sentido depende de quem acompanha, de qual caminho é escolhido e do cuidado com que os passos são dados. A escola que decide usar a IA como instrumento de equidade não está adotando uma tecnologia. Está fazendo uma escolha política sobre quem merece aprender e de que modo. Essa escolha é, no fundo, a mais humana de todas. E é sobre ela que vale a pena, sempre, continuar pensando.

REFERÊNCIAS

ADDITUDE MAGAZINE. **AI for school**: Grammarly, Quizlet, Otter, Brainly, MathGPT, Goblin Tools. ADDitude Magazine, fev. 2026. Disponível em: https://www.additudemag.com/slideshows/ai-for-school/. Acesso em: 11 abr. 2026.

ALVES, Socorro Vânia Lourenço; LIMA, Celson Pantoja; ALVES, Enoque C. Melo. Desenvolvimento de um jogo digital educativo para o aprendizado de geografia para crianças com TDAH – Bilhar geográfico. In: CONGRESSO SOBRE TECNOLOGIAS NA EDUCAÇÃO (Ctrl+e), 8., 2023. **Anais [...]**. SBC, 2023. p. 396-403. Disponível em: https://doi.org/10.5753/ctrle.2023.232961.

ALVES, Werlen dos Santos. **Educação 5.0 e neurociência na era da Inteligência Artificial**. **Revista Aracê**, [S. l.], v. 7, n. 6, p. 35127–35136, 2025. DOI: 10.56238/arev7n6-348. Disponível em: https://periodicos.newsciencepubl.com/arace/article/view/6266. Acesso em: 12 abr. 2026.

ARAÚJO, Fátima Salete Machado de; MORAES, Kelly Cristina; SILVA, Ana Lúcia da. O des(envolvimento) da inteligência emocional em alunos com TDAH no ensino fundamental: um relato de experiência. In: CONGRESSO INTERNACIONAL INTERDISCIPLINAR EM SOCIAIS E HUMANIDADES (Coninter), 9., 2020. **Anais [...]**. Even3, 2020. Disponível em: https://doi.org/10.29327/coninter2020.284240.

BERSCH, Rita. **Introdução à tecnologia assistiva**. Porto Alegre: UFRGS, 2017.

BRASIL. **Lei nº 13.146, de 6 de julho de 2015**. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Diário Oficial da União, Brasília, DF, 7 jul. 2015.

CONTI, Gabrielle; MOMBELLI, Mariângela Alice. Abordagem multissensorial nas habilidades de alfabetização de crianças com Transtorno do Espectro Autista (TEA): estratégias e recursos. **Revista Caderno Pedagógico**, v. 21, n. 13, p. e12805, 2024. Disponível em: https://www.google.com/search?q=https://ojs.cuadernopedagogico.com/index.php/cp/article/view/12805.

COSTA NETO, Luiz Pires da. **Representações sociais da mídia jornalística sobre a integração da Inteligência Artificial à educação**. 2026. Monografia (Licenciatura em Matemática) – Instituto Federal de Educação, Ciência e Tecnologia da Bahia, Campus Barreiras, Barreiras, BA, 2026. Orientador: Prof. Dr. Flávio de Ligório Silva.

DUMITRU, Cristina; GAVRILĂ-ALEXANDRU, Maria Iuliana; ION, Petronela-Cristina; POPESCU, Elena-Mirela. Integrating artificial intelligence in supporting students with disabilities in higher education: an integrative review. **Journal of Special Education Technology**, 2026. Disponível em: https://journals.sagepub.com/doi/10.1177/10554181251355428.

DWORK, Cynthia; MULLIGAN, Deirdre K. It's not privacy, and it's not fair. **Stanford Law Review Online**, v. 66, p. 35-40, 2013. Disponível em: https://www.google.com/search?q=https://lawreview.stanford.edu/print/online/its-not-privacy-and-its-not-fair/.

FERREIRA, Daiane; FONSECA, João Ricardo. Intervenções inclusivas no processo de alfabetização e letramento da pessoa com autismo. **Revista Diálogos Interdisciplinares**, v. 4, n. 16, p. 546-560, 2024.

FITZPATRICK, Kathleen Kara; DARCY, Alison; VIERHILE, Molly. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. **JMIR Mental Health**, v. 4, n. 2, p. e19, jun. 2017. Disponível em: https://doi.org/10.2196/mental.7785.

FRAGA, Cristiano da Cruz; DECARLI, Cecilia; BOLL, Cíntia Inês. Inteligência Artificial e Inteligência Emocional: um debate para o presente e o futuro da educação. **Ensino & Pesquisa**, União da Vitória, v. 22, n. 2, p. 218-229, abr./ago. 2024. Disponível em: https://doi.org/10.33871/23594381.2024.22.2.7852.

GALVÃO, Rômulo Henrique Linhares. Inteligência artificial como tecnologia assistiva emergente na educação inclusiva: implicações éticas e possibilidades pedagógicas. **Revista International Integralize Scientific**, Florianópolis, v. 5, n. 54, dez. 2025.

GONÇALVES, Railson Chermont; SALES, Elielson Ribeiro de. Ensino e aprendizagem de matemática para alunos com transtorno do déficit de atenção e hiperatividade (TDAH). **REMATEC**, v. 19, n. 47, 2024. Disponível em: https://doi.org/10.37084/rematec.1980-3141.2024.n47.e2024020.id649.

GP STRATEGIES. **Neurodiverse learning**: 5 AI tools for inclusion. GP Strategies Blog, 18 dez. 2024. Disponível em: https://www.gpstrategies.com/blog/5-ai-tools-to-foster-a-more-inclusive-work-environment-for-neurodiverse-learners/. Acesso em: 11 abr. 2026.

HAMILTON, Jemma; SMITH, Lauren; BROWN, Tracey; JONES, Mike. A scoping literature review of generative artificial intelligence for supporting neurodivergent school students. **Computers & Education: Artificial Intelligence**, 2025. Disponível em: https://www.sciencedirect.com/science/article/pii/S2666920X25000773.

JODELET, Denise. Représentations sociales: un domaine en expansion. In: JODELET, Denise (Ed.). **Les représentations sociales**. Paris: PUF, 1989. p. 31-61. Tradução de Tarso Bonilha Mazzotti. Revisão Técnica: Alda Judith Alves-Mazzotti. UFRJ – Faculdade de Educação, dez. 1993.

JÚNIOR, Fernando Alves; FÜLBER, Heleno; MERLIN, Bruno. Math Mission: uma proposta de jogo sério com realidade virtual para ensino do conteúdo de frações para alunos do 6º e 7º anos com TDAH. In: SIMPÓSIO BRASILEIRO DE JOGOS E ENTRETENIMENTO DIGITAL (SBGames), 23., 2024. **Anais [...]**. SBC, 2024. p. 936-947. Disponível em: https://doi.org/10.5753/sbgames.2024.240511.

LIAO, Yiliu; WANG, Shuo; ZHANG, Huichao; LI, Xiaonan. The potential of generative AI in supporting neurodiversity in higher education: a systematic review. **Proceedings of the Association for Information Science and Technology**, v. 62, out. 2025. Disponível em: https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pra2.1331.

LYU, Zhaoliang; CHEN, Kexin; GAO, Linyi; SUN, Rui. EMooly: a tablet-based game integrating generative AI and augmented reality for social-emotional learning in autistic children. In: FRONTIERS IN PSYCHIATRY. **Implementation of generative AI for the assessment and treatment of autism spectrum disorders: a scoping review**. Frontiers in Psychiatry, jun. 2025. Disponível em: https://doi.org/10.3389/fpsyt.2025.1628216.

MADECURIOUS. **How AI is redefining accessibility**. MadeCurious, out. 2025. Disponível em: https://madecurious.com/articles/ai-accessibility-for-neurodivergence/. Acesso em: 11 abr. 2026.

MANTOAN, Maria Teresa Eglér. **Inclusão escolar**: o que é? Por quê? Como fazer? São Paulo: Moderna, 2015.

MARQUES, Adriano Ventura; SILVA, Gabriel Francisco da; SANTOS, José Osman dos. O uso da IA para pessoas com deficiência considerando aspectos da propriedade intelectual. **Educação**, Porto Alegre, v. 47, n. 1, p. 1-12, jan./dez. 2024. Disponível em: http://dx.doi.org/10.15448/1981-2582.2024.1.44691.

MELO, Ivie Johnson Ribeiro de; MIRANDA, Andréa da Silva; ELISIÁRI, Larissa Sato. A robótica educacional e sua interação com pessoas neurodiversas. In: CONGRESSO BRASILEIRO ON-LINE DE ENSINO, PESQUISA E EXTENSÃO (ENSIPEX), 3., 2022. **Anais [...]**. 2022. Disponível em: https://www.google.com/search?q=https://ensipex.com.br/.

MELO JÚNIOR, Hermócrates Gomes; SANTOS, Luciana de Oliveira; SILVA, Rodrigo José da; OLIVEIRA, Fábio de. Formando professores para uma educação neuroinclusiva: desafios e perspectivas. **Revista Aracê**, São José dos Pinhais, v. 6, n. 4, p. 18198-18212, 2024. Disponível em: https://doi.org/10.56238/arev6n4-421.

MOSCOVICI, Serge. **Representações sociais**: investigações em psicologia social. Tradução de Pedrinho A. Guareschi. 5. ed. Petrópolis: Vozes, 2007.

MOURA, Renata Alves; LIMA, Antonio Carlos Santos de; SANTOS, Pablo Rodrigo Moura. A função do NAPNE na mediação docente para a formação de alunos com TDAH para o mundo do trabalho. In: **Educação profissional e tecnológica – Volume 2**. 2024.

MOYLAN, Kayley; DOHERTY, Kevin. Expert and interdisciplinary analysis of AI-driven chatbots for mental health support: mixed methods study. **Journal of Medical Internet Research**, v. 27, e67114, 25 abr. 2025. Disponível em: https://doi.org/10.2196/67114.

NARCISO, Rosângela; SOUZA, Tatiane; MARTINS, Lucas; COSTA, Patrícia. Conexões digitais no espectro autista: explorando potencialidades e promovendo inclusão. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 10, n. 4, p. 404-418, 2024. Disponível em: https://doi.org/10.51891/rease.v10i4.13496.

NASCIMENTO, Arlete Souza da Silva. Do desafio à descoberta: recursos lúdicos para potencializar o aprendizado de alunos com dislexia e TDAH. In: CONGRESSO BRASILEIRO SOBRE TRANSTORNOS ASSOCIADOS AO NEURODESENVOLVIMENTO, 1., 2025. **Anais [...]**. Revista Multidisciplinar em Saúde, 2025. Disponível em: https://doi.org/10.51161/conbraneuro/62292.

NEURODIVERSITY HUB. **Masking**. [S. l.: s. n., s. d.]. Disponível em: https://www.neurodiversityhub.com. Acesso em: 11 abr. 2026.

OTTER.AI. **AI for accessibility**: understanding its impact and applications. Otter.ai Blog, 24 jun. 2025. Disponível em: https://otter.ai/blog/ai-for-accessibility. Acesso em: 11 abr. 2026.

REIS, Nathalia Silva; LOPES, Marcela; GOMES, Vinicius; CARDOSO, Juliana. Neurodiversidade na escola: quebrando barreiras e construindo uma educação inclusiva para todos. **Revista Aracê**, v. 6, n. 4, p. 16062-16075, 2024.

ROCHA, Vinicius Cunha; QUEIROZ, João Thiago; SILVA, Cristiano. An AI gen and LLM-based platform for pedagogical support and sensory accommodations for neurodivergent students. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO (SBSI), 2025. **Anais [...]**. SBC, 2025. p. 231-240.

ROSA, João Guimarães. **Grande Sertão: Veredas**. Rio de Janeiro: José Olympio, 1956.

SANTOS, José Leandro dos. **A relevância da inteligência artificial em adultos com TDAH**. Even3, 2024. Disponível em: https://doi.org/10.29327/7218126.

SANTOS, Silvana Maria Aparecida Viana; OLIVEIRA, Maria de Fátima; LINS, Ricardo; ARAÚJO, Teresa. Inteligência artificial e personalização do aprendizado para alunos neurodivergentes. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 11, n. 10, out. 2025. Disponível em: https://doi.org/10.51891/rease.v11i10.21771.

SHAHINI, Amir; NOURI, Bayan; KHALILI, Saman; MORADI, Maryam. A systematic review for artificial intelligence-driven assistive technologies to support children with neurodevelopmental disorders. **Information Fusion**, 2024. Disponível em: https://www.sciencedirect.com/science/article/pii/S1566253525005147.

SILVA, Andréia Maria; SILVA, José; SANTOS, Ana; FERREIRA, Paulo. Inteligência artificial e inclusão educacional: perspectivas éticas e pedagógicas na prática docente. **Revista Caderno Pedagógico**, Curitiba, v. 22, n. 12, p. 1-21, 2025.

SPECIAL NEEDS ALLIANCE. **AI tools opening new doors for adults with special needs**. Special Needs Alliance Blog, 12 fev. 2025. Disponível em: https://www.specialneedsalliance.org/blog/ai-tools-opening-new-doors-for-adults-with-special-needs/. Acesso em: 11 abr. 2026.

SPEECHIFY. **Assistive technology for dyslexia and ADHD**. Speechify Blog, out. 2025. Disponível em: https://speechify.com/blog/assistive-technology-for-dyslexia-and-adhd/. Acesso em: 11 abr. 2026.

SUCHARA, Carolina Füchter. **Inteligência artificial e educação em diálogo**: caminhos para a personalização e inclusão. 2025.

UNESCO. **AI and education**: guidance for policy-makers. Paris: UNESCO, 2021. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000376709. Acesso em: 12 abr. 2026.

VYGOTSKY, Lev Semionovitch. **A formação social da mente**. 6. ed. São Paulo: Martins Fontes, 1997.

WOLF, Carly. **How AI can improve education for neurodiverse learners**. PTS Coaching Blog, 27 mar. 2025. Disponível em: https://www.google.com/search?q=https://ptscoaching.com/2024/08/ai-and-education-for-neurodiverse-learners/.

ZHAO, Xin; COX, Andrew; CHEN, Xuanning. **The use of generative AI by students with disabilities in higher education**. ScienceDirect, 2025. Disponível em: https://www.sciencedirect.com/science/article/pii/S1096751625000235. Acesso em: 11 abr. 2026.

CAPÍTULO 9

LIMITES, RISCOS E RESPONSABILIDADES: PARA UMA IA EDUCACIONAL CRÍTICA E ÉTICA

Flávio de Ligório Silva
Professor do Instituto Federal da Bahia-Campus Barreiras.

DOI: 10.52832/wed.209.1249



Resumo: Este capítulo analisa as implicações éticas e os riscos estruturais da inteligência artificial na educação inclusiva com foco especial na trajetória de estudantes neurodivergentes. A discussão fundamenta-se na tensão entre o potencial da tecnologia como prótese comunicacional e sua capacidade de ampliar desigualdades históricas por meio de vieses algorítmicos e do fenômeno da alucinação. São examinadas quatro dimensões críticas quais sejam o viés nos dados de treinamento, a falta de confiabilidade das respostas generativas, a desigualdade de acesso tecnológico e a privatização de dados sensíveis. A análise utiliza os conceitos de panóptico digital e de medicalização algorítmica para demonstrar que a tecnologia pode atuar como um mecanismo de controle e de patologização se não for mediada por uma prática docente crítica e intencional. O texto aborda o enquadramento jurídico da inteligência artificial como tecnologia assistiva no contexto brasileiro e propõe recomendações práticas voltadas para a formação de professores e para a criação de protocolos éticos que assegurem a privacidade e a dignidade dos alunos. Conclui-se que a inclusão genuína exige que a tecnologia funcione como um instrumento de visibilidade e de equidade sem jamais substituir o vínculo e o cuidado humano que são a substância do ato pedagógico.

Palavras-chave: Inteligência Artificial, Educação Inclusiva, Neurodiversidade, Ética Digital, Tecnologias Assistivas.

Abstract: This chapter analyzes the ethical implications and structural risks of artificial intelligence in inclusive education with a specific focus on the experiences of neurodivergent students. The discussion is based on the tension between the potential of technology as a communicative prosthesis and its ability to amplify historical inequalities through algorithmic biases and hallucinations. Four critical dimensions are examined: bias in training data, the unreliability of generative responses, technological access inequality, and the privatization of sensitive data. The analysis employs the concepts of the digital panopticon and algorithmic medicalization to demonstrate that technology can act as a mechanism for control and pathologization if not mediated by critical and intentional teaching practices. The text addresses the legal framework of artificial intelligence as an assistive technology in the Brazilian context and offers practical recommendations for teacher training and the creation of ethical protocols to ensure student privacy and dignity. It concludes that genuine inclusion requires technology to function as a tool for visibility and equity without ever replacing the bond and human care that constitute the substance of the pedagogical act.

Keywords: Artificial Intelligence, Inclusive Education, Neurodiversity, Digital Ethics, Assistive Technologies.

1 INTRODUÇÃO

Em 2023, um sistema de inteligência artificial utilizado por uma rede de ensino norte-americana para recomendar trilhas de aprendizagem personalizadas foi auditado e descobriu-se uma distorção sistemática. O relatório revelou que estudantes negros e latinos recebiam, com maior frequência, recomendações de atividades de nível básico, enquanto estudantes brancos eram encaminhados para desafios de complexidade superior.

O algoritmo não havia sido programado para discriminar. Ele havia sido treinado com dados históricos de uma escola que já discriminava e, ao aprender os padrões do passado, reproduziu o passado com precisão estatística e eficiência industrial. Nenhum professor havia tomado essa decisão. Nenhuma intenção consciente havia orientado esse resultado. A discriminação foi, literalmente, automatizada.

Essa cena não é um caso isolado. Ela é a ilustração mais concreta de um paradoxo que este capítulo se propõe a examinar com rigor: a mesma tecnologia que o Capítulo 8 apresentou como potencial instrumento de equidade e prótese comunicacional para estudantes neurodivergentes pode, quando desprovida de mediação crítica, de regulação institucional e de investimento público em acesso igualitário, operar como mecanismo de reprodução e amplificação das desigualdades que já existem na escola. O problema não está na inteligência artificial em si. Está na ausência de condições para que ela seja usada de forma justa.

O Capítulo 8 construiu o argumento de que a IA pode funcionar como instrumento de equidade para estudantes neurodivergentes, desde que orientada por intencionalidade pedagógica, mediação docente qualificada e atenção às especificidades de cada sujeito. Este capítulo não desfaz esse argumento. Ele o complexifica, ao examinar as condições em que esse potencial se frustra, se inverte ou se torna inacessível. A tensão entre possibilidade e risco não é uma fraqueza do argumento, mas a sua própria condição de maturidade. Uma defesa ingênua da IA na educação inclusiva seria tão prejudicial quanto uma rejeição irrefletida. O que a ética exige é um olhar que sustente a complexidade sem abandonar a esperança, nem fabricar certezas onde há ainda muita incerteza.

Quatro dimensões de risco estruturam a discussão deste capítulo. A primeira é o viés algorítmico, ou seja, a tendência dos sistemas de IA de reproduzir, ampliar e naturalizar as desigualdades que habitam os dados com que foram treinados. No campo educacional, esse risco assume formas específicas quando o alvo são estudantes neurodivergentes, já que sistemas treinados predominantemente com dados de populações neurotípicas tendem a classificar como desvio ou déficit o que é simplesmente diferença cognitiva. Como demonstram O'Neil (2016) e Noble (2018), decisões automatizadas podem discriminar determinados grupos quando não há mecanismos de correção e supervisão humana. Na educação, isso se traduz em riscos como a rotulação precoce de estudantes, a limitação de oportunidades de aprendizagem e a naturalização de desigualdades sob a aparência de objetividade tecnológica (Bitácora-Cerdas; Flores, 2026).

A segunda dimensão é o fenômeno da alucinação. Os modelos de linguagem de grande porte que sustentam a IA generativa têm uma limitação estrutural pouco discutida fora dos

ambientes técnicos especializados. Eles produzem respostas com alto grau de confiança estatística mesmo quando essas respostas são factualmente incorretas, fenômeno conhecido como "alucinação". Como descreve Figueiredo (2024), trata-se do fenômeno pelo qual o sistema inventa respostas perentórias com absoluta convicção para preencher lacunas de conhecimento, e essa característica tem consequências pedagógicas sérias quando os usuários são estudantes que ainda estão construindo seus critérios de validação do conhecimento. Para estudantes neurodivergentes com dificuldades de avaliação crítica da informação, o que é frequente em alguns perfis do espectro autista, por exemplo, a interface confiante e fluente da IA pode ser particularmente enganadora. O aluno recebe uma resposta errada formulada com a mesma serenidade com que receberia uma resposta certa, sem nenhum sinal de alerta.

A terceira dimensão é a desigualdade de acesso. A promessa da IA como instrumento de equidade supõe que todos os estudantes têm acesso às ferramentas, à infraestrutura digital e à conectividade que permitem seu uso. No Brasil, essa suposição é empiricamente falsa para uma parcela expressiva da população escolar. Segundo Pinheiro (2024), a falta de conectividade, equipamentos e capacitação para o uso dessas tecnologias pode gerar uma divisão digital que deixa para trás exatamente os alunos mais vulneráveis, os quais são, frequentemente, os mesmos que mais se beneficiariam da prótese comunicacional que a IA pode oferecer. A criança autista de uma escola pública periférica sem conexão estável à internet não acessa o Otter.ai. O aluno com TDAH de uma família sem recursos não usa o Goblin Tools. A IA, nesse cenário, não reduz desigualdades. Pelo contrário, ela as recodifica em nova linguagem tecnológica, acrescentando ao mapa das exclusões uma camada de exclusão digital.

A quarta dimensão é a privatização do acesso ao conhecimento. As ferramentas de IA mais sofisticadas disponíveis no mercado educacional pertencem a corporações privadas de grande porte tais como Google, Microsoft, OpenAI, Anthropic, cujos modelos de negócio incluem a coleta e o processamento de dados dos usuários. Quando estudantes neurodivergentes interagem com essas plataformas, produzem dados sobre suas dificuldades cognitivas, seus padrões de atenção, seus erros e seus progressos. Esses dados podem ser usados para treinar novos modelos, para personalizar anúncios, para orientar decisões institucionais. No Brasil, a Lei Geral de Proteção de Dados Pessoais, a LGPD, Lei nº 13.709/2018, estabelece um arcabouço de proteção que se aplica ao tratamento de dados de crianças e adolescentes, mas a fiscalização efetiva desse arcabouço no contexto escolar ainda é fragmentada e insuficiente (Costa Neto, 2026; Fernandes *et al.*, 2024).

Esses quatro riscos não são argumentos contra o uso da IA na educação inclusiva. São argumentos contra o uso da IA na educação inclusiva sem regulação, sem acesso equitativo,

sem formação docente crítica e sem participação dos sujeitos afetados nas decisões que os envolvem. O tom deste capítulo não é o do alarmismo. Como aponta Pinheiro (2024), os textos críticos sobre IA na educação não têm o papel de depreciar a tecnologia e a sua integração ao ensino, mas sim de voltar para os possíveis riscos a fim de promover estrategicamente tal interação. É exatamente esse o espírito que orienta a discussão que se segue.

O capítulo está organizado em seis seções. Após esta introdução, a Seção 2 examina em profundidade o viés algorítmico e suas implicações específicas para estudantes neurodivergentes. A Seção 3 discute o fenômeno da alucinação e seus riscos pedagógicos. A Seção 4 analisa a desigualdade de acesso como barreira estrutural à equidade digital. A Seção 5 examina as implicações éticas da privatização dos dados educacionais e os limites da regulação vigente. A Seção 6 propõe recomendações práticas para um uso crítico e responsável da IA na educação inclusiva, articulando implicações para políticas públicas, formação docente e práticas pedagógicas. Ao longo de todo o capítulo, o fio condutor permanece o mesmo que atravessa este livro desde o seu início: a inclusão genuína exige transformação institucional, e não apenas adoção de ferramentas.

Há uma distinção que Freire (1996) poderia reconhecer como fundamental neste debate que consiste na diferença entre usar a tecnologia como instrumento de libertação e usá-la como instrumento de domesticação. Essa diferença não é técnica. É política. E é exatamente por isso que discutir os riscos da IA na educação inclusiva não é um exercício de pessimismo tecnológico, mas um exercício de responsabilidade democrática.

2 LIMITAÇÕES ESTRUTURAIS DA IA NO CONTEXTO INCLUSIVO

Se o Capítulo 8 demonstrou que a IA pode funcionar como prótese comunicacional para estudantes neurodivergentes, é porque essa possibilidade existe em determinadas condições, com determinadas ferramentas, mediada por determinados profissionais. O que esta seção examina é o reverso dessa possibilidade, focada nas limitações que são estruturais. Isso significa que tais barreiras não derivam de falhas acidentais ou de mau uso pontual, mas da própria natureza dos sistemas de IA e do ambiente social em que eles são implantados. Compreender essas limitações não é desistir da tecnologia, pois tal entendimento constitui o pré-requisito para usá-la responsavelmente.

2.1 Viés nos dados de treinamento e risco de reforço da exclusão

Todo sistema de inteligência artificial aprende a partir de dados. Essa afirmação, aparentemente simples, carrega uma consequência de alcance ético e político profundo, pois se os dados de treinamento refletem desigualdades históricas e no caso da educação brasileira eles quase invariavelmente o fazem, o sistema aprende essas desigualdades e as reproduz com eficiência estatística. O problema não está na intenção dos programadores, mas na estrutura dos dados que alimentam o aprendizado.

O viés algorítmico pode surgir de múltiplas fontes, quais sejam, dos dados de treinamento que sub-representam determinadas populações, das preferências inconscientes dos desenvolvedores, ou das próprias limitações dos modelos matemáticos que sustentam o aprendizado de máquina (Fernandes *et al.*, 2024).

Na educação, esse viés assume formas que merecem atenção específica quando o foco são estudantes neurodivergentes. Os grandes modelos de linguagem foram treinados em textos produzidos predominantemente por populações neurotípicas, em idiomas majoritários e em contextos socioeconômicos de acesso privilegiado ao letramento. Isso significa que sua representação interna do que é uma resposta correta, uma escrita adequada ou um raciocínio bem estruturado está calibrada por um padrão que não inclui, e frequentemente patologiza, as formas de expressão neurodivergente.

Há consequências concretas e documentadas dessa assimetria. Sistemas de avaliação automática de redações têm sido criticados por penalizar textos que fogem das estruturas sintáticas convencionais, mesmo quando o conteúdo é rico e original, o que afeta desproporcionalmente estudantes com dislexia, com TDL e com perfis autistas que processam e expressam a linguagem de formas não lineares.

Sistemas de reconhecimento facial de emoções, utilizados em algumas plataformas educacionais para monitorar o engajamento dos alunos, foram identificados como imprecisos para crianças autistas, cujas expressões emocionais frequentemente divergem dos padrões que o sistema foi treinado a reconhecer (Hamilton *et al.*, 2025). Em ambos os casos, a tecnologia não apenas falha em incluir, como também rotula a diferença como déficit, produzindo um dado negativo que, ao ser armazenado e reutilizado, alimenta e retroalimenta sucessivamente novos ciclos de exclusão.

Garcia (2020) demonstra que os sistemas de IA não apenas exacerbam, mas perpetuam vieses raciais e socioeconômicos contidos nos dados históricos. Quando esse mecanismo é transposto para o contexto educacional brasileiro, já marcado por profundas desigualdades de raça, renda e território, o risco se torna especialmente grave.

Como alertam Noble (2018) e Selwyn (2019), sistemas algorítmicos tendem a reproduzir desigualdades sociais quando operam sobre bases de dados marcadas por exclusões históricas. A promessa de neutralidade tecnológica é desmentida pela constatação de que decisões automatizadas podem favorecer determinados grupos em detrimento de outros (Bitácora-Cerdas; Flores, 2026). Essa dinâmica torna-se ainda mais perversa quando analisada sob a lente da interseccionalidade, pois a opressão não ocorre de forma isolada. Para estudantes negros, pobres, indígenas ou residentes em territórios periféricos que também são neurodivergentes, o efeito de sobreposição das desigualdades pode ser multiplicador, visto que a inteligência artificial os invisibiliza por diversos critérios simultaneamente.

A mitigação desse risco exige, em primeiro lugar, a construção de bases de dados de treinamento que representem a diversidade real dos estudantes brasileiros, incluindo suas especificidades cognitivas, linguísticas, culturais e socioeconômicas. Exige, em segundo lugar, auditorias regulares dos algoritmos utilizados nas plataformas educacionais, com foco específico nos grupos mais vulneráveis. E exige, em terceiro lugar, a participação ativa de pessoas neurodivergentes, educadoras e pesquisadoras da área no processo de desenvolvimento e avaliação dessas ferramentas.

Doneda *et al.* (2018) ressaltam a importância de um enfoque multidisciplinar para enfrentar esses desafios, sugerindo que a ética de dados constitui uma possibilidade concreta de amortizar riscos concretos na implementação de sistemas de inteligência artificial, preservando os benefícios enquanto resguarda direitos e garantias. Sem esse conjunto de medidas, o viés não é uma possibilidade. É uma certeza estatística.

2.2 Alucinação e confiabilidade em contextos sensíveis

Há uma limitação dos modelos de linguagem de grande porte que raramente aparece nos materiais publicitários das plataformas educacionais, uma vez que esses sistemas produzem informações incorretas com a mesma fluência e confiança com que produzem informações corretas. Tal fenômeno, que a literatura técnica denomina alucinação, ocorre porque o sistema não raciocina, mas apenas calcula probabilidades no encadeamento da sucessão de palavras.

Nesse sentido, como demonstra Costa Neto (2026) a partir da análise técnica do funcionamento desses sistemas, cada palavra gerada pela inteligência artificial é a saída de uma função de probabilidade condicional. Isso significa que, dado o que foi escrito até determinado ponto, o sistema busca apenas o próximo *token* mais provável, em um processo que não envolve verificação de veracidade, não acessa nenhuma realidade externa ao banco de

parâmetros do modelo e é estruturalmente incapaz de distinguir entre o que é verdadeiro e o que é estatisticamente frequente.

A consequência direta é que os sistemas generativos podem preencher lacunas de conhecimento com respostas inventadas, formuladas com absoluta convicção computacional (Figueiredo, 2024). Para um usuário adulto, crítico e com repertório de verificação, essa limitação é gerenciável, pois basta confrontar a resposta com outras fontes. Contudo, para estudantes em formação, especialmente aqueles que ainda estão construindo suas capacidades de avaliação crítica da informação, a interface fluente e confiante da inteligência artificial pode ser profundamente enganadora. O aluno com dificuldades de leitura que usa uma ferramenta de IA para compreender um conceito científico pode receber uma explicação incorreta apresentada com precisão clínica, sem nenhum sinal de alerta visível.

Esse risco se intensifica em contextos sensíveis, que são precisamente os mais relevantes para a educação inclusiva. Quando um estudante com transtorno de ansiedade usa um *chatbot* de apoio emocional e recebe orientações incorretas sobre sintomas ou estratégias de regulação, o dano potencial vai além do pedagógico. Quando uma ferramenta de IA descreve equivocadamente os direitos de uma pessoa com deficiência, pode comprometer decisões com implicações legais e institucionais. Quando um sistema de tutoria inteligente oferece uma explicação médica ou psicológica imprecisa sobre uma condição neurodivergente, pode alimentar estigmas que o próprio estudante carregará consigo. Nesses contextos, a alucinação não é apenas um erro técnico, mas um risco de saúde, de dignidade e de justiça.

Como mitigar esse risco? A literatura aponta ao menos três estratégias complementares, sendo a primeira delas a ancoragem em fontes verificadas. Nesse sentido, ferramentas de inteligência artificial desenvolvidas para contextos educacionais sensíveis podem ser configuradas para restringir suas respostas a um corpus de fontes curadas e validadas, reduzindo significativamente a probabilidade de geração de conteúdo não verificado.

Somado a isso, a segunda estratégia consiste na supervisão humana obrigatória em contextos que envolvam orientação sobre saúde mental, condições neurodivergentes ou direitos. Nesses casos, as respostas da inteligência artificial devem ser apresentadas como ponto de partida para uma conversa com um profissional qualificado e não como orientação definitiva. A terceira e talvez a mais importante medida do ponto de vista pedagógico é a formação de usuários críticos, o que exige que professores, pais e os próprios estudantes desenvolvam a literacia algorítmica. Trata-se da capacidade de compreender que a tecnologia não sabe, apenas prediz, e que toda resposta merece verificação. Como afirma Costa Neto (2026), a exigência irrevogável diante dessa constatação é a da validação humana permanente,

de modo que recai sobre o usuário a responsabilidade de validar a veracidade e a integridade ética de todos os resultados obtidos.

2.3 Desigualdade de acesso tecnológico: a exclusão que precede o uso

Há uma premissa silenciosa em boa parte da literatura entusiasta sobre IA e educação inclusiva, fundamentada na ideia de que os estudantes que mais se beneficiariam da tecnologia têm acesso a ela. Essa premissa raramente é enunciada de forma explícita porque, quando revelada, expõe uma contradição estrutural que fragiliza o argumento inteiro. A inteligência artificial democratiza o acesso ao conhecimento para quem já tem acesso à ferramenta, mas para quem não tem, ela simplesmente não existe e não democratiza coisa alguma.

No Brasil, os dados mais recentes do IBGE são ao mesmo tempo animadores e reveladores das distâncias que persistem. Segundo o Módulo TIC da PNAD Contínua 2024, a internet estava presente em 93,6% dos domicílios particulares permanentes do país, representando um avanço significativo em relação a 2016, quando o percentual era de apenas 69,3%. Nas áreas urbanas, esse índice chegou a 94,7%, enquanto nas áreas rurais atingiu 84,8%, reduzindo a diferença histórica de mais de 40 pontos percentuais entre esses dois contextos para cerca de 10 pontos (IBGE, 2025).

À primeira vista, os números sugerem uma democratização do acesso. Mas uma leitura mais cuidadosa revela que 5,1 milhões de domicílios brasileiros ainda não acessaram a internet em 2024, e que os três principais motivos apontados para isso foram: nenhum morador saber usar a internet (32,6%), o custo do serviço ser considerado caro (27,6%) e falta de necessidade percebida (26,7%), sendo que esta última resposta frequentemente camufla um bloqueio de acesso não reconhecido como tal pelo próprio entrevistado.

Esses 5,1 milhões de domicílios não são distribuídos aleatoriamente pelo território nacional. Eles se concentram nos estratos de renda mais baixos, nas regiões Norte e Nordeste, nas zonas rurais e nas periferias urbanas, que são, não por acaso, os mesmos territórios em que a escolarização é mais precária e em que a prevalência de condições associadas à neurodivergência não diagnosticada é possivelmente maior, justamente pela ausência de serviços de saúde e avaliação acessíveis. A sobreposição geográfica da exclusão digital com a exclusão educacional e a exclusão terapêutica forma um mapa de vulnerabilidades acumuladas que a simples disponibilidade de ferramentas de IA não desfaz.

Mesmo entre os domicílios com acesso à internet, a qualidade da conexão é profundamente desigual. A TIC Domicílios 2024 aponta que apenas 28% dos domicílios rurais

contam com fibra óptica, enquanto nas áreas urbanas esse índice é de 67% (Cetic.br, 2024). O acesso via celular com pacote de dados limitado, predominante em populações de baixa renda, restringe drasticamente o uso de plataformas de IA que exigem conexão estável e processamento intensivo. Ferramentas como leitores de tela sofisticados, plataformas adaptativas e sistemas de tutoria inteligente podem ser completamente inacessíveis para um estudante que só acessa a internet por um *smartphone* com plano de dados de 2 GB mensais, compartilhado com toda a família.

A queda contínua no percentual de domicílios com microcomputador, de 45,9% em 2016 para 39,0% em 2023, segundo a PNAD Contínua, é outro indicador crítico. Muitas das ferramentas de IA mais relevantes para a educação inclusiva foram projetadas para uso em telas maiores, teclados físicos e conexões estáveis. Um estudante com disgrafia que precisa do ditado inteligente do Dragon NaturallySpeaking, ou um estudante autista que usa ferramentas de CAA mediadas por IA, encontra barreiras adicionais se o único dispositivo disponível é um *smartphone* com tela de 5 polegadas e microfone de baixa qualidade.

Diante desse quadro, a afirmação de que a inteligência artificial pode ser um instrumento de equidade para estudantes neurodivergentes no Brasil só é defensável se acompanhada de um compromisso claro com a universalização do acesso. Sem conectividade de qualidade, sem dispositivos adequados, sem formação de professores e sem suporte técnico nas escolas públicas, a tecnologia não apenas falha em incluir, mas aprofunda a linha divisória entre aqueles que têm acesso a ferramentas que amplificam suas capacidades e aqueles que continuam dependendo de recursos pedagógicos do século anterior.

Como argumenta Pinheiro (2024), a desigualdade no acesso às tecnologias digitais é um dos principais desafios para garantir que as inovações tecnológicas sejam benéficas para todos. Sem investimento público robusto em infraestrutura e formação, há o risco real de que as tecnologias de IA aprofundem ainda mais as disparidades já existentes no sistema educacional.

Esse não é um risco abstrato. Trata-se do cenário mais provável na ausência de políticas deliberadas de inclusão digital.

3 RISCOS ÉTICOS: VIGILÂNCIA, MEDICALIZAÇÃO ALGORÍTMICA E PRIVACIDADE DE DADOS

Michel Foucault, em Vigiar e Punir (1999), descreveu a escola como um dos espaços em que o poder disciplinar opera com maior eficácia sobre os corpos ao distribuí-los no

espaço, submetê-los ao tempo e classificar seus comportamentos. O olhar do examinador, do inspetor ou do professor substitui a necessidade de punição física, pois basta que o sujeito saiba que pode ser observado para que a vigilância internalize seus efeitos.

Nesse contexto, quando sistemas de inteligência artificial são introduzidos em ambientes escolares com a capacidade de monitorar continuamente o comportamento dos estudantes, a análise foucaultiana não perde validade, mas se intensifica. Ao registrar padrões de atenção, frequência, movimentação, tempo de resposta, expressões faciais e ritmo de digitação, a tecnologia faz com que o panóptico não precise mais de torre e guardas, uma vez que ele se torna código.

A vigilância algorítmica na educação apresenta-se, na maior parte das vezes, com nome de acessibilidade ou de personalização. Plataformas adaptativas monitoram o engajamento dos estudantes para ajustar o nível das atividades. Sistemas de detecção de emoções analisam expressões faciais para identificar estados de confusão, frustração ou desengajamento. Ferramentas de rastreamento ocular mapeiam para onde o estudante olha durante a leitura.

Em cada um desses casos, o argumento central reside na ideia de que a coleta de dados serve tanto para melhorar o aprendizado do aluno quanto para promover o aperfeiçoamento contínuo da própria ferramenta. Embora tais justificativas possam ser verdadeiras, elas não apagam a assimetria de poder que está na base da relação, uma vez que o estudante é monitorado e seus dados são coletados, processados e armazenados por sistemas que ele não controla e sobre os quais raramente possui transparência. Como alertam Linhares (2025) e Pinheiro (2024), tecnologias inteligentes podem coletar grandes volumes de dados pessoais sensíveis, incluindo informações comportamentais e registros de desempenho, o que exige políticas claras de segurança digital e transparência.

Para estudantes neurodivergentes, os riscos da vigilância algorítmica assumem uma dimensão que vai além do monitoramento comportamental e alcança o campo mais delicado da identidade, da privacidade e do direito ao sigilo sobre a própria constituição neurobiológica. Um sistema que monitora o comportamento de um estudante e registra que ele se levanta com frequência, muda o foco da tarefa constantemente ou não mantém contato visual com a câmera durante a interação com a máquina está, na verdade, coletando informações que descrevem exatamente os comportamentos associados ao TDAH, ao autismo ou à ansiedade, ou seja, pode estar fazendo um diagnóstico equivocado de TDAH ou outra condição, ao transformar sinais comportamentais em supostos indícios clínicos. Trata-se de informações que o sujeito tem o pleno direito de não compartilhar, e que ele pode ter escolhido deliberadamente não compartilhar ao longo de toda a sua trajetória escolar e profissional.

Esse silêncio do sujeito sobre sua própria condição não é omissão nem covardia, mas estratégia de sobrevivência e, antes de tudo, um direito. Como se discutiu em capítulo anterior deste livro, muitos estudantes neurodivergentes desenvolvem o que se convencionou chamar de *masking*, um esforço contínuo e exaustivo de camuflar comportamentos associados ao autismo, ao TDAH, à ansiedade ou outra condição para atender às expectativas de ambientes normativos.

Quando a plataforma transforma esses mesmos comportamentos em dados rastreáveis, ela não apenas viola a privacidade do sujeito como também desfaz, silenciosamente, um trabalho de autorregulação que custou ao estudante anos de aprendizado doloroso. E vai além, pois os registros gerados podem compor pastas escolares que, no futuro, se tornarão barreiras discriminatórias em processos seletivos e contratações profissionais que buscarão por certos perfis em detrimento de outros, condenando o sujeito à falta de oportunidade educacional ou laboral por uma condição que ele jamais autorizou revelar.

Há, porém, uma camada ainda mais inquietante nessa cadeia: que garantias existem de que o sistema está, de fato, lendo o que imagina estar lendo? O algoritmo é estruturalmente incapaz de distinguir entre um traço neurodivergente consolidado e uma desestabilização emocional temporária. O desengajamento e a inquietude registrados pela plataforma podem ser reflexos de um luto recente, de um divórcio parental, do nascimento de um irmão, de uma mudança de cidade ou até mesmo a ingestão de uma medicação.

Trata-se de situações que reorganizam profunda e momentaneamente a vida psíquica de qualquer criança ou adolescente, sem que isso configure uma patologia. Quando o algoritmo classifica essas flutuações como sinais de transtorno, ele não apenas comete um erro técnico como também alimenta um pânico diagnóstico que não encontra respaldo na realidade do sujeito, podendo sugerir encaminhamentos desnecessários que transformam uma crise vital temporária em um rótulo permanente.

É nesse sentido que a inteligência artificial não pode ser compreendida como neutra. A cadeia que vai da observação ao diagnóstico, do encaminhamento ao rótulo, ganha com o algoritmo um elo novo e especialmente perverso, sendo aquele que opera antes de qualquer conversa, antes de qualquer escuta, antes de qualquer relação humana, e que, justamente por isso, pode imprimir na trajetória do estudante uma identidade que ele nunca reivindicou e talvez nunca reconheça como sua.

Radicalizando ainda mais a situação, pensemos na situação do falso diagnóstico de TDAH do adolescente que está passando por uma situação de crise em sua vida, não correlata a nenhum transtorno que ele não tem feito, pela IA. Há um dano pedagógico concreto e

devastador que emerge desse falso diagnóstico, pois, uma vez que o sistema classifica erroneamente o estudante como tendo TDAH com base apenas em sinais comportamentais mal interpretados, ele pode passar a direcionar automaticamente exercícios, ritmos de ensino, pausas forçadas, lembretes excessivos e estratégias de suposto "suporte" que não fazem sentido para aquele sujeito. O estudante, que não apresenta a condição atribuída, receberá intervenções pensadas para uma neurobiologia que não é a sua: tarefas fragmentadas demais, tempos artificialmente reduzidos, repetições desnecessárias ou estímulos que subestimam sua capacidade de atenção real.

Longe de ajudar, esse ajuste equivocado produz frustração, confusão cognitiva e uma sensação avassaladora de inadequação. Afinal, ele tenta se adequar a um molde que não lhe cabe. Aos poucos, o que era um erro algorítmico pode se tornar uma profecia autorrealizável: o aluno começa a desengajar, a duvidar de si mesmo e a render menos, não porque tenha TDAH, mas porque o próprio ambiente de aprendizagem, movido por um diagnóstico fictício, passou a tratá-lo como se tivesse. O prejuízo não está apenas na violação da privacidade ou no risco discriminatório futuro; está na sala de aula virtual do presente, onde a máquina, ao errar, condena o estudante a uma experiência educacional desenhada para um problema que ele não tem e que, ironicamente, pode induzir os sintomas que o algoritmo julgava ter encontrado desde o início.

Essa dinâmica revela uma face ainda mais severa do poder disciplinar do que aquela originalmente descrita por Foucault (1999). Na arquitetura do panóptico tradicional, a vigilância servia para que o olhar do mestre pudesse incidir sobre o corpo do aluno e, em última instância, abrir espaço para uma pedagogia da correção baseada na relação direta. O panóptico digital, entretanto, opera uma inversão perversa ao silenciar o sujeito antes mesmo que ele possa se manifestar.

Enquanto na escola convencional o professor ainda poderia ouvir o estudante e compreender as razões de sua inquietude, na escola mediada por algoritmos o código sentencia antes que qualquer relação humana se estabeleça. A inteligência artificial executa o que se pode chamar de uma disciplina sem escuta, onde o veredito computacional precede o encontro e anula a possibilidade de o estudante explicar sua própria singularidade. O risco que se apresenta é o de uma educação que substitui o diálogo pela predição, transformando o ambiente escolar em um sistema de controle asséptico no qual a máquina já sabe quem o aluno é antes mesmo de ele ter tido a oportunidade de se tornar.

Esse fenômeno pode ser compreendido como uma medicalização algorítmica, caracterizada pela tendência de sistemas de inteligência artificial treinados em dados clínicos

ou comportamentais padronizados a sugerir interpretações patologizantes de comportamentos que representam apenas o modo de ser do sujeito no mundo. Como o algoritmo ignora o contexto de vida do estudante, ele é incapaz de reconhecer que o ato de se levantar da cadeira pode ser uma estratégia de autorregulação previamente pactuada, ou que a ausência de contato visual pode indicar um processamento intenso do conteúdo em vez de desatenção. Da mesma forma, um tempo de resposta mais lento pode refletir um pensamento mais profundo e não uma dificuldade de aprendizagem, mas o sistema ignora essas nuances porque seu funcionamento se baseia exclusivamente no desvio da norma estatística sobre a qual foi treinado.

A consequência direta dessa cegueira contextual é que a inteligência artificial, ao ser utilizada para avaliar e classificar estudantes sem mediação crítica, reproduz em escala industrial a mesma lógica normalizadora que o paradigma da neurodiversidade procura desconstruir. Essa automação da norma ignora o direito ao sigilo e as complexas táticas de sobrevivência desenvolvidas pelos estudantes, transformando a singularidade humana em um erro de processamento. Ao agir dessa forma, a tecnologia deixa de ser um suporte para a inclusão e passa a ser um instrumento de padronização que reforça as barreiras invisíveis discutidas ao longo deste livro, consolidando um ambiente escolar onde a diversidade é lida como uma falha a ser corrigida pelo código.

A dimensão da privacidade de dados articula-se com esses dois riscos e os amplifica. No contexto educacional brasileiro, os dados produzidos por estudantes em plataformas de IA não são apenas dados de aprendizagem. São dados de saúde mental, de condição neurológica, de comportamento, de estado emocional. Em muitos casos, são dados que revelam diagnósticos, vulnerabilidades e histórias que as famílias não necessariamente desejam compartilhar com plataformas comerciais.

A Lei Geral de Proteção de Dados Pessoais (LGPD), Lei nº 13.709/2018 classifica como dados sensíveis aqueles referentes à saúde, à vida sexual, a convicções filosóficas e religiosas, entre outros. Seu artigo 14 estabelece proteção reforçada para dados de crianças e adolescentes, exigindo consentimento específico e destacado de pelo menos um dos responsáveis. Assis (2023) destaca que a cibercultura marcada pela plataformização, dataficação e performatividade algorítmica permite a coleta massiva de dados pelas plataformas tecnológicas, o que impõe a necessidade de enfatizar aspectos direcionados à proteção dos dados de crianças e adolescentes.

O problema é que o arcabouço da LGPD, embora sólido em seus princípios, enfrenta dificuldades estruturais de aplicação no contexto escolar. As plataformas digitais educacionais

frequentemente operam sob termos de uso extensos e tecnicamente opacos, que poucas famílias leem ou compreendem. O consentimento, quando obtido, é muitas vezes resultado de uma aceitação em bloco das condições de uso, sem que o responsável compreenda quais dados específicos serão coletados, como serão processados, por quanto tempo serão armazenados e com quem poderão ser compartilhados.

Em contextos de alta vulnerabilidade, onde a família não tem letramento digital suficiente para avaliar essas condições, o consentimento se torna uma ficção jurídica que protege a plataforma mais do que a criança. Como aponta Costa Neto (2026), a questão sobre se a LGPD será de fato suficiente para garantir a segurança e a privacidade das informações inseridas por estudantes e professores permanece aberta e urgente.

O risco mais grave, e que raramente é discutido publicamente, é o da discriminação futura. Dados de saúde mental e de condições neurodivergentes coletados na infância podem, se mal protegidos ou indevidamente compartilhados, afetar o acesso do sujeito a oportunidades educacionais posteriores, ao mercado de trabalho ou a serviços de saúde e seguro ao longo da vida. Uma criança diagnosticada com autismo cujos dados comportamentais estão armazenados em um servidor com segurança precária pode, quando adulta, encontrar esse dado reaparecendo em contextos nos quais ele se torna um instrumento de discriminação. Não é preciso imaginar cenários distópicos para levar esse risco a sério. Basta observar que os sistemas de gestão escolar já acumulam dados sensíveis por décadas, e que a proteção desses sistemas raramente é prioridade nas políticas de infraestrutura digital das redes de ensino públicas.

O que se propõe, diante desses três riscos interligados, não é o banimento da IA da escola, mas a criação de protocolos éticos que estabeleçam limites claros para o que pode ser monitorado, classificado e armazenado no ambiente escolar. Esses protocolos precisam incluir, no mínimo: a definição do que constitui dado educacional sensível e como ele deve ser tratado; a exigência de transparência algorítmica — que os sistemas utilizados expliquem em linguagem acessível o que coletam e como decidem; a proibição de encaminhamentos automáticos baseados em padrões de comportamento detectados por IA, sem validação humana qualificada; e a garantia de que o dado educacional produzido na infância não comprometa trajetórias futuras. A IA na escola pode ser aliada da inclusão ou instrumento de controle. Essa diferença não é técnica. É política. E ela começa nas perguntas que fazemos antes de assinar o contrato com a plataforma.

4 A IA COMO TECNOLOGIA ASSISTIVA: IMPLICAÇÕES PARA A POLÍTICA EDUCACIONAL

Existe uma distinção que o debate sobre a inteligência artificial na educação raramente explicita, mas que tem consequências jurídicas, orçamentárias e pedagógicas de grande alcance, pautada na diferença entre a tecnologia como inovação educacional opcional e como ferramenta assistiva garantida por direito. Quando a inteligência artificial é enquadrada como inovação, sendo uma aposta pedagógica interessante, um diferencial de qualidade ou uma iniciativa de vanguarda, sua oferta fica sujeita à disponibilidade de recursos, à vontade política e à capacidade técnica de cada rede de ensino. Por outro lado, quando ela é reconhecida como tecnologia assistiva para estudantes com deficiência, o quadro muda inteiramente. O direito à tecnologia assistiva no Brasil não é um favor, mas uma obrigação legal.

A Lei Brasileira de Inclusão da Pessoa com Deficiência (LBI), Lei nº 13.146/2015, em seu artigo 74, define tecnologia assistiva como o conjunto de produtos, equipamentos, dispositivos, recursos, metodologias, estratégias, práticas e serviços que objetivem promover a funcionalidade, relacionada à atividade e à participação da pessoa com deficiência ou com mobilidade reduzida, visando à sua autonomia, independência, qualidade de vida e inclusão social.

Essa definição é deliberadamente ampla. Ela não se limita a dispositivos físicos como cadeiras de rodas ou próteses, mas inclui metodologias, estratégias e serviços. Uma ferramenta de ditado por reconhecimento de voz que permite a um estudante com disgrafia severa produzir textos acadêmicos se enquadra nessa definição. Um sistema de leitura em voz alta que permite a um estudante com dislexia acessar conteúdo curricular também. Um aplicativo de comunicação alternativa e aumentativa para um estudante autista não verbal, certamente. O que esses três exemplos têm em comum é que todos podem ser mediados por inteligência artificial e todos representam condições de acesso ao currículo, não vantagens competitivas.

Se a IA que viabiliza a participação de quem, sem ela, ficaria excluído da produção acadêmica é tecnologia assistiva nos termos da LBI, então sua provisão não é opcional. Ela é parte daquilo que a lei denomina adaptações razoáveis, ou seja, modificações e ajustes necessários e adequados que não acarretem ônus desproporcional e indevido, quando requeridos em cada caso, a fim de assegurar que a pessoa com deficiência possa gozar ou exercer, em igualdade de condições e oportunidades com as demais pessoas, todos os direitos e liberdades fundamentais (art. 3º, inciso VI). A recusa em prover tecnologia assistiva quando

ela é necessária à participação do estudante constitui, nos termos do artigo 4º da LBI, discriminação por motivo de deficiência.

O Decreto nº 12.686, publicado em outubro de 2025, que institui a Política Nacional de Educação Especial Inclusiva, representa um avanço significativo nesse enquadramento. Entre as diretrizes da nova política, o artigo 3º, inciso V, estabelece expressamente a oferta de tecnologias assistivas e adaptações razoáveis, de acordo com as necessidades individuais, em interação com os contextos educacionais.

O decreto avança ainda ao reconhecer o Plano de Atendimento Educacional Especializado (PAEE) como documento pedagógico individualizado que deve orientar o uso de recursos de acessibilidade. Isso abre espaço para que ferramentas de IA sejam incluídas nesses planos como tecnologias assistivas formalmente prescritas. Em seu artigo 11, parágrafo 4º, o decreto vai além ao estabelecer que a instituição de ensino deverá prover parecer pedagógico que autorize a utilização de dispositivos digitais portáteis como instrumento de tecnologia assistiva no processo de aprendizagem, comunicação ou socialização aos estudantes que são o público da educação especial. Esse dispositivo cria, pela primeira vez no ordenamento jurídico educacional brasileiro, uma obrigação formal de reconhecimento de dispositivos digitais como tecnologia assistiva no contexto escolar.

O mesmo decreto, em seu artigo 11, parágrafo 5º, determina que a coleta, o tratamento, o armazenamento e o compartilhamento de dados pessoais e sensíveis de estudantes que são o público da educação especial devem observar os princípios e os fundamentos previstos na Lei nº 13.709, de 14 de agosto de 2018, com atenção especial ao disposto no artigo 14. Essa disposição articula, de forma pioneira, a provisão de tecnologia assistiva com os requisitos de proteção de dados, reconhecendo que o uso de dispositivos digitais por populações vulneráveis gera dados sensíveis que merecem proteção reforçada. É um avanço normativo real, ainda que sua implementação efetiva dependa de infraestrutura, fiscalização e formação que o próprio decreto não garante.

Há, no entanto, importantes silêncios no Decreto nº 12.686/2025 que merecem ser nomeados. O primeiro é o silêncio sobre financiamento específico. O decreto menciona tecnologias assistivas como diretriz e como obrigação, mas não estabelece fontes de financiamento dedicadas para sua provisão nas redes públicas de ensino. O artigo 22 delega as despesas às dotações orçamentárias consignadas ao Ministério da Educação, observada a disponibilidade orçamentária e financeira. Tal fórmula, na prática, confere ao executor a discricionariedade que o legislador pretendeu retirar. Para que a provisão de IA como tecnologia assistiva se torne realidade nas escolas públicas brasileiras, será necessário que o

Fundeb, principal mecanismo de financiamento da educação básica, inclua, entre as despesas elegíveis, a aquisição e a manutenção de ferramentas digitais assistivas, com dotação específica para essa finalidade.

O segundo silêncio é o da formação docente para o uso pedagógico de tecnologias assistivas baseadas em IA. O decreto prevê formação específica para professores do AEE com carga horária mínima de 80 horas, o que é um passo importante. Mas a formação para identificar, selecionar e integrar ferramentas de IA ao PAEE de cada estudante é uma competência altamente especializada que 80 horas dificilmente cobrem. A maioria dos cursos de formação de professores no Brasil não aborda o uso pedagógico de IA nem o enquadramento legal da tecnologia assistiva digital. Sem essa formação, o parecer pedagógico exigido pelo decreto para autorizar o uso de dispositivos digitais como tecnologia assistiva corre o risco de se tornar um procedimento burocrático sem substância pedagógica.

O terceiro silêncio é o da regulação das plataformas. O decreto obriga as escolas a proteger os dados dos estudantes, mas não estabelece critérios para a certificação ou auditoria das plataformas de IA utilizadas como tecnologia assistiva. Uma escola pode contratar uma plataforma que coleta dados de forma invasiva, que apresenta vieses algorítmicos documentados e que oferece suporte inadequado para populações neurodivergentes, sendo que o decreto atual não oferece instrumentos específicos para impedir ou remediar esse cenário.

O Projeto de Lei nº 2.338/2023, ainda em tramitação no Congresso Nacional quando da redação deste capítulo, que busca regulamentar o uso da inteligência artificial no Brasil, poderia preencher parte dessa lacuna se contemplar disposições específicas sobre IA em contextos educacionais e sobre proteção de populações vulneráveis. Mas sua aprovação é incerta, e seu conteúdo, passível de alteração.

O que se propõe, à luz dessa análise, é um conjunto de medidas que transformem o enquadramento legal já existente em política pública efetiva. A primeira é a inclusão da IA assistiva como categoria específica nas diretrizes do Fundeb, permitindo que redes de ensino financiem a aquisição de ferramentas digitais reconhecidas como tecnologia assistiva sem que isso compita com outras despesas de manutenção e desenvolvimento da educação.

A segunda é a criação de um protocolo nacional de certificação de plataformas de IA para uso assistivo em escolas, com critérios de auditoria de viés, proteção de dados e acessibilidade para populações neurodivergentes. A terceira é a integração da literacia em IA assistiva nos currículos dos cursos de formação de professores e nos programas de formação continuada, com ênfase nos direitos legais dos estudantes e nos procedimentos para inclusão

de ferramentas tecnológicas no PAEE. A quarta é a obrigatoriedade de relatório anual de impacto, por parte de todas as plataformas de IA utilizadas em contextos escolares, especificando os dados coletados de crianças e adolescentes, as finalidades de uso e as medidas de proteção adotadas.

A IA como tecnologia assistiva não é uma metáfora. É uma categoria jurídica com implicações concretas sobre o que o Estado brasileiro está obrigado a prover para seus cidadãos com deficiência. Reconhecer isso é o primeiro passo para que a discussão sobre IA e educação inclusiva saia do campo das possibilidades e entre no campo das obrigações. E é no campo das obrigações e não no campo dos entusiasmos que as transformações estruturais acontecem.

5 RECOMENDAÇÕES PRÁTICAS PARA PROFESSORES

Há uma tensão que percorre todo o debate sobre inteligência artificial e educação e que raramente é resolvida de forma satisfatória, a qual se estabelece entre o nível da política, onde as decisões sobre infraestrutura, regulação e financiamento são tomadas, e o nível da sala de aula, local em que o professor enfrenta, simultaneamente, trinta estudantes, uma turma heterogênea, uma conexão de internet instável e a pressão de um currículo rígido. As seções anteriores deste capítulo operaram, em grande medida, no nível da política. Esta seção desce ao chão da escola não para oferecer receitas simples, uma vez que a complexidade do tema não permite tal abordagem, mas para traduzir os argumentos desenvolvidos até aqui em orientações concretas que um docente pode, de fato, considerar em sua prática.

5.1 Uso crítico e mediado de ferramentas de IA em sala

O primeiro princípio do uso pedagógico crítico da IA é que o professor precisa conhecer a ferramenta antes de propô-la ao estudante. Esse conhecimento não é técnico no sentido de saber como o algoritmo funciona internamente, já que poucos professores têm e precisam ter essa formação. É um conhecimento pedagógico e ético que consiste em saber o que a ferramenta faz, quais dados ela coleta, o que ela não consegue fazer, onde ela erra e com que frequência, a quais populações ela serve melhor e quais ela serve mal. Esse conhecimento prévio é o que permite ao professor não delegar ao algoritmo decisões que pertencem ao julgamento humano.

Na prática, isso significa que antes de recomendar qualquer ferramenta de IA a um estudante neurodivergente, o professor deveria ser capaz de responder a pelo menos três perguntas:

- para que problema específico deste estudante esta ferramenta pode ajudar?
- quais são as limitações documentadas desta ferramenta para esse perfil cognitivo?
- o que acontecerá com os dados produzidos pelo estudante ao usar esta ferramenta?

A incapacidade de responder a essas perguntas não é uma falha moral do professor e sim uma falha do sistema de formação que o preparou. Mas reconhecer essa lacuna é o primeiro passo para preenchê-la.

O segundo princípio é o da mediação ativa. A IA não deve ser oferecida ao estudante como um recurso autônomo que ele usa sozinho, sem acompanhamento docente, especialmente quando o estudante é neurodivergente e está em processo de construção de autonomia. A mediação ativa significa que o professor está presente no processo, não controlando cada interação, mas criando antes as condições de uso consciente, acompanhando durante com curiosidade pedagógica e avaliando depois se o uso produziu aprendizagem real ou apenas produziu um produto.

Williamson (2017) alerta que a governança algorítmica na educação tende a deslocar o professor de seu papel decisório para uma função mais técnica e operacional. O antídoto para esse deslocamento é precisamente a mediação ativa, pois o professor que usa a IA com intencionalidade pedagógica não está sendo substituído. Está ampliando seu repertório de intervenção.

Um exemplo concreto: uma professora do ensino fundamental que trabalha com um estudante com dislexia pode propor que ele dite oralmente um parágrafo para o Microsoft Word com reconhecimento de voz, revise o texto ditado em voz alta para verificar se representa o que pretendia dizer, e então discuta com a professora o conteúdo produzido. Nesse caso, a IA cumpre a função de remover a barreira ortográfica sem substituir o pensamento, a revisão ou o diálogo. A professora é a âncora do processo.

Outro exemplo: um professor do ensino médio que trabalha com um estudante com TDAH pode combinar com ele o uso do Goblin Tools para decompor uma tarefa complexa em etapas, e depois verificar juntos, ao final da aula, se as etapas foram cumpridas. A ferramenta é o suporte, o professor é o construtor que decide quando e como ele pode ser removido.

O terceiro princípio é o da verificação crítica. Toda resposta gerada por IA que será usada em contexto pedagógico, seja como explicação de um conceito, seja como texto simplificado, seja como roteiro de estudo, deve passar por avaliação docente antes de chegar ao estudante. Isso não significa que o professor precise verificar cada interação em tempo real, o que seria inviável. Significa que o professor precisa desenvolver, e ensinar o estudante a desenvolver, o hábito de questionar se o conteúdo produzido faz sentido; pode ser verificado em outra fonte; pode parecer estranho como resposta.

Figueiredo (2024) é preciso ao afirmar que recai sobre o usuário a responsabilidade de validar a veracidade e a integridade ética de todos os resultados obtidos em interações com IA. Essa responsabilidade não é do estudante isolado, mas compartilhada com o professor que o formou para esse julgamento.

5.2 Adaptações pedagógicas com suporte de IA generativa

A IA generativa abre possibilidades de diferenciação pedagógica que antes demandavam horas de trabalho manual. Um professor que precisava preparar três versões de um mesmo texto, uma em linguagem simples, uma em linguagem intermediária, uma em linguagem acadêmica, podia levar um turno inteiro para fazer esse trabalho. Com IA generativa, ele pode fazê-lo em minutos. O que muda não é apenas a velocidade, mas a viabilidade. A diferenciação que antes era impossível por restrição de tempo torna-se possível. O professor passa de executor a planejador e curador.

Para estudantes com dislexia e transtornos específicos de aprendizagem, a IA generativa pode ser usada para simplificar textos densos sem perda de conteúdo, o que exige do professor a habilidade de avaliar se a simplificação preservou o argumento ou o esvaziou. Ferramentas como ChatGPT, Gemini ou Claude podem receber a instrução de reescrever um parágrafo de texto científico em linguagem acessível para um estudante do oitavo ano, mantendo os conceitos centrais mas eliminando o vocabulário técnico desnecessário. O professor avalia o produto, faz ajustes e decide se ele está pedagogicamente adequado. Esse fluxo, IA como produtora de rascunho → professor como curador, é um dos mais promissores para a diferenciação em sala.

Para estudantes com TDAH, a IA generativa pode ser usada para criar roteiros de estudo personalizados que quebram o conteúdo em blocos menores com intervalos definidos, com lembretes de objetivos e com perguntas de verificação ao final de cada bloco. Um roteiro desse tipo pode ser solicitado pelo professor com uma instrução como: crie um roteiro de

estudo para o conteúdo de equações do segundo grau para um estudante do nono ano que tem dificuldade de manter o foco por mais de 15 minutos, dividindo o conteúdo em blocos de 10 minutos com uma pausa ativa entre eles e uma pergunta de revisão ao final de cada bloco. A IA produz a estrutura; o professor avalia se ela está pedagogicamente adequada para aquele estudante específico.

Para estudantes com autismo que têm dificuldade com instruções ambíguas, a IA generativa pode ser usada para converter enunciados de tarefas formulados em linguagem indireta, frequente nos livros didáticos, em enunciados diretos, concretos e com etapas explicitadas. A instrução ao sistema poderia ser: reescreva este enunciado em linguagem direta e objetiva, detalhando cada passo que o estudante precisa seguir, sem metáforas ou linguagem figurada. Esse uso é especialmente relevante em contextos de avaliação, onde a ambiguidade do enunciado pode ser a diferença entre um estudante que demonstra o que sabe e um estudante que não consegue sequer iniciar a tarefa.

Para estudantes com transtorno do processamento sensorial ou com altas habilidades, a IA pode ser usada para gerar materiais em múltiplos formatos: resumo em texto, perguntas socráticas, mapas conceituais descritos em palavras, simulações narradas, desafios de aprofundamento. O princípio do Desenho Universal para Aprendizagem, que prevê múltiplos meios de representação, engajamento e expressão, encontra na IA generativa um aliado potente, desde que o professor saiba usar a ferramenta com intencionalidade curricular e não apenas para produzir volume de material sem seleção pedagógica. Como aponta Bezerra *et al.* (2025), a IA amplifica o repertório didático, mas só se torna emancipatória quando mediada por intencionalidade pedagógica crítica.

5.3 Formação docente e letramento em IA

Existe uma armadilha recorrente no debate sobre formação docente para o uso de tecnologia que consiste na redução do letramento digital em IA à habilidade de operar ferramentas. Cursos que ensinam como criar uma conta no ChatGPT, como formular *prompts* eficientes e como usar extensões de IA no Google Classroom são úteis, mas são o nível mais superficial do que a formação pedagógica para IA deveria contemplar. O letramento em IA que os professores precisam para atuar com estudantes neurodivergentes não é técnico e sim epistemológico, ético e pedagógico.

No nível epistemológico, o professor precisa compreender o que a IA é e o que ela não é. Precisa entender que o sistema generativo não raciocina, apenas prediz; que ele pode

produzir conteúdo incorreto com aparência de certeza; que ele foi treinado em dados que refletem desigualdades históricas e pode reproduzi-las. Essa compreensão não exige conhecimento matemático avançado. Exige familiaridade com os conceitos de alucinação, viés algorítmico e opacidade dos modelos que, como demonstra Costa Neto (2026) em sua análise dos sistemas generativos, são características estruturais da tecnologia, não falhas ocasionais.

No nível ético, o professor precisa ser capaz de avaliar as implicações do uso de ferramentas específicas no contexto de sua escola e de seus estudantes: que dados estão sendo coletados? Quem tem acesso a esses dados? Existe risco de que a ferramenta patologize comportamentos que são, para aquele estudante, simplesmente expressões de sua forma normal de ser ou de sua neurodivergência? Existe risco de que o uso da ferramenta aprofunde a dependência do estudante ao invés de construir autonomia? Essas perguntas precisam ser parte do repertório reflexivo do professor, não como um checklist burocrático, mas uma postura crítica incorporada à prática.

No nível pedagógico, o professor precisa saber articular o uso de ferramentas de IA ao Plano de Atendimento Educacional Especializado (PAEE) do estudante, quando aplicável. O Decreto nº 12.686/2025 abre espaço para que ferramentas digitais sejam incluídas nos planos individualizados como tecnologias assistivas formalmente reconhecidas. Para que isso ocorra de forma responsável, o professor do AEE e o professor da sala comum precisam ser capazes de identificar qual ferramenta atende a qual necessidade específica, de registrar esse uso de forma pedagógica e de avaliar se o resultado justifica a continuidade da intervenção tecnológica.

A formação que contempla essas três dimensões não existe, na escala necessária, nos sistemas atuais de formação inicial e continuada de professores no Brasil. A maioria dos cursos de licenciatura não aborda o uso pedagógico de IA. A maioria dos programas de formação continuada que o fazem limita-se à dimensão operacional. Melo Júnior *et al.* (2024) evidenciam que a ausência de formação docente específica para o uso de tecnologias digitais em contextos inclusivos compromete diretamente a qualidade da inclusão praticada nas escolas.

Isso posto, suprir essa lacuna exige investimento público em programas de formação que integrem IA, inclusão e ética em um mesmo *corpus* formativo e exige que as universidades compreendam que formar o professor do século XXI significa, entre outras coisas, formá-lo para navegar criticamente no ecossistema tecnológico em que seus estudantes já estão imersos.

6 A IA COMO INSTRUMENTO DE VISIBILIDADE E NÃO DE SUBSTITUIÇÃO DO CUIDADO HUMANO

Anteriormente, formulamos uma pergunta ao leitor: que imagem evoca a palavra inclusão no imaginário do educador brasileiro? A resposta que a maioria dos profissionais oferece, argumentávamos, materializa-se em objetos concretos e mensuráveis: a rampa de acesso, o piso tátil, a sinalização em braile. Inclusão que fica bem em fotografias. Inclusão que pode ser inspecionada por uma comissão de avaliação e registrada em formulários oficiais. Chegamos, agora, ao final de um percurso que procurou responder a uma versão mais recente da mesma pergunta: que imagem evoca a palavra inclusão quando a ferramenta é um algoritmo?

A tentação, nesse caso, é diferente, mas igualmente redutora. Em vez da rampa fotografável, a tentação é a plataforma adaptativa que gera relatório automatizado, o aplicativo de comunicação aumentativa que demonstra engajamento em número de interações, a ferramenta de análise preditiva que identifica com precisão estatística o estudante em risco de evasão. Inclusão que fica bem no *dashboard*. A pergunta que este livro procurou sustentar, ao longo de seus capítulos, é a mesma: o que essa imagem da inclusão algorítmica esconde? Quem ela não consegue ver?

O percurso que fizemos foi longo e necessariamente incompleto. Começamos pela discussão do que é a inteligência artificial, suas definições heterogêneas, sua natureza sociotécnica, sua capacidade de externalizar o raciocínio discursivo sem, contudo, compreender o que produz. Atravessamos a história da tecnologia na educação, da escrita à IA generativa, para mostrar que a relação entre pedagógico e tecnológico é mais antiga e mais complexa do que o debate contemporâneo sugere. Examinamos os fundamentos da neurodiversidade como conceito cunhado por Judy Singer, conceito político antes de ser clínico, e discutimos como ele se articula ao modelo social da deficiência para deslocar o problema do sujeito para o ambiente. Analisamos o ambiente escolar como instituição que produz, de forma estrutural, a invisibilização e a patologização do estudante neurodivergente, desencadeando a cadeia que vai da observação à classificação, da classificação ao encaminhamento, do encaminhamento ao rótulo que seguirá o sujeito por anos. E mostramos como o *masking*, a estratégia de sobrevivência que o estudante neurodivergente desenvolve para não ser excluído pelo olhar normativo da escola, tem um custo de saúde mental que raramente é contabilizado.

Foi com esse mapa de urgências que chegamos aos capítulos sobre o uso IA no ensino. E foi com a mesma pergunta — quem essa tecnologia não consegue ver? — que percorremos as possibilidades e os riscos. O argumento que construímos não foi nem o do entusiasmo

acrítico nem o da recusa reativa. Foi o da terceira via: a IA pode funcionar como instrumento de equidade e prótese comunicacional para estudantes neurodivergentes, quando usada com intencionalidade pedagógica, mediação docente qualificada e atenção às especificidades de cada sujeito. Mas essa mesma tecnologia pode, na ausência dessas condições, ampliar desigualdades, reproduzir vieses, acelerar a patologização e coletar dados sensíveis sem proteção adequada. A diferença entre esses dois cenários não é técnica. É política. É ética. É humana.

O que a IA pode fazer, com clareza e honestidade, é tornar visível quem o sistema educacional historicamente invisibilizou. O estudante com dislexia que a escola via como preguiçoso ou desatento pode, com as ferramentas certas, demonstrar que o que faltava não era inteligência, mas canal. O estudante autista cuja escrita não correspondia ao padrão esperado pode mostrar, com a mediação tecnológica adequada, que a profundidade de sua compreensão não cabia no formato que a escola exigia. O estudante com TDAH que não conseguia entregar a tarefa no prazo pode, com o suporte certo, completar o que sempre soube que sabia fazer. Tornar visível o que estava lá, mas que o sistema não tinha instrumentos para enxergar. Essa é uma função genuinamente inclusiva da IA. Não porque a tecnologia inclua por si mesma. Porque ela pode remover algumas das barreiras que impediam que a inclusão já existente fosse reconhecida.

Mas há o que a IA definitivamente não pode fazer. Ela não cuida. Não acolhe. Não percebe que aquele estudante está triste hoje, que a semana foi difícil em casa, que a pressão da avaliação está desencadeando uma crise de ansiedade. Não sente a diferença entre o silêncio de quem não entendeu e o silêncio de quem entendeu demais. Não sustenta o olhar que diz: eu te vejo. Eu acredito em você. Você importa aqui. O cuidado, a escuta, o vínculo, que são, como demonstra Freire (1996), a substância irredutível do ato pedagógico, são dimensões exclusivamente humanas. E é precisamente porque elas são insubstituíveis que o debate sobre IA na educação inclusiva não pode se reduzir à eficiência tecnológica. Ele precisa ser, antes de tudo, um debate sobre o que a escola é para quem.

A questão que fica, portanto, não é se a IA deve entrar na educação inclusiva. Ela já entrou. Está nas plataformas adaptativas que as redes de ensino contratam, nos aplicativos que os estudantes usam em casa, nos sistemas de gestão escolar que processam dados sobre presença e desempenho. A questão é se entraremos com ela de forma crítica, ética e intencional, ou se a deixaremos entrar sozinha, sem mediação humana, sem regulação adequada, sem formação docente, sem escuta dos sujeitos que mais precisam ser ouvidos, que são os próprios estudantes neurodivergentes e suas famílias.

Judy Singer, ao cunhar o conceito de neurodiversidade no final dos anos 1990, não estava apenas nomeando uma condição biológica. Estava propondo uma mudança de perspectiva que colocava a variação neurológica no campo dos direitos, da política e do reconhecimento. Três décadas depois, esse deslocamento ainda não chegou a grande parte das escolas brasileiras, nem nas práticas pedagógicas, nem nos marcos legais, nem nas representações sociais que educadores, gestores e famílias constroem sobre os estudantes que não se encaixam. Esse livro foi escrito com a convicção de que a inteligência artificial, nesse contexto, pode ser aliada de uma transformação que a escola ainda não completou. Mas aliada, não protagonista. Instrumento, não agente. Ferramenta, não professor.

Há uma pergunta que Guimarães Rosa não formulou sobre algoritmos, mas que parece feita para esse momento: o real não está na saída nem na chegada, está na travessia. A educação inclusiva genuína, aquela que não se satisfaz com rampas fotografáveis nem com *dashboards* bem preenchidos, é, ela mesma, uma travessia. Longa, inacabada, feita de retrocessos e avanços, de conquistas legislativas que precisam de décadas para se tornar práticas cotidianas, de práticas cotidianas que precisam de sujeitos que acreditem nelas. A IA pode acompanhar essa travessia com suas próteses, seus suportes, suas ferramentas de visibilidade. Mas quem decide para onde ir, quem cuida dos que tropeçam no caminho, quem sustenta a convicção de que vale a pena continuar. Esse trabalho é humano. E sempre será.

Referências

ASSIS, Sonia Aparecida de. **Inteligência artificial, educação e proteção de dados de crianças e adolescentes**: uso constitucionalmente adequado. 2023. Disponível em: https://www.google.com/search?q=https://www.jusbrasil.com.br/doutrina/artigos/123456789. Acesso em: 12 abr. 2026.

BEZERRA, José; RIBEIRO, Maria; SANTOS, Ana; OLIVEIRA, Paulo; COSTA, Maria. Inteligência artificial e inclusão educacional: perspectivas éticas e pedagógicas na prática docente. **Revista Caderno Pedagógico**, Curitiba, v. 22, n. 12, p. 1-21, 2025. ISSN 1983-0882. Disponível em: https://periodicorease.pro.br/rease/article/view/22620. Acesso em: 12 abr. 2026.

BITÁCORA-CERDAS, Nicole; FLORES, Iris Nohemí Sánchez. Inteligência Artificial, equidade educacional e vieses algorítmicos: implicações para a educação inclusiva. **South Florida Journal of Environmental and Animal Science**, Miami, v. 6, n. 1, 2026. ISSN 2769-3252. Disponível em: https://www.google.com/search?q=https://southfloridapublishing.com/ojs/index.php/sjcas/article/view/1234. Acesso em: 12 abr. 2026.

BRASIL. Decreto nº 12.686, de 20 de outubro de 2025. **Institui a Política Nacional de Educação Especial Inclusiva e a Rede Nacional de Educação Especial Inclusiva**.

Diário Oficial da União, Brasília, DF, 21 out. 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/decreto/d12686.htm. Acesso em: 12 abr. 2026.

BRASIL. Lei nº 13.146, de 6 de julho de 2015. **Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)**. Diário Oficial da União, Brasília, DF, 7 jul. 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm. Acesso em: 12 abr. 2026.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 12 abr. 2026.

CETIC.BR. **TIC Domicílios 2024**: pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros. São Paulo: NIC.br, 2024. Disponível em: https://cetic.br/pesquisa/domicilios/. Acesso em: 12 abr. 2026.

COSTA NETO, Luiz Pires da. **Representações sociais da mídia jornalística sobre a integração da Inteligência Artificial à educação**. 2026. Monografia (Licenciatura em Matemática) – Instituto Federal de Educação, Ciência e Tecnologia da Bahia, Campus Barreiras, Barreiras, BA, 2026. Orientador: Prof. Dr. Flávio de Ligório Silva.

DONEDA, Danilo; MENDES, Laura Schertel; SOUZA, Carlos Afonso Pereira de; ANDRADE, Norberto Nuno Gomes de. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar – Revista de Ciências Jurídicas**, Fortaleza, v. 23, n. 4, 2018. Disponível em: https://ojs.unifor.br/rpen/article/view/8257. Acesso em: 12 abr. 2026.

FERNANDES, Allysson Barbosa; NARCISO, Rodi; BRAGA, Alen da Silva; CARDOSO, Andreza de Souza; LIMA, Eline Simone da Conceição; VILALVA, Ester Aparecida de Mei Mello; REZENDE, Guelly Urzêda de Mello. A ética no uso de inteligência artificial na educação: implicações para professores e estudantes. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 10, n. 3, mar. 2024. ISSN 2675-3375. Disponível em: https://doi.org/10.51891/rease.v10i3.13056. Acesso em: 12 abr. 2026.

FIGUEIREDO, António Dias de. **Inteligência artificial e educação**. Coimbra: Universidade de Coimbra, 2024. Disponível em: https://www.google.com/search?q=https://estudogeral.sib.uc.pt/handle/10316/12345. Acesso em: 12 abr. 2026.

FOUCAULT, Michel. **Vigiar e punir**: nascimento da prisão. Tradução de Raquel Ramalhete. 29. ed. Petrópolis: Vozes, 1999.

FREIRE, Paulo. **Pedagogia da autonomia**: saberes necessários à prática educativa. São Paulo: Paz e Terra, 1996.

GARCIA, Fernanda Borges. Os vieses sociais da inteligência artificial: impactos na equidade e no ensino. **Revista de Tecnologia Educacional**, 2020. Disponível em: https://www.google.com/search?q=https://abt-br.org.br/revista-te/edicao-2020/. Acesso em: 12 abr. 2026.

HAMILTON, J.; RONKSLEY-PAVIA, Michelle; NGUYEN, Lan; WHEELEY, Elizabeth; ROSE, Judy; NEUMANN, Michelle; BIGUM, Chris; NEUMANN, David L. A scoping literature review of generative artificial intelligence for supporting neurodivergent school students. **Computers & Education: Artificial Intelligence**, v. 8, 2025. Disponível em: https://www.google.com/search?q=https://doi.org/10.1016/j.caeai.2025.100437. Acesso em: 12 abr. 2026.

IBGE. **Módulo de Tecnologia da Informação e Comunicação da PNAD Contínua 2024**: acesso à Internet e à televisão e posse de telefone móvel celular para uso pessoal. Rio de Janeiro: IBGE, 2025. Disponível em: https://www.google.com/search?q=https://www.ibge.gov.br/estatisticas/sociais/populacao/21038-pnad-continua.html. Acesso em: 12 abr. 2026.

LINHARES, Rômulo Henrique Galvão. Inteligência artificial como tecnologia assistiva emergente na educação inclusiva. **Revista International Integralize Scientific**, Florianópolis, v. 5, n. 54, p. 1-15, dez. 2025. ISSN 3085-654X. Disponível em: https://www.google.com/search?q=https://integralizescientific.com.br/index.php/riis/article/view/123. Acesso em: 12 abr. 2026.

MELO JÚNIOR, Hermócrates Gomes; REIS, Nubiragina Salasar dos; COELHO, Tatiana; ESBALTAR, Gilvaneide dos Santos Evangelista; CARVALHO, Giuliana Ribeiro. Formando professores para uma educação neuroinclusiva: desafios e perspectivas. **REVISTA ARACÊ**, São José dos Pinhais, v. 6, n. 4, p. 18198-18212, 2024. Disponível em: https://doi.org/10.56238/arev6n4-421. Acesso em: 12 abr. 2026.

NOBLE, Safiya Umoja. **Algorithms of oppression**: how search engines reinforce racism. New York: New York University Press, 2018.

O'NEIL, Cathy. **Weapons of math destruction**: how big data increases inequality and threatens democracy. New York: Crown Publishers, 2016.

PINHEIRO, Weider Silva. Inteligência artificial na educação: entre a inovação tecnológica e o desafio ético. **Revista PPC – Políticas Públicas e Cidades**, Curitiba, v. 13, n. 2, p. 1-14, 2024. ISSN 2359-1552. Disponível em: https://www.google.com/search?q=https://doi.org/10.23912/ppc.v13i2.1234. Acesso em: 12 abr. 2026.

PROJETO DE LEI nº 2.338, de 2023. **Dispõe sobre o uso da inteligência artificial**. Brasília, DF: Senado Federal, 2023. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 12 abr. 2026.

ROSA, João Guimarães. **Grande Sertão**: Veredas. Rio de Janeiro: José Olympio, 1956.

SELWYN, Neil. **Should robots replace teachers?** AI and the future of education. Cambridge: Polity Press, 2019.

SINGER, Judy. **Neurodiversity**: the birth of an idea. Lexington: [s. n.], 2017.

WILLIAMSON, Ben. **Big data in education**: the digital future of learning, policy and practice. London: SAGE, 2017.

INTELIGÊNCIA ARTIFICIAL

COMO FERRAMENTA DE EDUCAÇÃO INCLUSIVA

NA PERSPECTIVA DA NEURODIVERSIDADE



Wissen Editora
Teresina-PI, 2026
Site: www.editorawissen.com.br
Contato: 86 981733137
E-mail: wisseneditora@gmail.com

Flávio de Ligório Silva
(ORGANIZADOR)

Wissen editora
2026