

POR UMA REVISÃO CRÍTICA DO USO DE INTELIGÊNCIA ARTIFICIAL NA EDUCAÇÃO

Rodrigo Esteves de Lima-Lopes*

RESUMO

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

Palavras-chave: Tecnosfera. Inteligência artificial. Tecnologia. Ensino. Dados.

* Livre Docente em Linguagem e Tecnologia, professor associado do Departamento de Linguística Aplicada do IEL/UNICAMP. Líder do grupo de pesquisa MiDiTeS (Mídia, Discurso, Tecnologia e Sociedade) e coordenador do PH₂D (Centro de Pesquisa em Pós-Humanismo e Humanidades Digitais), bolsista produtividade CNPq (processo 311099/2021-1) Suas pesquisas estão voltadas para a Linguística do Corpus, uso de dados nos estudos da linguagem, Linguagem e Tecnologia e Linguística Sistemico-Funcional. Universidade Estadual de Campinas/CNPq. Orcid: 0000-0003-3681-1553. E-mail: rll307@gmail.com.

FOR A CRITICAL REVIEW OF ARTIFICIAL INTELLIGENCE IN EDUCATION

ABSTRACT

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

Key-words: Technosphere. Artificial Inteligente. Tecnology. Education. Data.

POR UNA REVISIÓN CRÍTICA DEL USO DE LA INTELIGENCIA ARTIFICIAL EN LA EDUCACIÓN

RESUMEN

Este artículo reflexiona críticamente sobre las implicaciones de la introducción de los Large Language Models o Modelos de Lenguaje a Gran Escala, popularmente conocidos como Inteligencia Artificial, en el ámbito educativo. El ensayo comienza con una discusión sobre cómo se crean estos modelos. Parte de la definición del papel de los datos, reflexionando sobre el mito de su objetividad y cómo su selección y organización pueden generar suposiciones ideológicas y subjetivas. También se analiza el papel de la intervención humana en el proceso de entrenamiento, con el fin de desmitificar la perspectiva de que estos modelos parezcan inteligencias superiores y autónomas, pues, en realidad, se basan en modelos estadísticos de entrenamiento. A continuación, se debate el papel de los diferentes modelos de Inteligencia Artificial en nuestra sociedad, con especial énfasis en los conceptos de invisibilidad tecnológica y singularidad tecnológica, que implican la asimilación de la tecnología en la vida cotidiana, en contraposición al asombro momentáneo, pero sin la necesaria criticidad. Se avanza hacia una discusión sobre los LLMs como parte de nuestro sistema tecnológico-social, con énfasis en sus fundamentos ideológicos y prácticos. El artículo concluye con algunas reflexiones sobre los caminos a seguir si consideramos el uso de los LLMs en situaciones de enseñanza.

Palabras clave: Tecnosfera. Inteligencia Artificial. Tecnología. Enseñanza. Datos.

1 INTRODUÇÃO

A inteligência artificial (IA) tem gerado transformações significativas nas dinâmicas técnico-sociais. Embora ela seja um tema de pesquisa há pelo menos 70 anos, com o desenvolvimento dos chamados *Large Language Models* ou Grandes Modelos Linguísticos (doravante LLMs) ela se tornou algo ubíquo, ganhando destaque em nossas discussões cotidianas, na mídia de massa e em novas áreas da ciência e no ensino. Em tais discussões, ela tende a ser tratada controversamente; ora como uma ferramenta revolucionária, ora como uma ameaça.

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

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

Por conta deste contexto, tomo por objetivo refletir sobre os possíveis efeitos da introdução dos LLMs como componente integrante do currículo. Busco uma reflexão que abra a Caixa de Pandora representada por estes modelos técnicos, explorando seu treinamento e suas implicações sociais mais amplas. Faço-o por acreditar que a escola é o lugar no qual estas reflexões devam ocorrer, para que possamos, enquanto sociedade, entender os efeitos do uso destas ferramentas em uma perspectiva ampla. Ao buscar tal reflexão, concordo com Buckingham (2010, 2022), para quem a escola não é um lugar ultrapassado ou apartado do sistema tecnológico, mas sim o contexto ideal para discutir como a tecnologia

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

pode impactar nosso mundo. De fato, a escola tem sido, historicamente, alvo de diversos programas de introdução tecnológica que datam da introdução da imprensa e se intensificam no pós-guerra, especialmente pelo impacto de capital resultante da adoção deste processo por programas públicos e privados (Buckingham, 2010).

Este artigo é resultado de minha participação em dois eventos, nos quais tive a oportunidade de discutir os LLMs no contexto da educação: a *Semana de Letras*, promovida pelo Campus São José do Rio Preto da Universidade Estadual Paulista-UNESP, e o *CoPeD, Colóquio de Pesquisas Discentes*, promovido pelo Programa de Pós-Graduação em Filologia e Língua Portuguesa da Universidade de São Paulo-USP. Apesar da discussão ter seu foco principal nos impactos dessa implementação na área de Linguística Aplicada, acredito que ela pode ser facilmente extrapolada para outros contextos, uma vez que a pressão social para inclusão rápida, e, por vezes, sem que a devida reflexão seja realizada, parece-me algo generalizado. Busca-se, assim, trazer um olhar crítico a partir da tradição de pesquisas em tecnologia no âmbito da Linguística Aplicada. De fato, poucas são as discussões sobre a natureza de seu treinamento, dos dados que fazem parte dos motores geradores de respostas e, certamente, das implicações do uso de uma tecnologia na qual o professor, assim como o aluno, não têm poder de curadoria e seleção. Afinal, como colocam Gonsales e Kaufman (2023), tal tecnologia parece mudar o eixo das perspectivas do uso de Inteligência Artificial, especialmente se pensarmos que ela sugere uma concepção de utilização baseada na extração de padrões e de conhecimento da máquina e não em saber programá-la para poder-se, assim, extrair resultados desejados. Por outro lado, são comuns discussões sobre o uso de tal ferramenta em sala de aula, com foco principal em sua instrumentalidade.

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

Os LLMs são treinados. Em outras palavras, passam por um sistema de programação para as respostas, o qual é alcançado via modelos estatísticos. Este treinamento depende de uma tríade complementar: dados, modelos estatísticos e reforço humano (Brown *et al.*, 2020; Christiano *et al.*, 2023; Lee *et al.*, 2024; Liu *et al.*, 2019; Stiennon *et al.*, 2022). Os dados, por maiores ou mais compreensivos que o sejam, representam uma seleção naturalmente exclusiva (Gitelman, 2013). Além disso, as diferentes formas de manipulação e estruturação destes dados são ações subjetivas. Elas dependem de nossa concepção daquilo que seja aceitável como dados, algo ligado ao nosso momento histórico-científico e à área da ciência na qual um pesquisador se filia. Nenhum banco de dados representa a totalidade de um fenômeno, a não ser em contexto específico e, mesmo assim, as formas de coleta e análise podem estar embebidas por questões ideológicas diversas. Os modelos estatísticos representam a matematização dos dados, permitindo a tradução para uma linguagem que o computador compreenda. Isso é essencial

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

para que se criem grandes tabelas de co-ocorrência e co-relação que permitem ao programa metrificar a possibilidade de uma palavra ocorrer ao lado da outra. Desse processo nasce a possibilidade de a máquina construir frases e oferecer respostas a nossos questionamentos. Suas respostas são, por conseguinte, baseadas nesse processo de associação que parte das medidas extraídas de seu processamento. Contudo, esse processamento é incapaz de filtrar os resultados a partir daquilo que acreditamos ser cultural e socialmente aceitável. É nesse ponto que a correção humana entra em cena. Ela pode ser realizada de duas formas principais. A primeira utiliza corpora já etiquetados semanticamente, oferecendo à máquina exemplos previamente classificados. A segunda ocorre por meio de correção direta, na qual pessoas corrigem as respostas da máquina em situações específicas. Em alguns casos, o próprio usuário é questionado sobre qual resposta considera mais adequada — um modelo muito similar à lógica de estímulo e resposta.

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

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

2 DE ONDE VEM A INTELIGÊNCIA?

A percepção cotidiana da IA está enraizada em representações que a humanizam. A literatura, o cinema e a televisão nos apresentam robôs antropomórficos que possuem emoções e dilemas morais. Exem-

plos não nos faltam, Rick Deckard e sua busca pelo autoconhecimento em *Blade Runner*; a dupla de robôs em *Star Wars - C3PO* e sua insegurança traduzida em sua voz hesitante ou a comovente lealdade de R2D2 -; ou mesmo Data em *Star Trek*, que luta por aprender emoções como humor e medo. Talvez o computador Hal 9000, em *2001: Uma Odisseia no Espaço*, seja a mais humanizada das personagens do filme. Ele elimina seus inimigos, barganha por sua vida e tenta convencer os humanos, que estão prestes a desligá-lo, de que mudou e não mais incorrerá nos mesmos erros.

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

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

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

Dados podem ser definidos como processos de matematização do comportamento humano ou da natureza (Kitchin, 2014a, 2014b), construindo representações. Todas as inteligências artificiais partem de algum tipo de dado. Mídias sociais, por exemplo, mensuram não somente os cliques, mas também o tempo que gastamos em postagens, aquilo que escrevemos e repostamos, nossas reações e toda sorte de ação. Estes dados são cruzados com os oriundos de nossa lista de contatos, ajudando a construir um perfil nosso e da comunidade que realizamos (O'Neil, 2016). Tal cruzamento é responsável pela criação

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

das diversas bolhas, que nos cercam por conteúdos que reforçam comportamentos. O varejo, outro exemplo, colhe nossos dados de navegação para nos oferecer diferentes produtos e serviços, criando verdadeiros perfis de consumo, em uma perspectiva chamada de capitalismo de vigilância (Zuboff, 2019). A Google disponibilizou de forma anônima relatório de mobilidade urbana durante a pandemia de Covid-19 (Google, 2022), com dados colhidos a partir dos telefones que usam sistema operacional Android ou os aplicativos de mapa disponibilizados pela *big tech*. Tais dados, podiam ser cruzados com aqueles sobre infecção disponibilizados pelos órgãos públicos, criando mapas de mobilidade e alastramento do vírus.⁴ Nesta mesma época houve uma forte migração para sistemas digitais (automatizados e semi-automatizados) de gestão escolar, incluindo atos administrativos e pedagógicos. Tais sistemas de gestão monitoram atividades de alunos e professores, geram metadados e levantam questões sobre privacidade e autonomia.

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

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

Tal perspectiva, por conseguinte, rompe com a ideia de que dados são representativos e neutros (Boyd; Crawford, 2012). Eles podem ser grandes em volume e exaustivos em um universo limitado (Kaplan, 2015; Kitchin, 2014a), mas sempre deixarão de representar algum comportamento não mapeado. Tal realidade dá a qualquer análise quantitativa cores qualitativas. Como todas as atividades humanas,

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

dados são o resultado de operações limitadas e exibem uma perspectiva social e política intrínseca. Há uma escolha relacionada ao que olhar e como olhar; na forma como organizamos, disponibilizamos e interpretamos. Ao aceitar-se tal subjetividade, aceita-se que a interpretação não se inicia após a coleta e organização dos dados, ela está presente desde o planejamento da pesquisa.

2.1 Como as máquinas aprendem?

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

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

2.1.1 Pré-processamento de dados

Neste primeiro momento, os dados textuais são representados em espaços vetoriais, que buscam padrões sobre a combinação de palavras. Buscam-se os grupos nominais (NP) e os grupos verbais (VP) para observar quais itens lexicais estão *embedded*, ou melhor, embutidos em cada grupo. São processadas informações sobre composições de n-gramas,⁶ possibilitando a observação de quais combinações são mais raras e mais comuns por meio da inserção dos n-gramas em que uma palavra se apresenta como metainformação em um vetor (Bojanowski *et al.*, 2017). Nesta fase também são retiradas informações que não são necessárias para a construção do modelo. Por exemplo, dados retirados da internet costumam ter pedaços de códigos HTML⁷ e, por vezes, apresentar diferenças na codificação de caracteres,⁸ ambos causando problemas de leitura e nas estatísticas.

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

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

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

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

2.1.2 Arquitetura do modelo

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

2.1.3 Treinamento do modelo de linguagem

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

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

2.1.4 Validação humana e ajustes

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

Bowman e colaboradores (2015) introduzem um modelo de intervenção humana por meio de corpora. Seu objetivo é demonstrar como um grande corpus anotado por humanos, o SNLI (Stanford

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

Natural Language Inference) corpus, poderia auxiliar a máquina a realizar inferências sobre conteúdo. Tal preocupação nasce do fato de relações lógicas são problemas complexos de serem resolvidos no processo de treinamento. O SNLI é composto por 570.000 pares de sentenças, criados a partir das legendas de um banco de imagens. A anotação feita por seres humanos garante a confiabilidade do resultado. Cada par foi anotado para classificar a relação entre as frases em, pelo menos, uma de três categorias: implicação, contradição e neutralidade. Esse corpus anotado serve como base de verificação se o modelo de LLM consegue inferir de forma satisfatória as relações estabelecidas.

Estudos como Christiano e colaboradores (2023) e Ouyang e colaboradores (2022) e Stiennon e colaboradores (2022) desenvolveram metodologias nas quais humanos intervêm diretamente nos resultados da máquina, avaliando e classificando as respostas.¹⁰ Seres humanos teriam pares de respostas apresentadas para si, tais pares seriam avaliados conforme a preferência dos participantes de pesquisa, que, então, acabam por reforçar comportamentos que mostrem padrões similares. Ouyang et al. (2022), por seu turno, partem de uma ação de pré-treinamento com pares de sentenças anotadas por humanos. Após o processamento pelo LLM, os humanos voltam a avaliar estes pares, criando um modelo que reflete aquilo que os avaliadores preferem. Por fim, essas preferências são alimentadas no sistema à guisa de reforço. Por fim, Stiennon e colaboradores (2022) trabalham com resumos produzidos pelo sistema, sendo que tais resumos são julgados por humanos, que manifestam sua preferência. O reforço positivo faz com que as máquinas passem a seguir o padrão melhor avaliado.

3 REFLETINDO SOBRE A IA NA EDUCAÇÃO

3.1 Realmente algo novo?

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

A lista seria quase infinita, assim como seria impossível pensar em nossa vida contemporânea sem estas diversas Inteligências Artificiais. Isso nos leva a perguntar: por que não as percebemos? Por que

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

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

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

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

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

Quando uma nova tecnologia e suas expressões midiáticas¹¹ são introduzidas em nosso cotidiano, há um impacto na forma como vemos o mundo e nossas possibilidades de futuro. Segundo Coeckelbergh (2020a, p. 12), isso é motivado pela ideia de que vivemos uma *singularidade tecnológica*. Essa singularidade daria ao momento histórico em que vivemos um papel e um lugar únicos, responsável por um crescimento tecnológico exponencial que marca o progresso, promovendo uma transformação dramática. Cria-se um novo mundo, que nós não mais entendemos e

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

compreendemos como dantes. Estas tecnologias, por seu turno, nos são opacas, uma espécie de *Caverna de Platão* invertida: clara para aqueles que a criam, mas opaca para aqueles que estão fora e simplesmente a usam (Bridle, 2018).

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

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

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

3.2 LLMs: Questões para reflexão

Como professores da área de ciências humanas, carregamos conosco a opinião de que aquilo que praticamos e ensinamos é uma linguagem de compreensão do mundo em suas diversas instâncias. Ao mesmo tempo, temos certeza de que estas possibilidades são infinitas e não encerradas em um juízo

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

único e definitivo. Diferentemente desta perspectiva, aplicativos como os LLMs caminham em uma direção contrária: eles são dados estáticos. Por maiores que eles sejam, eles representam apenas aquilo que puderam matematizar, apenas aquilo que se corrigiu pela interpretação humana. De uma certa forma, os LLMs ainda são dependentes de nosso insumo, representando uma versão estatisticamente compreensiva, mas simbolicamente menor que a cognição humana. A IA pode mimetizar, todavia não recriar, a cognição humana, pois os processos são significativamente diferentes (Coeckelbergh, 2020a).

Gostaria de sugerir que vejamos a IA, assim como todos os processos tecnológicos que historicamente a precedem ou antecedem, na perspectiva da *Tecnosfera*, como definida por Santos (2008) e por Haff (2014). Para Santos (2008, p. 30) que estabeleceu o conceito nos anos 1990, a tecnosfera poderia ser definida como uma constante artificialização do ambiente, na qual a esfera natural seria substituída, gradativamente, por uma esfera técnica.¹³ Para o autor, alinhada a este conceito estaria a psicofera, ou seja, nossa atitude filosófica perante o mundo, incluindo certamente a tecnologia. Apesar de sua importância histórica, a definição de Santos (2008) separa a existência de um mundo natural de sua perspectiva técnica. Para o autor, a tecnologia é algo artificial e externo que transforma a relação do homem natural com seu mundo. Por conta disso, acredito que este conceito possa ser ampliado a partir do sugerido por Haff (2014), para quem a tecnosfera deve abranger todos os sistemas técnicos humanos, nossos processos tecnológicos compreensivamente, além de sistemas que tradicionalmente consideramos humanos, como as ONGs e entidades religiosas.

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

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

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

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

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

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

3.2.1 Variações e estilo

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

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

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

3.2.2 O papel do professor e do material

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

Os materiais distribuídos pelo MEC às escolas públicas de educação básica do país são escolhidos pelas escolas, desde que inscritos no PNLD e aprovados em avaliações pedagógicas coordenadas pelo Ministério da Educação e conta com a participação de Comissões Técnica específica, integrada por especialistas das diferentes áreas do conhecimento correlatas, cuja vigência corresponderá ao ciclo a que se referir o processo de avaliação (Brasil, 2024).

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

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

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

3.2.3 Privacidade

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

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

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

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

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

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

3.2.4 Direitos autorais e responsabilidade

Da mesma forma que jornalistas são responsáveis por seus textos, escritores e pesquisadores também o são. Mas ao utilizarmos materiais gerados por um LLM, quem são os responsáveis? Uma vez que as

respostas oferecidas por estes modelos são o resultado de uma matriz de dados, como regular o direito autoral?

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

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

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

Questiona-se, então, seriam estes modelos legalmente responsáveis pelas informações que nos oferecem? O ponto é polêmico. Por não haver criatividade simbólica, eles acabam sempre repetindo informações com as quais foram treinados, projetando, por conseguinte, a responsabilidade em suas fontes. Entretanto, é salutar pensar que a escolha de tais fontes é responsabilidade daqueles que desenham seu treinamento. Ou seja, se o LLM foi treinado com textos racistas ou misóginos, ele tenderá a repetir esses valores em seus resultados. Nota-se, também, que a correção humana pode não dar conta de todas as questões pretensamente objetivas sobre conteúdo, o que pode levar à falta de acuidade, algo que tem impacto direto no usuário, que confia na informação.

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

3.2.5 Meio ambiente, sustentabilidade e emprego

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

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

Há diversos trabalhos que têm como foco a segunda questão. Apesar de não haver esta ênfase ao se oferecerem os sistemas para uso, todos os LLMs são dependentes deste tipo de infraestrutura. Em outras palavras, LLMs podem contribuir sobremaneira para o aquecimento global graças ao seu consumo de energia e emissões. Bridle (2018) discute como os diferentes *data centers* — ou centrais que armazenam e distribuem nossos dados quando em servidores externos, ou nuvens — causam aquecimento no meio ambiente dos locais em que são instalados, além de seu grande consumo de energia para manutenção e de água para resfriamento. Em um estudo comparativo Strubell e colaboradores (2019), observam que o treinamento de um modelo de LLM pode chegar a consumir 626.155 CO₂ e lbs (Libras de Carbono), algo quase seis vezes superior a toda existência de um veículo movido motor de explosão alimentado por combustível fóssil (126.000 CO₂ e lbs) e quase 57 vezes a média de uma vida humana (11.023 CO₂ lbs). Já o treinamento de um modelo poderia custar até 3 milhões de dólares somente em termos da energia elétrica consumida.¹⁵

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

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

4 UM CAMINHO A SER PERCORRIDO

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

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

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

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

Deve-se ressaltar que os modelos humanos de cognição emergem de uma combinação única de experiências e emoções. Nossa capacidade de entender, responder a emoções e necessidades alheias

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

Ao pensarmos no ensino de linguagens — sendo o ensino de línguas algo certamente considerável neste aspecto —, o uso de LLMs pode gerar padronizações indesejadas. Tal fato é o resultado daquilo que a faz impressionar à primeira vista: os dados. Como eles são treinados por um grupo de pessoas que os selecionam — mesmo que grandes em quantidade —, eles têm uma tendência totalitária em sua gênese. Em outras palavras, se os dados são sempre finitos e excludentes, também o serão as representações de linguagem que estão neles presentes, incapazes de abarcar qualquer totalidade de comportamento social. A pseudo-autoridade delegada a estes sistemas pode levar a uma padronização nos resultados oferecidos, impactando diretamente as perspectivas inter e multiculturais no ensino de linguagens. Garantir esta diversidade vem da desmitificação desses resultados e da criação de modelos que reflitam nossa cultura. Mesmo sabendo que jamais conseguiremos refletir toda a diversidade cultural e de pensamento de um país como o Brasil em um modelo de LLM, mesmo assumindo que esta tarefa seria algo utópico, certamente um desenvolvimento local estaria mais próximo de garantir esta diversidade. Enquanto tais modelos ainda não são desenvolvidos, observar os resultados com criticidade de forma conjunta com alunos e integrada aos demais membros da comunidade escolar pode ser algo positivo.

Uma visão sobre a IA que não a integra na tecnosfera, mas a vê como uma singularidade, traz, em si, um paradoxo. A este ufanismo tecnológico é creditada uma garantia de personalização, inovação e aceleração de resposta. Entretanto, como ela na essência é o resultado de um banco de dados, de uma grande matriz e seus cruzamentos estatísticos possíveis e prováveis, até que ponto ela garante

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

resultados personalizados? Na verdade, o que se vê é uma grande padronização, nos quais resultados similares são distribuídos uniformemente pelas diversas consultas.

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

A equidade de acesso é outra discussão relevante (Miao; Holmes, 2023). Diferentes grupos sociais podem ter acesso limitado a tais ferramentas, seja pela conectividade, seja pela necessidade de hardware capaz de processar tais modelos, seja pelo pagamento de taxas ou assinaturas. Essa realidade leva à necessidade de criação de políticas de acesso que podem, inclusive, esbarrar em questões de soberania e política social mais ampla, uma vez que a intervenção humana está presente desde a construção do modelo até seu treinamento e correção. Talvez por isso o desenvolvimento de modelos de LLM nacionais, como dados de treinamento que reflitam um desenho específico para a língua portuguesa e visando refletir nossa cultura e diversidade, é imperativo neste processo. Mas, por que isso seria importante? Ao importarmos modelos criados em outros contextos políticos e culturais, estamos, necessariamente, trazendo a reboque representações políticas e semióticas, valores sociais e padrões de linguagem que podem levar situações de desvalorização de culturas locais.

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

Villén Flusser (2010), ao questionar o impacto da técnica em nossa sociedade, realiza uma importante reflexão sobre o papel que devemos dar ao ensino da escrita. O autor observa que, à medida que os

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

Agradecimentos

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

REFERÊNCIAS

- ADORNO, Theodor W. *Indústria cultural e sociedade*. São Paulo (SP): Paz e Terra, 2009.
- APEOESP. *Inteligência Artificial é instrumento auxiliar, jamais poderá substituir o professor!* São Paulo, abr. 2024. {Nota}. Acesso em: 7 jan. 2025.
- BALMANT, Ocimara. Inteligência artificial é bem-vinda na educação, mas precisa de regulação, defendem especialistas. *OESP*, São Paulo, 2024. Acesso em: 7 jan. 2025.
- BOJANOWSKI, Piotr *et al.* *Enriching Word Vectors with Subword Information*. arXiv, 2017. Disponível em: <https://arxiv.org/abs/1607.04606>. Acesso em: 29 jan. 2025.
- BOWMAN, Samuel R. *et al.* *A Large Annotated Corpus for Learning Natural Language Inference*. arXiv, 2015. Disponível em: <https://arxiv.org/abs/1508.05326>. Acesso em: 30 jan. 2025.
- BOYD, Danah; CRAWFORD, Kate. Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon. *Information Communication and Society*, [s. l.], v. 15, n. 5, p. 662–679, 2012.
- BRASIL. *PNLD (Programa Nacional do Livro didático)*. Ministério da Educação <http://portal.mec.gov.br/component/content/article?id=12391:pnld>, 2024. Disponível em: Acesso em: 12 fev. 2025. {Org{\~a}o P{\^u}blico}
- BRIDLE, James. *New Dark Age: Technology, Knowledge and the End of the Future*. London ; Brooklyn, NY: Verso, 2018.
- BROWN, Tom B. *et al.* *Language Models Are Few-Shot Learners*. arXiv, 2020. Disponível em: <https://arxiv.org/abs/2005.14165>. Acesso em: 7 jan. 2025.
- BUCKINGHAM, David. Cultura Digital, Educação Midiática e o Lugar da Escolarização. *Educação e Realidade*, [s. l.], v. 35, n. 3, p. 37–58, 2010.
- BUCKINGHAM, David. *Manifesto pela educação midiática*. tradução: José Ignacio Coelho Mendes Neto. São Paulo, SP: Sesc Sp, 2022.
- CAFARDO, Renata. Governo de SP vai revisar slides de aulas após erros e recomenda novo download para professores. *Estadão*, São Paulo, 2023a. Acesso em: 8 jan. 2025.

- CAFARDO, Renata. Tarcísio e material didático impresso em SP: entenda idas e vindas do governo. *Terra/Educar*, São Paulo, 2023b. Acesso em: 8 jan. 2025.
- CAMPBELL, Murray; HOANE, A. Joseph; HSU, Feng-hsiung. Deep Blue. *Artificial Intelligence*, [s. l.], v. 134, n. 1, p. 57–83, 2002. Acesso em: 10 jan. 2025.
- CANDIDO, Antonio. *Os parceiros do Rio Bonito*. 1ª edição. São Paulo, SP: Todavia, 2023.
- CHRISTIANO, Paul et al. *Deep Reinforcement Learning from Human Preferences*. arXiv, 2023. Disponível em: <https://arxiv.org/abs/1706.03741>. Acesso em: 7 jan. 2025.
- CLARKE, Michael. The Digital Revolution. In: *Academic and Professional Publishing*. [S. l.]: Chandos Publishing, 2012. p. 79–98. Acesso em: 4 abr. 2023.
- COECKELBERGH, Mark. *AI Ethics*. Cambridge, MA: The MIT Press, 2020a. (The MIT Press Essential Knowledge Series).
- COECKELBERGH, Mark. *Using Words and Things: Language and Philosophy of Technology*. [S. l.]: Routledge, 2020b.
- DARWIN, Charles. *Origem das espécies: ou A preservação das raças favorecidas na luta pela vida*. São Paulo: Ubu Editora, 2018.
- DEEPMIND. *AlphaGo. Google DeepMind*. Disponível em: <https://deepmind.google/research/breakthroughs/alphago/>, 2024. Acesso em: 11 jan. 2025.
- DEVLIN, Jacob et al. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv, 2019. Disponível em: <https://arxiv.org/abs/1810.04805>. Acesso em: 30 jan. 2025.
- FAWZI, Alhussein et al. Discovering Faster Matrix Multiplication Algorithms with Reinforcement Learning. *Nature*, [s. l.], v. 610, n. 7930, p. 47–53, 2022. Acesso em: 12 jan. 2025.
- FERNEDA, Gabriel; BERNARDES, Vinícius. SP abre mão de verba para material didático e usará só livro digital a partir do 6º ano. *CNN Brasil*, São Paulo, 2023. Acesso em: 8 jan. 2025.
- FLUSSER, Vilém. *A escrita: há futuro para a escrita?* São Paulo: Annablume, 2010.
- FLUSSER, Vilém. *Língua e Realidade*. São Paulo: Annablume, 2007a.
- FLUSSER, Vilém. *O Mundo Codificado*. São Paulo: Cosac Naif, 2007b.
- FREIRE, Paulo. *Conscientização*. Edição: 1ª ed. São Paulo - SP: Cortez, 2018.
- FREIRE, Paulo. *Extensão Ou Comunicação?* São Paulo: Paz & Terra, 2006.
- FREITAS, Aline; MACHADO, Livia. Governo de SP avalia utilizar inteligência artificial para 'aprimorar' conteúdo digital nas escolas estaduais São Paulo. *G1*, São Paulo, 2024. Acesso em: 8 jan. 2025.
- GIELOW, Igor. Professores apontam falhas nas aulas digitais de Tarcísio em SP. *Folha de S.Paulo*, São Paulo, 2023. Acesso em: 8 jan. 2025.
- GIGERENZER, Gerd. *How to Stay Smart in a Smart World: Why Human Intelligence Still Beats Algorithms*. Cambridge, Massachusetts; London, England: The MIT Press, 2022.
- GILMONT, Jean-François. Protestant Reformations and Reading. In: *A History of Reading in the West*. tradução: Lydia G. Cochrane. Amherst: University of Massachusetts Press, 1999. (Studies em Print Culture e the History of the Book). p. 213–237.
- GITELMAN, Lisa (org.). *"Raw Data" Is an Oxymoron*. Cambridge, Massachusetts: The MIT Press, 2013. (Infrastructure Series).
- GITELMAN, Lisa; JACKSON, Virginia. Introduction. In: *"Raw Data" Is an Oxymoron*. Cambridge, Massachusetts: The MIT Press, 2013. (Infrastructure Series). p. 1–14.

- GLOBO, O. É Retrógrada a Oposição Ao Uso de Inteligência Artificial Nas Escolas. RÍo de Janeiro, 2024. Acesso em: 7 jan. 2025.
- GONSALES, Priscila; KAUFMAN, Dora. IA na educação: da programação à alfabetização em dados. *ETD - Educação Temática Digital*, [s. l.], v. 25, p. e023032–e023032, 2023. Acesso em: 7 jan. 2025.
- GOOGLE. *COVID-19 Community Mobility Report*. *COVID-19 Community Mobility Report*. Disponível em: <https://www.google.com/covid19/mobility?hl=pt-PT>, 2022. Acesso em: 6 jan. 2023.
- HAFF, Peter. Humans and Technology in the Anthropocene: Six Rules. *The Anthropocene Review*, [s. l.], v. 1, n. 2, p. 126–136, 2014. Acesso em: 16 set. 2024.
- HALLIDAY, Michael Alexander Kirkwood; HASAN, Ruqaiya. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford: Oxford University Press, 1991.
- HAYLES, N. Katherine. Computing the Human. *Theory, Culture & Society*, [s. l.], v. 22, n. 1, p. 131–151, 2005. Acesso em: 11 set. 2024.
- HAYLES, N. Katherine. Unfinished Work: From Cyborg to Cognisphere. *Theory, Culture & Society*, [s. l.], v. 23, n. 7-8, p. 159–166, 2006. Acesso em: 12 set. 2024.
- HEIDEGGER, Martin. The Question Concerning Technology. In: *THE QUESTION CONCERNING TECHNOLOGY, AND OTHER ESSAYS*. Reissue editioned. New York/London: Garland Publishing, 1977. p. 3–36.
- HUNTER, Dave. *The Guitar Amp Handbook: Understanding Tube Amplifiers and Getting Great Sounds*. Updated editioned. [S. l.]: Backbeat, 2015.
- IBM. *Deep Blue* IBM. <https://www.ibm.com/history/deep-blue>, [s. d.]. Acesso em: 10 jan. 2025.
- JULIA, Dominique. Reading and the Counter-Reformation. In: *A History of Reading in the West*. tradução: Lydia G. Cochrane. Amherst: University of Massachusetts Press, 1999. (Studies em Print Culture e the History of the Book). p. 238–268.
- JUMPER, John *et al.* Highly Accurate Protein Structure Prediction with AlphaFold. *Nature*, [s. l.], v. 596, n. 7873, p. 583–589, 2021. Acesso em: 12 jan. 2025.
- KANTROWITZ, Alex. *He Helped Train ChatGPT. It Traumatized Him*. Medium <https://kantrowitz.medium.com/he-helped-train-chatgpt-it-traumatized-him-2ae58f5f91ec>, 2023. Acesso em: 17 fev. 2025.
- KAPLAN, Frédéric. A Map for Big Data Research in Digital Humanities. *Frontiers in Digital Humanities*, [s. l.], v. 2, 2015. Acesso em: 9 jun. 2020.
- KÄSSI, Otto; LEHDONVIRTA, Vili; STEPHANY, Fabian. How Many Online Workers Are There in the World? A Data-Driven Assessment. *Open Research Europe*, [s. l.], v. 1, p. 53, 2021. Acesso em: 17 fev. 2025.
- KITCHIN, Rob. Big Data, New Epistemologies and Paradigm Shifts. *Big Data & Society*, [s. l.], v. 1, n. 1-2, p. 1–12, 2014a.
- KITCHIN, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. Illustrated edição. Los Angeles, California: Sage Publications Ltd, 2014b.
- LEE, Tong King. Artificial Intelligence and Posthumanist Translation: ChatGPT versus the Translator. *Applied Linguistics Review*, [s. l.], 2023a. Acesso em: 12 set. 2024.
- LEE, Tong King. Distribution and Translation. *Applied Linguistics Review*, [s. l.], v. 14, n. 2, p. 369–390, 2023b. Acesso em: 12 set. 2024.
- LEE, Jinsook *et al.* The Life Cycle of Large Language Models in Education: A Framework for Understanding Sources of Bias. *British Journal of Educational Technology*, [s. l.], v. 55, n. 5, p. 1982–2002, 2024. Acesso em: 7 mar. 2025.

- LIMA-LOPES, Rodrigo Esteves de. Estamos apenas repetindo preconceitos? *Observatório da Imprensa*, [s. l.], 2014. Acesso em: 22 fev. 2022.
- LIMA-LOPES, Rodrigo Esteves de. Letramento de Dados e Suas Possibilidades Para a Educação Científica Em Linguagem. *Revista Leia Escola*, [s. l.], v. 23, n. 1, p. 62–81, 2023. Acesso em: 26 jun. 2023.
- LIMA-LOPES, Rodrigo Esteves de. Oficina de Introdução à Linguística Do Corpus: Relato de Experiência. In: OFICINA DE INTRODUÇÃO À LINGUÍSTICA DO CORPUS, 2016, João Pessoa, Brasil. *Anais do VIII Escola Brasileira de Linguística Computacional & XIII Encontro de Linguística de Corpus*. João Pessoa, Brasil: Editora Edgard Blücher, 2016. p. 18–26. Acesso em: 13 jul. 2020.
- LITWAK, Priscila; GAMA, Madson. Inteligência artificial revoluciona a forma de aprender: colégios adotam ferramenta para personalizar estudos. *O Globo*, RIO de Janeiro, 2024. Acesso em: 7 jan. 2025.
- LIU, Yinhan *et al.* RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv, 2019. Disponível em: <https://arxiv.org/abs/1907.11692>. Acesso em: 7 jan. 2025.
- LOURENÇO, Christine. Educação: inteligência artificial pode otimizar rotina pedagógica de instituições de ensino Exame. *Exame*, São Paulo, 2023. Acesso em: 7 jan. 2025.
- MANOVICH, Lev. Novas Mídias Como Tecnologia e Idéia: Dez Definições. In: *O Chip e o Caleidoscópio: Reflexões Sobre as Novas Mídias*. São Paulo: Senac, 2005. p. 25–50.
- MANOVICH, Lev. *Trending: The Promises and the Challenges of Big Social Data*. 2011. Paper. Acesso em: 12 jan. 2025.
- MCCARTHY, John *et al.* A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, [s. l.], v. 27, n. 4, p. 12–12, 2006. Acesso em: 13 jan. 2025.
- MCLUHAN, Marshall. *The Gutenberg Galaxy: The Making of Typographic Man*. Toronto: University of Toronto Press, 1962.
- MCLUHAN, Marshall. *Understanding Media: Extensions of Man*. Cambridge (US): MIT Press, 1964.
- MCNEILL, Zane. *Kenyan Workers Who Trained ChatGPT Demand Government Investigate Work Conditions*. Truthout. Disponível em: <https://truthout.org/articles/kenyan-workers-who-trained-chatgpt-demand-govt-investigate-work-conditions/>, 2023. Acesso em: 17 fev. 2025.
- MIAO, Fengchun; HOLMES, Wayne. *Guidance for Generative AI in Education and Research*. [S. l.]: UNESCO, 2023. Acesso em: 21 out. 2024.
- MOROZOV, Evgeny. *Big Tech: A ascensão dos dados e a morte da política*. São Paulo: Ubu Editora, 2018.
- MOROZOV, Evgeny. *The Net Delusion: The Dark Side of Internet Freedom*. 1st eded. New York: Public Affairs, 2011.
- O'NEIL, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. First editioned. New York: Crown, 2016.
- OUYANG, Long *et al.* *Training Language Models to Follow Instructions with Human Feedback*. arXiv, 2022. Disponível em: <https://arxiv.org/abs/2203.02155>. Acesso em: 7 jan. 2025.
- PAES, Antonio. Os Desafios Da Implementação Da IA Na Educação No Brasil Colunas Época NEGÓCIOS. *Epoca Negócios*, Rio de Janeiro, 2024. Acesso em: 7 jan. 2025.
- PERRIGO, Billy. *Exclusive: The \$2 Per Hour Workers Who Made ChatGPT Safer*. TIME. Disponível em: <https://time.com/6247678/openai-chatgpt-kenya-workers/>, 2023. Acesso em: 17 fev. 2025.
- POELL, Thomas; NIEBORG, David; VAN DIJCK, José. Platformisation. *Internet Policy Review*, [s. l.], v. 8, n. 4, 2019. Acesso em: 16 nov. 2022.
- PORTO, Douglas. Governo de SP volta atrás e diz que vai usar livros didáticos físicos do MEC. *CNN Brasil*, São Paulo, 2023. Acesso em: 8 jan. 2025.

- QUINTAIS, João Pedro. *Generative AI, Copyright and the AI Act*. Rochester, NY: Social Science Research Network, 2025. Disponível em: Acesso em: 13 fev. 2025. {{SSRN Scholarly Paper}}
- RADFORD, Alec; NARASIMHAN, Karthik. *Improving Language Understanding by Generative Pre-Training*. 2018. DisAcesso em: 7 jan. 2025.
- REN, Jie *et al.* *Copyright Protection in Generative AI: A Technical Perspective*. arXiv, 2024. Disponível em: <https://arxiv.org/abs/2402.02333>. Acesso em: 13 fev. 2025.
- SADOWSKI, Jathan. When Data Is Capital: Datafication, Accumulation, and Extraction. *Big Data & Society*, [s. l.], v. 6, n. 1, p. 2053951718820549, 2019. Disponível em: Acesso em: 23 maio 2024.
- SANTOS, Milton. *Técnica, Espaço, Tempo: Globalização e Meio Técnico-Científico*. 5ª edição. [S. l.]: Edusp, 2008.
- SCARTEZINI, Natalia. A fascitização da indignação: as manifestações de 2015 no Brasil. *Cadernos de Campo: Revista de Ciências Sociais*, [s. l.], n. 20, 2016. Disponível em: Acesso em: 17 jan. 2025.
- SEDUC. *Educação de SP oferece material digital inédito para professores da rede; assista ao vídeo*. Secretaria da Educação do Estado de São Paulo. Disponível em: <https://www.educacao.sp.gov.br/sala-futuro-educacao-de-sp-oferece-material-digital-inedito-para-professores-da-rede/>, 2023a. Acesso em: 8 jan. 2025.
- SEDUC. *Inteligência artificial a favor de alunos e professores: Redação Paulista agora tem apoio de assistente de correção virtual*. Secretaria da Educação do Estado de São Paulo. Disponível em: <https://www.educacao.sp.gov.br/inteligencia-artificial-favor-de-alunos-e-professores-redacao-paulista-agora-tem-apoio-de-assistente-de-correcao-virtual/>, 2023b. Acesso em: 8 jan. 2025.
- SILVA, Julio. Uso de IA nas escolas automatiza aprendizagem e impede a liberdade criativa dos alunos. *Jornal da USP*, São Paulo, 2024. Acesso em: 7 jan. 2025.
- SILVER, David *et al.* A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go through Self-Play. *Science*, [s. l.], v. 362, n. 6419, p. 1140–1144, 2018. Acesso em: 12 jan. 2025.
- SILVER, David *et al.* Mastering the Game of Go with Deep Neural Networks and Tree Search. *Nature*, [s. l.], v. 529, n. 7587, p. 484–489, 2016. Acesso em: 11 jan. 2025.
- SINGH, Gregory Matthew. *The Death of Web 2.0: Ethics, Connectivity and Recognition in the Twenty-First Century*. London/New York: Routledge Taylor & Francis Group, 2019.
- SØRBØE, Celina Myrann. Urban Uprisings between Revolutionary Openings and Reactionary Outcomes: Making Sense of the 2013 “June Days” in Brazil. *Urban Geography*, [s. l.], v. 44, n. 6, p. 1146–1165, 2023. Acesso em: 17 jan. 2025.
- STIENNON, Nisan *et al.* *Learning to Summarize from Human Feedback*. arXiv, 2022. Disponível em: <https://arxiv.org/abs/2009.01325>. Acesso em: 7 jan. 2025.
- STRUBELL, Emma; GANESH, Ananya; MCCALLUM, Andrew. *Energy and Policy Considerations for Deep Learning in NLP*. arXiv, 2019. Disponível em: <https://arxiv.org/abs/1906.02243>. Acesso em: 17 fev. 2025.
- TAYLOR, Wilson L. “Cloze Procedure”: A New Tool for Measuring Readability. *Journalism Quarterly*, [s. l.], v. 30, n. 4, p. 415–433, 1953. Acesso em: 30 jan. 2025.
- VASWANI, Ashish *et al.* *Attention Is All You Need*. arXiv, 2023. Disponível em: <https://arxiv.org/abs/1706.03762>. Acesso em: 7 jan. 2025.
- WEISER, Mark. The Computer for the 21st Century. *Mobile Computing and Communications Review*, [s. l.], v. 3, n. 3, p. 3–11, 1999. Disponível em: Acesso em: 10 jan. 2025.
- YADAV, Amit. *Introduction to OpenAI’s GPT Models. Biased-Algorithms*, 2024. Disponível em: Acesso em: 12 jan. 2025.

ZAHN, Max. *'Overlooked' Workers Who Train AI Can Face Harsh Conditions, Advocates Say*. ABC News. Disponível em: <https://abcnews.go.com/Business/overlooked-workers-train-ai-face-harsh-conditions-advocates/story?id=110303586>, 2024. Acesso em: 17 fev. 2025.

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

