

Estado da publicação: O preprint não foi submetido para publicação

Por uma Revisão Crítica do Uso de Inteligência Artificial na Educação

Rodrigo Esteves de Lima-Lopes

<https://doi.org/10.1590/SciELOPreprints.11450>

Submetido em: 2025-03-10

Postado em: 2025-03-17 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o endosso de:

Suzi Marques Spatti Cavalari (ORCID: <https://orcid.org/0000-0001-7748-8516>)

1 Por uma Revisão Crítica do Uso de Inteligência

2 Artificial na Educação

3 For a Critical Review of Artificial Intelligence in Education

4 Rodrigo Esteves de Lima-Lopes
5 <https://orcid.org/0000-0003-3681-1553>
6 Universidade Estadual de Campinas/CNPq

7
8 **Resumo:** Este artigo reflete criticamente sobre as implicações da introdução dos Large Language Models ou
9 Grandes Modelos Linguísticos, popularmente chamados de Inteligência Artificial, no contexto do ensino. O
10 ensaio se inicia com uma discussão sobre como estes modelos são criados. Parte-se de uma definição do papel
11 dos dados, refletindo sobre o mito de sua objetividade e como sua seleção e organização podem criar
12 assunções ideológicas e subjetivas. O papel do humano no processo de treinamento também é discutido para
13 desmitificar a perspectiva de que estes modelos sejam inteligências superiores e autônomas, pois, na verdade,
14 elas partem de modelos estatísticos de treinamento. Discute-se, então, o papel dos diferentes modelos de
15 Inteligência Artificial em nossa sociedade, com especial ênfase nos conceitos de apagamento tecnológico e
16 singularidade tecnológica, que implicam na assimilação tecnológica no cotidiano em contraponto como
17 maravilhamento momentâneo sem a necessária criticidade. Caminha-se para discutir os LLMs como parte de
18 nosso sistema tecnológico-social, com ênfase em seus assentimentos ideológicos e práticos. O artigo se encerra
19 com algumas reflexões sobre caminhos a serem seguidos se pensarmos no uso dos LLM em situações de
20 ensino.

21 **Palavras-chave:** Tecnosfera, Inteligência Artificial, Tecnologia, Ensino, Dados.

22 **Abstract:** This article reflects critically upon the implications of introducing Large Language Models,
23 popularly called Artificial Intelligence, in Brazilian Education. This essay begins with a discussion of how
24 these models are created. It starts with a definition of data and their role in such models, pointing to the
25 myth of its objectivity and how its selection and organisation might lead to ideological and subjective
26 assumptions. The need for human intervention in the training process is also discussed to demythologise the
27 idea that these models are superior and autonomous intelligences, as they are in fact based on statistical
28 training models. The role of the different Artificial Intelligence models in our society is then discussed. The
29 discussion relies on the concepts of technological invisibility and technological singularity, which imply
30 technological assimilation in everyday life over momentary wonder without the necessary criticality. Finally,
31 the article discusses LLMs as part of our techno-social system, emphasising their ideological significance and
32 practical impacts. The final section discusses some paths to be considered if we think about Large Language
33 Models presence in education situations.

34 **Key-words:** Technosphere, Artificial Intelligence, Technology, Education, Data.

35 1. Introdução

36 A inteligência artificial (IA) tem gerado transformações significativas nas dinâmicas
37 técnico-sociais. Embora ela seja um tema de pesquisa há pelo menos 70 anos, com o
38 desenvolvimento dos chamados *Large Language Models* ou Grandes Modelos Linguísticos
39 (doravante LLMs) ela se tornou algo ubíquo, ganhando destaque em nossas discussões

1 cotidianas, na mídia de massa e em novas áreas da ciência e no ensino. Em tais discussões,
2 ela tende a ser tratada controversamente; ora como uma ferramenta revolucionária, ora
3 como uma ameaça.

4 No contexto da educação pública do estado de São Paulo, a partir de 2023, a
5 Secretaria de Educação do Estado vem realizando diversos movimentos em favor da
6 transformação de processos tecnológicos associados ao cotidiano das escolas paulistas.
7 Inicialmente, as mudanças ocorreram pelo anúncio da não participação do Estado no
8 Programa Nacional do Livro Didático (Ferneda; Bernardes, 2023) em favor da preparação
9 por professores conteudistas de slides não impressos e disponíveis à comunidade escolar
10 apenas digitalmente (Seduc, 2023a). Houve críticas em relação ao conteúdo de tal material
11 (Gielow, 2023), assim como sua indisponibilidade em forma impressa, que poderia levar à
12 exclusão de milhares de alunos que não possuem infraestrutura computacional para leitura
13 e acompanhamento das matérias. Como resultado, o Governo do Estado anunciou a
14 revisão do material (Cafardo, 2023a), assim como sua disponibilização para professores e
15 alunos (Cafardo, 2023b). Posteriormente, críticas também levaram a uma revisão da
16 postura em relação à não participação no PNLD (Porto, 2023).

17 O anúncio da utilização de LLMs¹ ocorreu, em um primeiro momento, como
18 ferramenta corretora de redações (Seduc, 2023b) sendo, então, proposta sua utilização na
19 revisão dos referidos materiais produzidos pelo Estado (Freitas; Machado, 2024). Se por
20 um lado, parte da sociedade civil, encarou este uso com ressalvas (APEOESP, 2024; Silva,
21 2024), apoiando-se, principalmente, nas possibilidades de automação que
22 descaracterizariam o componente criativo e humano do aprendizado, por outro, houve
23 censura à posição crítica de professores em relação ao uso de tais ferramentas (Globo,
24 2024). O discurso comum carrega cores de revolução no ensino (Lourenço, 2023),
25 especialmente com a ideia de personalização (Litwak; Gama, 2024) e otimização. Este
26 discurso, normalmente, é veiculado por publicações que não são especializadas na área de
27 educação e reforçam as ideias de oportunidade e avanço. Poucos são os veículos de mídia
28 que têm um olhar crítico, mesmo que superficialmente (Balmant, 2024; Paes, 2024).

29 Por conta deste contexto, tomo por objetivo refletir sobre os possíveis efeitos da
30 introdução dos LLMs como componente integrante do currículo. Este artigo é resultado de
31 minha participação em dois eventos, nos quais tive a oportunidade de discutir os LLMs no
32 contexto da educação: a *Semana de Letras*, promovida pelo Campus São José do Rio Preto
33 da Universidade Estadual Paulista-UNESP e o *CoPeD, Colóquio de Pesquisas Discentes*,
34 promovido pelo Programa de Pós-Graduação em Filologia e Língua Portuguesa da
35 Universidade de São Paulo-USP. Apesar da discussão ter seu foco principal nos impactos
36 dessa implementação na área de Linguística Aplicada, acredito que ela pode ser facilmente
37 extrapolada para outros contextos, uma vez que a pressão social para inclusão rápida, e
38 por vezes sem que a devida reflexão seja realizada, parece-me algo generalizado. Busca-se,

1 ¹ Tende-se a utilizar o termo inteligência artificial (ou IA) em referência ao desenvolvimento de LLMs, quando, na
2 verdade, o termo se aplica a uma série de aplicações que já fazem parte de nossa vida cotidiana há décadas.

1 assim, trazer um olhar crítico a partir da tradição de pesquisas em tecnologia no âmbito
2 da Linguística Aplicada. De fato, poucas são as discussões sobre a natureza de seu
3 treinamento, dos dados que fazem parte dos motores geradores de respostas e, certamente,
4 das implicações do uso de uma tecnologia na qual o professor, assim como o aluno, não
5 tem poder de curadoria e seleção. Afinal, como colocam Gonsales e Kaufman (2023), tal
6 tecnologia parece mudar o eixo das perspectivas do uso de Inteligência Artificial,
7 especialmente se pensarmos que ela sugere uma concepção de utilização baseada na
8 extração de padrões e de conhecimento da máquina e não em saber programá-la para
9 poder-se, assim, extrair resultados desejados. Por outro lado, são comuns discussões sobre
10 o uso de tal ferramenta em sala de aula, com foco principal em sua instrumentalidade.

11 Convido o leitor a dividir um entendimento no qual aquilo a que chamamos
12 inteligência artificial, na verdade, é um conjunto de aplicações baseadas em dados. O
13 termo inteligência artificial surgiu durante um encontro de verão na Universidade de
14 Dartmouth em 1956, no qual pesquisadores definiram quais seriam os caminhos de
15 desenvolvimento na área naquele momento. Conforme a proposta original de pesquisa, o
16 segundo objetivo de desenvolvimento do projeto de IA seria: *2. Como um computador*
17 *poderia ser programado para usar uma linguagem* (McCarthy *et al.*, 2006, p. 12, tradução
18 minha).² Ou seja, aquilo que chamamos de IA procura a produção de linguagem como um
19 de seus primeiros objetivos.

20 Os LLMs são treinados. Em outras palavras, passam por um sistema de programação
21 para as respostas, o qual é alcançado via modelos estatísticos. Este treinamento depende
22 de uma tríade complementar: dados, modelos estatísticos e reforço humano (Brown *et al.*,
23 2020; Christiano *et al.*, 2023; Lee *et al.*, 2024; Liu *et al.*, 2019; Stiennon *et al.*, 2022). Os
24 dados, por maiores ou mais compreensivos que o sejam, representam uma seleção
25 naturalmente exclusiva (Gitelman, 2013). Além disso, as diferentes formas de manipulação
26 e estruturação destes dados são ações subjetivas. Elas dependem de nossa concepção
27 daquilo que seja aceitável como dados, algo ligado ao nosso momento histórico-científico e
28 à área da ciência na qual um pesquisador se filia. Nenhum banco de dados representa a
29 totalidade de um fenômeno, a não ser em contexto específico e, mesmo assim, as formas de
30 coleta e análise podem estar embebidas por questões ideológicas diversas. Os modelos
31 estatísticos representam a matematização dos dados, permitindo a tradução para uma
32 linguagem que o computador compreenda. Isso é essencial para que se criem grandes
33 tabelas de co-ocorrência e co-relação que permitem ao programa metrificar a possibilidade
34 de uma palavra ocorrer ao lado da outra. Desse processo nasce a possibilidade de a
35 máquina construir frases e oferecer respostas a nossos questionamentos. Suas respostas são,
36 por conseguinte, baseadas nesse processo de associação que parte das medidas extraídas de
37 seu processamento. Contudo, esse processamento é incapaz de filtrar os resultados a partir
38 daquilo que acreditamos ser cultural e socialmente aceitável. É nesse ponto que a correção
39 humana entra em cena. Ela pode ser realizada de duas formas principais, uma utiliza

1 ² Do original: *2. How Can a Computer be Programmed to Use a Language* (McCarthy *et al.*, 2006, p. 12).

1 corpora já etiquetados semanticamente — oferecendo à máquina exemplos — por correção
2 direta — pessoas corrigem a máquina nas questões apresentadas. Em alguns casos, o
3 usuário é questionado diretamente qual resposta acha mais adequada, um modelo muito
4 similar ao estímulo vs resposta.

5 Do ponto de vista social, pode-se observar que instâncias de inteligência artificial já
6 fazem parte de nosso cotidiano há tempos. De maneira geral, ela tende a ser apagada por
7 um processo de invisibilidade tecnológica (Hayles, 2005). Um fenômeno que parece estar
8 relacionado à incorporação de tecnologias a nossas tarefas diárias, inicialmente como
9 facilitadoras e, posteriormente, como principais executoras. Algo importante a ser
10 observado é que, ao desaparecer, uma tecnologia deixa de ser vista como tal, fazendo
11 parte, agora, de um ecossistema técnico-social. Quando uma tecnologia é inserida neste
12 ecossistema, ela tende a causar impactos em nossa percepção. É nesse ponto que a ideia de
13 singularidade tecnológica ganha força, pois projetamos nesta tecnologia em uma
14 perspectiva futurística e de revolução (Coeckelbergh, 2020a). Ou seja, como já colocado em
15 Lima-Lopes (2023), ao tomarmos uma tecnologia somente pela ideia de *nova*, estamos
16 simplesmente afirmando seu caráter histórico recente e deixando de lado sua perspectiva
17 política. Assim, em uma contingência óbvia, o novo, um dia, tornar-se-á invisível, e uma
18 nova singularidade tecnológica sempre surgirá. Ou seja, boa parte da questão para a qual
19 buscamos solução está em encontrar um modelo alternativo de percepção da IA. Tê-la
20 como uma tecnologia “must have” e como agente da modernidade inovadora, poderá não
21 ser a melhor abordagem. Qual caminho alternativo podemos percorrer?

22 Para realizar tal reflexão, este ensaio se inicia com uma discussão sobre definição de
23 inteligência artificial. Parte-se de uma observação de como os LLMs se encaixam na
24 paisagem tecnológica das Inteligências Artificiais, discutindo sua forma de treinamento e o
25 papel que os dados possuem em sua formação (ver seção 2). A relação dos LLMs no ensino
26 é o tema da seção 3, na qual é explorada a ideia tradicionalmente difundida em nossa
27 cultura midiática ou de massa, observando-se como ela parece moldar nossa percepção
28 sobre o que seja inteligência artificial. Destaca-se, também, sua presença em nosso
29 cotidiano sócio-tecnológico, observando-se questões críticas que vêm sendo discutidas pela
30 comunidade na seção 3.2. Por fim, na seção 4 encerro com propostas de discussão dos
31 LLMs, com pontos gerais que podem ser relevantes para o porvir desta tecnologia na
32 educação linguística e, talvez, em nossa sociedade.

33 2. De onde vem a inteligência?

34 A percepção cotidiana da IA está enraizada em representações que a humanizam. A
35 literatura, o cinema e a televisão nos apresentam robôs antropomórficos que possuem
36 emoções e dilemas morais. Exemplos não nos faltam, Rick Deckard e sua busca pelo
37 autoconhecimento em *Blade Runner*; a dupla de robôs em *Star Wars - C3PO* e sua
38 insegurança traduzida em sua voz hesitante ou a comovente lealdade de R2D2 -; ou

1 mesmo Data em *Star Trek*, que luta por aprender emoções como humor e medo. Talvez o
2 computador Hall 9000, em 2001 - *Uma Odisseia no Espaço*, seja a mais humanizada das
3 personagens do filme. Ele elimina seus inimigos, barganha por sua vida e tenta convencer
4 os humanos, que estão prestes a desligá-lo, de que mudou e não mais incorrerá nos mesmos
5 erros.

6 Percebemos, também, tais máquinas como superiores, capazes de realizar tarefas mais
7 competentemente que nós humanos. Pelo menos é o que fatos como as derrotas do
8 campeão mundial de xadrez Garry Kasparov para o *Deep Blue*, computador construído
9 pela IBM, e do campeão mundial de Go Lee Sedol para o *AlphaGo*, sistema desenvolvido
10 pela DeepMind, uma subsidiária da Google, parecem nos ensinar. Todavia, há algumas
11 diferenças e semelhanças importantes entre os dois eventos, além do tempo que os
12 separam. O *Deep Blue* foi um computador criado para uma função específica: jogar xadrez.
13 Sua capacidade de vencer derivou de sua força bruta de processamento (IBM, [s. d.]).
14 Como explicam Campbell et al. (2002), alguns fatores contribuíram para seu sucesso, em
15 especial um sistema de processamento paralelo (no qual diversos núcleos dão conta da
16 informação simultaneamente) maciço para sua época. Ele contava com um complexo
17 sistema de avaliação e busca de informações, um sistema híbrido que operava tanto pelo
18 hardware quanto pelo software. O treinamento partiu de dezenas de milhares de partidas,
19 as quais foram matematizadas. O computador realizava, então, um robusto processo de
20 busca, processamento e cálculo estatístico para realizar a melhor jogada.

21 *AlphaGo* também foi criado para uma função dedicada (DeepMind, 2024). Seu
22 treinamento foi baseado em uma combinação do aprendizado supervisionado por meio de
23 intervenção humana. Jogadores de Go aplicavam correções nos sistemas de escolha do
24 aplicativo, que eram baseados em buscas, para simular centenas de milhares de jogos
25 (Silver et al., 2016). Ou seja, a máquina tinha dados de partidas inseridas por humanos
26 que trabalham com sistema de reforço de comportamentos desejados no algoritmo, sendo
27 insumo para o cálculo de jogadas e respostas possíveis em simulações previstas pela
28 máquina. Diferentemente da máquina da IBM, a equipe da DeepMind utilizou a mesma
29 tecnologia de treinamento para construção de outros sistemas, como a aplicação em jogos
30 de tabuleiro como shogi e xadrez (Silver et al., 2018), predição e mapeamento de estrutura
31 de proteínas (Jumper et al., 2021) e melhoria de algoritmos complexos de treinamento
32 (Fawzi et al., 2022).³

33 Apesar das diferenças, as semelhanças são igualmente importantes. *AlphaGo* e *Deep*
34 *Blue* possuem muito em comum: ambos partem de dados.

35 Dados podem ser definidos como processos de matematização do comportamento
36 humano ou da natureza (Kitchin, 2014a, 2014b), construindo representações. Todas as
37 inteligências artificiais partem de algum tipo de dado. Mídias sociais, por exemplo,
38 mensuram não somente os cliques, mas também o tempo que gastamos em postagens,

1 ³ O *AlphaGo* usa uma tecnologia de treinamento muito similar à aplicada pelo *Chat GPT* e outras LLMs, que serão
2 discutidas a seguir.

1 aquilo que escrevemos e repostamos, nossas reações e toda sorte de ação. Estes dados são
2 cruzados com os oriundos de nossa lista de contatos, ajudando a construir um perfil nosso
3 e da comunidade que realizamos (O’Neil, 2016). Tal cruzamento é responsável pela criação
4 das diversas bolhas, que nos cercam por conteúdos que reforçam comportamentos. O
5 varejo, outro exemplo, colhe nossos dados de navegação para nos oferecer diferentes
6 produtos e serviços, criando verdadeiros perfis de consumo, em uma perspectiva chamada
7 de capitalismo de vigilância (Zuboff, 2019). A Google disponibilizou de forma anônima
8 relatório de mobilidade urbana durante a pandemia de Covid-19 (Google, 2022), com
9 dados colhidos a partir dos telefones que usam sistema operacional Android ou os
10 aplicativos de mapa disponibilizados pela *big tech*. Tais dados, podiam ser cruzados com
11 aqueles sobre infecção disponibilizados pelos órgãos públicos, criando mapas de mobilidade
12 e alastramento do vírus.⁴ Nesta mesma época houve uma forte migração para sistemas
13 digitais (automatizados e semi-automatizados) de gestão escolar, incluindo atos
14 administrativos e pedagógicos. Tais sistemas de gestão monitoram atividades de alunos e
15 professores, gerando metadados e levantam questões sobre privacidade e autonomia.

16 A integração da IA no cotidiano é evidente em fenômenos como a plataformização
17 (Poell; Nieborg; van Dijck, 2019). Plataformas digitais acumulam dados de usuários, desde
18 cliques até interações sociais, criando perfis detalhados que influenciam o comportamento
19 individual e coletivo. Entretanto, estes dados, como todos, possuem recortes ideológicos e
20 sociais. Ao incluir todos os usuários de um sistema ou aplicativo, está-se automaticamente
21 excluindo todos os que nele não interagem. A fortuna destes sistemas, assim como hábitos
22 de uso, determinam qual a qualidade e natureza dos dados colhidos em determinadas
23 regiões.

24 Para Gitelman e Jackson (2013) e Manovich (2011), os dados que servem como
25 insumo para pesquisas e para os diversos bancos de dados não existem naturalmente, eles
26 são gerados e refletem os objetivos que norteiam este processo. Isso significa que, em
27 diferentes contextos, materiais diversos podem ser considerados dados. As autoras usam a
28 literatura como exemplo: intelectuais da área de estudos literários, ou mesmo de outras
29 artes, têm dificuldades em encerrar seus objetos de pesquisa como dados. Entretanto, se
30 estes textos forem inseridos em um banco, serão tratados como tal. De fato, o que
31 consideramos como dados e como os organizamos é algo dependente da disciplina em que
32 trabalhamos e do momento histórico em que vivemos. Alguns fatores, como a forma de
33 armazenamento, as metodologias de compilação, as tecnologias de armazenamento de
34 disponibilização e até mesmo o que escolhemos observar, são elementos determinantes do
35 recorte. Tal perspectiva parece, então, romper com a ideia corrente de que algo possa ser
36 “descoberto”, que algo estava ali nos dados esperando para ser observado. Resultados são

1 ⁴ Importante colocar que muitos destes dados são coletados pelo uso das plataformas e autorizados pelo usuário ao
2 concordar com termos de uso. Discussões críticas sobre tal perspectiva podem ser explorados em Lima-Lopes (2023), no
3 qual discuto questões referentes ao letramento de dados e sua importância para a contemporaneidade e em Singh (2019),
4 que tece uma importante crítica a falácia da Web 2.0, sobre especialmente a pseudo liberdade e gratuidade que ela
5 representa.

1 construídos, condicionados e têm um claro recorte técnico-histórico-temporal ([Gitelman;](#)
2 [Jackson, 2013](#)).

3 Tal perspectiva, por conseguinte, rompe com a ideia de que dados são representativos
4 e neutros ([Boyd; Crawford, 2012](#)), eles podem ser grandes em volume e exaustivos em um
5 universo limitado ([Kaplan, 2015; Kitchin, 2014a](#)), mas eles sempre deixarão de representar
6 algum comportamento não mapeado. Tal realidade dá a qualquer análise quantitativa
7 cores qualitativas. Como todas as atividades humanas, dados são o resultado de operações
8 limitadas e exibem uma perspectiva social e política intrínseca. Há uma escolha
9 relacionada ao que olhar e como olhar; na forma como organizamos, disponibilizamos e
10 interpretamos. Ao aceitar-se tal subjetividade, aceita-se que a interpretação não se inicia
11 após a coleta e organização dos dados, ela está presente desde o planejamento da pesquisa.

12 **2.1 Como as máquinas aprendem?**

13 Se todas as IAs partem de dados para sua existência, qual seria sua origem? Como
14 eles alimentam os sistemas em LLM? Como esta subjetividade pode impactar seus
15 resultados?

16 Segundo o blog da OpenAI ([Yadav, 2024](#)), o treinamento do *Chat GPT*, que
17 provavelmente se estende a todos os modelos de LLM presentes na atualidade, parte de
18 grandes volumes de texto disponíveis publicamente na internet ou licenciados. Tais
19 volumes de texto abrangem áreas e idiomas diversos, constituindo, talvez, trilhões de
20 tokens (palavras diferentes). Esses textos são transformados em vetores, matemáticos que
21 representam os dados textuais de forma que o computador consiga extrair padrões. Tais
22 padrões são, posteriormente, re combinados de acordo com suas probabilidades. O
23 treinamento de um modelo de LLM segue algumas fases.⁵

24 **Pré-processamento de dados**

25 Neste primeiro momento, os dados textuais são representados em espaços vetoriais,
26 que buscam padrões sobre a combinação de palavras. Buscam-se os grupos nominais (NP)
27 e os grupos verbais (VP) para observar quais itens lexicais estão *embedded*, ou melhor,
28 embutidos em cada grupo. Buscam-se informações sobre composições de n-gramas,⁶
29 possibilitando a observação de quais combinações são mais raras e mais comuns por meio
30 da inserção dos n-gramas em que uma palavra se apresenta como metainformação em um
31 vetor ([Bojanowski et al., 2017](#)). Nesta fase também são retiradas informações que não são
32 necessárias para a construção do modelo. Por exemplo, dados retirados da internet

1 ⁵ Não é objetivo dar grandes detalhes técnicos do processo, a ideia central é tentar abrir a caixa de pandora e refletir
2 sobre as fases de treinamento.

3 ⁶ Em Linguística do Corpus, N-gramas são sequências de palavras em um texto definidas por “N”, que indica o número
4 de palavras ou considerados na sequência. Ou seja, 3-gramas (tri-gramas) seriam três palavras ou caracteres, 2-gramas
5 (bi-gramas) seriam sequências de duas palavras ou caracteres e assim por diante ([Lima-Lopes, 2016](#)).

1 costumam ter pedaços de códigos HTML⁷ e, por vezes, apresentar diferenças na codificação
2 de caracteres,⁸ ambos causando problemas de leitura e nas estatísticas.

3 **Arquitetura do modelo**

4 Os LLMs partem de um modelo estatístico/computacional chamado de *transformer*
5 definido por Vaswani e colaboradores (2023). Neste modelo de processamento, o foco está
6 nos conjuntos de n-gramas considerados estatisticamente importantes em uma sentença ou
7 frase, direcionando a atenção da máquina para eles. Esse tipo de forma de processamento
8 torna a resposta da máquina mais eficiente — ou seja, mais rápido e mais próximo de uma
9 associação reconhecível por um falante. Em outros trabalhos, como Brown e colaboradores
10 (2020) e Radford e Narasimhan (2018), este modelo é expandido para não mais precisar de
11 treinamento específico para cada tipo de tarefa que pedimos a um LLM, pode-se agora
12 gerar uma resposta a partir do treinamento com uma grande massa de textos. Ou seja, não
13 se precisa mais informar para o computador quais textos específicos de uma temática, o
14 cálculo estatístico tentará resolver a questão.

15 **Treinamento do modelo de linguagem**

16 A próxima etapa seria o treinamento, em que os dados são inseridos em um sistema
17 que os processa, representando o momento em que a máquina é efetivamente, é alimentada
18 com os textos limpos, e seus vetores.

19 A forma com que os LLMs contemporâneos são treinados parte da metodologia
20 definida por Devlin e colaboradores (2019), estabelecendo o BERT (*Bidirectional Encoder*
21 *Representations from Transformers*) como metodologia básica para estes contextos. Eles
22 partem do *transformer* (Vaswani *et al.*, 2023) desenvolvendo um sistema que “lê” as
23 palavras tanto da esquerda para direita como da direita para a esquerda, refinando o
24 sistema de estatística e de vetorização. Os dados são alimentados e o treinamento ocorre,
25 utilizando-se uma técnica chamada MLM (Masked Language Modeling). Essa técnica,
26 inspirada nos testes Cloze (Taylor, 1953)⁹ nos quais palavras aleatórias de um texto são
27 apagadas, resulta em vetores mais difusos, permitindo que a leitura em ambos os sentidos
28 preveja mais assertivamente as palavras. Então, o modelo é iniciado com os dados da fase
29 anterior utilizando-se dados etiquetados, os resultados são avaliados em seu desempenho e

1 ⁷ HTML (HyperText Markup Language) é uma linguagem utilizada para criar os textos que vemos na internet. Ela
2 funciona como uma série de instruções (ou marcações) em um arquivo texto simples, sendo responsáveis por permitir que
3 nos navegadores (como Firefox) leiam e formatem o que vemos. Como essas instruções (ou marcações) são escritas
4 diretamente no texto simples, muitas vezes eles continuam no material coletado, “sujando” o texto.

5 ⁸ A codificação é a tradução de caracteres que compõe um texto — como letras (a, b), números (10, 100), símbolos (&
6 %) e caracteres especiais (ê, ç) — em seqüências de números; cada uma delas é única e identifica o caractere no sistema.
7 Entretanto, há diversos padrões, baseados na língua, em fabricantes de computadores/programas e consórcios ou acordos
8 diversos. Quando lemos um texto no computador que não possui o mesmo padrão de caracteres que um texto, temos
9 erros de interpretação desses códigos, renderizando o texto ilegível, por vezes. Nos últimos anos, há um esforço de
10 padronização no sistema UTF-8, que suporta a maioria das línguas.

11 ⁹ Um teste Cloze demanda do aprendiz preencher lacunas em um texto, tradicionalmente utilizado em avaliações da
12 capacidade leitora em língua estrangeira.

1 refinados (Devlin *et al.*, 2019). Dados específicos podem ser utilizados para a validação de
2 resultados.

3 **Validação humana e ajustes**

4 A validação humana entra como parte integrante do processo de constituição de
5 LLMs, solucionando problemas que a máquina não conseguiria realizar sozinha. Por
6 exemplo, como um LLM poderia saber que uma combinação de palavras é algo tabu em
7 determinadas regiões do Brasil? É neste momento que a correção humana entra em cena,
8 sendo que ela pode se dar de duas formas: 1) Pela utilização de corpora semanticamente
9 anotados e 2) pela intervenção direta em respostas do modelo.

10 Bowman e colaboradores (2015) introduzem um modelo de intervenção humana por
11 meio de corpora. Seu objetivo é demonstrar como um grande corpus anotado por humanos,
12 o SNLI (Stanford Natural Language Inference) corpus, poderia auxiliar a máquina a
13 realizar inferências sobre conteúdo. Tal preocupação nasce do fato de relações lógicas
14 serem problemas complexos de serem resolvidos no processo de treinamento. O SNLI é
15 composto por 570.000 pares de sentenças, criados a partir das legendas de um banco de
16 imagens. A anotação feita por seres humanos garante a confiabilidade do resultado. Cada
17 par foi anotado para classificar a relação entre as frases em, pelo menos, uma de três
18 categorias: 1) implicação, 2) contradição e 3) neutralidade (sem relação). Esse corpus
19 anotado serve como base de verificação se o modelo de LLM consegue inferir de forma
20 satisfatória as relações estabelecidas.

21 Na segunda, estudos como Christiano e colaboradores (2023) e Ouyang e
22 colaboradores (2022) e Stiennon e colaboradores (2022) desenvolvem metodologias nas
23 quais humanos interveem diretamente nos resultados da máquina, valendo e classificando
24 as respostas.¹⁰ No caso do primeiro, seres humanos teriam pares de respostas apresentadas
25 para si, tais pares seriam avaliados conforme a preferência dos participantes de pesquisa,
26 que, então, acabam por reforçar comportamentos que mostrem padrões similares. Ouyang
27 *et al.* (2022), por seu turno, partem de uma ação de pré-treinamento com pares de
28 sentenças anotadas por humanos. Após o processamento pelo LLM, os humanos voltam a
29 avaliar estes pares, criando um modelo que reflete aquilo que os avaliadores preferem. Por
30 fim, essas preferências são alimentadas no sistema à guisa de reforço. Por fim, Stiennon e
31 colaboradores (2022) trabalham com resumos produzidos pelo sistema, sendo que tais
32 resumos são julgados por humanos, que manifestam sua preferência. O reforço positivo faz
33 com que as máquinas passem a seguir o padrão melhor avaliado.

1 ¹⁰ Apesar de haver diferenças técnicas na forma como essas metodologias se dão, neste ensaio o objetivo pode deixar tais
2 minúcias escaparem em favor de uma perspectiva mais geral.

1 3. Refletindo sobre a IA na educação

2 3.1 Realmente algo novo?

3 Podemos não perceber, mas a inteligência artificial está em todo lugar. Em nossa vida
4 cotidiana utilizamos processadores de textos e seus corretores ortográficos/gramaticais,
5 aplicativos de mobilidade urbana, motores de busca, mídias sociais, máquinas fotográficas
6 de telefone celular, relógios inteligentes para exercícios, sítios ou aplicativos de compras,
7 agregadores de notícias entre uma infinidade de outros dispositivos e aplicativos. Algumas
8 profissões há tempos trabalham com modelos de inteligência artificial. Entre elas estariam
9 os animadores, cujos programas completam movimentos ao lhes oferecermos coordenadas
10 de início e fim, músicos, que utilizam sistemas de correção de notas e distorções, analistas
11 financeiros, que contam com sistemas de previsão de mercado, meteorologistas, que partem
12 do cruzamento de dados diversos para a realização de previsão meteorológica, mecânicos,
13 que utilizam sistemas computadorizados para leitura de sensores e determinação da
14 regulação dos motores. As placas de nossos carros são automaticamente lidas na entrada
15 de estacionamentos ou por radares de trânsito, que podem aplicar multas automaticamente
16 se estivermos trafegando em horários ou locais restritos ao nosso tipo de veículo.

17 A lista seria quase infinita, assim como seria impossível pensar em nossa vida
18 contemporânea sem estas diversas Inteligências Artificiais. Isso nos leva a perguntar: por
19 que não as percebemos? Por que estes processos estão tão enraizados em nossa sociedade
20 que sequer nos damos conta de sua existência?

21 Isso corre porque deixamos de perceber a tecnologia como não integrada em nossa
22 sociedade (Weiser, 1999). Destarte, as tecnologias funcionais se tornam enraizadas e
23 desaparecem em segundo plano; como a computação, que parece se tornar onipresente e
24 cada vez mais integrada à vida cotidiana pela sua capacidade de realizar tarefas. Weiser
25 (1999) dá como exemplo os motores, que se tornaram uma parte integrante e despercebida
26 da vida de comunidades onde o acesso a sistemas de transporte mecânico ocorre. Como
27 seriam as grandes cidades brasileiras sem seus sistemas de metrô e ônibus?

28 De certa forma, Weiser (1999) parece próximo da definição de tecnologia apresentada
29 por Flusser (2007b), para quem ela é tudo aquilo que aumenta a nossa capacidade de
30 interagir com o mundo. Flusser (2007b) parte da ideia de alavanca: ela aumentaria a
31 habilidade do trabalhador e, com o uso e a construção técnico-social dela decorrente,
32 torná-la-ia um elemento indispensável. É nesse momento que a alavanca se funde ao braço
33 do homem, que planeja suas ações a partir de sua interação com esta alavanca. A
34 linguagem para o Flusser (2007a, 2010) é, assim, uma simbologia artificial que nos estende
35 e codifica o mundo exterior; a linguagem se funde com o homem e lhe dá caminhos de
36 interpretação do mundo. As representações, artificiais e manufaturadas pela língua, são o
37 código-fonte de outras linguagens, como a programação de computadores, a produção
38 audiovisual, etc.

1 Ou seja, as tecnologias, ao se tornarem invisíveis, se tornam parte integrante de nossa
2 realidade, a ponto de não mais imaginarmos nossa vida sem elas. Em seu exercício de
3 futurologia, Weiser (1999) observou que a ubiquidade dos computadores não se daria por
4 meio dos pesados computadores de mesa, mas mediante uma rede dispersa de dispositivos
5 menores. “Tabs”, “pads” e “boards” seriam termos usados pelo autor para definir pequenos
6 computadores que deveriam invadir o cotidiano, incorporados a nosso ambiente e
7 adaptados a tarefas e escalas específicas. Tais dispositivos, assim, desapareceriam na
8 periferia de nosso olhar, permitindo interações quase intuitivas, por vezes sem exigir
9 esforço ou atenção conscientes. Weiser (1999), assim, previu equipamentos como telefones
10 inteligentes, tablets, assistentes digitais que tocam nossas músicas prediletas ou mesmo
11 apagam a luz de nossas casas.

12 Algo que percebemos de comum entre Flusser e Weiser é que o impacto da tecnologia
13 em nossa sociedade parece ocorrer em ciclos. Existe um ecossistema de mídias e tecnologias
14 já sedimentado em nossa vida cotidiana, uma vez que a utilização destes dispositivos está
15 condicionada ao uso seu social, criando-se espaços técnico-sociais (Santos, 2008).

16 Quando uma nova tecnologia e suas expressões midiáticas¹¹ são introduzidas em nosso
17 cotidiano, há um impacto na forma como vemos o mundo e nossas possibilidades de
18 futuro. Segundo Coeckelbergh (2020a, p. 12), isso é responsável pela ideia de que vivemos
19 uma *singularidade Tecnológica*. Essa singularidade daria ao momento histórico em que
20 vivemos um papel e um lugar únicos, responsável por um crescimento tecnológico
21 exponencial que marca o progresso, promovendo uma transformação dramática. Cria-se
22 um novo mundo, que nós não mais entendemos e compreendemos como dantes. Estas
23 tecnologias, por seu turno, nos são opacas, uma espécie de *Caverna de Platão* invertida,
24 clara para aqueles que a criam, mas opaca para aqueles que estão fora, e simplesmente a
25 usam (Bridle, 2018).

26 Exemplos dessa singularidade são diversos. A tipografia nasce no século XV e foi
27 responsável por estabelecer uma indústria que permitira à informação ser reproduzida
28 infinitamente. Ela tem impacto direto na viabilização das casas editoriais, possibilitando, a
29 longo prazo, que livros, jornais e revistas existissem (McLuhan, 1962). Até mesmo nossa
30 tipografia digital seria um resultado a longo prazo dessa revolução. Do ponto de vista
31 político, a tipografia e os livros impressos foram imprescindíveis para a reforma protestante
32 (Gilmont, 1999) e a contrarreforma católica (Julia, 1999). Já o rádio sedimentou-se como
33 primeiro veículo de informação em tempo real, com papel relevante para a criação da
34 mídia de massa durante a Segunda Guerra e nos anos seguintes (Adorno, 2009). A
35 televisão para McLuhan (1964) era a grande novidade tecnológica, a expressão mais
36 contemporânea daquilo que ele chamava de informação pura, ou seja, a energia elétrica.

1 ¹¹ Nesse contexto, é essencial distinguir tecnologia de mídia. Enquanto a tecnologia representa o conhecimento técnico e
2 os dispositivos criados, a mídia constitui a expressão cultural desse conhecimento. As mídias sociais representam um ato
3 cultural derivado de plataformas tecnológicas, moldando comportamentos e interações humanas de maneira sistemática
4 (Manovich, 2005).

1 Desde a chamada *Revolução Digital* (Clarke, 2012), o reforço da ideia de
2 singularidade proposta por Coeckelbergh (2020a) tem se acentuado. A introdução de
3 computadores, a introdução de sistemas síncronos (como mensageiros) e assíncronos (como
4 e-mail), a bolha dos negócios na internet, os blogs-que ganharam até um universo próprio:
5 a blogsfera-, as mídias sociais e seu poder revolucionário — as jornadas de junho no Brasil
6 e seu resultado reacionário (Scartezini, 2016; Sørboe, 2023) — são casos desta
7 singularidade. Por trás dessa realidade, há a concepção de que a tecnologia possui um
8 desenvolvimento linear e, necessariamente, benéfico.¹² Aquilo que é mais “novo”, usado
9 aqui como contemporâneo, seria algo indefectivelmente melhor que a tecnologia de outrora.
10 Entretanto, há diversas discussões que não comungam desta perspectiva, um importante
11 exemplo seria a música (Hunter, 2015), área na qual diversos artistas reconhecem a
12 qualidade da tecnologia de fabricação de instrumentos e da utilizada em apresentações
13 musicais. Há a preferência por instrumentos produzidos por fabricantes específicos em
14 determinadas épocas, como um violino *Estradivário*, ou pela utilização de amplificadores
15 valvulados por guitarristas e baixistas.

16 Mas qual singularidade a Inteligência Artificial tem causado? Para Coeckelbergh
17 (2020a), ela poderia ser definida como a *explosão da inteligência*. Os computadores
18 finalmente suplantaram o ser humano, de que esta tecnologia chegou ao ponto de auto
19 melhorar-se, criando versões mais inteligentes de si mesma. Chega-se ao ponto crítico:
20 seriam sistemas como *Gemini*, *Chat GPT* ou *Co-Pilot* mais capazes que os seres humanos?

21 3.2 LLMs: Questões para reflexão

22 Como professores da área de ciências humanas, carregamos conosco a opinião de que
23 aquilo que praticamos, e ensinamos, é uma linguagem de compreensão do mundo em suas
24 diversas instâncias. Ao mesmo tempo, temos certeza de que estas possibilidades são
25 infinitas e não encerradas em um juízo único e definitivo. Diferentemente desta
26 perspectiva, aplicativos como os LLMs caminham em uma direção contrária: eles são
27 dados estáticos. Por maiores que eles sejam, eles representam apenas aquilo que puderam
28 matematizar, apenas aquilo que se corrigiu pela interpretação humana. De uma certa
29 forma, os LLMs ainda são dependentes de nosso insumo, representando uma versão
30 estatisticamente compreensiva, mas simbolicamente menor que a cognição humana. A IA
31 pode mimetizar, todavia não recriar, a cognição humana, pois os processos são
32 significativamente diferentes (Coeckelbergh, 2020a).

33 Gostaria de sugerir que vejamos a IA, assim como todos os processos tecnológicos que
34 historicamente a precedem ou antecedem, na perspectiva da *Tecnosfera*, como definida por
35 Santos (2008) e por Haff (2014). Para Santos (2008, p. 30) que estabeleceu o conceito nos
36 anos 1990, a tecnosfera poderia ser definida como uma constante artificialização do

1 ¹² Acredito que esta perspectiva seja fruto de uma interpretação superficial do trabalho de Darwin (2018), uma vez que o
2 próprio autor coloca no capítulo VII de seu livro, *Origem das espécies: ou A preservação das raças favorecidas na luta*
3 *pela vida*, que os processos evolutivos podem não ser lineares ou benéficos.

1 ambiente, na qual a esfera natural seria substituída, gradativamente, por uma esfera
2 técnica.¹³ Para o autor, alinhada a este conceito estaria a psicofera, ou seja, nossa atitude
3 filosófica perante o mundo, incluindo certamente a tecnologia. Apesar de sua importância
4 histórica, a definição de Santos (2008) separa a existência de um mundo natural de sua
5 perspectiva técnica. Para o autor, a tecnologia é algo artificial e externo que transforma a
6 relação do homem natural com seu mundo. Por conta disso, acredito que este conceito
7 possa ser ampliado a partir do sugerido por Haff (2014), para quem a tecnosfera deve
8 abranger todos os sistemas técnicos humanos, nossos processos tecnológicos
9 compreensivamente, além de sistemas que nós tradicionalmente consideramos humanos,
10 como as ONGs e entidades religiosas.

11 Haff argumenta que a tecnosfera se tornou um sistema dinâmico que interage com os
12 humanos. O autor sugere que devemos nos distanciar de uma perspectiva antropocêntrica
13 para uma abordagem que reconheça a tecnosfera como um sistema autônomo. Algo que
14 determina o comportamento humano tanto quanto por ele é influenciado. A utilização
15 destes dispositivos, assim, está condicionada ao uso seu uso social. Como cada grupo
16 utiliza os diversos dispositivos e tecnologias não é universal, criam-se espaços técnico-
17 sociais específicos (Santos, 2008) nos quais tais dispositivos e hábitos evoluem
18 conjuntamente (Hayles, 2005). Nas palavras de Hayles (2006):

19 As crenças e práticas culturais fazem parte dessa dinâmica co-evolutiva porque
20 influenciam como as ferramentas que são desenvolvidas e usadas, o que, por sua
21 vez, influencia quem somos enquanto organismos biológicos, o que, por sua vez,
22 retroalimenta a espiral co-evolutiva (Hayles, 2006, p. 169, tradução minha).¹⁴

23 Como a tecnologia e o ser humano constroem um processo de simbiose, afinal, ela
24 deve garantir a existência do homem para que sua própria função neste sistema seja
25 garantida. A evolução conjunta traz, a nós, a possibilidade de, enquanto sistemas menores
26 que interagem dentro desta grande tecnosfera, modificar como conjunto as funções e seus
27 usos.

28 Se pretendemos inserir, regular e constituir práticas de bom uso da IA em nosso fluxo
29 técnico-social, precisamos seriamente refletir sobre suas implicações éticas e
30 socioeconômicas, com especial foco em seus possíveis impactos de forma crítica e
31 desenvolver padrões éticos de uso. Assim, se pensarmos nos impactos da adoção dos LLMs
32 sem a devida reflexão no contexto técnico-social do ensino (com ênfase no ensino de
33 línguas), algumas questões precisam ser consideradas.

1 ¹³ Uma observação similar também pode ser compreendida a partir do trabalho de Antonio Candido (2023) em sua tese
2 de doutoramento de 1964. Em seu trabalho, Candido (2023) demonstra como a interação mais naturalizada da sociedade
3 caipira é gradativamente substituída por uma industrializada e mecanizada.

4 ¹⁴ Do original: *Cultural beliefs and practices are part of this co-evolutionary dynamic because they influence what tools*
5 *are made and how those tools are used, which in turn affects who we are as biological organisms, which then feeds back*
6 *into the co-evolutionary spiral* (Hayles, 2006, p. 169).

1 3.2.1 Variações e estilo

2 Ao imaginarmos que a linguagem possa se definir como um objeto de análise de
3 aprendizado múltiplos, LLMs poderiam ser responsáveis por excluir variedades linguísticas
4 que estejam sub representadas dentro dos corpora utilizados para treinamento. Algo
5 similar pode ser dito se pensarmos em termos das estruturas estilísticas representadas
6 neste modelo. Boa parte do processo de ensino de escrita, por exemplo, está centrado no
7 desenvolvimento de uma habilidade aplicável a contextos que, apesar de mapeáveis dentro
8 de nosso agir social, poderiam demandar uma capacidade pessoal de produção. Os
9 diferentes LLMs tendem a construir respostas já organizadas linguisticamente, causando a
10 padronização dos textos e produções linguísticas apresentadas.

11 Ou seja, sua utilização sem reflexão pode levar a uma desvalorização de variantes não
12 inseridas nestes modelos, e a padronização de usos tem termos de macro e microestruturas
13 textuais. Como já refleti em outra ocasião ([Lima-Lopes, 2014](#)), não se questiona que o
14 papel da escola seja ensinar a norma culta (ou variante de prestígio), uma vez que ela
15 possibilita o acesso a uma série de bens culturais e sociais. Questiona-se, sim, o papel dado
16 a outras variantes, que merecem ser valorizadas e mantidas como parte de nosso acervo
17 cultural. Além disso, o reforço de determinados padrões macro e micro textuais pode
18 alterar a rota da transformação e uso da língua, talvez fossilizando estruturas e
19 organizações potenciais ([Halliday; Hasan, 1991](#)) específicas ([Lee, 2023b, 2023a](#)). O uso de
20 ferramentas que reforçam um padrão pode, efetivamente, levar a uma diminuição no uso
21 de outras variedades, uma espécie de totalitarismo no uso. Se a língua é uma forma de
22 identidade social, tais modelos poderiam, assim, contribuir para o favorecimento de
23 determinadas variantes, desmerecendo ainda mais as que sejam menos prestigiosas em
24 determinados contextos.

25 Apesar de o exemplo estar centrado na educação linguística, um efeito de minha
26 formação inicial, é importante observar que a questão da sub-representação e da imposição
27 de uma interpretação sobre grupos sociais ou fenômenos naturais específicos pode ser algo
28 que também deva impactar outras áreas. Como discutir questões como aquecimento global
29 ou terraplanismo a partir de um modelo de LLM? Seus dados de treinamento têm validade
30 para discutir a questão?

31 3.2.2 O papel do professor e do material

32 Tecnologias de apoio ao professor e ao aprendiz existem há séculos, talvez a primeira
33 que nos venha à mente seja, exatamente, o livro didático. Estes livros, ao serem analisados
34 por programas como o PNLD (Plano Nacional do Livro Didático), ou por bancas/grupos
35 de professores ao definirem sua adoção, são julgados por critérios claros. No caso dos
36 materiais disponibilizados pelo Ministério da Educação no Brasil, podemos observar que:

37 Os materiais distribuídos pelo MEC às escolas públicas de educação básica do
38 país são escolhidos pelas escolas, desde que inscritos no PNLD e aprovados em
39 avaliações pedagógicas coordenadas pelo Ministério da Educação e conta com a
40 participação de Comissões Técnica específica, integrada por especialistas das

1 diferentes áreas do conhecimento correlatas, cuja vigência corresponderá ao ciclo
2 a que se referir o processo de avaliação (Brasil, 2024).

3 Há diversas camadas de verificação ética. Em nosso contexto contemporâneo, tais
4 materiais são produzidos por professores especialistas que assinam sua produção. Tal
5 assinatura os faz responsáveis técnicos pelos conteúdos e visões de mundo e ciência ali
6 presentes. As comissões/programas/bancas/grupos de professores que realizam as escolhas
7 têm suas justificativas baseadas em visões tecno-ideológicas do que seja a disciplina e seu
8 conteúdo. Há sempre a possibilidade de discordarmos, e podemos fazê-lo pela credibilidade
9 que atribuímos (ou não) a estes especialistas.

10 Quando atribuímos a modelos de LLM este papel, estamos, de fato, atribuindo esta
11 função a uma caixa de pandora. Apesar da abundância de dados que servem como base
12 para estes modelos, nunca teríamos certeza de quais seriam os conteúdos que compõe o
13 conjunto de fontes utilizadas em uma área do conhecimento. Como a base de treinamento
14 é sempre pouco explícita aos usuários, é impossível saber se o que está ali representa a
15 visão daquilo que gostaríamos de ensinar. Ao não sabermos qual a base de dados que
16 compõe o conjunto de temáticas, estamos, assim, dando uma *Carte Blanche* para o sistema
17 realizar a curadoria e a avaliação do que deve ser ensinado. Acrescente-se a isso a
18 necessidade do processo de formação dos aprendizes passar por uma perspectiva
19 comunicativa-dialética. Tal perspectiva parece não ocorrer nos LLMs, o que há ali é o
20 conteúdo estendido ao usuário, sem a negociação comunicacional típica do processo
21 comunicativo (Freire, 2006) e sem que haja o devido processo de conscientização (Freire,
22 2018).

23 Parece que neste contexto há um processo de desconstrução do papel de professores e
24 do material didático. O professor e sua formação têm parte de suas funções — de curador
25 dos materiais e métodos a serem utilizados — substituídas pelos LLMs valorizados pelas —
26 pretensas, mas não efetivas — exatidão e acuidade. O material assim, deixa de ser algo
27 planejado conforme a demanda e necessidade das diferentes turmas, mas pensado de forma
28 homogênea a partir de um conjunto de dados hipoteticamente completos, todavia
29 desconhecidos. Materiais estes que impõem uma visão única, destarte limitada, do
30 conteúdo. Nesse contexto, professores passam a ser meros aplicadores de materiais cujo
31 feitio e escolha não lhes cabe e a tão cobiçada personalização, tema principal de grande
32 parte das fechitizações que assistimos sobre IA, claramente, não acontece.

33 3.2.3 Privacidade

34 O uso de dispositivos digitais, principalmente os com grande capacidade de
35 penetração em nosso dia-a-dia, e sua relação com potenciais riscos para a privacidade dos
36 usuários é uma preocupação que já emerge desde os trabalhos de Weiser (1999). Todavia,
37 desde o início da década de 2000 há uma constante plataformização da sociedade (Poell;
38 Nieborg; van Dijck, 2019), sendo ela, como já observado, responsável por um maciço
39 processo de coleta de dados de indivíduos. Algo que torna esta preocupação ainda mais
40 premente. Estes dados podem incluir desde estatísticas de uso até dados pessoais e

1 comunicacionais, como os textos e materiais que postamos ou transferimos. Segundo
2 Zuboff (2019), tais dados são utilizados na construção de perfis cujo objetivo principal é
3 construir estratégias de oferecimento de produtos e de consumo.

4 A autora chama este processo de capitalismo de vigilância (Zuboff, 2019). Entre suas
5 características está a extração infinita de dados, algo muito similar ao que se praticava nos
6 séculos XVII e XX tendo como base os recursos naturais. Tal prática faz de nós reféns de
7 tais plataformas, uma vez que ceder os dados contratualmente durante o seu uso é a única
8 forma de utilização. Por trás do pseudo-ideia de personalização, está o reforço de
9 comportamentos por sistemas externos que monitoram e direcionam nossas ações,
10 orientando para o consumo, mais das vezes inconsciente. Um exemplo deste processo
11 seriam as mídias sociais e sua orientação em relação à polêmica. Como coloca Morozov
12 (2011, 2018), tais empresas incentivam o embate direto entre indivíduos e grupos com
13 posturas sociais e políticas desalinhadas. Mesmo podendo monitorar, por exemplo, perfis e
14 grupos que promovem ódio e preconceito, tais empresas escolhem não o fazer, uma vez que
15 a discussão e o enfrentamento geram tráfego de rede e, por conseguinte, lucro.

16 A partir de 2020, o processo de plataformização acelerou-se devido ao distanciamento
17 social imposto pela pandemia. Tal fato fez com que toda a sorte de documentos oficiais de
18 nossas escolas — incluindo nossos planejamentos e planos de aula —, assim como de
19 nossos alunos — seus trabalhos, negociações para fazimento de tarefas e hábitos — fossem
20 fartamente mapeados. Outro ponto relevante é que, ao aceitarmos utilizar os diversos
21 LLMs, eles coletam os dados textuais ali inseridos, utilizando-os para treinamento futuro.

22 Para autores como Sadowski (2019), este fato seria um fenômeno muito próximo
23 daquilo que poderíamos chamar de *extração de dados*. Para o autor, a coleta de dados é
24 impulsionada por um ciclo de acumulação de capital, apoiado em um universo no qual
25 tudo pode ser caracterizado como dados. Tais dados, assim, influenciam não somente
26 modelos de negócio, como ações governamentais e desenvolvimento social. Cria-se, assim,
27 um imperativo de acumulação de dados por todas as fontes e meios possíveis,
28 infinitamente. Sadowski (2019) defende que estas práticas devem ser vistas como
29 extrativistas, uma vez que têm pouco, ou nenhuma consideração compensatória, às fontes.
30 Por conseguinte, ao considerarmos os dados como uma forma de capital, a dadificação da
31 sociedade deve ser vista como um regime político e econômico.

32 No caso do Brasil, carecemos de sistemas regulatórios. Os dados colhidos nos LLMs e
33 outras plataformas precisam de regras claras a respeito da utilização do que ali é colhido,
34 ou estaremos correndo riscos graves de exposição de aprendizes e professores. Caso não
35 criemos tais parâmetros, corremos o risco de nos transformarmos numa colônia de
36 extração, mas ao invés de madeira, pedras, metais preciosos e açúcar, serão dados.

37 3.2.3.1 Direitos autorais e responsabilidade

38 Da mesma forma que jornalistas são responsáveis por seus textos, escritores e
39 pesquisadores também o são. Mas ao utilizarmos materiais gerados por um LLM, quem são

1 os responsáveis? Uma vez que as respostas oferecidas por estes modelos são o resultado de
2 uma matriz de dados, como regular o direito autoral?

3 Quintais (2025) e Ren e colaboradores (2024) discutem a questão dos direitos
4 autorais, direitos de uso e os LLMs. Como colocado pelos autores, os diversos modelos
5 utilizam textos, imagens e, mais recentemente, peças de audiovisuais como fonte de seu
6 treinamento. Entre as principais preocupações, está a forma de proteção desses materiais.
7 Esta proteção poderia percorrer dois caminhos, o primeiro relacionado ao armazenamento
8 dos dados protegidos no modelo de treinamento — afinal os dados deste material estão ali
9 guardados mesmo que matematizados — e das possíveis reproduções não autorizadas, uma
10 vez que estes materiais podem — resultantes dos cálculos que geram as respostas — estar
11 presentes na linguagem gerada por estas máquinas.

12 Duas perguntas emanam deste tipo de reflexão: 1) As respostas produzidas pelos
13 LLMs poderiam, assim, ser objeto de direitos autorais? e 2) Como fornecer a compensação
14 devida aos produtores que serviriam como base de geração dos dados? Os autores sugerem
15 soluções diferentes. Ren e colaboradores (2024) desenvolvem um complexo sistema de
16 proteção a ser aplicado em textos e imagens, desenvolvendo códigos que podem ajudar
17 autores a melhor gerenciar e garantir que os créditos sejam a eles atribuídos. Já Quintais
18 (2025) sugere uma solução na qual devem ser desenvolvidas instâncias legais de controle.
19 Entre suas sugestões estariam a abertura dos modelos de treinamento, assim como dos
20 dados neles utilizados. Essa abertura poderia gerar impostos específicos a serem cobrados
21 dos produtos gerados por LLMs, levando ao devido repasse. Apesar de os autores não
22 citarem tal solução, é ainda imprescindível que as fontes de cada produção sejam citadas
23 no *output* sejam disponibilizadas para o usuário. Mesmo havendo um movimento nesta
24 direção por alguns aplicativos, eles tendem a trazer ligações — exemplos seriam sítios
25 como enciclopédias online — onde mais informações poderiam ser encontradas em caráter
26 igualmente genérico.

27 Discutir estas questões sobre como aplicar questões relacionadas a direitos autorais
28 está fora do escopo deste trabalho. Contudo, há algumas implicações importantes para o
29 trabalho de profissionais da educação. Primeiramente, parte do processo de formação
30 intelectual de nossos alunos prevê o reconhecimento das fontes externas utilizadas na
31 execução de seus trabalhos e projetos. É claro que a forma de reconhecimento muda
32 conforme o momento escolar em que estamos; entretanto, ele sempre está presente. A
33 utilização de LLMs indiscriminadamente pode, por conseguinte, agregar formas de
34 conhecimento sem que sua fonte seja identificada, dado à ferramenta o valor de um
35 oráculo de modernidade.

36 Questiona-se, então, seriam estes modelos legalmente responsáveis pelas informações
37 que nos oferecem? O ponto é polêmico. Por não haver criatividade simbólica, eles acabam
38 sempre repetindo informações com as quais foram treinados, projetando, por conseguinte, a
39 responsabilidade em suas fontes. Entretanto, é salutar pensar que a escolha de tais fontes é
40 responsabilidade daqueles que desenham seu treinamento. Ou seja, se o LLM foi treinado

1 com textos racistas ou misóginos, ele tenderá a repetir esses valores em seus resultados.
2 Nota-se, também, que a correção humana pode não dar conta de todas as questões
3 pretensamente objetivas sobre conteúdo, o que pode levar à falta de acuidade, algo que
4 tem impacto direto no usuário, que confia na informação.

5 Em conclusão, ao usarmos tais modelos, estamos agregando em nossos resultados
6 estes valores, podemos estar reproduzindo, mesmo que desavisadamente, perspectivas
7 ideológicas que não correspondem a valores que gostaríamos de agregar em nossa sala de
8 aula.

9 3.2.4 Meio ambiente, sustentabilidade e emprego

10 Como qualquer sistema técnico inserido dentro de nossa tecnosfera, os sistemas de
11 inteligência artificial, logicamente os LLMs, podem trazer preocupações importantes
12 impactos ambientais. Segundo Coeckelbergh (2020a), tais problemas podem se manifestar
13 de duas formas. Na primeira, poderíamos incorrer no que o autor chama de solucionismo
14 tecnológico, ou seja, em um projeto de solução que delega à máquina a capacidade de
15 calcular e sugerir possíveis saídas para o problema. Tal postura teria implicações
16 importantes para as próximas gerações, especialmente no campo político, uma vez que os
17 LLMs tendem a ser gerenciados pelas chamadas *Big Techs* que se encontram no Vale do
18 Silício. Assim, estaríamos delegando tais decisões a um único ator político. Uma possível
19 consequência seria a criação de uma prótese para decisões que desconsidera elementos não-
20 humanos (como os interesses de outras espécies) e renega à inexistência diversas questões
21 políticas nestas decisões. Estas, aliás, seriam tomadas tendo como filtro um princípio
22 cultural único.

23 A segunda estaria relacionada ao impacto no meio ambiente criado por esses modelos.
24 Há uma tendência dos sistemas de informática a reforçar uma ideia de intangibilidade em
25 sua infraestrutura (Bridle, 2018), a qual se manifestaria em uma série de metáforas que
26 promovem sua compreensão como algo etéreo (Coeckelbergh, 2020b). Exemplos seriam
27 designações como “nuvem”, “aplicativo”, “processamento compartilhado”. O que tais
28 metáforas estão, efetivamente, fazendo é criar a ideia de que tais sistemas existem fora de
29 nossa materialidade, fazendo-nos imaginar que sua operação seria algo independente de
30 questões como aquecimento de maquinário, fluxo de água para resfriamento, obras de
31 instalação, consumo de energia elétrica, relações trabalhistas e diversas outras.

32 Há diversos trabalhos que têm como foco a segunda questão. Apesar de não haver
33 esta ênfase ao se oferecerem os sistemas para uso, todos os LLMs são dependentes deste
34 tipo de infraestrutura. Em outras palavras, LLMs podem contribuir sobremaneira para o
35 aquecimento global graças ao seu consumo de energia e emissões. Bridle (2018) discute
36 como os diferentes *data centers* — ou centrais que armazenam e distribuem nossos dados
37 quando em servidores externos, ou nuvens — causam aquecimento no meio ambiente dos
38 locais em que são instalados, além de seu grande consumo de energia para manutenção e
39 de água para resfriamento. Em um estudo comparativo Strubell e colaboradores (2019),
40 observam que o treinamento de um modelo de LLM pode chegar a consumir 626.155 CO₂

1 e lbs (Libras de Carbono), algo quase seis vezes superior a toda existência de um veículo
2 movido motor de explosão alimentado por combustível fóssil (126.000 CO₂ e lbs) e quase
3 57 vezes a média de uma vida humana (11.023 CO₂ lbs). Já o treinamento de um modelo
4 poderia custar até 3 milhões de dólares somente em termos da energia elétrica consumida.¹⁵

5 Algo também pouco discutido na grande mídia é a relação que as empresas
6 responsáveis pelos LLMs possuem com os trabalhadores responsáveis pelo seu treinamento.
7 Estudos como os de Kassi e colaboradores (2021) demonstram que os trabalhadores
8 empregados por empresas responsáveis por IA podem enfrentar condições de trabalho
9 deveras precárias e cerceadoras. Exemplos seriam trabalhadores no Kenia que estariam
10 recebendo salários máximos de U\$ 2,00 a hora para o treinamento e manutenção destes
11 sistemas (Perrigo, 2023), fato que tem causado movimentos trabalhistas de reparação
12 (McNeill, 2023) e melhoria de condições. Em muitos casos, há relatos de impactos
13 negativos na saúde mental (Kantrowitz, 2023; Zahn, 2024), pois estes trabalhadores — que
14 operam por empreitada (*gig workers*) sem contatos de trabalho fixos ou direitos de
15 qualquer natureza — são expostos a conteúdos possivelmente danosos e contam somente
16 com sua rede pessoal de apoio.

17 4. Um caminho a ser percorrido

18 Talvez o que tenha discutido até agora tenha deixado uma perspectiva de sabor acre
19 e desacreditada em relação às potencialidades dos LLMs no contexto do ensino,
20 especialmente na forma como eles têm sido adotados. Se pensarmos a IA dentro de seu
21 papel na tecnosfera em relação à escola, caberá decidirmos:

- 22 1. Qual seu papel na paisagem tecnológica da escola?
- 23 2. Qual seu limite de atuação nesta paisagem tecnológica?

24 Desenvolver uma perspectiva crítica, logo, é algo importante se quisermos avaliar e,
25 talvez, definir o lugar dos modelos de LLM na educação. Não acredito em banimentos
26 tecnológicos em sala de aula, pois minha experiência como professor tem me mostrado que
27 este tipo de atitude perante o processo tecnológico tende a levar a uma busca por novos
28 caminhos de utilização, formas de burlar a proibição. Criam-se, então, espaços não
29 supervisionados que, muitas vezes, podem levar ao uso acrítico. Seria impossível pensar na
30 tarefa a que este artigo se propõe se não se considerar alguns princípios: 1) todo processo
31 tecnológico precisa de regulação; 2) a cognição simbólica é uma característica humana não
32 podendo ser substituída e 3) os modelos computacionais são dependentes das ações
33 humanas e as refletem.

34 Enquanto educadores, deveríamos traduzir nossos princípios éticos em
35 regulamentação de uso. A regulamentação tem um papel essencial, permitindo que os

1 ¹⁵ O cálculo é dependente do tipo de treinamento, podendo variar consideravelmente dependendo das tecnologias e
2 infraestrutura utilizadas. Nos resultados apresentados, tais números são referentes a modelos NAS (*Neuro Architecture*
3 *Search*) (Strubell; Ganesh; McCallum, 2019, p. 04).

1 usuários não naveguem por águas inseguras (Miao; Holmes, 2023). Apesar de haver a
2 necessidade de criação de políticas globais, soluções locais podem remediar a questão
3 enquanto esta discussão se estabelece. Isso seria essencial para evitar situações de
4 acriticidade ou questões sociais e legais que podem resultar do inadvertido uso destas
5 plataformas. A natureza dos resultados oferecidos por elas poderia resultar na utilização de
6 materiais e ideias sem o reconhecimento de sua fonte, algo importante no contexto de
7 formação de nossos alunos, ou mesmo na construção de representações sociais em
8 desacordo com as perspectivas educacionais que por ventura sigamos.

9 Deve-se ressaltar que os modelos humanos de cognição emergem de uma combinação
10 única de experiências e emoções. Nossa capacidade de entender, responder a emoções e
11 necessidades alheias ou impostas pelo mundo têm um papel fundamental na construção de
12 nossa identidade e de nosso conhecimento (Coeckelbergh, 2020a; Gigerenzer, 2022). Para
13 autores como Gigerenzer (2022), é justamente esta capacidade de construção simbólica e
14 reestruturação do conhecimento a partir daquilo que vivemos e observamos *in loco* e em
15 tempo real que nos torna “*inteligentes em um mundo esperto*”.¹⁶ Valorizar nossa
16 capacidade quase imediata de reação e adaptação à realidade, parte da compreensão do
17 mundo e da sociedade como elemento em constante transformação. Um dos principais
18 papéis da escola, neste prisma, seria mostrar como o ato comunicativo (Freire, 2006) que
19 representa a educação deve incluir não somente o conhecimento enciclopédico, mas
20 também formas de conscientização e transformação da sociedade (Freire, 2018). Algo
21 sobremaneira relevante na formação de cidadãos, que tomam para si o conhecimento
22 letrado. Assim, se assumirmos que mais poder computacional faz as máquinas mais rápidas
23 (Gigerenzer, 2022) e mais dados talvez as façam menos imprecisas, estamos delegando a
24 elas seu efetivo papel em nossa sociedade. Os LLMs — e os demais sistemas de IA —
25 podem oferecer rapidez em ações pré-programadas, ou algoritmos, realizando tarefas que
26 possam ser cansativas para a cognição humana. Essa esfera permite a minimização de erros
27 e a possibilidade de crescimento na escala na observação de fenômenos sociais e naturais.
28 Algumas áreas das Humanidades Digitais — como a Linguística do Corpus, o mapeamento
29 e criação de mapas personalizados, entre diversas outras — seriam impensáveis sem a
30 presença do computador e, agora, podem se beneficiar das LLMs. Pois ela nos ajudaria a
31 construir e desconstruir perspectivas situadas em determinadas interpretações da
32 sociedade, por meio de uma análise extensiva que poderia ser impossível se feita
33 manualmente.

34 Ao pensarmos no ensino de linguagens — sendo o ensino de línguas algo certamente
35 considerável neste aspecto —, o uso de LLMs pode gerar padronizações indesejadas. Tal
36 fato é o resultado daquilo que a faz impressionar à primeira vista: os dados. Como eles são
37 treinados por um grupo de pessoas que os selecionam — mesmo que grandes em
38 quantidade —, eles têm uma tendência totalitária em sua gênese. Em outras palavras, se

1 ¹⁶ Uma tradução direta e adaptada para o português do título do livro de Gigerenzer (2022): *How to stay smart in a*
2 *smart world*.

1 os dados são sempre finitos e excludentes, também o serão as representações de linguagem
2 que estão neles presentes, incapazes de abarcar qualquer totalidade de comportamento
3 social. Uma vez que a pseudo-autoridade delegada a estes sistemas pode levar a uma
4 padronização nos resultados oferecidos, sendo que as perspectivas inter e multiculturais no
5 ensino de linguagens podem sofrer impactos importantes. Garantir esta diversidade vem da
6 desmitificação desses resultados e da criação de modelos que reflitam nossa cultura. Mesmo
7 sabendo que jamais conseguiremos refletir toda a diversidade cultural e de pensamento de
8 um país como o Brasil em um modelo de LLM, mesmo assumindo que esta tarefa seria
9 algo utópico, certamente um desenvolvimento local estaria mais próximo de garantir esta
10 diversidade. Enquanto tais modelos ainda não são desenvolvidos, observar os resultados
11 com criticidade de forma conjunta com alunos e integrada aos demais membros da
12 comunidade escolar pode ser algo positivo.

13 Uma visão sobre a IA que não a integra na tecnosfera, mas a vê como uma
14 singularidade, traz, em si, um paradoxo. A este ufanismo tecnológico é creditada uma
15 garantia de personalização, inovação e aceleração de resposta. Entretanto, como ela na
16 essência é o resultado de um banco de dados, de uma grande matriz e seus cruzamentos
17 estatísticos possíveis e prováveis, até que ponto ela garante resultados personalizados? Na
18 verdade, o que se vê é uma grande padronização, nos quais resultados similares são
19 distribuídos uniformemente pelas diversas consultas.

20 Não podemos deixar de lado os impactos ambientais e no emprego. Apesar de as
21 empresas responsáveis por estes modelos enfatizarem a ideia de intangibilidade e o caráter
22 etéreo destes sistemas por meio de metáforas, eles produzem abundância de carbono e
23 consomem energia para sua manutenção. Se pensarmos que parte da formação das
24 próximas gerações está centrada em um processo de preservação e respeito ao ambiente,
25 tais modelos correm na contramão. Além disso, boa parte dos processos de intervenção
26 humana ocorre em situações de subemprego em países em desenvolvimento, em uma
27 relação laboral que, em alguns casos, pode desafiar a legalidade. Tal realidade parece
28 confirmar as visões de Heidegger (1977), para quem a tecnologia é fruto de nossa forma de
29 enquadrar o mundo, normalmente visto como um repositório potencial para nossos saberes
30 e estrutura técnica. Entretanto, Heidegger (1977) coloca que um dos perigos desta
31 percepção é ver o ser humano como parte desta reposição.

32 A equidade de acesso é outra discussão relevante (Miao; Holmes, 2023). Diferentes
33 grupos sociais podem ter acesso limitado a tais ferramentas, seja pela conectividade, seja
34 pela necessidade de hardware capaz de processar tais modelos, seja pelo pagamento de
35 taxas ou assinaturas. Essa realidade leva à necessidade de criação de políticas de acesso
36 que podem, inclusive, esbarrar em questões de soberania e política social mais ampla, uma
37 vez que a intervenção humana está presente desde a construção do modelo até seu
38 treinamento e correção. Talvez por isso, o desenvolvimento de modelos de LLM nacionais,
39 como dados de treinamento que reflitam um desenho específico para a língua portuguesa e
40 visando refletir nossa cultura e diversidade, é imperativo neste processo. Mas, por que isso

1 seria importante? Ao importarmos modelos criados em outros contextos políticos e
2 culturais, estamos, necessariamente, trazendo a reboque representações políticas e
3 semióticas, valores sociais e padrões de linguagem que podem levar situações de
4 desvalorização de culturas locais.

5 O computador é uma excelente ferramenta para analisar dados, padronizá-los e fazer
6 crescer nossa observação escalarmente, mas a criatividade é uma habilidade
7 intrinsecamente humana. Valorizar a agentividade de nossos alunos e garantir a
8 diversidade de perspectivas deve ser um princípio a ser preservado. Por isso, devemos
9 integrar as LLMs, assim como todas as tecnologias, de forma que complemente, mas não
10 substitua, a capacidade de criação simbólica do ser humano. Estabelecer limites no uso de
11 dispositivos, como períodos sem telas, para fomentar a reflexão e a criatividade, são
12 essenciais na construção de um ambiente de uso que permita uma visão ampla e crítica.
13 Particularmente, acredito que a utilização desses modelos deva ser complementar em
14 relação às tarefas propostas pelo professor e pelos manuais didáticos escolhidos, sempre em
15 uma perspectiva coadjuvante em relação à produção dos alunos e sua interação crítica e
16 social. Pensar em modelos locais e em soluções que respeitem o meio ambiente e a
17 dignidade do emprego também são temas que precisamos trazer à baila. Em suma, tratar
18 os LLMs por uma perspectiva da singularidade não contribui para a discussão de seu papel
19 em nossa sociedade.

20 Villén Flusser (2010), ao questionar o impacto da técnica em nossa sociedade, realiza
21 uma importante reflexão sobre o papel que devemos dar ao ensino da escrita. O autor
22 observa que, à medida que os computadores e mídias digitais se tornam mais comuns em
23 nossa sociedade, a natureza da escrita tende a se modificar. O autor propõe que, ao invés
24 de pensarmos que a escrita e a criatividade humana estejam no fim, elas passam
25 por uma transformação. Dessa forma, o ato de escrever deveria ir além do simples registro,
26 ele agora integra a paisagem tecnológica. Um dos aspectos centrais do seu argumento é
27 uma crítica ao que ele chama de mecanização humana. Para Flusser, ao delegarmos à
28 máquina o papel de organizar nossos pensamentos e nossa criatividade, estamos abrindo
29 mão dos processos que nos tornam humanos. Então, qual seria o futuro da escrita? O
30 autor infere que ela ainda tem o papel de ser a base na qual outras tecnologias, como os
31 computadores e, por conseguinte, as LLMs. Ela serviria então como o código-fonte dessas
32 tecnologias, o lugar onde todas nascem. Se caminarmos por esta seara, precisamos
33 redimensionar o ensino das diferentes linguagens para tomar este papel, em uma
34 resistência à aplicação submissa com cores vazias de modernidade e mantendo esta, e
35 outras tecnologias, visíveis e perceptíveis para melhor decidirmos sua função na escola.

1 Agradecimentos

2 O autor gostaria de agradecer ao CNPq (processo 311099/2021-1) pelo financiamento
3 desta pesquisa e à Cátia Lassalvia pelas ricas observações em versões anteriores deste
4 ensaio.

5 Sobre o autor

6 Rodrigo Esteves de Lima Lopes é Livre Docente em Linguagem e Tecnologia,
7 professor associado do Departamento de Linguística Aplicada do IEL/UNICAMP. Líder
8 do grupo de pesquisa MiDiTeS (Mídia, Discurso, Tecnologia e Sociedade) e coordenador
9 do PH₂D (Centro de Pesquisa em Pós-Humanismo e Humanidades Digitais). Suas
10 pesquisas estão voltadas para a Linguística do Corpus e o uso de dados nos estudos da
11 linguagem.

12 Declaração de contribuição dos autores

13 Declaro que este artigo é integralmente de minha autoria.

14 Declaração de conflito de interesse

15 Declaro não haver conflitos de interesse.

16 Declaração de disponibilidade de dados da pesquisa

17 Todo o conjunto de dados de apoio aos resultados deste estudo foi publicado no
18 próprio artigo.

19 Referências

- 20 ADORNO, Theodor W. **Indústria cultural e sociedade**. São Paulo (SP): Paz e Terra, 2009.
- 21 APEOESP. **Inteligência Artificial é instrumento auxiliar, jamais poderá substituir o**
22 **professor!** São Paulo, Nota, 18 abr. 2024. Disponível em:
23 [http://www.apeoesp.org.br/publicacoes/professor/inteligencia-artificial-e-instrumento-](http://www.apeoesp.org.br/publicacoes/professor/inteligencia-artificial-e-instrumento-auxiliar-jamais-podera-substituir-o-professor/)
24 [auxiliar-jamais-podera-substituir-o-professor/](http://www.apeoesp.org.br/publicacoes/professor/inteligencia-artificial-e-instrumento-auxiliar-jamais-podera-substituir-o-professor/). Acesso em: 7 jan. 2025.
- 25 BALMANT, Ocimara. Inteligência artificial é bem-vinda na educação, mas precisa de
26 regulação, defendem especialistas. **OESP**, São Paulo, 29 out. 2024. Disponível em:
27 [https://www.estadao.com.br/educacao/inteligencia-artificial-educacao-regulacao-](https://www.estadao.com.br/educacao/inteligencia-artificial-educacao-regulacao-especialistas/)
28 [especialistas/](https://www.estadao.com.br/educacao/inteligencia-artificial-educacao-regulacao-especialistas/). Acesso em: 7 jan. 2025.
- 29 BOJANOWSKI, Piotr *et al.* **Enriching Word Vectors with Subword Information.**, 2017.
30 Disponível em: <http://arxiv.org/abs/1607.04606>. Acesso em: 29 jan. 2025.

- 1 BOWMAN, Samuel R. *et al.* **A Large Annotated Corpus for Learning Natural Language**
2 **Inference.**, 2015. Disponível em: <http://arxiv.org/abs/1508.05326>. Acesso em: 30 jan. 2025.
- 3 BOYD, Danah; CRAWFORD, Kate. Critical Questions for Big Data: Provocations for a
4 Cultural, Technological, and Scholarly Phenomenon. **Information Communication and**
5 **Society**, [s. l.], v. 15, n. 5, 5, p. 662-679, 2012. Disponível em:
6 <http://dx.doi.org/10.1080/1369118X.2012.678878>.
- 7 BRASIL. PNL D (**Programa Nacional do Livro didático**). Ministério da Educação, 2024.
8 Disponível em: <http://portal.mec.gov.br/component/content/article?id=12391:pnld>.
9 Acesso em: 12 fev. 2025.Orgão Público
- 10 BRIDLE, James. **New Dark Age: Technology, Knowledge and the End of the Future.**
11 London ; Brooklyn, NY: Verso, 2018.
- 12 BROWN, Tom B. *et al.* **Language Models Are Few-Shot Learners.**, 2020. Disponível em:
13 <http://arxiv.org/abs/2005.14165>. Acesso em: 7 jan. 2025.
- 14 CAFARDO, Renata. Governo de SP vai revisar slides de aulas após erros e recomenda
15 novo download para professores. **Estadão: Educação**, São Paulo, 9 jul. a 2023. Disponível
16 em: [https://www.estadao.com.br/educacao/governo-de-sp-vai-revisar-slides-de-aulas-apos-](https://www.estadao.com.br/educacao/governo-de-sp-vai-revisar-slides-de-aulas-apos-erros-e-recomenda-novo-download-para-professores/)
17 [erros-e-recomenda-novo-download-para-professores/](https://www.estadao.com.br/educacao/governo-de-sp-vai-revisar-slides-de-aulas-apos-erros-e-recomenda-novo-download-para-professores/). Acesso em: 8 jan. 2025.
- 18 CAFARDO, Renata. Tarcísio e material didático impresso em SP: entenda idas e vindas
19 do governo. **Terra/Educar: Educação**, São Paulo, 8 jul. b 2023. Disponível em:
20 [https://www.terra.com.br/noticias/educacao/tarcisio-e-material-didatico-impresso-em-sp-](https://www.terra.com.br/noticias/educacao/tarcisio-e-material-didatico-impresso-em-sp-entenda-idas-e-vindas-do-governo,d05b246ce7f0b41e84ffbde84ca5fa1fhv0vat1y.html)
21 [entenda-idas-e-vindas-do-governo,d05b246ce7f0b41e84ffbde84ca5fa1fhv0vat1y.html](https://www.terra.com.br/noticias/educacao/tarcisio-e-material-didatico-impresso-em-sp-entenda-idas-e-vindas-do-governo,d05b246ce7f0b41e84ffbde84ca5fa1fhv0vat1y.html). Acesso
22 em: 8 jan. 2025.
- 23 CAMPBELL, Murray; HOANE, A. Joseph; HSU, Feng-hsiung. Deep Blue. **Artificial**
24 **Intelligence**, [s. l.], v. 134, n. 1, p. 57-83, 2002. Disponível em:
25 <https://www.sciencedirect.com/science/article/pii/S0004370201001291>. Acesso em: 10 jan.
26 2025.
- 27 CANDIDO, Antonio. **Os parceiros do Rio Bonito**. 1ª ediçãoed. São Paulo, SP: Todavia,
28 2023.
- 29 CHRISTIANO, Paul *et al.* **Deep Reinforcement Learning from Human Preferences.**, 2023.
30 Disponível em: <http://arxiv.org/abs/1706.03741>. Acesso em: 7 jan. 2025.
- 31 CLARKE, Michael. The Digital Revolution. *In: Academic and Professional Publishing*. [S.
32 l.]: Chandos Publishing, 2012. p. 79-98. Disponível em:
33 <https://www.sciencedirect.com/science/article/pii/B9781843346692500044>. Acesso em: 4
34 abr. 2023.
- 35 COECKELBERGH, Mark. **AI Ethics**. Cambridge, MA: The MIT Press, 2020a. (The MIT
36 Press Essential Knowledge Series).
- 37 COECKELBERGH, Mark. **Using Words and Things: Language and Philosophy of**
38 **Technology**. [S. l.]: Routledge, 2020b.
- 39 DARWIN, Charles. **Origem das espécies: ou A preservação das raças favorecidas na luta**
40 **pela vida**. São Paulo: Ubu Editora, 2018.

- 1 DEEPMIND. **AlphaGo**. Google DeepMind, 2024. Disponível em:
2 <https://deepmind.google/research/breakthroughs/alphago/>. Acesso em: 11 jan. 2025.
- 3 DEVLIN, Jacob *et al.* **BERT: Pre-training of Deep Bidirectional Transformers for**
4 **Language Understanding.**, 2019. Disponível em: <http://arxiv.org/abs/1810.04805>. Acesso
5 em: 30 jan. 2025.
- 6 FAWZI, Alhussein *et al.* Discovering Faster Matrix Multiplication Algorithms with
7 Reinforcement Learning. **Nature**, [s. l.], v. 610, n. 7930, p. 47–53, 2022. Disponível em:
8 <https://www.nature.com/articles/s41586-022-05172-4>. Acesso em: 12 jan. 2025.
- 9 FERNEDA, Gabriel; BERNARDES, Vinícius. SP abre mão de verba para material
10 didático e usará só livro digital a partir do 6º ano. **CNN Brasil: política**, São Paulo, 8 jan.
11 2023. Disponível em: [https://www.cnnbrasil.com.br/nacional/sp-abre-mao-de-verba-para-](https://www.cnnbrasil.com.br/nacional/sp-abre-mao-de-verba-para-material-didatico-e-usara-so-livro-digital-a-partir-do-6o-ano/)
12 [material-didatico-e-usara-so-livro-digital-a-partir-do-6o-ano/](https://www.cnnbrasil.com.br/nacional/sp-abre-mao-de-verba-para-material-didatico-e-usara-so-livro-digital-a-partir-do-6o-ano/). Acesso em: 8 jan. 2025.
- 13 FLUSSER, Vilém. **A escrita: há futuro para a escrita?** São Paulo: Annablume, 2010.
- 14 FLUSSER, Vilém. **Língua e Realidade**. São Paulo: Annablume, 2007a.
- 15 FLUSSER, Vilém. **O Mundo Codificado**. São Paulo: Cosac Naif, 2007b.
- 16 FREIRE, Paulo. **Conscientização**. Edição: 1ªed. São Paulo - SP: Cortez, 2018.
- 17 FREIRE, Paulo. **Extensão Ou Comunicação?** São Paulo: Paz & Terra, 2006.
- 18 FREITAS, Aline; MACHADO, Livia. Governo de SP avalia utilizar inteligência artificial
19 para 'aprimorar' conteúdo digital nas escolas estaduais | São Paulo. **G1: G1**, São Paulo, 17
20 abr. 2024. Disponível em: [https://g1.globo.com/sp/sao-paulo/noticia/2024/04/17/governo-](https://g1.globo.com/sp/sao-paulo/noticia/2024/04/17/governo-de-sp-avalia-utilizar-inteligencia-artificial-para-aprimorar-conteudo-digital-nas-escolas-estaduais.ghtml)
21 [de-sp-avalia-utilizar-inteligencia-artificial-para-aprimorar-conteudo-digital-nas-escolas-](https://g1.globo.com/sp/sao-paulo/noticia/2024/04/17/governo-de-sp-avalia-utilizar-inteligencia-artificial-para-aprimorar-conteudo-digital-nas-escolas-estaduais.ghtml)
22 [estaduais.ghtml](https://g1.globo.com/sp/sao-paulo/noticia/2024/04/17/governo-de-sp-avalia-utilizar-inteligencia-artificial-para-aprimorar-conteudo-digital-nas-escolas-estaduais.ghtml). Acesso em: 8 jan. 2025.
- 23 GIELOW, Igor. Professores apontam falhas nas aulas digitais de Tarcísio em SP. **Folha de**
24 **S.Paulo: Educação**, São Paulo, 16 ago. 2023. Disponível em:
25 [https://www1.folha.uol.com.br/cotidiano/2023/08/aulas-digitais-de-tarcisio-em-sp-sao-](https://www1.folha.uol.com.br/cotidiano/2023/08/aulas-digitais-de-tarcisio-em-sp-sao-alvo-de-criticas.shtml)
26 [alvo-de-criticas.shtml](https://www1.folha.uol.com.br/cotidiano/2023/08/aulas-digitais-de-tarcisio-em-sp-sao-alvo-de-criticas.shtml). Acesso em: 8 jan. 2025.
- 27 GIGERENZER, Gerd. **How to Stay Smart in a Smart World: Why Human Intelligence**
28 **Still Beats Algorithms**. Cambridge, Massachusetts ; London, England: The MIT Press,
29 2022.
- 30 GILMONT, Jean-François. *Protestan Reformations and Reading. In: A History of Reading*
31 **in the West**. tradução: Lydia G. Cochrane. Amherst: University of Massachusetts Press,
32 1999. (Studies em Print Culture e the History of the Book). p. 213–237.
- 33 GITELMAN, Lisa (org.). **'Raw Data' Is an Oxymoron**. Cambridge, Massachusetts: The
34 MIT Press, 2013. (Infrastructures Series).
- 35 GITELMAN, Lisa; JACKSON, Virginia. Introduction. *In: 'Raw Data' Is an Oxymoron*.
36 Cambridge, Massachusetts: The MIT Press, 2013. (Infrastructures Series). p. 1–14.
- 37 GLOBO, O. É Retrógrada a Oposição Ao Uso de Inteligência Artificial Nas Escolas. RIO
38 de Janeiro, 23 abr. 2024. Disponível em:
39 <https://oglobo.globo.com/opiniao/editorial/coluna/2024/04/e-retrograda-a-oposicao-ao->

- 1 [uso-de-inteligencia-artificial-nas-escolas.ghtml?utm_source=pocket_shared](#). Acesso em: 7
2 jan. 2025.
- 3 GONSALES, Priscila; KAUFMAN, Dora. IA na educação: da programação à alfabetização
4 em dados. **ETD - Educação Temática Digital**, [s. l.], v. 25, p. e023032–e023032, 2023.
5 Disponível em:
6 <https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/8666522>. Acesso em: 7
7 jan. 2025.
- 8 GOOGLE. **COVID-19 Community Mobility Report**. COVID-19 Community Mobility
9 Report, 2022. Disponível em: <https://www.google.com/covid19/mobility?hl=pt-PT>.
10 Acesso em: 6 jan. 2023.
- 11 HAFF, Peter. Humans and Technology in the Anthropocene: Six Rules. **The**
12 **Anthropocene Review**, [s. l.], v. 1, n. 2, p. 126–136, 2014. Disponível em:
13 <http://journals.sagepub.com/doi/10.1177/2053019614530575>. Acesso em: 16 set. 2024.
- 14 HALLIDAY, Michael Alexander Kirkwood; HASAN, Ruqaiya. **Language, Context and**
15 **Text: Aspects of Language in a Social-Semiotic Perspective**. Oxford: Oxford University
16 Press, 1991.
- 17 HAYLES, N. Katherine. Computing the Human. **Theory, Culture & Society**, [s. l.], v. 22,
18 n. 1, p. 131–151, 2005. Disponível em:
19 <http://journals.sagepub.com/doi/10.1177/0263276405048438>. Acesso em: 11 set. 2024.
- 20 HAYLES, N. Katherine. Unfinished Work: From Cyborg to Cognisphere. **Theory, Culture**
21 **& Society**, [s. l.], v. 23, n. 7–8, p. 159–166, 2006. Disponível em:
22 <https://journals.sagepub.com/doi/10.1177/0263276406069229>. Acesso em: 12 set. 2024.
- 23 HEIDEGGER, Martin. The Question Concerning Technology. *In: THE QUESTION*
24 *CONCERNING TECHNOLOGY, AND OTHER ESSAYS*. Reissue editioned. New
25 York/London: Garland Publishing, 1977. p. 3–36.
- 26 HUNTER, Dave. **The Guitar Amp Handbook: Understanding Tube Amplifiers and**
27 **Getting Great Sounds**. Updated editioned. [S. l.]: Backbeat, 2015.
- 28 IBM. **Deep Blue | IBM.**, [s. d.]. Disponível em: <https://www.ibm.com/history/deep-blue>.
29 Acesso em: 10 jan. 2025.
- 30 JULIA, Dominique. Reading and the Counter-Reformation. *In: A History of Reading in*
31 **the West**. tradução: Lydia G. Cochrane. Amherst: University of Massachusetts Press,
32 1999. (Studies em Print Culture e the History of the Book). p. 238–268.
- 33 JUMPER, John *et al.* Highly Accurate Protein Structure Prediction with AlphaFold.
34 **Nature**, [s. l.], v. 596, n. 7873, p. 583–589, 2021. Disponível em:
35 <https://www.nature.com/articles/s41586-021-03819-2>. Acesso em: 12 jan. 2025.
- 36 KANTROWITZ, Alex. **He Helped Train ChatGPT. It Traumatized Him**. Medium, 2023.
37 Disponível em: [https://kantrowitz.medium.com/he-helped-train-chatgpt-it-traumatized-](https://kantrowitz.medium.com/he-helped-train-chatgpt-it-traumatized-him-2ae58f5f91ec)
38 [him-2ae58f5f91ec](https://kantrowitz.medium.com/he-helped-train-chatgpt-it-traumatized-him-2ae58f5f91ec). Acesso em: 17 fev. 2025.
- 39 KAPLAN, Frédéric. A Map for Big Data Research in Digital Humanities. **Frontiers in**
40 **Digital Humanities**, [s. l.], v. 2, 2015. Disponível em:

- 1 http://www.frontiersin.org/Digital_Humanities/10.3389/fdigh.2015.00001/full. Acesso em:
2 9 jun. 2020.
- 3 KÄSSI, Otto; LEHDONVIRTA, Vili; STEPHANY, Fabian. How Many Online Workers
4 Are There in the World? A Data-Driven Assessment. **Open Research Europe**, [s. l.], v. 1,
5 p. 53, 2021. Disponível em: <https://open-research-europe.ec.europa.eu/articles/1-53/v4>.
6 Acesso em: 17 fev. 2025.
- 7 KITCHIN, Rob. Big Data, New Epistemologies and Paradigm Shifts. **Big Data & Society**,
8 [s. l.], v. 1, n. 1–2, p. 1–12, 2014a. Disponível em:
9 [http://bds.sagepub.com/content/1/1/2053951714528481%5Cnhttp://bds.sagepub.com/
10 content/spbds/1/1/2053951714528481.full.pdf](http://bds.sagepub.com/content/1/1/2053951714528481%5Cnhttp://bds.sagepub.com/content/spbds/1/1/2053951714528481.full.pdf).
- 11 KITCHIN, Rob. **The Data Revolution: Big Data, Open Data, Data Infrastructures &**
12 **Their Consequences**. Illustrated edição. Los Angeles, California: Sage Publications Ltd,
13 2014b.
- 14 LEE, Tong King. Artificial Intelligence and Posthumanist Translation: ChatGPT versus
15 the Translator. **Applied Linguistics Review**, [s. l.], 2023a. Disponível em:
16 <https://www.degruyter.com/document/doi/10.1515/applirev-2023-0122/html>. Acesso em:
17 12 set. 2024.
- 18 LEE, Tong King. Distribution and Translation. **Applied Linguistics Review**, [s. l.], v. 14, n.
19 2, p. 369–390, 2023b. Disponível em:
20 <https://www.degruyter.com/document/doi/10.1515/applirev-2020-0139/html>. Acesso em:
21 12 set. 2024.
- 22 LEE, Jinsook *et al.* The Life Cycle of Large Language Models in Education: A Framework
23 for Understanding Sources of Bias. **British Journal of Educational Technology**, [s. l.], v. 55,
24 n. 5, p. 1982–2002, 2024. Disponível em:
25 <https://bera-journals.onlinelibrary.wiley.com/doi/10.1111/bjet.13505>. Acesso em: 7 mar.
26 2025.
- 27 LIMA-LOPES, Rodrigo Esteves de. Estamos apenas repetindo preconceitos? **Observatório**
28 **da Imprensa: Debates**, [s. l.], 8 jul. 2014. Disponível em:
29 [http://www.observatoriodaimprensa.com.br/jornal-de-debates/_ed806_estamos_apenas_
30 repetindo_preconceitos/](http://www.observatoriodaimprensa.com.br/jornal-de-debates/_ed806_estamos_apenas_repetindo_preconceitos/). Acesso em: 22 fev. 2022.
- 31 LIMA-LOPES, Rodrigo Esteves de. Letramento de Dados e Suas Possibilidades Para a
32 Educação Científica Em Linguagem. **Revista Leia Escola**, [s. l.], v. 23, n. 1, p. 62–81, 2023.
33 Disponível em: <https://zenodo.org/record/8021649>. Acesso em: 26 jun. 2023.
- 34 LIMA-LOPES, Rodrigo Esteves de. Oficina de Introdução à Linguística Do Corpus: Relato
35 de Experiência. *In*: VIII ESCOLA BRASILEIRA DE LINGUÍSTICA
36 COMPUTACIONAL & XIII ENCONTRO DE LINGUÍSTICA DE CORPUS, 2016, João
37 Pessoa, Brasil. **Anais Do VIII Escola Brasileira de Linguística Computacional & XIII**
38 **Encontro de Linguística de Corpus**. João Pessoa, Brasil: Editora Edgard Blücher, 2016. p.
39 18–26. Disponível em: [10.5151/sosci-viiiible-xiii-elc-05_artigo_02](https://doi.org/10.5151/sosci-viiiible-xiii-elc-05_artigo_02). Acesso em: 13 jul. 2020.

- 1 LITWAK, Priscila; GAMA, Madson. Inteligência artificial revoluciona a forma de
2 aprender: colégios adotam ferramenta para personalizar estudos. **O Globo: Rio;Bairros**, Rio
3 de Janeiro, 28 set. 2024. Disponível em:
4 [https://oglobo.globo.com/rio/bairros/noticia/2024/09/28/inteligencia-artificial-](https://oglobo.globo.com/rio/bairros/noticia/2024/09/28/inteligencia-artificial-revoluciona-a-forma-de-aprender-colegios-adotam-ferramenta-para-personalizar-estudos.ghtml)
5 [revoluciona-a-forma-de-aprender-colegios-adotam-ferramenta-para-personalizar-](https://oglobo.globo.com/rio/bairros/noticia/2024/09/28/inteligencia-artificial-revoluciona-a-forma-de-aprender-colegios-adotam-ferramenta-para-personalizar-estudos.ghtml)
6 [estudos.ghtml](https://oglobo.globo.com/rio/bairros/noticia/2024/09/28/inteligencia-artificial-revoluciona-a-forma-de-aprender-colegios-adotam-ferramenta-para-personalizar-estudos.ghtml). Acesso em: 7 jan. 2025.
- 7 LIU, Yinhan *et al.* **RoBERTa: A Robustly Optimized BERT Pretraining Approach.**, 2019.
8 Disponível em: <http://arxiv.org/abs/1907.11692>. Acesso em: 7 jan. 2025.
- 9 LOURENÇO, Christine. Educação: inteligência artificial pode otimizar rotina pedagógica
10 de instituições de ensino | Exame. **Exame: Educação**, São Paulo, 31 ago. 2023. Disponível
11 em: [https://exame.com/bussola/educacao-inteligencia-artificial-pode-otimizar-rotina-](https://exame.com/bussola/educacao-inteligencia-artificial-pode-otimizar-rotina-pedagogica-de-instituicoes-de-ensino/)
12 [pedagogica-de-instituicoes-de-ensino/](https://exame.com/bussola/educacao-inteligencia-artificial-pode-otimizar-rotina-pedagogica-de-instituicoes-de-ensino/). Acesso em: 7 jan. 2025.
- 13 MANOVICH, Lev. Novas Mídias Como Tecnologia e Idéia: Dez Definições. *In: O Chip e o*
14 **Caleidoscópio: Reflexões Sobre as Novas Mídias**. São Paulo: Senac, 2005. p. 25–50.
- 15 MANOVICH, Lev. **Trending: The Promises and the Challenges of Big Social Data.** paper,
16 2011. Disponível em: <https://www.manovich.net>. Acesso em: 12 jan. 2025.
- 17 MCCARTHY, John *et al.* A Proposal for the Dartmouth Summer Research Project on
18 Artificial Intelligence, August 31, 1955. **AI Magazine**, [s. l.], v. 27, n. 4, 4, p. 12–12, 2006.
19 Disponível em: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904>.
20 Acesso em: 13 jan. 2025.
- 21 MCLUHAN, Marshall. **The Gutenberg Galaxy: The Making of Typographic Man.**
22 Toronto: University of Toronto Press, 1962.
- 23 MCLUHAN, Marshall. **Understanding Media: Extensions of Man.** Cambridge (US): MIT
24 Press, 1964.
- 25 MCNEILL, Zane. **Kenyan Workers Who Trained ChatGPT Demand Government**
26 **Investigate Work Conditions.** Truthout, 2023. Disponível em:
27 [https://truthout.org/articles/kenyan-workers-who-trained-chatgpt-demand-govt-](https://truthout.org/articles/kenyan-workers-who-trained-chatgpt-demand-govt-investigate-work-conditions/)
28 [investigate-work-conditions/](https://truthout.org/articles/kenyan-workers-who-trained-chatgpt-demand-govt-investigate-work-conditions/). Acesso em: 17 fev. 2025.
- 29 MIAO, Fengchun; HOLMES, Wayne. **Guidance for Generative AI in Education and**
30 **Research.** [S. l.]: UNESCO, 2023. Disponível em:
31 <https://unesdoc.unesco.org/ark:/48223/pf0000386693?locale=en>. Acesso em: 21 out. 2024.
- 32 MOROZOV, Evgeny. **Big Tech: A ascensão dos dados e a morte da política.** São Paulo:
33 Ubu Editora, 2018.
- 34 MOROZOV, Evgeny. **The Net Delusion: The Dark Side of Internet Freedom.** 1st eded.
35 New York: Public Affairs, 2011.
- 36 O'NEIL, Cathy. **Weapons of Math Destruction: How Big Data Increases Inequality and**
37 **Threatens Democracy.** First editioned. New York: Crown, 2016.
- 38 OUYANG, Long *et al.* **Training Language Models to Follow Instructions with Human**
39 **Feedback.**, 2022. Disponível em: <http://arxiv.org/abs/2203.02155>. Acesso em: 7 jan. 2025.

- 1 PAES, Antonio. Os Desafios Da Implementação Da IA Na Educação No Brasil | Colunas |
2 Época NEGÓCIOS. **Epoca Negócios**, Rio de Janeiro, 24 nov. 2024. Disponível em:
3 <https://epocanegocios.globo.com/colunas/coluna/2024/11/os-desafios-da-implementacao->
4 [da-ia-na-educacao-no-brasil.ghtml](https://epocanegocios.globo.com/colunas/coluna/2024/11/os-desafios-da-implementacao-da-ia-na-educacao-no-brasil.ghtml). Acesso em: 7 jan. 2025.
- 5 PERRIGO, Billy. **Exclusive: The \$2 Per Hour Workers Who Made ChatGPT Safer.**
6 **TIME**, 2023. Disponível em: <https://time.com/6247678/openai-chatgpt-kenya-workers/>.
7 Acesso em: 17 fev. 2025.
- 8 POELL, Thomas; NIEBORG, David; VAN DIJCK, José. Platformisation. **Internet Policy**
9 **Review**, [s. l.], v. 8, n. 4, 2019. Disponível em: <https://policyreview.info/node/1425>. Acesso
10 em: 16 nov. 2022.
- 11 PORTO, Douglas. Governo de SP volta atrás e diz que vai usar livros didáticos físicos do
12 MEC. **CNN Brasil: política**, São Paulo, 17 ago. 2023. Disponível em:
13 <https://www.cnnbrasil.com.br/politica/governo-de-sp-volta-atras-e-diz-que-vai-usar-livros->
14 [didaticos-fisicos-do-mec/](https://www.cnnbrasil.com.br/politica/governo-de-sp-volta-atras-e-diz-que-vai-usar-livros-didaticos-fisicos-do-mec/). Acesso em: 8 jan. 2025.
- 15 QUINTAIS, João Pedro. **Generative AI, Copyright and the AI Act.** Rochester, NY, 2025.
16 Disponível em: <https://papers.ssrn.com/abstract=4912701>. Acesso em: 13 fev. 2025.SSRN
17 Scholarly Paper
- 18 RADFORD, Alec; NARASIMHAN, Karthik. **Improving Language Understanding by**
19 **Generative Pre-Training.**, 2018. Disponível em:
20 <https://www.semanticscholar.org/paper/Improving-Language-Understanding-by->
21 [Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035](https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035). Acesso em:
22 7 jan. 2025.
- 23 REN, Jie *et al.* **Copyright Protection in Generative AI: A Technical Perspective.**, 2024.
24 Disponível em: <http://arxiv.org/abs/2402.02333>. Acesso em: 13 fev. 2025.
- 25 SADOWSKI, Jathan. When Data Is Capital: Datafication, Accumulation, and Extraction.
26 **Big Data & Society**, [s. l.], v. 6, n. 1, p. 2053951718820549, 2019. Disponível em:
27 <https://doi.org/10.1177/2053951718820549>. Acesso em: 23 maio 2024.
- 28 SANTOS, Milton. **Técnica, Espaço, Tempo: Globalização e Meio Técnico-Científico.** 5^a
29 edição. [S. l.]: Edusp, 2008.
- 30 SCARTEZINI, Natalia. A fascistização da indignação: as manifestações de 2015 no Brasil.
31 **Cadernos de Campo: Revista de Ciências Sociais**, [s. l.], n. 20, 20, 2016. Disponível em:
32 <https://periodicos.fclar.unesp.br/cadernos/article/view/8019>. Acesso em: 17 jan. 2025.
- 33 SEDUC. **Educação de SP oferece material digital inédito para professores da rede; assista**
34 **ao vídeo.** Secretaria da Educação do Estado de São Paulo, 2023a. Disponível em:
35 <https://www.educacao.sp.gov.br/sala-futuro-educacao-de-sp-oferece-material-digital->
36 [inedito-para-professores-da-rede/](https://www.educacao.sp.gov.br/sala-futuro-educacao-de-sp-oferece-material-digital-inedito-para-professores-da-rede/). Acesso em: 8 jan. 2025.
- 37 SEDUC. **Inteligência artificial a favor de alunos e professores: Redação Paulista agora tem**
38 **apoio de assistente de correção virtual.** Secretaria da Educação do Estado de São Paulo,
39 2023b. Disponível em: <https://www.educacao.sp.gov.br/inteligencia-artificial-favor-de->

- 1 [alunos-e-professores-redacao-paulista-agora-tem-apoio-de-assistente-de-correcao-virtual/](#).
2 Acesso em: 8 jan. 2025.
- 3 SILVA, Julio. Uso de IA nas escolas automatiza aprendizagem e impede a liberdade
4 criativa dos alunos. **Jornal da USP: Atualidades**, São Paulo, 24 abr. 2024. Disponível em:
5 [https://jornal.usp.br/radio-usp/uso-de-ia-nas-escolas-automatiza-aprendizagem-e-impede-](https://jornal.usp.br/radio-usp/uso-de-ia-nas-escolas-automatiza-aprendizagem-e-impede-a-liberdade-criativa-dos-alunos/)
6 [a-liberdade-criativa-dos-alunos/](#). Acesso em: 7 jan. 2025.
- 7 SILVER, David *et al.* A General Reinforcement Learning Algorithm That Masters Chess,
8 Shogi, and Go through Self-Play. **Science**, [s. l.], v. 362, n. 6419, p. 1140–1144, 2018.
9 Disponível em: <https://www.science.org/doi/10.1126/science.aar6404>. Acesso em: 12 jan.
10 2025.
- 11 SILVER, David *et al.* Mastering the Game of Go with Deep Neural Networks and Tree
12 Search. **Nature**, [s. l.], v. 529, n. 7587, p. 484–489, 2016. Disponível em:
13 <https://www.nature.com/articles/nature16961>. Acesso em: 11 jan. 2025.
- 14 SINGH, Gregory Matthew. **The Death of Web 2.0: Ethics, Connectivity and Recognition**
15 **in the Twenty-First Century**. London/New York: Routledge Taylor & Francis Group,
16 2019.
- 17 SØRBØE, Celina Myrann. Urban Uprisings between Revolutionary Openings and
18 Reactionary Outcomes: Making Sense of the 2013 “June Days” in Brazil. **Urban**
19 **Geography**, [s. l.], v. 44, n. 6, p. 1146–1165, 2023. Disponível em:
20 <https://doi.org/10.1080/02723638.2022.2055925>. Acesso em: 17 jan. 2025.
- 21 STIENNON, Nisan *et al.* **Learning to Summarize from Human Feedback.**, 2022.
22 Disponível em: <http://arxiv.org/abs/2009.01325>. Acesso em: 7 jan. 2025.
- 23 STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. **Energy and Policy**
24 **Considerations for Deep Learning in NLP.**, 2019. Disponível em:
25 <http://arxiv.org/abs/1906.02243>. Acesso em: 17 fev. 2025.
- 26 TAYLOR, Wilson L. “Cloze Procedure”: A New Tool for Measuring Readability.
27 **Journalism Quarterly**, [s. l.], v. 30, n. 4, p. 415–433, 1953. Disponível em:
28 <https://journals.sagepub.com/doi/10.1177/107769905303000401>. Acesso em: 30 jan. 2025.
- 29 VASWANI, Ashish *et al.* **Attention Is All You Need.**, 2023. Disponível em:
30 <http://arxiv.org/abs/1706.03762>. Acesso em: 7 jan. 2025.
- 31 WEISER, Mark. The Computer for the 21st Century. **Mobile Computing and**
32 **Communications Review**, [s. l.], v. 3, n. 3, p. 3–11, 1999. Disponível em:
33 <https://dl.acm.org/doi/10.1145/329124.329126>. Acesso em: 10 jan. 2025.
- 34 YADAV, Amit. **Introduction to OpenAI’s GPT Models**. Biased-Algorithms, 2024.
35 Disponível em: [https://medium.com/biased-algorithms/introduction-to-openais-gpt-](https://medium.com/biased-algorithms/introduction-to-openais-gpt-models-1de0d73fa5b5)
36 [models-1de0d73fa5b5](#). Acesso em: 12 jan. 2025.
- 37 ZAHN, Max. **’Overlooked’ Workers Who Train AI Can Face Harsh Conditions, Advocates**
38 **Say**. ABC News, 2024. Disponível em: [https://abcnews.go.com/Business/overlooked-](https://abcnews.go.com/Business/overlooked-workers-train-ai-face-harsh-conditions-advocates/story?id=110303586)
39 [workers-train-ai-face-harsh-conditions-advocates/story?id=110303586](#). Acesso em: 17 fev.
40 2025.

- 1 ZUBOFF, Shoshana. **The Age of Surveillance Capitalism: The Fight for a Human Future**
- 2 **at the New Frontier of Power**. Main ediçãoed. [S. l.]: Profile Books, 2019.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.