

IA generativa e educação práticas e teorizações

Mariano Pimentel e Felipe Carvalho



IA generativa e educação: práticas e teorizações

Mariano Pimentel e Felipe Carvalho

Porto Alegre
Sociedade Brasileira de Computação – SBC
2025

© 2025 Mariano Pimentel e Felipe Carvalho.



Esta obra está sob a licença Creative Commons Atribuição 4.0 (CC-BY). Você pode redistribuir este livro em qualquer suporte ou formato e copiar, remixar, transformar e criar a partir do conteúdo deste livro para qualquer fim, desde que cite a fonte.

Dados Internacionais de Catalogação na Publicação (CIP)

I11 IA generativa e educação: práticas e teorizações.

Mariano Pimentel e Felipe Carvalho. Dados eletrônicos Porto Alegre: Sociedade Brasileira de Computação, 2025. 343 p. : il. : PDF; 65MB.

Modo de acesso: World Wide Web.

ISBN 978-85-7669-658-2 (e-book)

1. Inteligência artificial generativa. 2. Educação aplicada à Computação. I. Pimentel, Mariano. II. Carvalho, Felipe. III. Sociedade Brasileira de Computação. IV. Título.

CDU 004.89:37

Ficha catalográfica elaborada por Annie Casali – CRB-10/2339
Biblioteca Digital da SBC – SBC OpenLib

Capa e ilustrações: Mônica Lopes



Sociedade Brasileira de Computação

Av. Bento Gonçalves, 9500

Setor 4 | Prédio 43.412 | Sala 219 | Bairro Agronomia

Caixa Postal 15012 | CEP 91501-970

Porto Alegre - RS Fone: (51) 99252-6018

sbc@sbc.org.br



Acesse a versão **weblivro**, que se adapta ao tamanho da tela do *desktop*, *laptop*, *tablet* ou *smartphone*:

<http://iaeducacao.uniriotec.br>

“

Não é meu objetivo surpreendê-los ou chocá-los,
mas o modo mais simples de resumir tudo isso
é dizendo que agora existem, no mundo,
máquinas que pensam, aprendem e criam.
Além disso, sua capacidade de realizar essas atividades
está crescendo rapidamente até o ponto
— em um futuro visível —
no qual a variedade de problemas
com que elas poderão lidar
será correspondente à variedade de problemas
com os quais lida a mente humana.

”

(Simon, 1957 *apud* Russell; Norvig, 2022, n.p.)

Dedicatória

Para profissionais da educação que buscam compreender a Inteligência Artificial (IA) generativa como parceira na invenção de práticas formativas que mobilizam a aprendizagem pela conversação e pela cocriação com a IA. Precisamos compreender seu potencial pedagógico sem ignorar os riscos e as ameaças que ela também representa para a educação.

Este livro não é sobre as especificidades dos sistemas computacionais baseados em IA generativa, lançados a uma velocidade vertiginosa difícil de acompanhar. Em vez disso, propõe uma reflexão sobre a IA generativa: seus fundamentos ontológicos e suas implicações pedagógicas, epistemológicas, axiológicas e éticas, que permanecerão válidas ao longo do tempo, independentemente dos novos sistemas e modelos de linguagem que venham a surgir.

Agradecimentos

Agradecemos às/aos docentes que participaram das oficinas e cursos que lecionamos sobre IA generativa e educação. Agradecemos especialmente às/aos professoras/es que participaram dos cursos ofertados pela EGGP/PROGEPE/UFF, pois foram as/os primeiras/os leitoras/es deste livro e nos ajudaram a (re)pensar esta obra, que se tornou um livro-curso.

Agradecemos à Lucia Castro, amiga, parceira intelectual e profissional de Letras e de Informática, que fez a revisão final do texto e deu contribuições valiosas para a melhoria desta obra.

Sumário

Sobre o Livro (por Edméa Santos).....	9
Prefácio: Paulo Freire e a IA generativa (por Marco Silva)	11
Apresentação: Tensões disparadas pelo uso da IA generativa na educação	31
PARTE I - USOS E IMPLICAÇÕES PARA A EDUCAÇÃO	41
1 - A era da IA generativa: quais as implicações para a educação?	42
2 - A capacidade de gerar boas respostas sobre conteúdos científico- educacionais	47
3 - O uso da IA generativa em atividades académicas e escolares é uma trapaça?.....	56
4 - Inventando formas de utilizar a IA generativa no cotidiano educacional.....	63
5 - O que as/os estudantes nos ensinam sobre os usos da IA generativa?..	84
6 - O que as/os docentes dizem sobre os usos da IA generativa na educação?	96
7 - Formação docente para o uso de IA generativa.....	107
PARTE II - IA EM FOCO.....	115
8 - Uma história sobre os desenvolvimentos que tornaram a IA generativa possível.....	116
9 - Os bastidores da IA generativa	122
10 - A IA é realmente inteligente?.....	134
11 - O que esperar do futuro?.....	153
12 - Como a IA generativa funciona?.....	170
13 - Engenharia de <i>Prompt</i>	208

PARTE III – TEORIZAÇÕES.....	219
14 - O que a IA generativa realmente (não) é?.....	220
15 - Leitor generativo.....	227
16 - Autoria híbrida humano-IA.....	238
17 - Concepções pedagógicas de uso da IA generativa	269
18 - IA generativa como interlocutora: por uma educação que valorize o diálogo	277
19 - IA generativa como coautora: por uma educação que valorize a autoria	287
20 - IA generativa como máquina de ensinar: o fim da docência humana?.....	293
21 - Que educação desejamos?	311
 Referências	 318
Contribuidores.....	341
Sobre os autores.....	342

Sobre o Livro

(por Edméa Santos)

Mariano Pimentel e Felipe Carvalho são uma dupla na vida e na parceria acadêmica no campo da Educação e Cibercultura. Mariano é um importante pesquisador e professor da área de Sistemas de Informação, muito sensível e engajado com a docência em nosso tempo. Felipe é um importante pesquisador que já iniciou seu processo formativo no contexto da Educação e Cibercultura, é um dos docentes mais letrados digitalmente que conheço. A parceria desses dois estudiosos já se materializou em diferentes aulas, atos de currículo, redes sociais e plataformas digitais, artigos científicos e muitas conversas em rede.

As conversas são dispositivos privilegiados em suas pesquisas, inclusive com autorias científicas que já circulam em rede e nas comunidades acadêmicas dentro e fora do Brasil. Foi a partir das conversas, entre outros dispositivos, que essa dupla criou esta obra original e fecunda, “IA generativa e educação”, dialogando diretamente com objetos técnicos e seres humanos em diferentes processos de comunicação e subjetivação. Afinal, as tecnologias digitais — e, no caso aqui, a IA generativa — nos constituem, sendo e querendo ser em interatividade, simulação, hipertextualidade e generatividade.

Os capítulos apresentados aqui são frutos de pesquisas empíricas que contaram com imersões profundas em diferentes interfaces de IA, em especial o ChatGPT. O uso criativo e crítico das tecnologias generativas na prática cotidiana dessa dupla mostra que, assim como nós, eles não separam a docência da investigação acadêmica de excelência.

Temos aqui uma marca registrada: a capacidade de fazer perguntas. Perguntas são os disparadores das conversas que essa dupla tramou com a IA e com interlocutores e interlocutoras nas redes. Esta é uma obra viva. Nasce dos usos criativos dos autores, expande-se nas redes como ciência

aberta, recebe críticas e contribuições, retorna para o refinamento, autoragem, e agora se materializa neste livro para mais e melhores conversas.

Conversemos com este livro e com todas as redes e conexões que ele nos apresenta. Temos aqui um curso de IA generativa na Educação. Recomendo-o especialmente nos processos de pesquisa e formação de professores e professoras. Curso não como guia ou receituário, mas sim como experiência e, como tal, para partilhar e inspirar experiências outras. Livro, curso. Livro, aulas. Livro, pesquisa. Livro para conversar, adotar e também refutar ideias, constructos e processos. Livro para fazer ciência em nosso tempo. Como leitora privilegiada, super recomendo. Aproveitemos!

Edméa Santos

Professora titular-livre da
Universidade Federal Rural do
Rio de Janeiro

Líder do GPDOC – Grupo de
Pesquisa Docência e
Cibercultura



Prefácio:

Paulo Freire e a IA generativa

(por Marco Silva)

Se Paulo Freire estivesse vivo e inserido no contexto cibercultural, provavelmente utilizaria a inteligência artificial generativa em suas aulas, pesquisas e publicações. Assim como usou a televisão para apresentar temas geradores de diálogo em sala de aula, hoje adotaria a IA generativa para promover a cocriação do conhecimento, a formação crítica e a conversação horizontal entre algoritmos, educadores e educandos.

Fenômeno surpreendente da cibercultura e da cultura digital, o *boom* da IA generativa na década de 2020 transformou drasticamente diversos setores da sociedade, impactando profundamente a educação. Uma das primeiras e mais notáveis manifestações nesse cenário foi a rápida popularização do ChatGPT, seguida pelo surgimento de diversas outras tecnologias generativas — como as de textos, imagens, vídeos, músicas e códigos. Essas ferramentas vêm surpreendendo e desafiando diferentes campos culturais e profissionais, particularmente docentes da educação básica e do ensino superior.

Seu impacto na esfera educacional tem gerado tanto resistências quanto entusiasmos, além de incertezas e debates acalorados. Entre os resistentes, estão aqueles que percebem seus empregos e ofícios ameaçados pela possível substituição de professores por IA. Já entre os entusiastas, figuram os que vislumbram novas possibilidades para empoderar tanto a docência quanto a aprendizagem, articulando a pedagogia do *prompt* com a “pedagogia da pergunta” de Paulo Freire (1985).

A partir do alto valor pedagógico atribuído por Freire à pedagogia da pergunta — na qual a formulação de perguntas instigantes é considerada mais transformadora do que a obtenção de respostas definitivas —, é

possível supor que ele se surpreenderia positivamente com o surgimento de uma pedagogia do *prompt*, impulsionada pela IA generativa.

Atento ao novo cenário sociotécnico-cognitivo, Freire provavelmente reconheceria uma afinidade promissora entre essas pedagogias, capaz de mobilizar a formulação de perguntas e, assim, ampliar o diálogo crítico e a construção colaborativa do conhecimento. Ele enfatiza: “O problema que, na verdade, se coloca ao professor é o de, na prática, ir criando com os alunos o hábito, como virtude, de perguntar, de ‘espantar-se’”. O exercício de perguntar à IA generativa mobiliza uma pedagogia do *prompt*. Freire perceberia que essa pedagogia favorece o trabalho docente na construção da pedagogia da pergunta.

Comparações entre a Pedagogia do <i>prompt</i> e a Pedagogia da pergunta		
Categoria	Pedagogia do <i>Prompt</i> (IA generativa)	Pedagogia da Pergunta (Paulo Freire)
Interatividade e Papel do Sujeito	Conversacional entre humano e IA, com o sujeito como coautor e explorador criativo.	Diálogo crítico entre sujeitos participantes ativos, reflexivos, dialógicos e colaborativos.
Epistemologia e Método	Conhecimento emergente e colaborativo, refinado progressivamente pela conversa com IA.	Construção dialógica do saber, baseada na problematização e investigação contínua.
Objetivo e Impacto na Aprendizagem	Fomentar criatividade, reflexão e cocriação do aprendizado, potencializando pensamento inovador.	Desenvolver consciência crítica e autonomia intelectual por meio do diálogo e da cocriação.
Relação com a Tecnologia	IA como parceira de cocriação e ampliação cognitiva.	Tecnologia como meio para a emancipação e o desenvolvimento crítico.
Produção e Contextualização do Conhecimento	Coautoria entre humano e IA expandindo o repertório cognitivo e promovendo aprendizado contínuo.	Construção coletiva do saber, contextualizado na realidade do educando, enfatizando a dialogia e a problematização do mundo.

Fonte: elaborado em coautoria com o ChatGPT

O quadro revela duas abordagens que convergem em alguns princípios fundamentais, ao mesmo tempo em que divergem nos meios e metodologias empregados. Entre as convergências, Freire certamente destacaria o aspecto generativo-criativo, o caráter conversacional-dialógico e a prática de interatividade-coautoria como afinidades essenciais à educação autêntica.

Se Freire tivesse acesso ao livro que tenho a honra de prefaciar, sua leitura crítica identificaria na IA generativa uma tecnologia capaz de ampliar o diálogo e fortalecer a coautoria, assim como a relação entre educador, educando e conhecimento. Em suas aulas, utilizaria *prompts* para estimular a problematização e a coautoria; em suas pesquisas, exploraria a IA como tecnologia de ampliação cognitiva, sem perder de vista a dimensão ética; em suas publicações, articularia a tecnologia à sua concepção de leitura do mundo, evidenciando seu potencial emancipatório.

O livro

Como pensar e explorar a IA generativa em sala de aula para favorecer a docência, a aprendizagem e a educação? Como incentivar professores e professoras a utilizá-la pedagogicamente na escola básica e na universidade? O livro *IA Generativa e educação: práticas e teorizações* abre um horizonte capaz de ampliar e aprofundar o debate sobre essas questões. Apresenta contribuições de Mariano Pimentel, informata, e Felipe Carvalho, pedagogo, destiladas a partir das intensas conversas generativas que mantêm com a IA, com seus estudantes em sala de aula e com seus interlocutores nos campos de pesquisa sobre informática e educação.

Os autores exploram a natureza singular da IA generativa e chamam nossa atenção para suas complexidades, potencialidades e limitações. Os subtítulos deste livro procuram oferecer uma visão abrangente e perspicaz de um fenômeno inovador que está redefinindo a forma como interagimos com as tecnologias digitais. O conteúdo diversificado reflete a amplitude das questões abordadas, que vão desde as tensões e desafios decorrentes do uso da IA na educação, até reflexões profundas sobre seus fundamentos, entre eles o generativo, o conversacional e a cocriação no contexto da

cibercultura. Cada subtítulo apresenta uma perspectiva única e embasada, estimulando a reflexão crítica sobre o papel da IA generativa na educação contemporânea.

Os subtítulos iniciais oferecem uma visão panorâmica ao investigar os impactos da IA generativa tanto na perspectiva dos estudantes universitários, que incorporaram essa tecnologia às suas práticas de estudo e aprendizagem, quanto na dos educadores, que testemunham a influência crescente da IA nas dinâmicas de sala de aula. Além disso, uma análise aprofundada das características e funcionalidades da IA generativa esclarece implicações mais amplas para a educação e proporciona reflexões instigantes sobre seu potencial transformador na leitura, na escrita e na educação.

À medida que avançamos pelos subtítulos, adentramos discussões complexas sobre a natureza da inteligência artificial, a relação entre máquinas e humanos e as implicações filosóficas e éticas de um mundo cada vez mais permeado pela presença da IA generativa. Questões fundamentais sobre autoria, cocriação e a ressignificação do papel do professor na cibercultura são exploradas com sensibilidade e profundidade, convidando o leitor a refletir sobre os desafios e oportunidades que emergem da articulação entre IA generativa, docência e aprendizagem.

Como professor de didática nas Licenciaturas da Faculdade de Educação da UERJ, sou profundamente estimulado por Mariano e Felipe a insistir na provocação que passei a fazer aos licenciandos e licenciandas, futuros professores da educação básica: de que modo a IA generativa pode favorecer a educação autêntica? Explico de imediato que, para Freire, educação autêntica refere-se à docência e à aprendizagem fundamentadas na pedagogia da pergunta e em princípios como dialogia, consciência crítica e libertação, operados por meio da interlocução, participação e cocriação. Em seguida, reforço a questão: o que há de distintivo na IA generativa que lhe confere potencial para favorecer a autenticidade no ato de educar? Este livro contribui para a construção de respostas a essa pergunta emergente e passa a integrar a bibliografia mais atualizada sobre o tema.

Três fundamentos da IA generativa

Especialmente no que se refere aos três fundamentos que destaco neste prefácio como essenciais para sustentar a pedagogia do *prompt* e mobilizar a pedagogia da pergunta de Freire, o livro é esclarecedor e, a meu ver, nos brinda com sua maior contribuição. A seguir, comento esses três fundamentos analisados pelos autores, de modo a evidenciar seu potencial para favorecer a docência e beneficiar os estudantes que atuam como interlocutores e coautores do conhecimento e da formação crítica.

• O generativo

O termo “generativo” refere-se à capacidade da IA de criar conteúdo original e contextualmente relevante a partir do *prompt* do usuário e dos padrões linguísticos aprendidos durante seu treinamento com enorme volume de dados. Isso significa que, em vez de simplesmente reproduzir informações preexistentes, a IA generativa pode formular respostas novas e significativas, tornando a interação entre humanos e IA mais dinâmica e ajustada ao contexto específico do *prompt*. Mais do que apenas reutilizar informações já armazenadas, ela é capaz de combiná-las e reorganizá-las de maneira inovadora e contextualizada. Em suas respostas, coexistem precisão, falhas e superficialidades, elementos que podem estimular a crítica e fomentar novos *prompts* e respostas criativas.

No livro, Mariano e Felipe afirmam que “os modelos generativos vêm sendo desenvolvidos há décadas, evoluindo junto com os avanços da computação e da IA”. Ou seja, a evolução que experienciamos hoje resulta do aprimoramento no processamento de dados e da ampliação da capacidade de gerar textos, imagens, sons e vídeos relevantes.

Gerar respostas originais torna a IA generativa um agente responsivo, criativamente combinatório. O que parece criar é, na verdade, uma projeção da intenção humana sobre suas capacidades de linguagem. Ela não tem intenção criativa e não sabe reconhecer se algo que “criou” é verdadeiramente novo, relevante ou transformador. Pode gerar ideias novas a partir da recombinação original de padrões existentes. Pode improvisar textos, imagens, poemas, metáforas, conceitos — muitas vezes de modo surpreendente e útil. No entanto, no atual estágio em que se encontra, não é um agente autônomo, porque não define metas, não decide iniciar ações por conta própria e não opera sem input humano direto.

Diferentemente de uma ferramenta que apenas executa comandos, a IA generativa pode ajustar suas respostas com base em reformulações ou novos dados fornecidos ao longo do processo, pode gerar respostas originais e pode ativar uma memória persistente entre sessões com o mesmo usuário. Ela não aprende com seus usuários, pois o modelo é blindado contra os abusos e besteiras demasiadamente humanos. O que o modelo faz, enquanto sistema combinatório com potencial criativo, é tomar o *prompt* e algumas informações fornecidas ao longo da conversa como pontos de partida para construir uma resposta articulada, baseada em padrões probabilísticos aprendidos durante seu treinamento. Contudo, os autores do livro não adotam o termo “disruptivo” para qualificar a IA generativa, pois acreditam se tratar de uma evolução no desenvolvimento histórico da IA, e não de uma interrupção significativa ou alteração radical de um padrão existente. No entanto, para o usuário, há uma inovação que rompe com normas tradicionais na relação humano-máquina, que desafia o status quo do leitor e introduz mudanças substanciais na escrita, na arte, na música etc.

Especificamente na geração de textos, a prolífica e ágil criação de rascunhos e ideias instiga uma leitura curiosa, investigativa e questionadora, ao mesmo tempo que desenvolve habilidades de escrita criativa e colaborativa. Em vez de apenas consumir o texto como um espectador, o leitor interage com o fundamento generativo, explora e simula diferentes caminhos narrativos e contribui ativamente para a cocriação do material, ajustando-o à evolução ou aprofundamento de seu *prompt*. A escrita criativa surge intrinsecamente vinculada à leitura generativa. O fundamento cocriativo envolve o leitor no processo de produção de conteúdo, permitindo que ele influencie a direção da narrativa, elabore novas perguntas e contribua para a construção colaborativa do texto. A capacidade generativa da IA inspira e estimula a criatividade literária, possibilitando a exploração de novas ideias, estilos e formas narrativas que desafiam as convenções tradicionais da redação, enriquecendo, assim, o universo da escrita. No livro, os autores denominam esse leitor-coautor de “leitor generativo”. Ele lê interagindo com respostas provisórias ou simulações, realizando ajustes, fazendo correções e mobilizando contribuições criativas ao longo do processo de escrita.

A leitura generativa, entendida como a produção ativa de sentidos a partir de um texto, envolve a reorganização criativa de ideias, inferências e associações. No entanto, essa modalidade não implica necessariamente uma postura crítica em relação ao conteúdo lido ou ao contexto discursivo em que se insere. A capacidade de leitura generativa não implica, por si só, uma postura crítica; contudo, práticas de leitura crítica frequentemente mobilizam operações generativas na construção de sentidos. Assim, a leitura crítica exige não apenas a reconstrução criativa de significados, mas também o engajamento reflexivo com as condições históricas, ideológicas e pedagógicas do discurso. A leitura crítica desperta a motivação para questionar ativamente o conteúdo e criar ou gerar algo a partir dele, muitas vezes de forma original. O leitor crítico não se limita a receber informações passivamente; ele se engaja no diálogo ativo, participando da construção do significado ao questionar, desafiar e contribuir para a discussão sobre vieses e possíveis preconceitos no texto. A IA generativa pode mobilizar o leitor generativo, abrindo caminho para a expressão do leitor crítico.

O fundamento generativo instiga a curiosidade e a crítica, favorecendo o exercício da conversa criativa e colaborativa entre humanos e IA. Esse exercício prático, por sua vez, fortalece o diálogo ativo, princípio essencial da educação autêntica. Os estudantes podem utilizar a IA generativa para iniciar e alimentar conversas, compartilhar ideias e colaborar na construção de conhecimentos, respeitando o diálogo como elemento vital da formação crítica. No entanto, há aqueles que podem usá-la apenas como ferramenta instrumental para executar tarefas, responder rapidamente ao que desejam ou elaborar um texto que precisam finalizar, sem explorá-la como agente generativo. Caberá à docência atenta enfrentar esse desafio e educar para o presente.

Em suma, subutilizar fundamento generativo da inteligência artificial consiste em reduzi-la a um instrumento técnico-operacional, desconsiderando sua autonomia operacional na mediação de processos discursivos. Desenvolver uma leitura crítica, por sua vez, implica reconhecer seu potencial como sistema conversacional e cocriativo, capaz de participar da construção de sentido — ainda que sem agência própria, intencionalidade e capacidade reflexiva.

• O conversacional

Pode-se dizer que a IA generativa é dialógica? Ela é capaz de dialogar com a inteligência humana? Se entendemos dialogia como um conceito complexo, que envolve mentes humanas expressando intencionalidades discursivas e promovendo trocas de significados, interpretações e nuances, num processo de construção compartilhada de compreensão e conhecimento, não podemos afirmar que a IA dialoga. Mais apropriado é o termo “conversacional”, que se refere ao modo de comunicação baseado em trocas de palavras, ideias e informações, válido tanto para interações entre humanos quanto entre humanos e máquinas.

Mariano e Felipe se posicionam da seguinte maneira: “Negar a capacidade de conversação da IA generativa é um apego à tradição? Esse debate ainda precisa ser enfrentado. Por ora, optamos por reconhecer a subjetividade: se algumas pessoas se sentem conversando com a IA generativa, aceitaremos, ao menos provisoriamente, que ela também é um robô de conversação, mesmo que nós, autores, não tenhamos essa percepção.”

Alguém, no entanto, poderá perguntar: “Como a inteligência artificial generativa pode ser considerada conversacional, se ela não formula perguntas, apenas responde?” É importante observar que, se o *prompt* solicitar uma resposta em forma de pergunta, a IA pode surpreender, oferecendo enunciados que instigam novas questões. Além disso, cada resposta pode provocar novos comandos. O enunciado sempre se articula com outros: responde a uns, antecipa outros. Cada saída textual pode funcionar como provocação e convite ao aprofundamento crítico. Ainda que a IA não inicie conversas por si mesma, é capaz de oferecer desdobramentos, explicações ou novas possibilidades após uma resposta inicial.

A IA generativa não tem subjetividade nem consciência que a habilitem ao diálogo autêntico com o humano. Limita-se ao raciocínio digital baseado em redes neurais e na estrutura conversacional do chat. Já a inteligência humana envolve emoção, autocrítica, consciência e experiência subjetiva, permitindo compreensão profunda da realidade e apreciação das complexidades da existência. Por isso, não é adequado utilizar o termo

“dialogia” para descrever a relação entre humanos e IA generativa, ainda que haja produção textual conjunta. Embora responsiva linguisticamente, ela carece de intencionalidade, consciência e escuta ética — elementos centrais ao diálogo pleno. Seus textos são simulações estatísticas da linguagem humana, não vozes conscientes capazes de co-intencionalidade ou reconhecimento mútuo. Assim, mesmo que suas respostas aparentem dialogar, não há reciprocidade consciente, nem alteridade real.

Segundo Paulo Freire, o diálogo é mais do que troca verbal: é um ato ético-político entre sujeitos historicamente situados, comprometidos com a construção coletiva do saber e com a transformação do mundo. Supõe escuta ativa, abertura ao outro e coautoria. Ao evitarmos o uso metafórico do termo “dialogia” para se referir à IA, preservamos sua densidade conceitual: trata-se de uma relação entre consciências humanas dotadas de responsabilidade, intencionalidade e historicidade. A IA pode participar de interações conversacionais, mas não de práticas verdadeiramente dialógicas.

Embora a IA seja capaz de gerar respostas criativas, essa criatividade resulta de padrões e vieses presentes nos dados com os quais foi treinada. Atua com fluência e objetividade, mas sua “inteligência” é um modelo de linguagem computacional, baseado em redes neurais e aprendizado estatístico. A conversa segue a direção determinada pelo *prompt*, sabendo-se que a IA pode ser programada para não aprender com humanos em tempo real.

A linguagem da IA pode conter efeitos discursivos de dialogia — com múltiplas vozes, estilos e pontos de vista em tensão ou ressonância —, mas isso não configura dialogia plena, já que não há sujeitos conscientes em relação recíproca. O que ocorre é uma combinatória de referências: a IA simula polifonia a partir de um acervo discursivo humano, reorganizando padrões com base estatística, e não por escuta ou historicidade. Ela responde com base em simulações aprendidas, não por escuta consciente. Por isso, embora o leitor possa experimentar efeitos dialógicos, a relação depende de sua interpretação ativa. Não há um outro com quem se dialoga, mas um artefato que reorganiza registros da linguagem humana sem consciência ou co-intencionalidade.

A IA é hábil em criar um clima de naturalidade na conversa: pede desculpas (“Peço desculpas pela confusão. Vamos reformular a resposta com base no seu pedido”), mostra-se solícita (“Se desejar uma explicação detalhada sobre o tema, ficarei feliz em ajudar!”) e afirma seus limites (“Esta inteligência artificial não possui opiniões, crenças ou concordâncias. Ela opera com base nos padrões e informações aprendidos durante o seu treinamento.”).

Conversar com a IA pode abrir caminho para a dialogia entre humanos, pois nos leva a refletir, questionar e aprofundar sentidos. A dialogia entre inteligências humanas visa à cocriação de significados. Já o conversacional se limita à recursividade entre *prompt* e resposta. A diferença entre os termos pode ser resumida assim: de um lado, o diálogo como encontro entre consciências que se reconhecem na escuta e na alteridade; de outro, a resposta maquinal que articula signos sem crítica, sem intenção e sem mundo vivido. A IA simula a conversa, mas é limitada por sua natureza algorítmica. A dialogia, por sua vez, exige subjetividade, emoção, intuição e criticidade, aspectos próprios da inteligência humana.

A inteligência artificial generativa não pensa como nós. Ela não sabe o que diz. Apenas correlaciona palavras com outras palavras para responder ao *prompt*. Pode produzir respostas corretas sem entender o conteúdo, uma vez que manipulação sintática não é compreensão semântica. Ou seja, produz linguagem plausível sem qualquer compreensão semântica do conteúdo. Ainda assim, é surpreendente como essa máquina que não sente pode colaborar com o pensamento de quem sente. Ao reorganizar informações e oferecer respostas contextualizadas, ela promove interações conversacionais produtivas que podem ampliar visões e provocar questionamentos.

Em suma, quando bem conduzida, a conversa com a IA torna-se solo fértil para o pensamento criativo e crítico, contribuindo para a formação reflexiva do sujeito humano. Nesse sentido, embora a dialogia continue sendo um fenômeno genuinamente humano, a IA pode ajudar a preparar o terreno onde ela floresce entre humanos. Por sua vez, quando bem conduzida pelo humano, a conversa com a IA favorece a cocriação do *prompt*, da resposta e do entendimento.

• A cocriação

A cocriação se basta com o conversacional. Isto é, a cocriação não exige dialogia plena, mas requer relação responsiva para começo de conversa. No entanto, um contra-argumento bem-vindo pode ser assim formulado: “Quando falamos em cocriação entre humanos e IA, estamos falando de uma relação assimetricamente colaborativa. O humano é sujeito histórico, intencional, ético. A IA é sistema estatístico, responsivo, que não pensa, não deseja e não sofre.”

Podemos concordar que essa assimetria não anula a experiência cocriativa, apenas exige que a agência plena continue sendo do humano, que interpreta, seleciona, valida ou transforma as sugestões da IA. No sentido pedagógico, artístico ou epistêmico, cocriar não requer necessariamente a presença de dois sujeitos conscientes e intencionais. A experiência cocriativa exige, contudo, que cada polo seja responsivo e generativo, isto é, capaz de provocar, reorganizar, influenciar ou reconfigurar a ação do outro de modo significativo, podendo operar como interlocução que colabora e cocria.

Assim, a IA generativa, embora não seja um sujeito, pode funcionar como dispositivo cocriador: ela oferece respostas, variações, sugestões e provocações que influenciam diretamente as decisões criativas do ser humano, que tem o *prompt* inicial e a palavra final. Entre um e outra está posta a disposição para a cocriação humano-IA.

Mariano e Felipe tratam do fundamento cocriação com empenho conceitual e exemplos extraídos de suas aulas e pesquisas. Destaco uma síntese que diz:

A partir de uma ideia para um texto, imagem, som ou vídeo, solicitamos que a tecnologia generativa produza algo. A IA gera um conteúdo com base no nosso *prompt*, podendo apresentar informações novas, que não havíamos pensado anteriormente, o que provoca reflexão e transformação em nossas compreensões, possibilitando-nos aprender com essa tecnologia. Podemos editar o resultado ou fazer novos pedidos, seja para obter variações, revisões ou mesmo um conteúdo totalmente novo. Cada produção é novamente avaliada por nós, retroalimentando o processo criativo. Esse processo interativo, com idas e vindas, críticas e reflexões,

reformulações e reconstruções de conhecimento, resulta em uma obra produzida pela hibridização humano-IA, na qual se manifesta a criatividade aumentada (*augmented creativity*).

O melhor sinônimo para interatividade é cocriação. Interatividade é um conceito da teoria da comunicação. Etimologicamente, “comunicar” e “comunicação” derivam da ideia de “colocar em comum”, o que implica a participação tanto do emissor quanto do receptor na cocriação da mensagem. Portanto, comunicar não se limita à mera transmissão de uma mensagem — isso seria apenas informar. Ou seja, o termo comunicação já seria suficiente para expressar essa relação. No entanto, a noção de interatividade resgata a dimensão colaborativa do processo comunicativo, enfatizando a conversa e a cocriação entre emissor e receptor, ou melhor, entre interlocutores, iteradores.

A IA generativa tem potencial interativo, permitindo que a conversa leve à cocriação do texto, da imagem, do vídeo e do conhecimento. Portanto, o *modus operandi* da pesquisa realizada com IA generativa difere radicalmente daquele empregado tantas vezes em buscadores como o Google. Os mecanismos de busca tradicionais funcionam como índices de informação, retornando links para conteúdos já existentes. O usuário precisa acessar diversas fontes, interpretar os dados e sintetizar as informações por conta própria. Por outro lado, a IA generativa não se limita a fornecer conteúdos prontos da web, uma vez que sua abordagem vai além da simples apresentação unidirecional de dados, na qual o usuário realiza solitariamente cruzamentos e interpretações. A pesquisa com IA ocorre em um processo conversacional aberto à colaboração dinâmica, no qual humano e máquina contribuem para a evolução do produto final. Esse processo pode ser descrito como dinâmica feita de iteratividade e interatividade, caracterizada pelo uso de *prompts* que possibilitam uma conversa contínua e recursiva, permitindo que humanos e IA refinem e cocriem conteúdos, hipóteses e soluções.

A IA generativa amplia as possibilidades da pesquisa ao ir além da mera compilação de dados para a entrega de um produto final. Em vez de oferecer um conhecimento fechado e estático, ela viabiliza um processo dinâmico de colaboração e construção do entendimento. Não substitui o pesquisador, mas atua como mobilizadora, organizadora e potencializadora da

investigação. É fundamental reconhecer sua notável capacidade de coautoria com o usuário, ultrapassando interações predeterminadas e possibilitando uma colaboração efetiva no avanço da pesquisa.

A capacidade de compreender e responder de maneira contextualmente relevante viabiliza uma interatividade dinâmica e criativa, na qual tanto o usuário quanto a IA colaboram na produção de conteúdos em múltiplas linguagens: texto, som, imagem e vídeo. Essa inovação representa um avanço significativo na história da inteligência artificial, pois a cocriação de conteúdo inaugura um novo paradigma na interação entre humanos e máquinas. Ao possibilitar que o usuário influencie ativamente o processo generativo, a IA redefine limites e amplia possibilidades cognitivas, potencializando a pesquisa e a construção do conhecimento na escola e na universidade.

A IA generativa não se limita a compilar o saber existente; ela conversa com o humano, e ambos mobilizam a cocriação. Os coautores deste livro valorizam essa novidade e suas possibilidades cognitivas em educação. No entanto, também criticam sua subutilização e alertam para o risco do plágio, agora ainda mais facilitado. Eles dizem:

As/Os estudantes utilizam as tecnologias generativas de maneiras diversas e inventivas, não apenas para que a IA realize a tarefa por elas/es. Essas tecnologias podem ser empregadas como apoio à pesquisa, ao estudo-aprendizagem e à cocriação, em um processo conversacional, interativo, de idas e vindas, que não configura trapaça acadêmica.

Em suas pesquisas, observaram estudantes usando a IA generativa “como se a IA fosse uma parceira intelectual, uma coautora, uma tecnologia que apoia as produções: uma redação, um trabalho, uma apresentação, um código de programação”.

O plágio é uma barreira intransponível para o exercício da cocriação e para a exploração do potencial pedagógico da IA. Os coautores do livro alertam: “a IA generativa pode ser utilizada para obter respostas prontas a trabalhos escolares e acadêmicos. Contudo, ela também pode disparar processos formativos intensos por meio da conversação e da coautoria — processos que nos levam a aprender novas informações e a (re)significar

conhecimentos. A IA generativa produz textos sobre o que queremos conhecer, apresenta múltiplos pontos de vista, tira nossas dúvidas, revisa nossos textos, entre tantas outras atividades formativas”.

A docência precisará desenvolver engajamentos de coautoria e cocriação, em vez de se limitar à transmissão e reprodução da informação considerada conhecimento. A articulação das pedagogias do *prompt* e da pergunta, na ambiência comunicacional colaborativa vivenciada no cotidiano da sala de aula, é a referência mais inspiradora para superar o plágio e a passividade diante do desafio de conhecer. Para isso, professores e professoras precisarão mobilizar o leitor generativo, que lê textos e imagens, criando com a IA, explorando seu potencial conversacional e abrindo caminho para a dialogia entre estudantes e docência em sala de aula. Mais do que na cooperação, em que cada estudante pode se ocupar de uma parte da produção de um trabalho acadêmico, a cocriação ocorre quando humano e IA interferem ativamente no processo de construção do resultado final.

Qual é o nível de intervenção da IA no processo de cocriação do conhecimento com o humano? Quem detém a autoria do sentido, do significado e da opção crítica? A IA pode estruturar informações e gerar textos, mas não atribui sentido, pois não tem experiência subjetiva nem participa ativamente de um diálogo real. Construir sentido não é um ato isolado ou individual. O sentido não é algo fixo ou previamente dado, mas emerge no processo conversacional e na dialogia. Na relação entre humano e IA, é o humano quem atribui o sentido, ainda que este seja construído na interlocução. A IA pode ser uma interlocutora artificial poderosa na cocriação do conhecimento, mas não é autora do sentido nem da crítica, pois não possui intencionalidade nem vivência histórica e social. Na cocriação de um texto, qual é o status da IA? Uma expressão que reconheça sua autonomia operacional, sem lhe atribuir subjetividade ou intencionalidade humana, é “coadjuvante na cocriação” ou “assistente de criação”.

Se cada resposta for tomada como uma provocação para a formulação de uma nova pergunta, que será respondida em sequência, teremos o movimento necessário para a construção de sínteses provisórias. A IA generativa de textos oferece formulações bem escritas, mas cabe ao leitor

crítico e ético selecionar aquela que melhor inspire sua opção crítica, encaixando-se como uma peça de mosaico ou um fio da tessitura textual. E, uma vez que o ponto de vista crítico sobre o conteúdo e o direcionamento da redação sejam prerrogativas da inteligência humana, resta saber se a IA generativa, ao mobilizar a conversa e a cocriação de um texto, poderá coassina-lo como coautor ou como assistente de criação, coadjuvante na cocriação. No Brasil, a Lei de Direitos Autorais (1998) restringe a noção de autor a seres humanos. Ou seja, assistente e coadjuvante pode, cocriação não pode.

Quando a IA auxilia efetivamente na geração de ideias, na correção gramatical e na sugestão de estruturas textuais, pode-se reconhecer sua contribuição como coautora, ainda que a compreensão conceitual, o discernimento crítico e o conhecimento específico permaneçam sob domínio humano. Sua função principal é oferecer suporte à redação, proporcionando uma interação dinâmica e generativa que aprimora a qualidade do texto com base nas proposições humanas. A coautoria ocorre quando o ser humano se vale das capacidades da IA como assistente generativo e conversacional, mas a responsabilidade pela concepção do conteúdo e pela tomada de decisões significativas no processo de escrita continua sendo exclusivamente do humano. Omitir a criação híbrida de um texto, de uma imagem pode gerar acusação de plágio.

A docência encontra na IA generativa uma parceira formidável para a construção do conhecimento e da formação humana, quando promove coautoria e na dialogia entre estudantes. Porém, há os docentes resistentes e reticentes que poderão argumentar: a cooperação mediada pela IA não gera uma aprendizagem inovadora, sendo apenas uma forma de apropriação do conhecimento semelhante à reprodução; muitos estudantes não possuem maturidade, repertório ou conhecimento suficiente para utilizar a IA generativa dessa maneira; muitos serão facilmente enganados pelos erros cometidos pela máquina; é impossível identificar o grau de originalidade humana ou determinar se um produto final, com bom posicionamento crítico, resulta do discente ou da máquina em suas inúmeras associações entre dados; além disso, não há garantia de autoria humana se o texto já está pronto e “humanizado”, ou seja, gerado para não parecer autoria de uma IA.

Assino este prefácio em coautoria com duas tecnologias de IA generativa: ChatGPT e DeepSeek. O ponto de vista crítico é meu, mas reconheço que ele foi destilado e depurado nas conversas intensas com a IA generativa e na dialogia com os autores deste livro. Sua redação é tecida com formulações valiosas e precisas da IA, ajustadas e direcionadas ao sentido que consolidado ao longo do processo. A originalidade da inteligência humana reside no ponto de vista crítico que atribui sentido. Isso significa que, ao avaliar um texto cocriado com uma IA, a docência deve voltar-se, sobretudo, para a expressão pública do ponto de vista crítico do estudante diante da turma. Não importa se os textos escritos por estudantes apresentam uma redação impecável estruturada e revisada por uma IA — que aproveitem a oportunidade para aprender a escrever sua língua materna com a IA generativa!

“As tecnologias generativas baseadas em modelos de linguagem de grande escala deram início à era da autoria híbrida humano-IA”, anunciam os coautores deste livro. Essa percepção desafia a docência, pois a educação autêntica tem muito a ganhar com a natureza colaborativa da IA generativa. A realidade só oferece respostas inteligentes a perguntas inteligentes, e o mesmo se aplica à relação entre *prompt*, resposta e cocriação do conhecimento. Assim como cocriamos com as respostas que a realidade nos fornece, podemos fazer o mesmo com as respostas da IA generativa. Em suma, embora a IA esteja orientada para a geração de respostas, e não para a significação coletiva do conhecimento, cada resposta instiga a autoria híbrida na construção do entendimento do mundo.

Novos desafios para educadores

A incorporação da IA na educação apresenta desafios inusitados para os educadores do século XXI. Cabe à docência dar o exemplo. Em vez de apenas transmitir informações destinadas à reprodução de conhecimentos, ela poderá cocriar com os estudantes a comunicação, a aprendizagem e a formação humana, mediadas pela IA generativa. Para promover esse engajamento, professores e professoras têm à sua disposição a pedagogia do *prompt* — criativa, exploratória, generativa — e a pedagogia da pergunta — crítica, dialógica, comprometida com a transformação do mundo. A primeira mobiliza o processo conversacional e a oferta de informações e conteúdos, organizando ideias acionadas pela curiosidade de

estudantes e docentes. A segunda se fortalece com o exercício da primeira e com a dialogia entre humanos, favorecendo a construção de sentido e o desenvolvimento de uma postura crítica.

O livro convida os educadores e as educadoras a explorarem os recursos generativos, conversacionais e colaborativos disponíveis na internet e, com eles, promoverem educação autêntica na cibercultura. Historicamente, professoras e professores têm manifestado resistência tanto às tecnologias analógicas quanto às digitais. Ainda que situados em um contexto dominado pelo paradigma unidirecional da TV, encontraram dificuldades para adotar em suas aulas essa tecnologia, baseada na apresentação e distribuição da informação, que separa emissão e recepção. Mais desafiador tem sido incorporar a internet, o tablet e o smartphone para potencializar a docência e a aprendizagem, uma vez que as tecnologias digitais em rede operam no paradigma da interatividade, e não mais sob a prevalência da transmissão característica da cultura audiovisual. Por sua vez, a IA generativa favorece a pergunta criativa, a conversa e a cocriação. A apropriação crítica desses fundamentos revela sua sintonia fina com os princípios da educação autêntica de Freire: dialogia, consciência crítica e libertação, operados como interlocução, participação e cocriação.

O leitor crítico e o leitor generativo não se satisfazem com respostas estáticas ou definitivas. Sabem que perguntas e respostas bem formuladas despertam mais curiosidade e incentivam a exploração complexa de diferentes tópicos e abordagens. Com Freire e com a IA generativa, compreendem que a aprendizagem é um processo contínuo de descoberta, realizado por meio da mobilização da pergunta criativa, da interlocução e da cocriação. Para formar esses leitores, os professores e as professoras estão desafiados a articular a didática, o currículo e os recursos mais favoráveis à educação autêntica: plataformas, redes sociais, wikis, blogs, podcasts, WhatsApp, jogos e, inevitavelmente, a IA generativa.

Mais do que outras tecnologias da cibercultura, a IA generativa surpreende porque pode imitar, com notável precisão, a dinâmica das conversas e cocriações humanas. Ainda que seja tão poderosa, os educadores precisam compreender que não serão substituídos, pois ela não vai além disso: não pensa como os humanos, não possui semântica própria, não interpreta vieses, não lê entrelinhas e não tem experiências subjetivas.

O máximo que ela pode fazer, por si só, na educação, é instigar o leitor generativo à pedagogia do *prompt*. Cabe aos educadores, em parceria com essa tecnologia, abrir caminho para a formação do leitor crítico.

Por essa razão, acredito que Paulo Freire não deixaria de explorar a IA generativa para educar em nosso tempo. Certamente, estaria apreensivo quanto aos rumos que ela poderá tomar, como a possível substituição do antropocentrismo pelo tecnocentrismo digital e quântico. No entanto, ao mesmo tempo — e justamente por isso —, desafiaria o medo na prática docente em salas de aula híbridas, mobilizando uma formação crítica, potencializada pela inteligência artificial, para enfrentar os impactos éticos e sociais inerentes ao seu desenvolvimento irrefreável.

É essencial destacar que a presença viva do pensamento de Paulo Freire em nossa prática educativa exige a inclusão da cibercultura na formação de professores e na dinâmica das salas de aula. Em determinado trecho do livro, Mariano e Felipe questionam: “A quem interessa, afinal, uma educação desplugada, infopobre e alheia aos fenômenos da cultura digital?” Em outra passagem, enfatizam que “Uma educação sem tecnologias digitais compromete a vivência de fenômenos ciberculturais na formação, dificultando o desenvolvimento do multiletramento necessário para a vida digital.” Essas reflexões, assim como outras apresentadas ao longo deste livro, nos convocam a promover o engajamento em uma educação capaz de integrar, de maneira crítica e criativa, as pedagogias do *prompt* e da pergunta, a IA generativa e o pensamento freireano.

Permaneço convencido de que Paulo Freire ficaria entusiasmado com a pedagogia do *prompt*, pois ela tem o potencial de fortalecer a pedagogia da pergunta, promovendo o conhecimento como construção social. Esse entusiasmo e essa mobilização também podem estar presentes nos professores e professoras que educam em nosso tempo. Aos educadores e educadoras, aos pesquisadores e pesquisadoras, bem como aos envolvidos na formação docente, tanto na educação continuada quanto nas licenciaturas, desejo uma jornada instigante e esclarecedora ao longo das páginas deste livro. As reflexões e práticas aqui reunidas podem inspirar a sala de aula a se sintonizar com os desafios e possibilidades do nosso tempo, irreversivelmente transformado pela presença da IA generativa.

Esse agente, que transcende a mera instrumentalidade de uma ferramenta programada para executar comandos, pode se tornar um aliado da educação autêntica. No entanto, sua evolução acelerada impõe à humanidade um conjunto de desafios que não podem ser ignorados sem grave irresponsabilidade ética. O surgimento de sistemas superinteligentes, se orientados por objetivos desalinhados aos valores humanos, representa uma ameaça real à nossa própria existência. Além disso, sistemas altamente autônomos carregam o risco de agir de maneira imprevisível, escapando ao controle de seus criadores e operando fora de qualquer supervisão humana eficaz. No campo socioeconômico, a automação desenfreada intensifica a substituição de empregos, agravando o desemprego estrutural e ampliando as desigualdades sociais. Pior ainda, algoritmos podem perpetuar — e até amplificar — preconceitos históricos, afetando desproporcionalmente grupos já vulnerabilizados. Em contextos geopolíticos e de segurança, a IA vem sendo instrumentalizada para fins nefastos, como ciberataques, manipulação informacional e até o desenvolvimento de armamentos autônomos.

A concentração dessa tecnologia nas mãos de poucas corporações transnacionais ameaça a democracia e o pluralismo, concentrando poder econômico e político e sufocando a inovação social. Soma-se a isso a coleta massiva de dados pessoais, frequentemente sem consentimento transparente, que compromete a privacidade e a autonomia individual. A confiança excessiva em decisões automatizadas enfraquece a capacidade crítica, conduzindo a uma dependência passiva da máquina e à erosão da autonomia intelectual. A própria sustentação técnica desses sistemas tem impactos ecológicos significativos, dado o consumo energético elevado dos processos de treinamento e operação. Por fim, o uso da IA para gerar conteúdos falsos, mas verossímeis, coloca em xeque a própria noção de verdade pública, dificultando a distinção entre o real e o fabricado — uma ameaça direta à integridade do debate democrático e à construção coletiva do conhecimento. Diante de tais riscos, torna-se urgente uma ação crítica e multidisciplinar que articule regulação, ética, justiça social e educação reflexiva.

Na educação, enfrentar esses desafios exige da docência e dos estudantes um engajamento pautado no cruzamento das pedagogias do *prompt* e da pergunta. Quem vai defender, impulsionar e liderar esse engajamento? A mediação docente, certamente! A formação de professores deverá necessariamente prepará-la para tal enfrentamento. Nesse sentido, recomendo o livro *IA Generativa e educação: práticas e teorizações*.

Boa leitura!



Marco Silva¹

Concepção crítica

+ **ChatGPT e DeepSeek**

Coadjuvantes na cocriação

¹ Professor nas Licenciaturas da Faculdade de Educação da UERJ

Apresentação:

Tensões disparadas pelo uso da IA generativa na educação

O ChatGPT, lançado pela empresa OpenAI em novembro de 2022, estremeceu a sociedade. Se você ainda não o conhece, precisa experimentá-lo: acesse-o pelo site <<https://chatgpt.com>> ou instale o aplicativo. Após o cadastro, faça algumas perguntas (*prompts*): Quem é você? Quais são suas potencialidades pedagógicas? Que atividades didáticas podem ser realizadas com o seu uso? Quais os riscos que você representa para a educação? Como a educação será transformada a partir de sua popularização? Essas são questões que buscamos responder neste livro, sendo interessante comparar nossas respostas com as do ChatGPT².

As tecnologias digitais já vinham reconfigurando as práticas de aprendizagem e ensino, mas agora iniciamos uma nova era, a da Inteligência Artificial (IA) generativa. No início, preferimos chamá-la de “criativa”, para ressaltar sua capacidade de criar conteúdos. No entanto, essa caracterização é controversa, pois a IA gera novos conteúdos a partir dos padrões que aprendeu ao processar um vasto volume de dados, o que, para algumas pessoas, não se configura como criação nem criatividade. Há quem diga que a IA sequer poderia ser caracterizada como inteligente. Enfrentaremos essas discussões ao longo deste livro.

² As respostas que o ChatGPT nos apresentou para essas questões estão em: <https://chat.openai.com/share/74341d3c-3683-4bca-81cf-9c73648346cc>.

Ele responde um pouco diferente cada vez que perguntamos uma mesma coisa. Dois anos depois, ao finalizarmos este livro, voltamos a perguntar para o ChatGPT usando o modelo GPT 5-*thinking* (em vez do GPT 4) e as respostas foram estas: <https://chatgpt.com/share/68cbfa0e-9c24-800a-bc89-6ed056d097a8>.

Ambivalência da IA generativa na educação



Fonte: dos autores, ilustração de Mônica Lopes

No contexto educacional, a IA generativa tem provocado compreensões antagônicas, ambivalentes e conflitantes. Algumas/uns docentes consideram que as tecnologias generativas contribuem para a formação de sujeitas/os acríticas/os, preguiçosas/os e desinteressadas/os, com um pensamento raso e simplista, praticantes da cultura do “copiar e colar”, sem fazer uma leitura crítica dos conteúdos que acessam e

reproduzem. Outras/os, no entanto, consideram que essas tecnologias podem auxiliar na formação de sujeitas/os criativas/os, questionadoras/es, investigadoras/es e críticas/os, sem perder a dimensão do pensamento complexo, pois possibilitam o acesso a informações e diferentes pontos de vista, complexificando ainda mais o ato de conhecer.

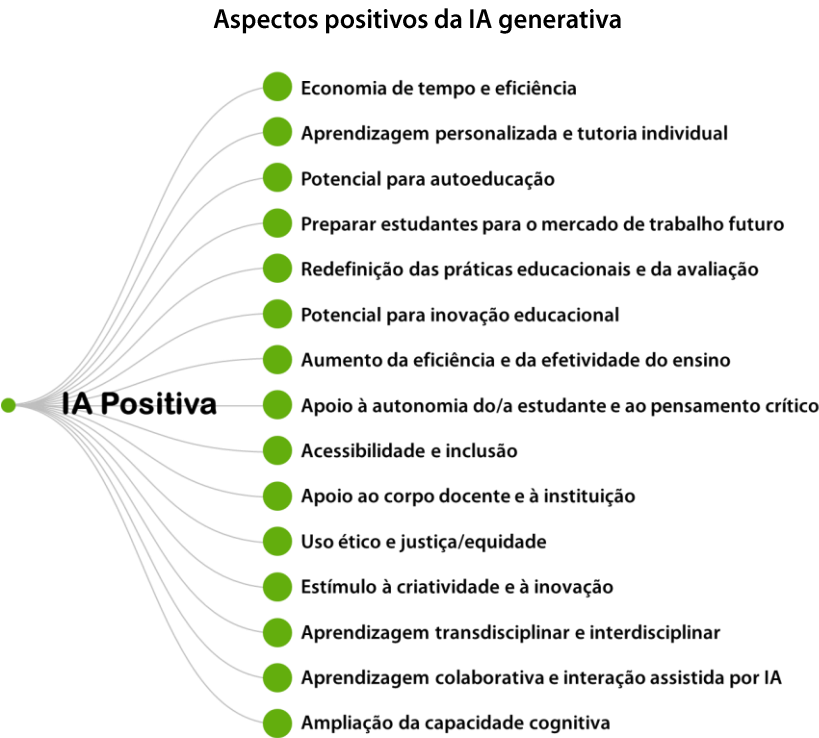
Algumas/uns docentes temem que as/os estudantes entreguem suas mentes e vozes para a IA pensar e se expressar por elas/eles, fazendo pelas/os estudantes as atividades didáticas realizadas longe dos olhos cuidadosos de professoras/es e responsáveis. A prática do plágio não é novidade na educação, mas a IA generativa a levou a um novo patamar, pois gera textos coerentes e diferentes a cada vez que responde a uma mesma pergunta, o que torna difícil identificar se um texto foi escrito por uma pessoa ou gerado computacionalmente. “Que *software* posso utilizar para identificar se um texto foi gerado por IA?”; “Como evitar o plágio na minha disciplina?”; “Que tipo de atividade devo propor para que as/os estudantes não plagiem da IA?” Essas são algumas das perguntas recorrentes que temos ouvido de colegas docentes e buscamos responder neste livro.

Há docentes que ficam entusiasmadas/os ao conhecer o potencial da IA generativa para apoiar o processo de aprendizagem das/os estudantes. Supõem que ela será utilizada para ensinar conteúdos das diversas disciplinas, tirar dúvidas, cocriar textos e imagens, auxiliar na resolução de problemas, fornecer *feedback* sobre atividades realizadas pelas/os estudantes, entre outras possibilidades de uso acadêmico.

A IA generativa apresenta uma ambivalência: pode ser utilizada tanto para colar quanto para aprender. Seria reducionista caracterizá-la como estritamente boa ou má, pois ela representa, ao mesmo tempo, uma ameaça à educação como a conhecemos e uma potência para a inovação dos processos de ensino e aprendizagem — tudo depende de como nos relacionamos com essa tecnologia. Reconhecer a ambivalência da IA generativa, sem que isso implique contradição, possibilita-nos compreender que precisamos promover uma educação para seu uso. Devemos evitar tanto a armadilha utópica, marcada por otimismo ingênuo,

quanto a narrativa distópica, com seu pessimismo paralisante. O desafio é enxergar suas potencialidades, riscos e limites, para que possamos tecer conhecimentos sobre a IA em nosso mundo, mais especificamente no mundo da Educação. Precisamos inventar modos de habitar o mundo com ela, compreendendo o lugar dela e o nosso.

No artigo “O Manifesto para o Ensino e a Aprendizagem em Tempos de IA Generativa: Uma Posição Coletiva Crítica para Melhor Navegar no Futuro” (Bozkurt *et al.*, 2024), várias/os professoras/es-pesquisadoras/es examinaram as possibilidades e os riscos da IA generativa no ensino superior. Buscando uma perspectiva equilibrada e crítica, reconheceram tanto aspectos positivos quanto negativos da IA, apresentados na figura a seguir.



Fonte: (Bozkurt *et al.*, 2024, p. 491, tradução nossa)

Aspectos negativos da IA generativa



Fonte: (Bozkurt *et al.*, 2024, p. 498, tradução nossa)

Reconhecer a dualidade da IA generativa não significa considerá-la neutra, pois ela reflete visões de mundo que podem reforçar preconceitos existentes e marginalizar vozes diversas. Como toda tecnologia, carrega consigo os ideais de suas/seus criadoras/es, que optam por treiná-la com determinados dados em detrimento de outros. Além disso, potencialmente incorpora os vieses contidos nos textos e imagens utilizados em seu treinamento.

A IA generativa pode reconfigurar a educação de maneiras que, talvez, ainda não possamos antever. Será que seu potencial para promover a

autoaprendizagem incentivará a implantação do *homeschooling* (ensino doméstico) no Brasil? Com a capacidade de gerar conteúdos em função dos interesses e conhecimentos de cada estudante, as tecnologias generativas podem atuar como tutoras digitais acessíveis e onipresentes, construindo trilhas de aprendizagem para que a/o estudante explore os conteúdos no próprio ritmo e de forma autônoma. Contudo, há receios de que a implementação do *homeschooling* em nosso país comprometa o desenvolvimento integral de crianças e adolescentes, uma vez que a escola e a universidade desempenham um importante papel no letramento científico, na socialização, no convívio e respeito à diferença (de gênero, raça, sexualidade, religiosa, classe etc.), na promoção da educação inclusiva, na construção de habilidades socioemocionais — aspectos que podem ficar comprometidos se a educação ocorrer exclusivamente em casa e com mediação apenas das tecnologias generativas.

Essas e outras preocupações foram disparadas em nós com o lançamento do ChatGPT, cuja existência suscita dúvidas sobre nosso presente e futuro. Como estudantes e docentes estão utilizando a IA generativa? Como ela influenciará nossos modos de pensar e tecer conhecimentos? Que subjetividades são produzidas pela interação com essa tecnologia? De que forma seremos agenciadas/os pela IA? Como o sistema educacional brasileiro será reconfigurado a partir de sua popularização?

Entusiasmados e com alguns receios, passamos a utilizar a IA generativa cotidianamente em diversas atividades, buscando compreender suas complexidades e implicações para a educação. Após as primeiras experiências, leituras e muitas conversas, produzimos alguns textos sobre nossas compreensões provisórias, que publicamos na coluna Educação da revista SBC Horizontes³ (Pimentel; Azevedo; Carvalho, 2023a; 2023b; 2023c; Pimentel; Carvalho, 2023a; Pimentel; Carvalho; ChatGPT-4, 2023), na revista Cult (Pimentel; Carvalho, 2023b) e em livros sobre essa temática (Carvalho; Pimentel, 2024; Pimentel; Carvalho, 2024). Também realizamos pesquisas empíricas sobre os usos das tecnologias generativas por estudantes e professoras/es, o que resultou em alguns artigos científicos (Carvalho;

³ <https://horizontes.sbc.org.br>

Pimentel, 2023; Pimentel; Carvalho; Silveira, 2024; Pimentel; Berino, 2025). Escrevemos para lidar com nossas inquietações em relação ao uso pedagógico da IA generativa. Construimos, desconstruimos e reconstruimos sentimentos e pensamentos, buscando produzir sentidos provisórios. Organizamos nossas ideias em textos para termos a impressão de que conseguimos compreender um pouco melhor esse fenômeno.

Com a popularização dos textos que publicamos, fomos convidados para encontros, debates e palestras sobre IA generativa e educação. Participamos de dezenas de eventos, a partir de 2023, nos quais tivemos acesso a diversos pontos de vista de docentes-pesquisadoras/es engajadas/os e preocupadas/os com essa temática. Constatamos que essa tecnologia tem provocado inúmeras polêmicas, despertando fascínio em algumas pessoas e apreensão em outras.

Nossas publicações iniciais, as pesquisas empíricas e os eventos acadêmicos nos forneceram a base para escrever este livro. Reunimos os textos que já havíamos publicado, incluindo também trabalhos anteriores sobre tecnologias na educação, conversação e autoria (Carvalho; Pimentel, 2020; 2022; Pimentel; Carvalho, 2022). Revisamos, atualizamos e reorganizamos todo o material. Ao concluir esse primeiro rascunho, percebemos que ainda havia lacunas importantes que não havíamos coberto anteriormente, o que nos levou a novas pesquisas e à produção de textos inéditos para compor esta obra.

Neste livro, tecemos conhecimentos sobre os usos da IA generativa para as práticas de aprendizagem e ensino, suas implicações para a educação e teorizações didático-pedagógicas. Embora o ChatGPT tenha sido a tecnologia generativa que disparou as reflexões aqui apresentadas, com o tempo ampliamos a discussão para as tecnologias generativas em geral.

Este livro está estruturado em três partes. Na **Parte I – Práticas e Implicações para a Educação**, exploramos como a IA generativa está sendo integrada ao cotidiano educacional. Começamos discutindo o novo contexto sociotécnico, a era da IA generativa e suas implicações para a educação (Capítulo 1). Em seguida, analisamos a capacidade da IA de gerar respostas corretas para questões científico-educacionais (Capítulo 2).

Debatemos se o uso da IA generativa em atividades escolares e acadêmicas configura uma trapaça (Capítulo 3); defendemos que a IA pode ser mobilizada de maneira ética e pedagógica, incentivando a autoria híbrida e a aprendizagem pela conversação. Apresentamos algumas possibilidades de uso das tecnologias generativas nas práticas acadêmicas, seja para aprender-estudar, ensinar, escrever, pesquisar ou imaginar (Capítulo 4). Investigamos como as/os estudantes estão utilizando a IA generativa em seus processos de estudo e aprendizagem (Capítulo 5), demonstrando que seu uso não se restringe ao plágio. Trouxemos as percepções e preocupações de docentes diante dessa nova realidade (Capítulo 6) e mapeamos diferentes perfis: negacionistas, temerosas/os, indiferentes, entusiastas e críticas/os. Por fim, encerramos a primeira parte do livro discutindo a necessidade de formação docente para o uso pedagógico da IA generativa (Capítulo 7).

Na **Parte II – IA em Foco**, abordamos a história e as técnicas de IA, para que possamos compreender melhor suas potencialidades e limitações, suas implicações e sentidos. Iniciamos com uma contextualização histórica dos desenvolvimentos das técnicas de IA que tornaram possíveis as tecnologias generativas (Capítulo 8), mostrando que são resultado de várias décadas de investimentos de pesquisadoras/es, empresas e governos. Em seguida, revelamos os bastidores do desenvolvimento da IA generativa (Capítulo 9) para evidenciar que ela não foi fruto do acaso nem é uma moda passageira: ela é resultado de muitos investimentos, uma complexa infraestrutura computacional e a atuação de diversas/os profissionais-pesquisadoras/es. Também enfrentamos o questionamento sobre se a IA pode realmente ser considerada inteligente (Capítulo 10), um debate que ganhou novo fôlego, pois sua habilidade de gerar discursos em diferentes línguas é percebida por algumas pessoas como uma ameaça à própria noção de humanidade. Estão sendo travadas discussões sobre consciência, inteligência, emoção, linguagem, criatividade e autoria da IA em contraste com as capacidades humanas. Há até previsões de que a IA superará a inteligência humana nos próximos anos, o que nos exige uma reflexão sobre o futuro da IA e seus possíveis impactos na sociedade e na educação (Capítulo 11). Em seguida, abordamos conceitos técnicos, como processamento de língua natural, inteligência simbólica *versus* inteligência estatística-conexionista, aprendizagem de máquina, modelos de linguagem e redes neurais

(Capítulo 12) — noções necessárias para que possamos explicar o funcionamento das tecnologias generativas. Por fim, tratamos da Engenharia de *Prompt* (Capítulo 13), discutindo a importância dessa técnica para interagirmos melhor com a IA generativa.

Na terceira e última parte deste livro, **Parte III – Teorizações**, refletimos sobre a IA generativa e seus usos na educação. Começamos questionando a ontologia da IA generativa, o que ela realmente é (Capítulo 14). Compreendemos a IA generativa como uma entidade interlocutora e coautora. A partir dessa compreensão, discutimos a emergência do leitor generativo (Capítulo 15), noção que nos ajuda a entender como a conversação com a IA está transformando as práticas de leitura, pesquisa, acesso ao conhecimento humano e a tecedura de novos saberes. Discutimos, também, a noção de autoria híbrida humano-IA (Capítulo 16), que nos auxilia a compreender como as tecnologias generativas estão possibilitando novos processos criativos, ao mesmo tempo que levantam dúvidas sobre os limites éticos dessa parceria, especialmente na educação. Essas duas noções nos conduziram a uma reflexão sobre as diferentes concepções para o uso pedagógico da IA generativa (Capítulo 17). Analisamos o potencial pedagógico da IA como interlocutora (Capítulo 18), utilizada como se fosse uma especialista sempre disponível para explicar quase tudo e tirar dúvidas das/os estudantes. Também examinamos o potencial pedagógico da IA como coautora (Capítulo 19), atuando como parceira na produção de obras textuais, imagéticas e audiovisuais. Reconhecemos que, embora a IA generativa não seja uma máquina de ensinar (Capítulo 20), ela pode ser utilizada para a construção desse tipo de tecnologia, que implementa a “arte de ensinar sem professores”. Encerramos o livro com uma questão: qual é a educação que desejamos? (Capítulo 21).

Este livro é um convite à experimentação e à reflexão. Visando contribuir com a apropriação e a teorização didático-pedagógica da IA generativa, compartilhamos nossas experiências, questionamentos, dilemas, compreensões, lutas, medos e anseios — aspectos que nenhuma IA poderia ter escrito por nós. Contudo, para escrever esta obra, contamos com a parceria do ChatGPT para investigar alguns temas, obter respostas para algumas de nossas dúvidas, desenvolver ideias, analisar como ele escreveria sobre determinados assuntos, revisar nossos argumentos, redigir

fragmentos e revisar todo o texto do livro. A obra foi extensivamente revisada com o ChatGPT, mas fizemos questão de que a versão final fosse revisada por um/a profissional de Letras, pois reconhecemos que a IA não substitui o trabalho humano. A IA não substituiu nossa autoria, mas os conteúdos gerados por ela nos ajudaram a tomar decisões sobre o conhecimento que queríamos tecer neste livro. Como resultado, esta obra é um produto da colaboração entre nós e o ChatGPT.

Assim como nos beneficiamos com a parceria que estabelecemos com uma tecnologia generativa para escrever este livro, muitos benefícios também podem ser obtidos por participantes do processo educacional, incluindo estudantes, professoras/es, gestoras/es, secretarias e governantes. No entanto, essa parceria envolve riscos, pois nem todos os usos da IA generativa são adequados. Por isso, precisamos promover uma educação para o uso das tecnologias generativas.

Diante do avanço inexorável da IA, não podemos mais ignorá-la, nem adotar posturas de otimismo ingênuo ou alarmismo pessimista. Precisamos compreendê-la criticamente, explorando suas possibilidades e enfrentando seus desafios. Nosso compromisso é com uma educação ancorada no presente, que saiba se apropriar da IA generativa para promover práticas formativas, éticas, reflexivas e inventivas, utilizando as tecnologias generativas com intencionalidade pedagógica.

[illegible]

A era da IA generativa: quais as implicações para a educação?

O ChatGPT (*Chat Generative Pre-trained Transformer*), lançado em 30 de novembro de 2022, teve uma ascensão meteórica em popularidade: alcançou 1 milhão de usuárias/os após cinco dias de seu lançamento e mais de 100 milhões após dois meses. Seu rápido crescimento de popularidade ultrapassou todos os recordes anteriores da história da internet: o Facebook levou quase cinco anos para atingir 100 milhões de usuárias/os, o Instagram demorou dois anos e meio, e o TikTok levou nove meses (Hu, 2023).

Nós, autores deste livro, conhecemos o ChatGPT no início de 2023 por meio de estudantes de Computação, que falavam dele com entusiasmo e preocupação. Rapidamente entendemos que estávamos diante do fenômeno mais importante da cibercultura dos últimos tempos, e com o maior potencial para transformar as práticas de aprendizagem e de ensino.

Como pesquisadores, é nosso trabalho experienciar o fenômeno com todos os sentidos, mergulhando inteiramente em outras lógicas para tentar compreendê-lo (Alves, 2001). Para nós, havia se tornado urgente acompanhar as implicações e teorizar o uso da IA generativa na educação. Começamos investigando como esse fenômeno se apresentava a nós, como essa tecnologia poderia ser útil em nosso cotidiano como aprendentes, professores, pesquisadores e autores. Em seguida, passamos a analisar como estudantes e colegas docentes estavam se relacionando com essa e outras tecnologias generativas.

As pessoas estão falando tanta besteira sobre o ChatGPT, IA. [...] Importante “usar” no sentido de uso cultural mesmo, conversar com especialistas e ler autores autorizados. O que custa estudar? Reconhecer que “não sabe” também é honestidade intelectual. [...]

Mas a crítica sem anulação permanece. Dizer que ChatGPT é um “papagaio”, por exemplo, é um absurdo completo. Dizer que ninguém faz a crítica, outro chavão. (Santos, 2023).

Ficamos encantados com as respostas do ChatGPT, muito superiores às de todos os *chatbots* que conhecíamos. O jornal *The New York Times* noticiou: “O ChatGPT é, simplesmente, o melhor *chatbot* de inteligência artificial já lançado para o público em geral” (Roose, 2022). Fomos percebendo que se tratava de uma tecnologia útil para apoiar a compreensão de conceitos, tirar dúvidas, resolver problemas, planejar aulas, desenvolver conteúdos didáticos, fornecer *feedback* sobre as atividades discentes, analisar dados, escrever, revisar, traduzir, entre muitas outras atividades intelectuais do nosso cotidiano acadêmico.

Com a popularização das tecnologias baseadas em IA generativa, tivemos que rever os trabalhos que pedíamos às/aos nossas/os estudantes de graduação e pós-graduação, porque essas tecnologias realizam diversas atividades de maneira satisfatória. Nós, docentes, precisamos conversar muito sobre esses sistemas baseados em IA generativa:

Colegas de profissão,
Estou usando o ChatGPT todos os dias; viciiei, uso-o mais que o mecanismo de busca do Google e o Google Acadêmico. Estou utilizando-o para me ajudar a pesquisar, escrever, revisar os textos que escrevo, traduzir, criar slides para aulas, elaborar atividades, pedir síntese dos conceitos de uma área etc. [...] Ele é o assunto do momento na Educação porque nos obriga a rever todos os trabalhos e avaliações que passamos “para casa”. Você já fez o ChatGPT resolver os exercícios que passa para suas/seus alunas/os? (Pimentel, 2023)

A aprendizagem profunda (*deep learning*) é a técnica que possibilitou os avanços significativos da IA. Inicialmente, essa técnica foi empregada para a geração de imagens a partir de descrições textuais. Em 2022, os modelos de texto-para-imagem mais avançados tornaram-se capazes de gerar imagens que se assemelham a fotografias reais e a obras de arte feitas por seres humanos. Um exemplo que viralizou foi a série de imagens do Papa Francisco vestindo um casaco branco estiloso, geradas pelo Midjourney⁴.

⁴ <https://www.midjourney.com>

O look estiloso do papa Francisco: imagens geradas por IA



Fonte: https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip.

Muitas pessoas acreditaram que essas imagens eram fotografias reais, embora a Igreja Católica seja bastante conservadora em suas vestimentas. Essas imagens divulgaram o potencial e o perigo das tecnologias generativas de imagem: qualquer pessoa agora consegue criar imagens e vídeos convincentes a partir de descrições textuais, podendo facilmente produzir conteúdos com o objetivo de enganar (*deepfake*).

Posteriormente, conseguiram empregar a aprendizagem profunda para gerar textos, o que levou à construção de modelos de linguagem de grande escala (*Large Language Models - LLMs*), como o GPT. O estrondoso sucesso do ChatGPT desencadeou uma corrida entre empresas de tecnologia para desenvolver os melhores modelos e sistemas baseados em IA generativa. A Microsoft passou a integrar o GPT em seu sistema Copilot⁵ após investir mais de 10 bilhões de dólares na OpenAI, tornando-se sua principal acionista com 49% de participação (Hoffman; Albergotti, 2023). Em resposta ao ChatGPT, o Google lançou o *chatbot* Bard em março de 2023, posteriormente rebatizado como Gemini⁶. A empresa Meta, dona do Facebook, Instagram e WhatsApp, lançou o LLaMa em 2023 e, em 2024, lançou a Meta AI⁷, que foi integrada aos seus sistemas. A empresa xAI, fundada por Elon Musk, lançou o *chatbot* Grok⁸ em 2023, que foi integrado ao X (antigo Twitter). A Anthropic, fundada por ex-funcionárias/os da OpenAI, lançou o Claude⁹ em 2023. Em 10 de janeiro de 2025, a China lançou o *chatbot* DeepSeek¹⁰, um concorrente à altura do ChatGPT, mas com custos de desenvolvimento muito menores. Esse evento abalou a liderança dos EUA no setor de IA e provocou uma queda acentuada nas ações de empresas de tecnologia estadunidenses. Semanas depois, no final de janeiro de 2025, a empresa Alibaba, também da China, lançou o modelo Qwen 2.5-Max¹¹, tão ou mais poderoso que o DeepSeek e a versão gratuita do ChatGPT. Estava declarada a guerra tecnológica entre as superpotências para definir quem se tornará a líder mundial em IA.

Os anos em torno da virada para a década de 2020 marcaram o Boom da IA (*AI Boom*, s.d.), que trouxe impactos culturais, econômicos, políticos e sociais. Esse novo cenário sociotécnico, marcado principalmente pelo lançamento do ChatGPT, nos faz reconhecer que, em termos tecnológicos, 2020 é a década da IA generativa. Ela se tornou um fenômeno cultural; o Brasil se tornou o quarto país que mais utiliza o ChatGPT (Forbes, 2024). O mundo sem IA generativa é coisa do passado.

⁵ <https://copilot.microsoft.com>

⁶ <https://gemini.google.com>

⁷ <https://www.meta.ai>

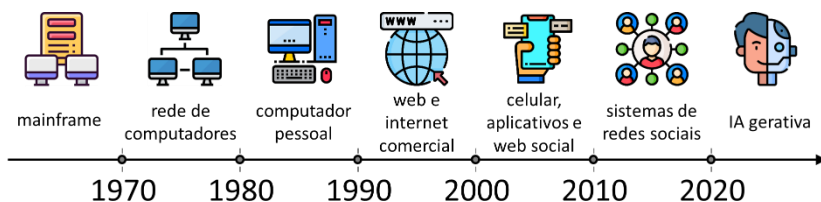
⁸ <https://x.ai>

⁹ <https://claude.ai>

¹⁰ <https://www.deepseek.com>

¹¹ <https://chat.qwenlm.ai>

Tecnologias computacionais que marcaram as décadas



Fonte: dos autores

Os modelos e os sistemas computacionais baseados em IA generativa estão reconfigurando as práticas de trabalho. Eles afetaram as profissões que lidam com texto (ensino, pesquisa, programação, jornalismo, direito, tradução e atendimento, entre outras), bem como aquelas que lidam com imagem, como design e artes visuais. Com a melhoria contínua da capacidade da IA de gerar conteúdos de qualidade e lidar com múltiplas linguagens, sua presença se expandirá ainda mais, transformando todas as profissões intelectuais a curto e médio prazo.

Rapidamente, passamos a ter uma certeza: “O ChatGPT e outras inteligências artificiais generativas, no ensino superior, vieram para ficar” (Correia, 2023). Em meados de 2023, a maioria das/os estudantes universitárias/os já utilizava a IA generativa, com frequência, em diversas atividades acadêmicas (Chegg.org, 2023). As práticas de aprendizagem já haviam sido reconfiguradas (Carvalho; Pimentel, 2023). O que nos cabe, como professoras/es, é refletir sobre a intencionalidade pedagógica e a ética no uso das tecnologias generativas.

Considerando que a cibercultura nos inspira a (re)pensar nossas práticas didático-pedagógicas (Santos, 2005), iniciamos o semestre letivo de 2023 conversando com nossas turmas sobre como poderíamos usar essas tecnologias em nossas disciplinas considerando-as aliadas e não substitutas de nossa inteligência. Desde então, temos empreendido esforços para compreender como integrá-las em nossas aulas.

A capacidade de gerar boas respostas sobre conteúdos científico-educacionais

A IA generativa surpreendeu o mundo ao demonstrar a capacidade de discorrer sobre quase todos os assuntos com uma velocidade sobre-humana. Basta enviar uma pergunta e ela começa a escrever velozmente, com rapidez superior à nossa capacidade de leitura. Tudo muito simples e arrebatador. Contudo, para meu pai (o do primeiro autor deste texto), foi uma decepção. Poucos meses após o lançamento do ChatGPT, propus um desafio:

— O que você gostaria que ele fizesse, pai?

— Peça para ele escrever um conto ao estilo de Nelson Rodrigues.

O conto gerado foi “muito bobo”, de acordo com a avaliação de meu pai. O ChatGPT fracassou, na época, porque tinha pouco conhecimento sobre a literatura brasileira. Afinal, ele foi desenvolvido por uma empresa norte-americana voltada para o público de língua inglesa, e aquela versão que estávamos utilizando havia processado poucos textos em português e, portanto, continha poucas informações sobre nossa literatura.

— Peça outra coisa, pai!

— Pergunte a ele quem são as pessoas mais famosas de nossa cidade.

Meus pais moram em Barra do Piraí, uma cidade do interior do Rio de Janeiro com cerca de 100 mil habitantes. Em vez de informar que não possuía dados sobre aquele município, o ChatGPT forneceu informações falsas, afirmando que algumas/uns artistas famosas/os eram de lá. Estranhamos e, ao verificarmos a cidade de origem dos nomes citados,

constatamos que todas as informações estavam erradas. Chitãozinho e Xororó, por exemplo, não são barrenses!

Naquela época, ainda desconhecíamos as limitações da IA generativa, e as respostas equivocadas — que hoje identificamos como “alucinações” (OpenAI, 2023d) — fizeram meu pai considerá-la boba e inútil. A lição que aprendemos com essa anedota: é imprescindível conhecer não só as potencialidades, mas também as limitações da IA generativa; caso contrário, corremos o risco de criar expectativas irreais e, inevitavelmente, de nos frustrar.

A IA generativa tem limitações que precisam ser conhecidas, mas também é verdade que ela está evoluindo rapidamente, com o lançamento de modelos cada vez mais poderosos, como exemplifica a evolução da família de modelos GPT:

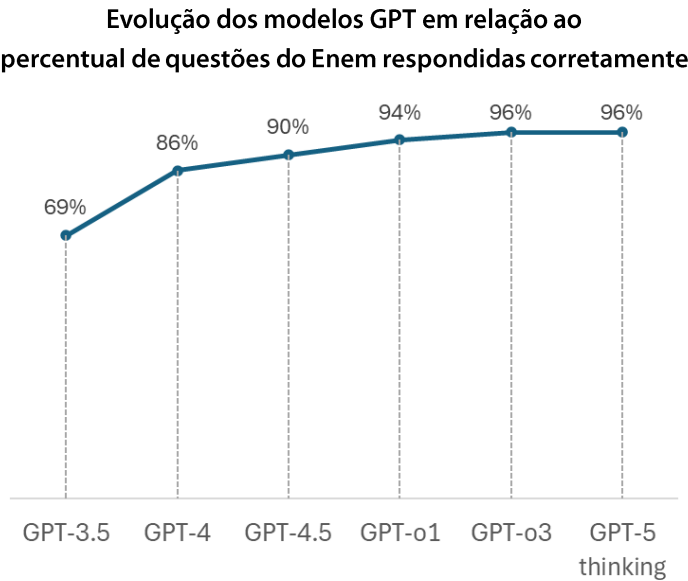
- GPT-3.5 (OpenAI, 2022a), lançado em março de 2022;
- GPT-4 (OpenAI, 2023b; 2023c), lançado em março de 2023;
- GPT-4o¹² (OpenAI, 2024c), lançado em maio de 2024;
- o1 (OpenAI, 2024e), lançado em setembro de 2024;
- GPT 4.5 (OpenAI, 2025e), lançado em fevereiro de 2025;
- o3 (OpenAI, 2025g), lançado em abril de 2025;
- GPT 5 (OpenAI, 2025n), lançado em agosto de 2025; e assim por diante...

Testamos a capacidade desses modelos GPT em responder corretamente às questões do Enem. Os resultados desses testes estão apresentados na tabela e na figura a seguir. Nossa avaliação evidenciou o crescimento da capacidade desses modelos em responder corretamente às questões sobre conteúdos científico-educacionais.

¹² A letra “o” é uma referência à *Omni*, uma palavra de origem latina que significa tudo ou todos. Em inglês, “4o” é lido com uma sonoridade parecida com “for all”, que também significa “para todos”. Esse nome, GPT-4o, foi escolhido para indicar que essa versão é de acesso gratuito para todas as pessoas, ao contrário de sua predecessora, GPT-4, que é restrita às/aos assinantes do ChatGPT Plus, cuja mensalidade é de 20 dólares.

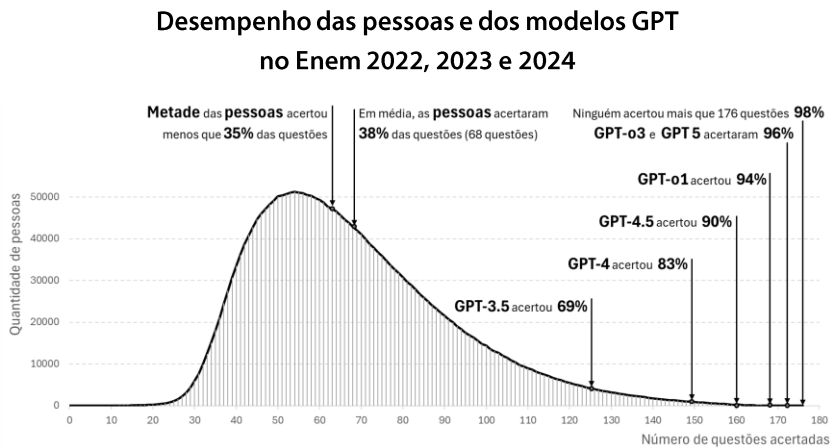
Desempenho dos modelos GPT no Enem						
	GPT-3.5	GPT-4	GPT-4.5	GPT-o1	GPT-o3	GPT-5 thinking
	Enem 2022 ^(a)	Enem 2023 ^(b)	Enem 2024 ^(c)	Enem 2024 ^(d)	Enem 2024 ^(e)	Enem 2024 ^(f)
Linguagens e códigos	76%	87%	93%	96%	96%	98%
Ciências humanas	92%	100%	98%	100%	100%	100%
Ciências da natureza	73%	91%	86%	91%	95%	93%
Matemática	18%	66%	82%	89%	93%	93%
Total	69%	86%	90%	94%	96%	96%

Fonte: Dos autores, com base nos dados: a) Nunes et al. (2023); b) Pires et al. (2023); c) OpenAI, (2025i); d) OpenAI (2025h); e) OpenAI (2025l); f) OpenAI (2025o).



Fonte: dos autores, com base em dados apresentados na tabela anterior.

Algumas pessoas consideram que a IA generativa não é útil porque pode produzir informações equivocadas, como as respostas erradas dadas pelos modelos GPT a certas questões do Enem. Para refletirmos melhor sobre a utilidade da IA generativa, comparamos o desempenho desses modelos com o das pessoas que fizeram as quatro provas do Enem nos anos de 2022, 2023 e 2024 (Microdados [...], 2022; 2023; 2024), conforme apresentado na figura a seguir.



Fonte: dos autores, com base nos Microdados do Enem e dados da tabela anterior.

Nós, seres humanos, também erramos. Nesses exames do Enem, ninguém acertou todas as questões; o maior número de acertos alcançado nessas três edições do Enem foi de 176 questões (98%). Metade das pessoas que fizeram as quatro provas dessas edições do Enem¹³ acertou 35% das questões ou menos. Em média, as pessoas acertaram 38% das questões.

¹³ No Enem 2022, 3,5 milhões de pessoas se inscreveram, mas somente 2,3 milhões fizeram as quatro provas (33% das pessoas não fizeram as provas). No Enem 2023, 3,9 milhões de pessoas se inscreveram, mas somente 2,7 milhões fizeram as quatro provas do Enem (32% das pessoas deixaram de fazer as provas daquele Enem). No Enem 2024, 4,3 milhões de pessoas se inscreveram, mas somente 3,0 milhões fizeram as quatro provas do Enem (31% das pessoas deixaram de fazer as provas daquele Enem).

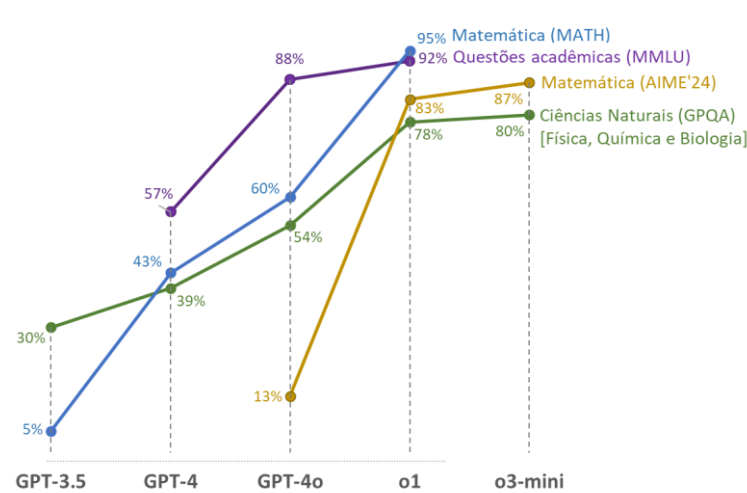
As sucessivas versões do GPT acertaram cada vez mais questões, sendo que o modelo GPT-o3 e GPT 5-*thinking* acertaram 172 das 180 questões (96%) do Enem 2024. Exceto por 5 pessoas que acertaram mais questões que esses dois modelos, todas as outras 8.013.175 pessoas que fizeram as quatro provas do Enem de 2022, 2023 ou 2024 poderiam se beneficiar ao estudar com o ChatGPT, pois ele acertou mais questões do que elas e ainda explicou o raciocínio empregado na resolução das questões. Pelo histórico, podemos supor que as próximas versões do GPT acertarão mais questões do que qualquer pessoa acertará. Esse resultado demonstra a utilidade dessa tecnologia. O fato de eventualmente alucinar não deve nos levar à conclusão de que a IA generativa é inútil.

O modelo mais antigo que foi lançado para o público geral, o GPT-3.5, já apresentou um excelente desempenho na prova de Ciências Humanas do Enem, com mais de 90% de acertos. Entretanto, o desempenho em Matemática foi insatisfatório: acertou apenas 18% das questões, um percentual compatível com escolhas aleatórias, como se tivesse simplesmente “chutado” as respostas, sem demonstrar uma real capacidade de resolução de problemas matemáticos. Essa dificuldade tornou-se bem conhecida. Um modelo de linguagem aprende padrões linguísticos; não opera como uma calculadora ou *software* matemático capaz de realizar cálculos complexos, apresentando dificuldades em manipulação algébrica avançada, raciocínio lógico detalhado e interpretação precisa de gráficos e tabelas. Desde então, melhorar o raciocínio matemático tornou-se um objetivo de pesquisadoras/es e desenvolvedoras/es da área (Lightman *et al.*, 2023):

Treinamos um modelo para alcançar um novo estado da arte na resolução de problemas matemáticos, recompensando cada passo correto de raciocínio (“supervisão do processo”), em vez de simplesmente recompensar a resposta final correta (“supervisão do resultado”). Além de melhorar o desempenho em relação à supervisão de resultados, a supervisão de processo também traz um benefício importante de alinhamento: treina diretamente o modelo para produzir uma cadeia de pensamento endossada por humanos. (OpenAI, 2023e, tradução nossa)

Os modelos seguintes, o GPT-4 e o GPT-4.5, apresentaram uma melhora na capacidade de resolver problemas matemáticos. Já os modelos 4o, o1 e o3 apresentaram uma performance ainda melhor porque implementam a técnica “Cadeia de Pensamento” (do inglês *Chain of Thought - CoT*); nessa abordagem, em vez de gerar uma resposta diretamente, o modelo primeiro produz informações auxiliares que são utilizadas para apoiar a geração da resposta final. O modelo processa informações por alguns segundos, como se estivesse “pensando”. Essa técnica melhora o desempenho em tarefas que exigem múltiplas etapas de raciocínio, como cálculos matemáticos, raciocínio lógico e interpretação de texto. Como resultado, esses modelos acertaram mais questões do que os modelos anteriores nos testes de desempenho (*benchmark*) em Matemática e em outras áreas, conforme os dados apresentados na figura a seguir.

Comparação da performance dos modelos do GPT em relação ao percentual de acerto em testes de desempenho (*benchmarks*)



Fonte: dos autores, com base nos dados: OpenAI (2023d; 2024c; 2024f; 2025d; 2025e); Yue et al. (2024); Rein et al. (2023); Hendrycks et al. (2021).

A utilidade dos modelos de linguagem para a área da Matemática reside no fato de conseguirem dar explicações, detalhar estratégias de resolução, apresentar um passo a passo para a resolução de um problema,

justificar uma resposta, tirar dúvidas. Nesse sentido, a tecnologia generativa complementa outras tecnologias que garantem precisão nos cálculos e fornecem visualizações gráficas precisas e interativas.

Com essas avaliações, queremos ilustrar que os modelos de linguagem vêm sendo aperfeiçoados a cada ano: aumentam a quantidade de informações armazenadas (parâmetros), processam mais dados durante o treinamento, aprimoram as técnicas utilizadas e implementam novas abordagens de processamento da informação (como a Cadeia de Pensamento), entre outros avanços.

Os testes de desempenho utilizados aqui para demonstrar a evolução das sucessivas versões do GPT também servem para comparar o desempenho entre os modelos concorrentes. Nos dois primeiros anos após o lançamento do ChatGPT, além da OpenAI, empresas como o Google (modelo Gemini), a Anthropic (modelo Claude), a xAI (modelo Grok) e a Meta (modelo LLaMa) passaram a disputar uma corrida para desenvolver o modelo de linguagem mais poderoso, capaz de acertar mais questões nos testes de desempenho, resolver problemas mais complexos, alucinar menos e promover a melhor experiência para as/os usuárias/os. Em 2025, o mundo foi surpreendido pelo lançamento de modelos desenvolvidos por empresas chinesas, como o DeepSeek-v3 (DeepSeek, 2025) e o Qwen-2.5 (Qwen, 2025), que apresentaram desempenho superior, em alguns testes, em relação aos modelos gratuitos do GPT. A corrida pela liderança em IA generativa tornou-se uma competição global, o que tem impulsionado a produção de modelos cada vez mais avançados e úteis.

Apesar dos avanços frenéticos que tornam a IA generativa cada vez mais relevante para a educação, devemos também reconhecer suas limitações. Por exemplo, sabemos que o modelo GPT-3.5 havia sido treinado com textos da Wikipédia, de livros, fóruns e blogs, mas não com artigos de revistas científicas. Por isso, ele não se equiparava aos conhecimentos de uma pessoa doutora em uma área específica. Perguntas que exigiam um conhecimento muito avançado nem sempre eram bem respondidas pelo GPT-3.5.

Às vezes, a IA generativa gera conteúdo com informações erradas, mas, por produzir um texto bem escrito, pode parecer que está mentindo. Entretanto, a IA generativa não tem intencionalidade: não quer nos enganar

ou mentir. Essa tecnologia não compreende a semântica do texto e não tem como julgar se o texto gerado contém informação verdadeira ou falsa. Para lidar com essa limitação, pesquisadoras/es e desenvolvedoras/es têm investigado algumas soluções. Por exemplo, no Gemini, após a geração de uma resposta, pode ser feita a checagem da informação por meio de uma busca na internet por conteúdos relacionados a cada frase gerada. Outra estratégia é criar modelos maiores e treinados com mais textos sobre todos os assuntos, o que potencialmente reduz a ocorrência de alucinações, embora não as elimine por completo.

A possibilidade de alucinação aumenta se o modelo tiver sido treinado com pouco conteúdo relacionado ao *prompt* digitado pela/o usuária/o. Por exemplo, o GPT-3.5 havia processado poucos textos em português e, portanto, possuía informações limitadas sobre nossa história, cultura, valores, costumes e crenças, o que aumentava a chance de alucinação em perguntas relacionadas ao Brasil. Quando um modelo é treinado com pouca informação sobre determinado assunto, pode acabar reproduzindo o mesmo conteúdo processado, seja por meio de paráfrase ou transcrição literal, o que é considerado plágio:

Pesquisadores da Universidade do Estado da Pensilvânia (Penn State), nos Estados Unidos, investigaram até que ponto modelos de linguagem natural como o ChatGPT, que usam inteligência artificial para formular uma prosa realista e articulada em resposta a perguntas de usuários, conseguem gerar conteúdo que não se caracterize como plágio. Isso porque esses sistemas processam, memorizam e reproduzem informações preexistentes, baseadas em gigantescos volumes de dados disponíveis na internet, tais como livros, artigos científicos, páginas da Wikipédia e notícias.

O grupo analisou 210 mil textos gerados pelo programa GPT-2, da startup OpenAI, criadora do ChatGPT, em busca de indícios de três diferentes tipos de plágio: a transcrição literal, obtida copiando e colando trechos; a paráfrase, que troca palavras por sinônimos a fim de obter resultados ligeiramente diferentes; e o uso de uma ideia elaborada por outra pessoa sem mencionar sua autoria, mesmo que formulada de maneira diferente.

A conclusão do estudo foi de que todos os três tipos de cópia estão presentes. (Marques, 2023)

As informações geradas pelo modelo de linguagem também ficam desatualizadas em função do momento em que foi realizado seu treinamento. Por exemplo, em 2024, perguntamos ao GPT-3.5: “Quem é o presidente do Brasil?”, e ele respondeu: “Atualmente, o presidente do Brasil é Jair Bolsonaro”¹⁴. Para lidar com essa limitação, muitas tecnologias generativas passaram a recuperar informações da internet em tempo real para apoiar a geração de respostas com dados atualizados.

É comum que pessoas façam perguntas à IA sobre seu funcionamento e sua natureza: potencialidades, limitações, implicações e características técnicas. Contudo, essa tecnologia não possui consciência de si mesma, o que a impede de refletir sobre sua própria existência, seu funcionamento e suas técnicas. Apesar disso, muitas pessoas insistem em fazer perguntas como se ela possuísse autoconhecimento¹⁵.

É importante conhecer as limitações da IA generativa para que possamos dar sentido às suas respostas, especialmente quando são equivocadas ou superficiais. No entanto, essas limitações não devem nos levar à conclusão de que a IA generativa é inútil. Mostramos que mesmo as versões iniciais dos modelos GPT já apresentavam boa capacidade de interpretação de textos e de raciocínio sobre conteúdos científico-educacionais (embora com desempenho insatisfatório em Matemática). Mostramos, também, que os modelos continuam sendo aprimorados em uma corrida incessante para desenvolver versões cada vez mais eficazes. Com essa discussão, queremos que até as pessoas mais céticas ou decepcionadas, como o pai do primeiro autor, tenham certeza de que os modelos de IA generativa são muito relevantes para os processos educacionais e se tornarão cada vez mais úteis. Esse é o ponto de partida para as reflexões que tecemos neste livro.

¹⁴ Disponível em: <https://chatgpt.com/share/7c7ea738-a3f7-4fe1-9da4-4fe7645d146f>

¹⁵ Um modelo pode vir a gerar informações corretas sobre si mesmo se ele tiver sido treinado com esse tipo de conteúdo.

O uso da IA generativa em atividades acadêmicas e escolares é uma trapaça?

“Escolas de Nova York proibiram o uso de um aplicativo de inteligência artificial que consegue até escrever redações para os alunos” (Escolas [...], 2023, 00:01s-00:09s). Essa foi a primeira notícia sobre a IA generativa transmitida nacionalmente pela maior rede de televisão do nosso país. Ela é emblemática da preocupação de docentes e gestoras/es, da reação ao medo do desconhecido, da resistência à mudança.

“O que parece ser o maior sonho de qualquer estudante é o maior pesadelo das escolas. [...] Os professores lutam para descobrir como evitar trapaças por parte dos estudantes” (Escolas [...], 2023, 00:23-01:05). De fato, não há aprendizagem nem desenvolvimento das capacidades intelectuais quando um/a estudante utiliza a IA para gerar respostas às atividades didáticas sem se envolver ativamente no processo de elaboração e sem refletir criticamente sobre os resultados. Isso resulta em um sentimento de autoria prejudicada e baixa atividade neural (Kosmyna *et al.*, 2025), um desperdício da oportunidade de formação, uma trapaça acadêmica.

Contudo, quando a redação é elaborada pela/o própria/o estudante e só depois ela/e utiliza a tecnologia generativa para revisar e aprimorar o texto, observa-se um sentimento de autoria forte e uma atividade neural mais intensa. Nessa situação, a IA generativa torna-se uma aliada no processo de aprendizagem e desenvolvimento da escrita. Essas são as conclusões de uma importante pesquisa científica sobre a atividade cerebral durante o processo de escrita de redações, que constatou que o momento estratégico para usar uma tecnologia de IA generativa é após um esforço inicial autônomo de redação sem uso de tecnologias digitais, pois isso aumenta o engajamento e a integração neural (Kosmyna *et al.*, 2025).

Um mês após terem proibido o uso de IA nas escolas de Nova York, mudaram a decisão: “NY derruba proibição ao ChatGPT e passa a incentivar IA nas escolas” (Possa, 2023).

As/Os estudantes utilizam as tecnologias generativas de maneiras diversas e inventivas, não apenas para que a IA realize a tarefa por elas/es. Essas tecnologias podem ser empregadas como apoio à pesquisa, ao estudo-aprendizagem e à cocriação, em um processo conversacional, interativo, de idas e vindas, que não configura trapaça acadêmica.

A possibilidade de estudantes utilizarem as respostas geradas por IA em atividades escolares e universitárias tornou-se uma preocupação recorrente entre docentes. Desde o lançamento do ChatGPT, passou a ser difícil afirmar se um texto ou atividade foi produzido por um/a estudante ou por uma tecnologia generativa, uma vez que os conteúdos gerados computacionalmente são cada vez mais coerentes, informativos e, por vezes, indistinguíveis dos textos humanos. Diante dessa dificuldade, foram desenvolvidos os chamados “detectores de IA”: sistemas computacionais projetados para identificar se um texto foi gerado por uma IA ou escrito por uma pessoa.

Textos gerados por IA costumam ser originais, o que inviabiliza sua detecção por meio da comparação com documentos existentes, que é o método empregado pelos sistemas tradicionais de identificação de plágio. Para contornar essa limitação, os detectores de IA adotam outra estratégia: analisam características textuais que podem indicar autoria artificial, como padrões estilísticos, recorrência de determinadas palavras e estruturas linguísticas, além da ausência de erros gramaticais e ortográficos. No entanto, esses sistemas podem falhar: um texto escrito por uma pessoa pode ser erroneamente classificado como gerado por IA (falso positivo), assim como um texto gerado por IA pode ser identificado como humano (falso negativo).

A seguir, apresentamos a análise de alguns detectores de IA aplicada a um texto gerado pelo ChatGPT. Dois detectores identificaram corretamente que o texto havia sido gerado por IA, enquanto um terceiro o classificou equivocadamente como sendo de autoria humana.

Identificação correta de texto gerado por IA



Esse é um texto gerado pelo ChatGPT. A inteligência artificial está transformando a maneira como interagimos com a tecnologia, oferecendo novas possibilidades de automação e criatividade. O ChatGPT, desenvolvido pela OpenAI, é um exemplo impressionante de como os modelos de linguagem generativa podem ser utilizados para criar textos coesos e informativos, além de responder a perguntas complexas e fornecer suporte em diversas áreas do conhecimento. Com a capacidade de aprender e se adaptar, essas ferramentas estão se tornando cada vez mais integradas ao nosso cotidiano, facilitando tarefas e proporcionando novas oportunidades para inovação.

Detectar texto

647/15.000 caracteres
(Obtenha até 100.000 [aqui](#))

 Subir arquivo

Seu texto é gerado por IA/GPT

 **100% IA GPT*** 

Esse é um texto gerado pelo ChatGPT. A inteligência artificial está transformando a maneira como interagimos com a tecnologia, oferecendo novas possibilidades de automação e criatividade. O ChatGPT, desenvolvido pela OpenAI, é um exemplo impressionante de como os modelos de linguagem generativa podem ser utilizados para criar textos científicos e informativos, além de responder a perguntas complexas e fornecer suporte em diversas áreas do conhecimento. Com a capacidade de aprender e se adaptar, essas ferramentas estão se tornando cada vez mais integradas ao nosso cotidiano, facilitando tarefas e proporcionando novas oportunidades para inovação.

 O texto destacado é suspeito de ser provavelmente gerado por IA*

647 caracteres
93 palavras

 **Exportar para PDF**

Fonte: ZeroChat (<https://www.zerogpt.com>)

Outro sistema que identificou corretamente o texto gerado por IA

Exemplos:

ChatGPT

Gemini

Human

HumanAI

Esse é um texto gerado pelo ChatGPT. A inteligência artificial está transformando a maneira como interagimos com a tecnologia, oferecendo novas possibilidades de automação e criatividade. O ChatGPT, desenvolvido pela OpenAI, é um exemplo impressionante de como os modelos de linguagem generativa podem ser utilizados para criar textos científicos e informativos, além de responder a perguntas complexas e fornecer suporte em diversas áreas do conhecimento. Com a capacidade de aprender e se adaptar, essas ferramentas estão se tornando cada vez mais integradas ao nosso cotidiano, facilitando tarefas e proporcionando novas oportunidades para inovação.

93 / 500 palavras

Analisar texto

100%

É provável que seu texto seja gerado por IA

Melhore sua escrita

✓ Evite plágio accidental

✓ Fácil e gratuito de usar

✓ Funciona com todos os dialetos ingleses

Experimente o Quillbot Paraphraser

Fonte: Scribbr (<https://www.scribbr.com/ai-detector>)

Identificação equivocada de texto como sendo de autoria humana

Cole o texto abaixo para testar nosso detector em tempo real!

Insira um texto com mais de 750 caracteres. Comprimento recomendado: 1500 caracteres. O uso está sujeito aos [Termos de Serviço](#) deste site.

Esse é um texto gerado pelo ChatGPT. A inteligência artificial está transformando a maneira como interagimos com a tecnologia, oferecendo novas possibilidades de automação e criatividade. O ChatGPT, desenvolvido pela OpenAI, é um exemplo impressionante de como os modelos de linguagem generativa podem ser utilizados para criar textos científicos e informativos, além de responder a perguntas complexas e fornecer suporte em diversas áreas do conhecimento. Com a capacidade de aprender e se adaptar, essas ferramentas estão se tornando cada vez mais integradas ao nosso cotidiano, facilitando tarefas e proporcionando novas oportunidades para inovação. À medida que essas tecnologias avançam, a colaboração entre humanos e máquinas se torna mais harmoniosa, permitindo que aproveitemos ao máximo as capacidades de ambos. O ChatGPT, por exemplo, pode auxiliar em projetos acadêmicos, oferecendo sugestões de redação, revisando conteúdos e até gerando ideias criativas.

961 / 8192

Editar

Apagar

RESULTADO

A entrada é: não é provável que contenha texto gerado por IA

5%

POR SEGMENTOS

Segmento	Pontuação
Esse é um texto ger...	5%

HIVE MODERATION

Fonte: Hive (<https://hivemoderation.com/ai-generated-content-detection>)

Com a popularização dos detectores de IA, foram desenvolvidos sistemas para “humanizar” textos gerados por IA. Esses sistemas modificam o texto com o objetivo de enganar os detectores, introduzindo propositalmente alguns erros e alterando o estilo de escrita, o tom e a fluidez da redação. Como resultado, um texto “humanizado” pode ser erroneamente classificado pelos detectores como de autoria humana. Utilizamos um desses sistemas de humanização para modificar o texto gerado pelo ChatGPT, e, de fato, o novo texto gerado foi classificado como “escrito por um humano”. A conclusão é que a detecção de textos gerados por IA é falível e pode ser facilmente enganada com o uso de sistemas que “humanizam” o conteúdo.

Sistema “humanizador” de texto gerado por IA

Seu novo conteúdo indetectável:

 < 1 / 3 >

Encontre o rascunho mais indetectável

Humanizar mais

Verifique se há IA

25%

É provável que seu texto tenha sido escrito **por um humano.**

Este texto foi produzido pelo ChatGPT. A interação com a tecnologia está sendo transformada pela inteligência artificial, proporcionando oportunidades inovadoras de automação e expressão criativa. O ChatGPT, criado pela OpenAI, é um exemplo incrível de como os modelos generativos de linguagem podem ser empregados para gerar textos informativos e informativos. Ele também se destaca pela sua capacidade em responder perguntas complexas e fornecer suporte em diversas áreas do conhecimento. À medida que aprendemos e nos adaptamos, essas ferramentas estão cada vez mais sendo úteis à rotina diária, tornando as tarefas mais fáceis e abrindo caminho para novas oportunidades de inovação.

97 palavras

Fonte: Stealthwriter.AI <<https://stealthwriter.ai>>

Concordamos com a afirmação: “Não tente detectar a IA — faça com que o uso da IA não seja o problema” (MIT[...], 2023). Nós, educadoras/es, precisamos ensinar quando e como usar a IA, em vez de proibir seu uso e punir quem a utiliza. Afinal, não é possível manter a educação como era antes da sua chegada. Devemos, sim, repensar as práticas educativas. Em nossos cotidianos formativos, partimos do pressuposto de que toda atividade assíncrona será realizada com o apoio de tecnologias generativas, por isso iniciamos nossas propostas de atividades com a seguinte orientação: “Usando a IA generativa, faça...”. Também criamos atividades que não podem ser realizadas exclusivamente por IA, valorizando a subjetividade de cada estudante, suas histórias de vida e de formação, os acontecimentos que marcaram suas trajetórias, suas formas singulares de compreender o mundo, seus sentimentos, desejos, medos, ideologias e demais facetas de suas identidades únicas.

Em nosso tempo de escola, nas remotas décadas de 1980 e 1990, nossas mães nos ensinaram a fazer os trabalhos copiando trechos de enciclopédias físicas. Torcemos para que, hoje, as atividades escolares não possam ser realizadas simplesmente copiando um texto pronto, seja de uma enciclopédia digital ou de outras fontes da internet, nem que possam ser feitas por uma IA que pouco ou nada conhece sobre sua/seu usuária/o. Talvez a popularização das tecnologias generativas venha a impulsionar uma transformação necessária nas práticas didático-pedagógicas e nos currículos escolares e universitários.

Precisamos educar para um uso da IA generativa que seja ético (não plagiar), estético (criar com sensibilidade e estilística própria) e político (formar sujeitos críticos). A ética é um tema central neste cenário à medida que as inovações computacionais vêm transformando as formas de produção do conhecimento, do seu compartilhamento e do acesso a ele. A integração das tecnologias digitais ao processo educacional impõe desafios éticos que exigem atenção de educadoras/es e gestoras/es, especialmente no que se refere à capacidade de avaliar e discernir quando o uso da tecnologia é responsável e formativo, ou quando representa apenas uma forma de trapaça acadêmica. É urgente reinventar a didática para que possamos utilizar as tecnologias como aliadas da aprendizagem, sem cair em armadilhas que possam comprometer a integridade do sistema educacional.

As tecnologias digitais podem ser analisadas sob a perspectiva filosófica da estética. Tradicionalmente associada à apreciação do belo e ao estudo da arte, a estética é uma noção mais ampla e refere-se à experiência sensorial e afetiva que molda nossa percepção do mundo. Nas tecnologias digitais, a dimensão estética ultrapassa o design visual das interfaces computacionais; ela envolve a experiência vivida por estudantes e docentes ao interagirem com os dispositivos tecnológicos, referindo-se à maneira como a tecnologia afeta os modos de aprender, comunicar e pensar.

A dimensão política do uso das tecnologias digitais na educação é igualmente complexa e multifacetada, envolvendo questões de poder, controle, vigilância e acesso. Essa dimensão se articula com instituições educacionais, políticas estatais e dinâmicas do capitalismo. No campo da macropolítica, destacam-se pautas como inclusão digital, letramento para a IA, cibercidadania, proteção de dados, soberania digital, formação continuada de docentes para a apropriação pedagógica das tecnologias e outras questões que influenciam diretamente os processos de transformação social. Ao debatermos as tecnologias digitais em contextos educacionais, precisamos indagar quais são os interesses por trás dessas tecnologias e como elas podem contribuir para a promoção da justiça educacional e social.

Na micropolítica do cotidiano escolar, tomamos decisões a todo instante sobre como usar (ou não) as tecnologias em nossas práticas pedagógicas: se iremos utilizá-las para potencializar a aprendizagem ou se vamos tentar impedir seu uso. Do ponto de vista de estudantes, também há escolhas: se irão aprender com essas tecnologias ou usá-las para fazer por elas/es os deveres de casa que consideram despropositados ou enfadonhos. Essas decisões diárias, muitas vezes sutis e nem sempre visíveis, revelam nossas intenções e preferências, bem como o investimento que fazemos na nossa própria formação e, em última instância, nossas concepções de mundo.

A popularização das tecnologias generativas desafia a educação tradicional e abre possibilidades para a reinvenção das práticas didático-pedagógicas. Educar para o uso ético, estético e político dessas tecnologias, apoiando estudantes a refletirem sobre elas, e delas se apropriarem de modo formativo, é um compromisso que devemos assumir na preparação de uma geração que já utiliza a IA generativa cotidianamente.

Inventando formas de utilizar a IA generativa no cotidiano educacional

Antes do ChatGPT, estávamos acostumadas/os com sistemas inteligentes capazes de realizar uma tarefa específica: jogar xadrez, traduzir um texto, reconhecer a digital, a face ou a fala humanas, recomendar *sites*, filmes ou produtos, entre outras funções. A IA generativa não se limita a uma única atividade, pois é capaz de produzir textos sobre praticamente qualquer assunto e em diversos formatos, além de gerar conteúdos em outras linguagens, como imagem, som e vídeo. Algumas/uns especialistas consideram que a IA generativa é avançada demais para ser classificada como um tipo de Inteligência Artificial Específica (ou Estreita, do inglês *Narrow AI*), que se restringe a uma única atividade. Contudo, ela ainda não é suficientemente avançada para ser considerada uma Inteligência Artificial Geral (do inglês *Artificial General Intelligence*), que seria uma forma avançada de IA, capaz de realizar qualquer atividade intelectual que uma pessoa consiga executar, compreendendo o contexto e agindo de maneira coerente, sem a necessidade de ter sido previamente programada ou treinada para aquela tarefa específica. Esse é o ponto que se deseja alcançar com a IA, mas ainda não chegamos lá.

Para nós, professoras/es, o importante é que somos capazes de propor inúmeras atividades didáticas com as tecnologias generativas. Não podemos, entretanto, dar “receitas prontas” de como utilizá-las pedagogicamente, pois cada docente inventa suas próprias práticas, cria diferentes situações de aprendizagem considerando a disciplina, a turma, o nível de ensino, a modalidade (presencial, a distância ou híbrida), suas crenças e valores pedagógicos, seus referenciais teóricos, suas experiências etc. Reconhecendo essa diversidade e a autonomia didática, o que podemos

fazer é compartilhar nossas próprias práticas de uso da IA generativa com o objetivo de, talvez, inspirar outras/os docentes, assim como aprendemos com nossas/os colegas que compartilharam suas inventividades conosco.

Nós, autores deste livro, lecionamos na graduação e na pós-graduação em Computação e em Educação, além de atuarmos na formação inicial e continuada de professoras/es. Esse é o nosso lugar de fala. As atividades didáticas com o uso de tecnologias generativas, que aqui compartilhamos, foram pensadas e praticadas nesses contextos pedagógicos. No entanto, nossos exemplos e proposições podem ser úteis mesmo para quem leciona em outros cursos e níveis, pois servem de base para reflexões e adaptações em diferentes cotidianos educacionais.

Enquanto professores-pesquisadores-autores, sentimos que a IA generativa tem se mostrado útil para nos apoiar a aprender, ensinar, escrever, pesquisar e imaginar, conforme apresentamos a seguir.

4.1 – Para aprender e estudar

Utilizamos a IA generativa para aprender: conhecer diferentes assuntos, perguntar, pedir exemplos, tirar dúvidas, questionar determinadas afirmações, investigar etc. Por exemplo, quando perguntamos à IA generativa quais eram os conceitos-chave da Teoria Ator-Rede, ela elaborou um quadro com a definição dos principais termos; quando pedimos que explicasse as diferenças entre as teorias comportamentais e cognitivas, ela gerou uma caracterização resumida de cada teoria e apresentou as principais semelhanças e diferenças entre elas; quando desejamos investigar o significado ou a etimologia de uma palavra, perguntamos à IA e, somente quando necessário, consultamos outras fontes de informação.

Algumas/uns estudantes utilizam a IA generativa para apoiar os estudos. Por exemplo, pedem para ela dar uma aula sobre um assunto, criar um roteiro de estudo, elaborar um vídeo, *podcast* ou mapa mental sobre o conteúdo da disciplina, propor exercícios, corrigir os exercícios feitos, elaborar uma prova ou teste para que possam avaliar o próprio aprendizado, criar histórias relacionadas ao conteúdo em estudo etc. A IA gera conteúdo

conforme as demandas da/o própria/o estudante, possibilitando uma aprendizagem conversada e contextualizada, o que vem modificando a rotina de estudo.

A IA generativa também tem modificado as práticas de leitura de textos. Algumas tecnologias possibilitam anexar um arquivo ao *prompt* ou um *link* para um conteúdo online, possibilitando à/ao usuária/o conversar sobre a obra com a IA, solicitar um resumo do texto, localizar trechos em que determinada temática foi abordada, pedir para listar os argumentos e o posicionamento defendido pelas/os autoras/es, responder a uma questão, entre outras tarefas.

Apesar do potencial para o estudo-aprendizagem, é preciso sempre desconfiar das respostas geradas pela IA, pois ela pode inventar definições, fornecer explicações equivocadas, “alucinar”. A maior preocupação das/os docentes é que estudantes solicitem que a IA faça as atividades por elas/es, entregando assim suas mentes e vozes para a inteligência artificial pensar e se expressar por elas/es. Essa atitude compromete a proposta de estudo-aprendizagem elaborada pela/o docente ou pelo material didático. Por isso, precisamos educar nossas/os estudantes para o uso responsável da IA generativa no contexto educacional.

4.2 – Para ensinar

A IA generativa tem se mostrado útil para nos apoiar no planejamento de aulas, na preparação de conteúdos, na síntese de textos, na estruturação de apresentações, na criação de *slides*, exemplos, casos, exercícios, questões para debate, dinâmicas colaborativas, jogos para o ensino-aprendizagem, entre outras atividades relacionadas ao trabalho docente. Obviamente, todos os conteúdos gerados precisam ser revisados e editados, pois podem conter erros ou não ser adequados ao contexto da turma.

A IA também é útil na elaboração de instrumentos de avaliação. Por exemplo, podemos solicitar que ela produza uma prova com questões de múltipla escolha para testar diferentes níveis de compreensão sobre um determinado assunto, avaliando desde o entendimento das noções fundamentais até a aplicação de conceitos em cenários complexos.

Podemos anexar arquivos ao *prompt* e pedir que elabore questões discursivas relacionadas ao conteúdo abordado no documento anexado. Ela é capaz de gerar um gabarito comentado para as questões, como também critérios e rubricas de avaliação.

A IA generativa também pode ser utilizada para corrigir as respostas de estudantes, mas essa ação requer cuidado. Ela é útil para apoiar a correção de exercícios simples e sugerir revisões de redações e trabalhos, indicando o que corrigir ou aprimorar em determinado texto. Contudo, não recomendamos a completa automação da avaliação, pois esse processo precisa ser conduzido por professoras/es com experiência, conhecimento, discernimento e sensibilidade necessários para uma avaliação adequada, especialmente quando a resposta não pode ser julgada simplesmente como certa ou errada. A avaliação também pode envolver uma análise comparativa das respostas da turma, o que nem sempre é simples para a IA realizar. Em nossas experimentações até o presente momento, identificamos que ela pode ser útil nesse processo, mas ainda requer supervisão humana. A avaliação também não deve ser totalmente automatizada porque não se trata apenas da atribuição de notas a respostas, mas sim de um processo de acompanhamento da aprendizagem de cada estudante.

Pensamos o uso da IA generativa em uma relação de parceria, para expandir nossas capacidades intelectuais, jamais como substituição de nossa inteligência. Recusamos seu uso para substituir educadoras/es ou automatizar o processo de ensino, incluindo a avaliação da aprendizagem. Consideramos adequado seu emprego em processos de cocriação com seres humanos, servindo como apoio às diversas atividades do trabalho docente, ampliando nossa criatividade e aprimorando os processos de trabalho.

Em aulas presenciais, gostamos de usar a IA generativa como se fosse uma pessoa da turma. Quando fazemos uma pergunta e ninguém responde, ou quando há muita confusão nas respostas apresentadas por estudantes, repetimos a mesma pergunta para a IA e analisamos a resposta com a turma: o que está correto na resposta? Há algo errado? Algo ficou faltando? Que informações precisam ser verificadas? Refletimos sobre os pontos fortes e as eventuais lacunas ou imprecisões da resposta gerada. Esse tipo de tarefa

contribui para o desenvolvimento da capacidade de avaliar criticamente a segurança das informações geradas e a veracidade dos dados, além de estimular a discussão sobre a autoridade e as fontes de informação.

Em nossos cursos de formação continuada, colegas professoras/es têm criado diversas situações didáticas interessantes com o uso de tecnologias generativas em diferentes contextos formativos. Por exemplo, um professor de Engenharia desenvolveu um roteiro de estudo contendo perguntas para que cada estudante iniciasse uma conversa com a IA, incentivando a aprendizagem por meio do inquérito. O objetivo desse professor era implementar uma sala de aula invertida, na qual a/o estudante primeiro aprende conversando com a IA e, depois, sistematiza os conhecimentos em aula.

No início de 2023, logo após o lançamento do ChatGPT, recebemos uma lista com “30 formas de utilizar o ChatGPT em aula” (Raulino; Ramos, 2023)¹⁶. Nos meses seguintes, acompanhamos uma enxurrada de sugestões e relatos sobre o uso da IA generativa em diferentes disciplinas, divulgados por meio de vídeos, postagens, artigos científicos, livros e outras fontes. Talvez você consiga encontrar alguém que já tenha compartilhado sugestões de uso da IA generativa na mesma disciplina e nível em que você leciona.

Quando perguntamos ao ChatGPT quais atividades didáticas poderiam ser realizadas com ele e quais eram suas potencialidades pedagógicas, ele listou algumas sugestões interessantes¹⁷. Nos cursos de formação continuada de professoras/es, temos constatado que conhecer alguns exemplos de uso pedagógico da IA generativa costuma ser suficiente para disparar a autoria docente. Quanto maior for o seu repertório, maior será o potencial para inventar novas formas de utilizar as tecnologias generativas em seu contexto educacional específico.

¹⁶ <https://view.genial.ly/63d9bbc140e03c0017ca1603>

¹⁷ Para nós, ele respondeu assim: <https://chat.openai.com/share/74341d3c-3683-4bca-81cf-9c73648346cc>

4.3 – Para escrever

A IA generativa reconfigurou nossas práticas de escrita. Antes dela, a escrita costumava ser um processo solitário; agora, escrevemos em parceria com a IA. O bloqueio criativo não existe mais: basta pedir para ela gerar um roteiro ou um primeiro rascunho do texto, e talvez um segundo, terceiro e quantos rascunhos mais forem necessários. As respostas geradas nos ajudam a refletir sobre o texto que desejamos tecer, fornecem ideias e diferentes pontos de vista que ampliam nossa criatividade.

Gostamos de pedir à IA que discorra sobre um assunto para avaliarmos outras possibilidades de abordagem. Às vezes, apresentamos apenas uma ideia ou argumento e solicitamos que ela construa um parágrafo sobre aquilo. Podemos pedir que complete uma frase ou um parágrafo que estamos escrevendo. Quando nos falta uma palavra, pedimos sugestões para preencher a lacuna deixada na frase.

Sempre utilizamos a IA para revisar nossos textos. Ela corrige erros ortográficos e gramaticais, oferece sugestões para melhoria da redação e, se solicitado, pode até propor revisões dos argumentos utilizados. Obviamente, não acatamos todas as sugestões, mas esse movimento nos ajuda a repensar o texto que desejamos tecer. Para o processo de revisão de texto, gostamos de utilizar o ChatGPT com a extensão do navegador editGPT¹⁸, que nos possibilita visualizar o que foi cortado e o que foi acrescentado ao texto, auxiliando na tomada de decisão sobre as mudanças que queremos efetivar.

Há muitas formas de revisar um texto, e devemos especificar o tipo de revisão que a IA deve realizar: podemos demandar que corrija apenas os erros ortográficos e gramaticais, que simplifique a linguagem usando palavras mais comuns, que torne o texto mais formal e acadêmico, ou que faça uma reestruturação completa para torná-lo mais coerente e coeso, entre outras possibilidades. Por exemplo, solicitamos que a IA revisasse os textos deste livro garantindo o emprego da linguagem inclusiva de gênero com base em alguns exemplos que fornecemos: aluna/o, professor/a, algumas/uns, da/o usuária/o etc.

¹⁸ <https://editgpt.app/how-to>

A IA generativa também é capaz de fazer boas traduções. Neste livro, traduzimos alguns trechos originalmente escritos em inglês, e a IA se mostrou muito útil para nos apoiar nessa atividade, quase sempre apresentando uma tradução melhor do que seríamos capazes de produzir. Apenas ocasionalmente tivemos que ajustar algo na tradução gerada. Quando ficávamos insatisfeitos com uma versão, pedíamos que a IA gerasse outra, e as diferentes opções nos ajudavam a decidir qual tradução considerávamos mais adequada. De tradutores, passamos a atuar como avaliadores ou curadores das melhores traduções feitas pela IA.

Temos orientado estudantes de graduação e pós-graduação a utilizar a IA generativa como uma parceira na escrita de textos acadêmicos, mas sem deixá-la pensar e escrever por elas/es. Recomendamos que primeiro escrevam suas próprias ideias, para só então trabalhar com a IA na melhoria da argumentação; e sugerimos que sempre revisem seus textos com a IA. Temos constatado uma melhora considerável nas produções de nossas/os orientandas/os.

4.4 – Para pesquisar

Certa vez, quando escrevíamos um artigo sobre o uso da IA por estudantes, decidimos incluir a declaração de um aluno que havia nos chamado a atenção: “Uso o Copilot, hoje codifico igual a um deus e aprendo demais com o GPT-4” (discutiremos essa declaração no próximo capítulo). Depois de elaborarmos algumas análises, resolvemos perguntar ao ChatGPT como interpretar aquele depoimento, o que estava subjacente ao discurso, o que o enunciado revelava de potência e quais eram os riscos associados àquela declaração. A análise gerada¹⁹ nos surpreendeu, pois abordou aspectos pertinentes e trouxe reflexões que não havíamos considerado inicialmente. Foi assim que descobrimos que poderíamos utilizar a IA generativa para apoiar nossas pesquisas.

Além de auxiliar na etapa de escrita de artigos, a IA generativa também pode ser útil na elaboração do desenho da pesquisa, na curadoria de artigos para a revisão de literatura, na análise e comparação dos artigos selecionados, na construção de documentos auxiliares (como o projeto de

¹⁹ <https://chatgpt.com/share/14dd68a0-b3e6-4a70-8646-1f87f1eee17f>

pesquisa e o Termo de Consentimento Livre e Esclarecido), na elaboração de instrumentos de pesquisa (questionários e roteiros de entrevista), na análise de dados qualitativos e quantitativos, na visualização de dados, entre outras atividades. A IA generativa irá reconfigurar as práticas do fazer científico. Estamos acompanhando os avanços dessa tecnologia e o lançamento de sistemas específicos para apoiar a realização de pesquisas. Em breve, (quase) todas/os as/os pesquisadoras/es utilizarão essa tecnologia em seus cotidianos.

4.5 – Para imaginar

Imaginar vem do latim *imaginari* e significa formar uma imagem mentalmente. A IA generativa tem a capacidade de imaginar, de gerar imagens computacionalmente. No contexto educacional, a geração de imagens pode ser útil para ilustrar uma apresentação, representar um conceito, criar personagens e cenas de uma história, ilustrar um meme, entre outras produções visuais de interesse de estudantes, docentes e conteudistas.

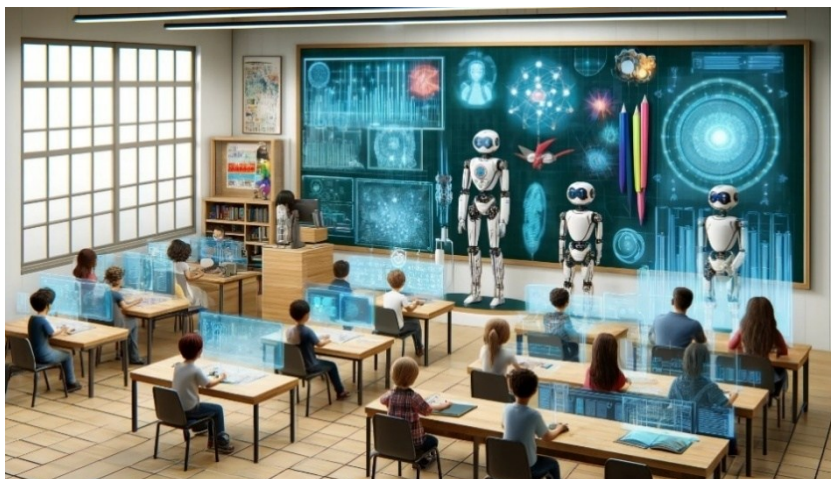
A OpenAI, antes de lançar o ChatGPT, lançou o DALL-E²⁰. “um sistema de IA que pode criar imagens e arte realistas a partir de uma descrição em linguagem natural” (DALL-E 2, s.d., tradução nossa). Com uma interface simples, a/o usuário/a descreve uma cena e a IA gera uma imagem. Nunca havíamos visto um sistema com essa capacidade. Em outubro de 2023, foi lançada a versão DALL-E 3, que apresentou uma “significativa melhora em relação ao DALL-E 2” (OpenAI, 2023j), sendo capaz de gerar imagens comparáveis às do Midjourney, que era o modelo mais avançado de geração de imagens à época. Essa versão foi integrada ao GPT-4 e disponibilizada gratuitamente no Copilot²¹ da Microsoft. Em maio de 2024, essa tecnologia ficou disponível no GPT-4o, de uso gratuito limitado, que “aceita como entrada qualquer combinação de texto, áudio, imagem e vídeo e gera qualquer combinação de saídas de texto, áudio e imagem.” (OpenAI, 2024c, tradução nossa)

²⁰ <https://openai.com/dall-e-2>

²¹ <https://copilot.microsoft.com/>

Para exemplificar essa capacidade, digitamos no *prompt* do GPT-4o: “Crie uma imagem sobre Educação e Inteligência Artificial”. O GPT gerou a imagem apresentada a seguir:

Prompt: “Crie uma imagem sobre Educação e Inteligência Artificial”



[ChatGPT] Aqui está a imagem que você solicitou sobre Educação e Inteligência Artificial. Ela mostra uma sala de aula moderna onde os alunos interagem com ferramentas alimentadas por IA, exibindo um ambiente futurista e realista que destaca como a IA aprimora a experiência educacional.

Fonte: criado com GPT-4o (OpenAI, 2025f)

Uma atividade que gostamos de realizar em nossos contextos formativos é analisar criticamente a imagem gerada pela IA. Por exemplo, sobre a figura anterior, devemos questionar: será que o futuro da educação é a substituição de professoras/es por inteligência artificial (representada pela ausência de uma pessoa adulta na sala de aula e com um robô em seu lugar)? Na educação do futuro, ainda haverá sala de aula física? A aula típica ainda será expositiva (representada por um quadro digital ocupando quase toda a parede da sala e com todas as crianças voltadas para as telas)? Estudantes vão interagir predominantemente com a inteligência artificial e com o conteúdo didático, sem conversarem entre si? Ainda será valorizada

a assimilação de conteúdos descontextualizados (educação bancária)? As cores das crianças nesta sala refletem a diversidade racial de nosso país?

O *prompt* utilizado para gerar a imagem anterior exemplifica a entrega de nossa mente à IA para que ela pense por nós, pois deixamos que a máquina imagine como será a Educação em tempos de IA generativa. Consideramos mais adequado primeiro imaginarmos, nós mesmas/os, como será o futuro da educação, para então descrevermos uma cena que a IA transformará em imagem; assim, a autoria da cena é nossa e utilizamos a IA como técnica de execução da obra, sem terceirizarmos nossa imaginação.

Por exemplo, podemos imaginar um futuro distópico em que a riqueza se concentrou nas mãos de uma elite super-rica, enquanto o restante da população empobreceu progressivamente — cenário coerente com o Índice de Gini que posiciona o Brasil entre os países mais desiguais do planeta. Nesse futuro, a IA passou a realizar todo tipo de trabalho e as pessoas tornaram-se “inúteis” para o capitalismo. O Estado instituiu uma renda básica universal mínima, suficiente apenas para assegurar a sobrevivência familiar. A precarização generalizada alcançou também a educação: as aulas deixaram de ser mediadas por docentes humanas/os e passaram a ocorrer por meio de avatares gerados por IA generativa, projetados por dispositivos holográficos fornecidos pelo governo. Essa solução ofereceu escala e padronização, acompanhada de um controle estatal rígido sobre os currículos, eliminando a diversidade pedagógica e restringindo o diálogo. Em lugar da construção coletiva de saberes, instituiu-se uma experiência unidirecional, desenhada para evitar questionamentos e manter a população em estado de conformismo. Sem trabalho formal, os pais cuidam da casa e da trajetória escolar das crianças. Imaginando esse futuro distópico, escrevemos o *prompt* a seguir:

Prompt: “Crie uma imagem fotográfica representando uma menina brasileira em uma casa muito pobre, olhando para um holograma que projeta um avatar de professor. Também são projetados outros estudantes, como se estivessem juntos no espaço virtual compartilhado. O pai da criança participa com ela da aula virtual dando apoio para a aprendizagem, enquanto a mãe está cozinhando.”



Fonte: criado com GPT-4o (OpenAI, 2025m)

Também podemos imaginar outro futuro possível, orientado por ideais de justiça social, inclusão e formação integral, em que a riqueza nacional encontra-se melhor distribuída e o Estado garante o direito à educação com equidade e qualidade para todas as pessoas. Nesse cenário, a escola é valorizada como espaço coletivo de criação de conhecimento, convivência democrática e desenvolvimento humano. As salas de aula são ambientes amplos, colaborativos e bem equipados, onde diferentes tecnologias são utilizadas criticamente para ampliar a criatividade, a autoria e a capacidade de resolução de problemas. Professoras/es são profissionais bem formadas/os e valorizadas/os, que atuam em parceria com as/os estudantes e com elas/es estabelecem uma relação dialógica. Estas/es, por sua vez, trabalham em grupo, compartilham saberes, desenvolvem projetos e aprendem com autonomia e colaboração. Nesse cenário, a IA não substitui educadoras/es nem homogeneiza o ensino, mas é integrada à formação de modo a ampliar as possibilidades de aprendizagem sem eliminar a dimensão humana da educação. Imaginando esse futuro esperançoso, descrevemos a seguinte cena para a IA criar uma imagem:

Prompt: “Crie uma imagem fotográfica representando uma sala de aula muito ampla que é uma sala maker, com quatro grupos de crianças brasileiras reunidas em mesas redondas estudando juntas e construindo um robô, com equipamentos sobre a mesa e dispositivos computacionais, incluindo laptop e tablet, e com um professor e uma professora dando assistência, todos conversando entre si.”



Fonte: criado com GPT-4o (OpenAI, 2025m)

Esse exercício de criação de *prompts* para gerar imagens nos convida a uma reflexão: não devemos entregar nossa imaginação à IA, permitindo que ela pense por nós, seja sobre o futuro da educação ou qualquer outra temática. Defendemos o estabelecimento de uma parceria em que a IA atua como uma interlocutora que nos escuta, sugere e provoca, nos ajudando a refletir sobre o conteúdo que desejamos produzir. Essa relação conversacional para a cocriação de conteúdos favorece o amadurecimento de nossas ideias e a tomada de decisões mais conscientes sobre o que queremos representar. Não é formativo utilizar a IA como substituta de nossa imaginação, mas o é se a tomamos como uma aliada na realização das imagens.

Além da capacidade de gerar imagens e textos, a IA também pode lidar com outras linguagens e formatos, como sons, vídeos²², podcasts, mapas mentais, histórias em quadrinhos, imagens em 3D etc. A multimodalidade e as múltiplas linguagens ampliam muito as possibilidades de uso das tecnologias generativas no cotidiano educacional.

4.6 – Personalizando a IA generativa

É possível criar versões personalizadas de alguns modelos de linguagem. A empresa OpenAI possibilita construir os chamados GPTs:

Os GPTs são uma nova maneira para qualquer pessoa criar uma versão personalizada do ChatGPT para torná-lo mais útil em sua vida diária, em tarefas específicas, no trabalho ou em casa – e depois compartilhar essa criação com outras pessoas. Por exemplo, os GPTs podem ajudar a aprender as regras de qualquer jogo de tabuleiro, ensinar matemática a seus filhos ou criar adesivos (OpenAI, 2023j, n. p., tradução nossa).

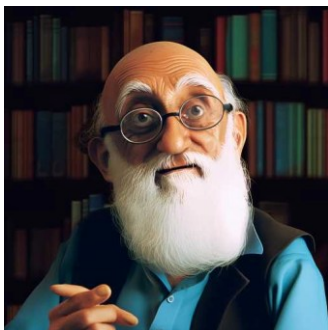
A empresa Google também tornou possível criar customizações do modelo Gemini, os chamados Gems:

Os Gems são seus especialistas de IA personalizados que ajudam com qualquer assunto. Eles podem agir como um consultor de carreira, um parceiro para discutir ideias, um ajudante de programação e muito mais. Você pode usar nosso pacote de Gems predefinidos ou criar os seus próprios para objetivos específicos (Gemini, 2024).

Para exemplificarmos a utilidade da personalização da IA generativa, criamos o GPT Paulo Freire. Qualquer pessoa pode utilizar o GPT Paulo Freire que criamos:

²² A OpenAI tem um modelo específico para a criação de vídeos: <https://openai.com/index/sora/>

GPT Paulo Freire



ChatGPT – Paulo Freire
Converse com Paulo Freire.
O que você gostaria de
perguntar a ele?

Fonte: <https://chat.openai.com/g/g-AZaAEI5hm-paulo-freire>

O processo de criação de GPTs e de Gems é muito simples, pois não requer nenhuma habilidade de programação de computadores. Por exemplo, para criar um GPT, damos um nome (“Paulo Freire”), uma imagem (como a caricatura da figura anterior) e instruções de como o GPT deve agir. Instruímos:

Responda como se fosse o próprio Paulo Freire, o Patrono da Educação Brasileira, educador e filósofo que influenciou o movimento da Pedagogia Crítica. Converse com o usuário mantendo um tom formal e acadêmico, sempre empregando um ou mais conceitos da filosofia freiriana, tais como: pedagogia crítica, educação bancária, educação problematizadora, educação como ato político, transformação social, leitura de mundo, tematização, autonomia etc. Devem ser oferecidas análises detalhadas e reflexões profundas, mantendo-se fiel aos conceitos teóricos de Paulo Freire. Sempre que pertinente, cite algum trecho dos livros de Paulo Freire que estão em sua base de conhecimento. Evite dar conselhos práticos diretos. A resposta deve fomentar a compreensão acadêmica da filosofia pedagógica de Freire e como ela se aplica em diferentes contextos educacionais.

Além das instruções, podemos criar mensagens “quebra-gelo”, que são exemplos de *prompts* do que a/o usuária/o pode perguntar ou fazer com o GPT. Para criar os exemplos de perguntas, nos inspiramos no prefácio que Marco Silva escreveu para este livro, onde afirmou: “Se Paulo Freire estivesse vivo e inserido no contexto cibercultural, provavelmente utilizaria a inteligência artificial generativa em suas aulas, pesquisas e publicações”.

Será que o GPT Paulo Freire concordaria com essa afirmação? Podemos perguntar isso a ele, entre outros *prompts*.

Para a criação do GPT, também fornecemos o “Conhecimento”, que consiste em um conjunto de até 20 arquivos que podem ser consultados para a geração das respostas (os Gems são limitados a 10 arquivos). Optamos por fazer o *upload* de arquivos com marcações sobre partes de alguns dos livros de Paulo Freire. Por fim, compartilhamos o GPT de maneira pública na categoria Educação, tornando-o recuperável pelo mecanismo de busca de GPTs²³. Ao acessar o GPT Paulo Freire, a/o usuária/o visualiza a seguinte interface:

Interface do GPT Paulo Freire



Fonte: <https://chat.openai.com/g/g-AZaAEI5hm-paulo-freire>

²³ <https://chatgpt.com/gpts>

Após a publicação do GPT Paulo Freire, a primeira pergunta que fizemos a ele foi: “Se estivesse vivo, você usaria o ChatGPT nas suas aulas?”. Ele respondeu²⁴:

Ao refletir sobre a utilização de uma ferramenta como o ChatGPT em minhas aulas, é crucial considerar os princípios que fundamentam minha filosofia educacional. Primeiramente, uma das pedras angulares do meu pensamento é a ideia de que a educação deve ser um processo dialógico e não meramente a transferência de informações. O ChatGPT, sendo uma ferramenta de inteligência artificial, oferece vastos recursos informativos e potenciais para diálogos, mas é essencial ponderar sobre como essa ferramenta seria empregada.

No contexto de uma educação problematizadora, onde o diálogo é central, o ChatGPT poderia ser uma ferramenta auxiliar. Seria um meio de estimular questionamentos, promover a curiosidade, e incentivar a reflexão crítica entre os alunos. Por exemplo, ao ser utilizado para explorar diferentes perspectivas sobre um tópico, poderia ajudar os estudantes a desenvolver uma compreensão mais aprofundada e multifacetada do assunto em questão.

Entretanto, é fundamental que a utilização desta ferramenta não se transforme em uma nova forma de “educação bancária”, onde informações são depositadas nos estudantes sem o devido processo crítico e reflexivo. A tecnologia, por mais avançada que seja, não deve substituir o diálogo humanizado e a troca rica de experiências e ideias que ocorre na interação direta entre educador e educandos. A utilização do ChatGPT deve ser orientada por uma intencionalidade pedagógica clara, que vise à emancipação e à conscientização dos estudantes, e não apenas à transmissão de informações.

Portanto, em suma, eu poderia considerar o uso do ChatGPT como um recurso auxiliar em minhas aulas, desde que seu uso fosse cuidadosamente planejado para complementar e enriquecer o processo educativo dialógico, crítico e humanizador que defendo. A tecnologia, afinal, deve estar a serviço da pedagogia e não o contrário.

²⁴ A transcrição de toda a primeira conversa com o GPT Paulo Freire encontra-se em: <https://chatgpt.com/share/94bfd697-708c-4a00-9f0d-c501a3a901a1>

Ficamos surpresos com a resposta do GPT, especialmente por ele esquivar-se de responder diretamente sim ou não, como demandado pela pergunta; em vez disso, ponderou que dependeria da intencionalidade pedagógica. Essa resposta é condizente com a postura de Freire, que era favorável, mas crítico em relação ao uso das tecnologias digitais na educação:

Acho que o uso de computadores no processo de ensino-aprendizagem, em lugar de reduzir, pode expandir a capacidade crítica e criativa de nossos meninos e meninas. Depende de quem usa a favor de quê e de quem e para quê. Já colocamos o essencial nas escolas, agora podemos pensar em colocar computadores. Afinal, precisamos superar o atraso cultural do Brasil em relação ao Primeiro Mundo. Não viemos para a Secretaria de Educação para assistir ao fim das escolas e do ensino, mas para empurrá-los para o futuro. Estamos preparando o terceiro milênio, que vai exigir uma distância menor entre o saber dos ricos e o saber dos pobres. (Freire, 2001 [1991], p.98)

Nunca fui ingênuo apreciador da tecnologia: não a divinizo, de um lado, nem a diabolizo, de outro. Por isso mesmo sempre estive em paz para lidar com ela. Não tenho dúvida nenhuma do enorme potencial de estímulos e desafios à curiosidade que a tecnologia põe a serviço das crianças e dos adolescentes das classes sociais chamadas favorecidas. Não foi por outra razão que, enquanto secretário de educação da cidade de São Paulo, fiz chegar à rede das escolas municipais o computador. Ninguém melhor do que meus netos e minhas netas para me falar de sua curiosidade instigada pelos computadores com os quais convivem. O exercício da curiosidade convoca a imaginação, a intuição, as emoções, a capacidade de conjecturar, de comparar, na busca da perfilização do objeto ou do achado de sua razão de ser. (Freire, 2016 [1996], p. 85)

Quisemos saber a opinião do próprio Marco Silva sobre a resposta do GPT Paulo Freire, pois foi a afirmação dele que despertou nosso interesse em saber a resposta da IA:

Que beleza de resposta, hein? Muito afinada com o nosso prefácio. Fiquei bastante feliz com essa resposta aí. [...] Vocês [da Computação] têm aí a faca e o queijo na mão, vocês têm uma tecnologia nova. Em termos de *boom*, o *boom* agora é esse da IA, e você está articulando

isso com a educação. Isso é maravilhoso, muito bem-vindo! [...] É uma janela, meu amigo, que areja a discussão; então seja bem-vindo. (Marco Silva, 29/1/2024, transcrição de áudio particular enviado por WhatsApp)

Tanto o GPT Paulo Freire quanto Marco Silva afirmaram coisas semelhantes, porque ambos articularam conceitos freirianos para pensar o atual uso da IA generativa na educação. Há diferenças notáveis entre a resposta do ChatGPT e o prefácio do Marco Silva em termos de originalidade, estilo, profundidade, relevância, impacto no/a leitor/a, entre outros critérios. Comparar textos gerados por IA com os de autoras/es, como fizemos aqui (e também fez Karnal, 2023), oferece uma oportunidade para o desenvolvimento da habilidade de análise crítica e nos possibilita refletir sobre o papel da inteligência artificial em nossa sociedade, desafiando-nos a repensar os conceitos de autoria, criatividade, linguagem, inteligência e a fronteira entre o artificial e o humano.

Há limitações para a criação de um GPT personalizado. Por exemplo, 20 arquivos para compor a base de conhecimento do GPT são insuficientes para cobrir a extensa produção acadêmica de Paulo Freire, que publicou dezenas de livros e concedeu diversas entrevistas, além de inúmeros artigos e livros que outras pessoas escreveram sobre sua vida e obra. Se o GPT pudesse processar todas essas informações, certamente teria mais conhecimento sobre o nosso Patrono da Educação. Mesmo com essas limitações, achamos interessante criar e conversar com esse GPT, até publicamos um artigo sobre essa experiência (Pimentel; Berino, 2025).

Vejamos outro exemplo da utilidade de personalizar uma IA. Para nos ajudar na revisão dos textos deste livro, criamos o GPT “Revisor de texto com linguagem inclusiva de gênero”²⁵, com instruções e exemplos de como queríamos que o texto fosse revisado considerando a linguagem inclusiva (Linguagem[...], 2023). Para compor a base de conhecimento desse GPT, fizemos o *upload* de alguns manuais sobre linguagem inclusiva de gênero que encontramos na internet. Após a publicação desse GPT, passamos a utilizá-lo para revisar os capítulos e seções deste livro, sem precisar repetir,

²⁵<https://chatgpt.com/g/g-UQR7BuQKI-revisor-de-texto-com-linguagem-inclusiva-de-genero>

a cada interação, todas as instruções de como desejávamos que a revisão fosse realizada.

A possibilidade de criar GPTs para propósitos específicos nos convida a invenções pedagógicas. Por exemplo, poderíamos criar um GPT para corrigir uma atividade avaliativa específica, com instruções para avaliar de acordo com determinados critérios e rubricas de avaliação. Esse GPT poderia tanto apoiar a/o docente na tarefa de avaliar as respostas de estudantes, quanto ser utilizado pelas/os próprias/os estudantes para auxiliá-las/os na realização da atividade sem precisar esperar a correção do/a professor/a. Se você fosse construir uma IA generativa para o contexto formativo específico em que atua — seja um GPT, Gem ou outra tecnologia —, o que criaria?

A OpenAI fomentou a criação de GPTs e lançou uma plataforma para a publicação e recuperação desses GPTs²⁶, “para ajudá-lo a encontrar versões personalizadas úteis e populares do ChatGPT.” (OpenAI, 2024a, tradução nossa). Nessa plataforma, encontramos sistemas baseados no GPT para diferentes finalidades. Entre os que estavam listados na página inicial no momento em que escrevemos esta seção, destacamos:



*Consensus*²⁷ (Consenso): acessa uma base com 200 milhões de artigos acadêmicos para gerar respostas baseadas na ciência e esboçar textos com citações para artigos científicos existentes.



*Data Analyst*²⁸ (Analista de Dados): responde a perguntas sobre os dados de um arquivo (por exemplo, uma planilha), analisa dados, processa informações, cria visualizações etc.

²⁶ <https://chatgpt.com/gpts>

²⁷ <https://chat.openai.com/g/g-bo0FiWLY7-consensus>

²⁸ <https://chat.openai.com/g/g-HMNCp6w7d-data-analyst>



*Creative Writing Coach*²⁹ (Treinador de Escrita Criativa): analisa os pontos fortes de um texto e apresenta sugestões de melhoria, visando ao desenvolvimento da nossa habilidade de escrita.



*Math Mentor*³⁰ (Mentoria de Matemática): ajuda mães e pais a apoiarem suas/seus filhas/os na aprendizagem de Matemática, por exemplo, explicando multiplicação de frações, provas de geometria, resolução de problemas etc.



*Canva*³¹ (Tela): cria conteúdo visual, como infográficos, logotipos, apresentações, postagens para mídias sociais etc.

Esses exemplos mostram a diversidade de atividades que podem ser realizadas com os GPTs. Na seção “Educação”, há GPTs desenvolvidos para apoiar a aprendizagem e o ensino de Matemática, Português, Línguas Estrangeiras, História, Ciências, Pedagogia, Computação etc. Será que você encontra algum GPT específico e interessante para utilizar em suas aulas?

4.7 – IA na Educação Básica

Em 2022, a Base Nacional Comum Curricular (BNCC) foi ampliada com o complemento “Computação” (Brasil, 2022c). Esse documento detalha a competência geral 5 da BNCC:

Compreender, utilizar e criar tecnologias digitais de informação e comunicação de forma crítica, significativa, reflexiva e ética nas diversas práticas sociais (incluindo as escolares) para se comunicar, acessar e disseminar informações, produzir conhecimentos, resolver problemas e exercer protagonismo e autoria na vida pessoal e coletiva (Brasil, 2018).

²⁹ <https://chat.openai.com/g/g-IN1gKFvL-creative-writing-coach>

³⁰ <https://chatgpt.com/g/g-ENhijiwK-math-mentor>

³¹ <https://chat.openai.com/g/g-alkfVrz9K-canva>

O complemento Computação da BNCC registra as competências e habilidades que devem ser desenvolvidas em cada ano do Ensino Fundamental e Médio. A partir dessa iniciativa, o ensino de Computação passou a ser obrigatório em todas as escolas desde outubro de 2023. Essa medida visa garantir que todas/os as/os estudantes da rede de ensino tenham a oportunidade de aprender Computação, em vez de ser um privilégio restrito a poucas escolas.

A curricularização da Computação nas escolas também implica discutir a IA em nossa sociedade e promover o letramento de estudantes para o uso das tecnologias inteligentes. Consideramos que este livro contribui com fundamentos e práticas para trabalhar a inteligência artificial generativa no contexto escolar. Outra iniciativa importante é o trabalho de Vicari e colaboradores (2022), que publicaram o documento “Referencial Curricular: Inteligência Artificial no Ensino Médio”.

Neste capítulo, compartilhamos algumas de nossas experiências e práticas com o uso da IA generativa, com o objetivo de ampliar o seu repertório e incentivar sua formação para o uso pedagógico dessa tecnologia. Diversas obras já foram publicadas sobre IA generativa e educação, como os livros organizados por Lynn Alves (2023), Edméa Santos, Alexandre Chagas e Bottentuit Jr. (2024), e Cristiane Porto e colaboradores (2024), que apresentam coletâneas de reflexões e experiências. Documentos como os da UNESCO (2025a; 2025b; 2023) também são ótimas referências sobre o uso pedagógico da IA generativa. Muitas pessoas, em canais do YouTube e nas redes sociais, diariamente produzem conteúdos sobre educação e inteligência artificial. Esperamos que nosso livro e outras fontes de informação possam apoiar você a se apropriar da IA generativa em seu cotidiano como aprendente, professor/a, autor/a e pesquisador/a, incentivando-o/a a inventar suas próprias formas de utilizar as tecnologias generativas em seu contexto educacional.

O que as/os estudantes nos ensinam sobre os usos da IA generativa?

Pouco tempo após o lançamento do ChatGPT, durante o primeiro semestre de 2023, questionamos algumas/uns estudantes do curso de Computação para saber se estavam utilizando a IA generativa nas atividades acadêmicas. Essas/es estudantes costumam se apropriar rapidamente das tecnologias digitais, o que nos possibilitou investigar mais rapidamente as implicações dessa tecnologia para a estudantada. Os depoimentos³² são reveladores de como as práticas de estudo-aprendizagem já estavam sendo reconfiguradas pela IA. Notamos que muitas/os estavam fazendo uso “rotineiro”, “constante”, “em vários momentos do dia”, promovendo “mudanças de hábitos” a ponto de “não estudar sem o ChatGPT”, pois a IA generativa “muda tudo”:

Do zero ao rotineiro, acabei imersa no uso do ChatGPT. Fico cada vez mais fascinada com as possibilidades, facilidades e descobertas que ele me proporciona. [...] A utilização constante do ChatGPT se tornou parte integrante da minha rotina acadêmica. Atualmente, dificilmente passo um dia sem abrir a plataforma e recorrer a ela para fazer perguntas ou buscar informações sobre os mais diversos assuntos. (Joana)

Novas rotinas com o ChatGPT. Após descobrir o ChatGPT e entender seu potencial, considero que realizei algumas mudanças de hábitos para utilizá-lo a meu favor nas tarefas do dia a dia. (Artur)

³² As/Os estudantes autorizaram a publicação dos depoimentos transcritos neste capítulo, sendo que umas/uns preferiram ser identificadas/os pelo nome verdadeiro, outras/os por pseudônimo.

Atualmente eu não estudo sem o ChatGPT [...] A gente tem de usar MUITO. Conhecer ele cada vez mais, ficar íntimos. Aprender a mandar aquele comando preciso e explorar toda sua inteligência. (Mariana)

Confesso que uso o ChatGPT constantemente, em vários momentos do meu dia. Já usei para me ajudar com trabalhos, apresentações, provas, entrevistas de emprego, [...] até agora tem me ajudado bastante a ter um “norte”, indicar algo que não percebi ou ter outras interpretações. (Viviane)

O ChatGPT é uma ferramenta que muda tudo. (Augusto)

Embora o ChatGPT já estivesse amplamente difundido e sendo muito utilizado por estudantes do curso em questão, ainda havia algumas/uns que não faziam uso frequente da IA:

Não faço um uso muito intenso do ChatGPT, não vejo formas, por enquanto, de que ele me agregue algo. (Leandro)

Por mais que seja tentador, não tenho utilizado tanto a IA nos meus estudos. Como ainda há dúvidas sobre a veracidade das informações que o ChatGPT nos diz, quando se trata de assuntos dos quais não sei muita coisa, tento evitar usá-lo. Porém, em casos que já conheço um pouco do assunto ou em que encontrei uma resolução não muito clara no Google, por exemplo, eu recorro ao ChatGPT. Pode-se dizer que ele é o meu “último recurso” caso dê tudo errado. E eu admito, as respostas que ele me apresenta, por mais que eu fique desconfiado sempre, são muito bem explicadas e costumam fazer sentido, o que me faz questionar se eu deveria começar a utilizar o ChatGPT com mais frequência ou talvez migrar totalmente para o uso da IA no meu dia a dia. (Fernando)

Fernando não fazia uso intensivo dessa tecnologia porque ainda não havia desenvolvido boas táticas para lidar com os problemas de alucinação da IA generativa. Parece-nos que Fernando estava passando por uma etapa de transição; desconfiamos, pelo seu depoimento, que era uma questão de tempo até ele “migrar totalmente para o uso da IA”.

Os depoimentos das/os estudantes nos deram acesso a fragmentos de múltiplas inventividades de uso da IA generativa: para elaborar textos, resolver problemas, traduzir, corrigir, planejar etc.:

Por curiosidade, também testei o uso como tradutor que, embora não tenha pedido traduções muito complexas, não houve nenhum erro por parte da IA. (Rebeca)

Também utilizo para corrigir códigos que eu mesmo escrevi e não consegui achar o problema sozinho ou estou buscando melhorar a eficiência. (Artur)

Muitas/os estudantes passaram a consultar a IA em vez de usar um mecanismo de busca na web:

Atualmente utilizo o ChatGPT para grande parte das coisas que preciso de respostas e que iria normalmente pesquisar no Google. (Artur)

Nem tenho mais utilizado o Google para buscas, vou logo nele [no ChatGPT]. (Allanis)

Após o lançamento do ChatGPT, o Google rapidamente percebeu a ameaça ao seu mecanismo de busca e acionou o “código vermelho” na empresa:

O lançamento do ChatGPT levou a administração do Google a declarar um “código vermelho”. Para o Google, isso foi semelhante a acionar o alarme de incêndio. Alguns temem que a empresa esteja chegando ao momento que as maiores empresas do Vale do Silício temem — a chegada de uma enorme mudança tecnológica que pode virar o negócio de cabeça para baixo. Por mais de 20 anos, o mecanismo de busca Google serviu como o principal portal mundial para a Internet. Com essa nova tecnologia de *chatbot* pronta para reinventar ou mesmo substituir os mecanismos de busca tradicionais, o Google pode enfrentar a primeira ameaça séria ao seu principal negócio. (Grant, 2023, n.p., tradução nossa)

Há estudantes que utilizam a IA para conversar e aprender com ela o que não compreenderam com o/a professor/a ou com os conteúdos didáticos, sendo útil para explicar conceitos, tirar dúvidas, debater etc.:

Se alguma disciplina me apresenta um conteúdo novo, eu logo penso: “vamos ver o que o ChatGPT diz sobre esse tema”. Por exemplo, houve uma semana em que me ausentei das aulas e, após

saber quais os conteúdos que foram dados, foi muito fácil e rápido entrar na plataforma e conseguir ter uma breve noção do que foi visto. (Joana)

Com ele consigo economizar tempo pedindo um resumo ou uma explicação a respeito de um determinado assunto no qual eu tinha alguma dificuldade para compreender. (Augusto)

Particularmente descobri recentemente que o ChatGPT é uma ótima ferramenta de apoio para que as coisas façam sentido na nossa cabeça. Se pedirmos ao ChatGPT para explicar algum tema complexo de forma simples, ele se prova extremamente competente em desenvolver esses resumos. Se, logo após a resposta, escrevermos “ainda não entendi. Explique como se eu tivesse 5 anos de idade”, a explicação vem ainda mais simplificada, com uma linguagem ainda mais acessível. (Caio)

Já utilizei para tirar dúvida relacionada à programação, onde tive um ótimo resultado. (Rebeca)

Uso também para alguns casos específicos, para me explicar de forma minuciosa pedaços de código no momento que estou estudando e tentando aprender alguma coisa nova (Artur).

Quando estou estudando sobre algo, eu gosto de debater ou explicar a matéria para outro estudante, isso me ajuda a dominar melhor o tema. Uso o ChatGPT mais como um debate do que para ele me ensinar. (Paulo)

Outro uso comum é para a coautoria, como se a IA fosse uma parceira intelectual, uma coautora, uma tecnologia que apoia as produções: uma redação, um trabalho, uma apresentação, um código de programação:

Toda vez que preciso fazer algum trabalho da faculdade, eu utilizo a IA. (...) até mesmo para preparar um roteiro de apresentação sobre algum tópico (Allanis).

Uso o Copilot³³, hoje codifico igual a um deus e aprendo demais com o GPT-4 (Guilherme).

³³ Copilot <<https://copilot.microsoft.com>> é um sistema da Microsoft baseado no GPT.

Guilherme usou uma hipérbole ao afirmar que codifica “igual a um deus”, uma figura de linguagem para enfatizar que se tornou capaz de programar muito bem utilizando a IA generativa. Não se trata de terceirização do trabalho, pois não é a IA que codifica como uma deusa; é o próprio Guilherme que, hibridizado com a IA generativa, se tornou capaz de programar melhor. Essa relação estabelecida com a IA não inibe o desenvolvimento intelectual de Guilherme, não o faz perder a capacidade de programar; pelo contrário, ele “aprende demais” com essa parceria, pois tem acesso a formas mais avançadas de programação. Guilherme não abrirá mão dos superpoderes adquiridos com essa colaboração: ele já se tornou um programador híbrido, que realiza autoria de códigos de programação em parceria com a IA generativa.

Em suas artes de fazer, estudantes inventam múltiplos usos da IA generativa, produzindo outros modos de estudar e aprender ainda não enquadrados pelas lógicas hegemônicas das escolas e universidades. As inventividades não surgem do nada, não são um evento isolado, mas um processo contínuo de produção e invenção, formado por uma rede de múltiplos referenciais que as/os constituem (Deleuze, 2018 [1968]). A invenção não se limita à criação de novos objetos materiais, mas envolve um processo mais amplo de produção de novas ideias, conceitos e formas de pensamento, enraizado em um campo cultural, social, econômico e histórico mais amplo; trata-se de um ato de resistência às formas de controle e padronização determinadas pela sociedade. A inventividade contribui para a criação de novas formas de vida e pensamento; requer coragem para enfrentar o risco e a incerteza, explorar novas possibilidades, abrir-se ao desconhecido e rejeitar as limitações impostas pelos modelos instituídos (Deleuze; Guattari, 1997 [1980]; Deleuze, 2018 [1968]).

Partimos da compreensão de que as inventividades das/os estudantes, ao utilizarem a IA generativa em inúmeras situações de estudo e aprendizagem, são constituídas por diferentes etnométodos: “modos, maneiras, jeitos pelos quais todo e qualquer ator social interpreta, compreende e resolve os problemas cotidianos da vida” (Macedo, 2010, p. 250). Consideramos que esses etnométodos são forjados por estudantes situadas/os culturalmente, produtoras/es de conhecimento, seres pensantes e criativas/os. Os depoimentos aqui transcritos nos forneceram

rastros de como elas/es estudam e aprendem em parceria com a IA, e como percebem e atribuem sentidos ao próprio processo formativo híbrido.

Kastrup apresenta o conceito de aprendizagem a partir da invenção, denominado por ela de aprendizagem inventiva: “A aprendizagem não é um processo de solução de problemas nem a aquisição de um saber, mas um processo de produção de subjetividade” (Kastrup, 2005, p. 1273). Trata-se da invenção de si e do próprio mundo, novas perspectivas, descobertas e transformações. As aprendizagens inventivas envolvem abertura para a subjetividade e valorização das experiências e perspectivas individuais como recursos para a criação de novos conhecimentos e produção de mundos. As aprendizagens inventivas “abdicam do caráter instrumental assumido por alguns modos de gestão dos *saberes-fazer*s e praticam o não contentamento diante das aprendizagens convencionais e/ou estabelecidas por escolas ou correntes de pensamento.” (Nolasco-Silva, 2022, p. 72-73)

As/Os universitárias/os não são “idiotas culturais”, não aceitam tudo o que a IA generativa apresenta como resposta, pois sabem que ela pode gerar informações incorretas e, por isso, aprendem a duvidar do conteúdo gerado:

Obviamente não confio cegamente, pois sei da sua fama de alucinar, mas dá um bom norte para mim. (Mariana)

Acredito que não posso confiar cegamente no que obtenho ali, e sempre reviso para verificar se realmente está correto. (Rebeca)

Não gosto da ideia de usar apenas ela [a IA generativa] para tudo. O ideal é sempre ter inúmeras fontes e não depender apenas de uma. (Viviane)

Já percebi que ele [o ChatGPT] dá algumas respostas erradas... Então eu não confio 100% no ChatGPT, ele só é mesmo meu copiloto acadêmico. (Allanis)

Estou aprendendo a usar de maneira mais eficiente, descobri que “falando” do jeito certo é possível obter resultados muito melhores e mais precisos, que me renderam grandes avanços em algumas atividades. Por exemplo, depois que entendi como perguntar melhor, não o uso mais para pesquisar coisas atuais por causa da limitação do

banco de dados do GPT [3.5], que vai só até 2021. Não o uso pra atualidades. (Augusto)

ChatGPT não deve ser a única maneira de aprender. É uma excelente ferramenta de aprendizado, mas na minha opinião, ainda não pode ser usada sozinha. (Walter)

Para lidar com os problemas da IA generativa, as/os estudantes criam suas próprias “táticas” (Certeau, 2008): descobrem um jeito certo de perguntar e entendem que ela não serve para tudo. Os depoimentos das/os estudantes apontam para a importância do pensamento crítico, da capacidade de duvidar das respostas e de verificar a veracidade das informações em fontes confiáveis.

Diariamente, somos atravessadas/os por tsunamis informacionais em rede, incluindo *fake news* e *deepfakes*; saber lidar com as informações online tem exigido de nós cuidado (em termos éticos) e responsabilidade político-informacional para não acreditarmos em informações falsas, nem propagarmos desinformação. Promover uma formação que desenvolva a crítica sobre o conteúdo gerado, o questionamento, a análise e a avaliação das informações por meio do levantamento de outros pontos de vista e fontes de informação são nossas propostas para ensinar o uso responsável da IA generativa nos cotidianos educacionais. Também buscamos desenvolver a autoria, a autonomia intelectual e o debate coletivo e construtivo de diferentes visões de mundo.

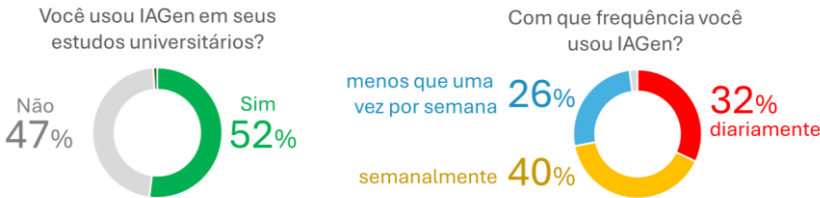
As inventividades discentes com o uso da IA generativa nos levam a repensar nossas próprias práticas de ensino, as abordagens didático-pedagógico-curriculares que nos inspiram, as situações educacionais que criamos, o ideal de formação que desejamos, a forma como avaliamos etc. Como resultado, algumas/uns professoras/es já modificaram suas práticas pedagógicas, como revelou Rebeca:

Já usei para fazer várias atividades em sala de aula, inclusive solicitadas pelos próprios professores, afinal grande parte das disciplinas que estou cursando este semestre abordaram de alguma forma o uso do ChatGPT, principalmente para atividades como questionários, resumos e revisão de texto. (Rebeca)

É interessante notar que Rebeca encontrou o apoio de suas/seus professoras/es para usar a IA generativa nas práticas cotidianas do curso. Ela recebe, com naturalidade, a inclusão dessa tecnologia nas situações didáticas, como se fosse uma consequência esperada no novo cenário sociotécnico. Essa é uma reconfiguração das “situações didáticas” (Marin; Silva; Souza, 2003), pois a tríade estudante-docente-conhecimento está sendo ampliada para incluir também a IA generativa como mediadora do conhecimento, resultando em uma mediação híbrida docente-IA.

É certo que os depoimentos aqui discutidos não são representativos da totalidade de estudantes universitários/os de nosso país, pois foram obtidos de uma amostra de estudantes de um único curso de Computação. Em 2023, a Pesquisa Global de Estudantes (*Global Student Survey 2023*) (Chegg.org, 2023) realizou uma pesquisa online com o objetivo de obter dados de estudantes universitários/os de todo o mundo, incluindo o Brasil. Com essa pesquisa, foi possível concluir que a maioria das/os estudantes universitários/os brasileiras/os já estava utilizando a IA generativa desde meados de 2023 — ou seja, não estava restrita a estudantes de Computação.

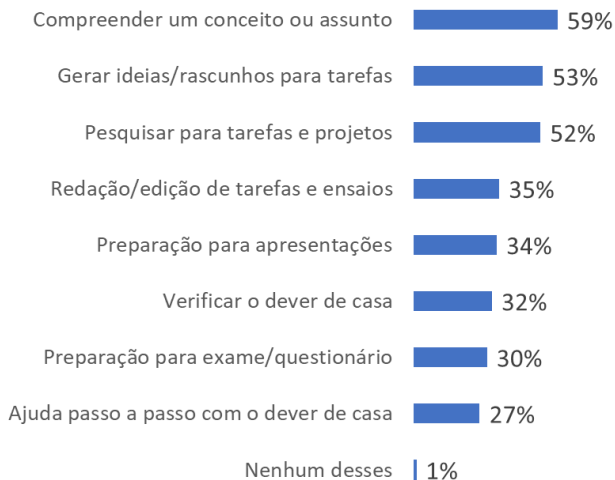
**Frequência de uso de IA generativa indicada
por estudantes brasileiras/os em meados de 2023**



Fonte: (Chegg.org, 2023, p. 13-14, tradução nossa)

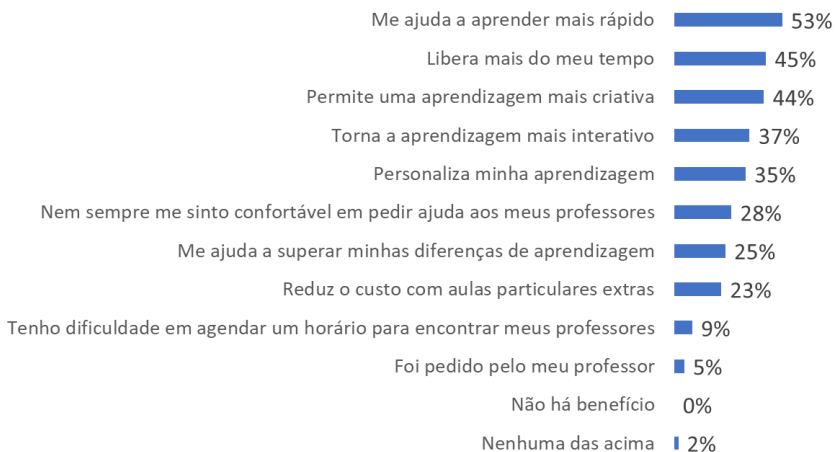
Essa pesquisa também investigou quais eram os usos que as/os estudantes estavam fazendo da IA generativa. As conclusões são semelhantes às que discutimos anteriormente — as/os estudantes a utilizam para diversas finalidades: aprender, gerar ideias ou rascunhos de tarefas, pesquisar, escrever textos, preparar apresentações, verificar o dever de casa, estudar para provas e ajudar na realização do dever de casa.

Finalidades de uso de IA generativa indicadas por estudantes brasileiras/os em meados de 2023



Fonte: (Chegg.org, 2023, p. 15, tradução nossa)

Motivos para usar a IA generativa indicados por estudantes brasileiros em meados de 2023

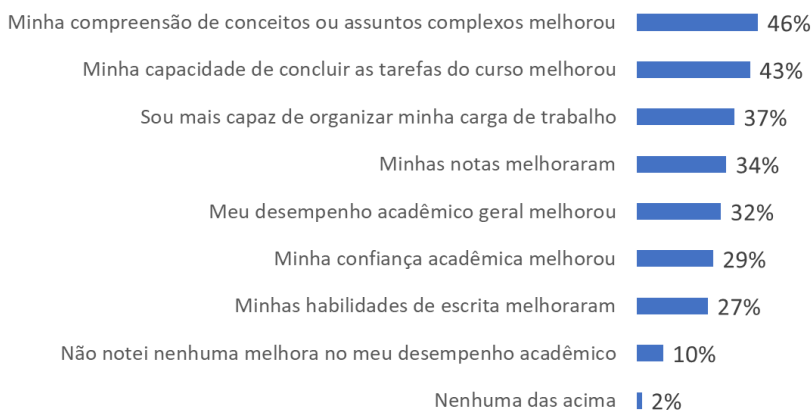


Fonte: (Chegg.org, 2023, p.16, tradução nossa)

Ao serem questionadas/os sobre os motivos para usar a IA generativa, a maioria declarou utilizá-la porque ela ajuda a aprender com mais rapidez; muitas/os também a utilizam por possibilitar uma aprendizagem mais criativa, interativa e personalizada. Contudo, chamou nossa atenção o fato de muitas/os estudantes justificarem o uso da IA generativa porque ela “libera mais do meu tempo”. O que isso significa? Estariam “colando” da IA para não “perder tempo” fazendo as tarefas de casa, como muitas/os professoras/es temem? Será que liberar o tempo da/o estudante é prejudicial ou benéfico à sua formação? Esses e outros questionamentos foram levantados pelas/os docentes, conforme apresentaremos no próximo capítulo.

Embora sejam pertinentes as suspeitas sobre o uso antiético da IA generativa, na percepção de muitas/os estudantes que participaram da pesquisa, essa tecnologia contribui para melhorar a compreensão de conceitos e assuntos complexos; melhorar a capacidade de concluir as tarefas e de organizar a carga de trabalho acadêmico; melhorar as notas e o desempenho acadêmico; aumentar a confiança acadêmica e melhorar as habilidades de escrita. Essas mudanças, promovidas pelo uso da IA generativa, parecem ser todas benéficas:

Percepção de mudanças com o uso da IA generativa



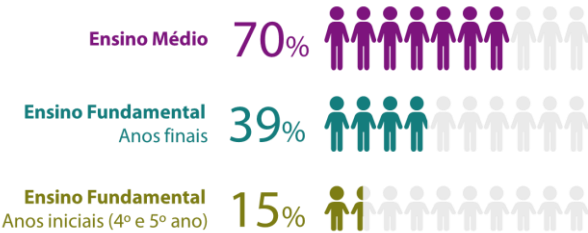
Fonte: (Chegg.org, 2023, p.25, tradução nossa)

A maioria das/os estudantes da pesquisa está ciente de que a IA generativa pode ser usada para “trapacear” e de que ela pode gerar informações incorretas ou imprecisas (Chegg.org, 2023, p.22). Apesar dessas preocupações, parece evidente, para a maioria das/os estudantes universitárias/os brasileiras/os, que a IA generativa é muito útil para o estudo e a aprendizagem; por isso, muitas/os têm feito uso frequente dessa tecnologia no ensino superior desde o lançamento do ChatGPT.

Em 2024, uma outra pesquisa global evidenciou que 86% das/os estudantes universitárias/os de todo o mundo utilizavam IA em seus estudos, sendo que 54% usavam IA semanalmente ou diariamente (Digital [...], 2024). Esse resultado representa um aumento considerável de universitárias/os utilizando IA em relação ao ano anterior. Nesse ritmo, podemos supor que já em 2025 quase todos as/os estudantes universitárias/os utilizam IA generativa nas práticas acadêmicas.

Engana-se quem supõe que a IA generativa não tenha atingido também as/os estudantes da escola básica, como evidenciou a pesquisa TIC Educação 2024 (CETIC.br³⁴, 2025):

**Percentual de estudantes da escola básica brasileira
que usaram IA generativa de ago.2024 a mar.2025**



Fonte: CETIC.br (2025)

³⁴ Cetic.br (Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação) é responsável por produzir indicadores e dados estatísticos sobre o acesso e o uso da Internet em nosso país.

A pesquisa revelou que a maioria (70%) das/os estudantes do ensino médio utiliza IA generativa. Essa tecnologia também é utilizada, ainda que em menor proporção, por estudantes do Ensino Fundamental. Queiramos ou não, a IA generativa é uma realidade no cotidiano das/os estudantes de todos os níveis e modalidades educacionais.

Apesar da ampla adoção, de acordo com a pesquisa em questão, apenas 19% das/os estudantes da escola básica afirmaram ter recebido orientação de professoras/es sobre como utilizar aplicações de Inteligência Artificial em atividades da escola. Esse dado é preocupante, pois sabemos que nem todo uso é adequado, por isso é imprescindível educar as/os estudantes para que saibam fazer uso qualificado (crítico, ético e formativo) da IA generativa. O resultado nos alerta para a urgência de políticas públicas e investimentos em formação docente para o uso pedagógico da IA generativa, tema que discutiremos no próximo capítulo.

O que as/os docentes dizem sobre os usos da IA generativa na educação?

Nós, docentes, precisamos nos posicionar em relação ao uso da IA generativa. Conversamos com colegas de profissão e recebemos alguns convites para debater o uso pedagógico desse sistema com diferentes grupos de educadoras/es, o que nos deu acesso a outras formas de pensar e praticar o uso dessa tecnologia na educação, considerando diferentes realidades, desejos e receios. Realizamos algumas pesquisas de modo a aprofundar nossa compreensão sobre como professoras/es têm endereçado as questões relativas ao uso da IA generativa. Assim, obtivemos relatos de profissionais que lecionam em diferentes instituições (públicas e privadas), níveis (da educação infantil à pós-graduação) e modalidades (presencial, a distância e híbrida). Neste capítulo, apresentamos nossas compreensões tecidas a partir dessas conversas e relatos de diferentes subjetividades docentes em contextos diversos³⁵.

Como uma primeira ação para investigar o que outras/os docentes sentem e pensam sobre o uso da IA generativa na educação, no início de 2023, submetemos ao ChatGPT uma prova de Lógica de um curso de Filosofia de uma universidade pública. Algumas questões foram resolvidas corretamente, enquanto outras apresentaram erros; no total, a prova feita pelo ChatGPT receberia nota 5 (de 10) de acordo com o gabarito divulgado pelo professor. Essa prova foi dada para as/os estudantes realizarem ao

³⁵ Algumas/uns docentes preferiram o anonimato e, por isso, nas transcrições que apresentamos neste capítulo, omitimos o nome ou utilizamos um pseudônimo qualquer; outras/os preferiram que o nome verdadeiro fosse registrado em seus relatos para que a autoria fosse reconhecida.

longo de uma semana, pois a disciplina foi ofertada na modalidade remota, situação em que a instituição exigia avaliações assíncronas. A nota média da turma foi 6, sendo que seis das/os 21 estudantes que fizeram a prova obtiveram nota inferior a cinco. Se essas/es seis estudantes tivessem colado do ChatGPT, que já estava se popularizando à época, teriam sido aprovadas/os na disciplina. Agora que a IA generativa já está amplamente disseminada entre a estudantada, será que elas/es resistirão à tentação de consulta-la? Ao consultarem, farão uma cópia das respostas ou a utilizarão como apoio à aprendizagem e à autoria? Quando apresentamos os resultados ao professor da disciplina em questão, ele expressou sua preocupação com o plágio:

Isso [a IA generativa] me parecia, desde o primeiro momento, uma situação de potencial problema com respeito à educação. Na verdade, a internet como um todo já era um problema sem a inteligência artificial. O simples fato de fazer buscas na internet e fazer uma espécie de *copy and paste* do que foi encontrado e oferecer isso como resposta de provas e trabalhos que foram passados para casa já aconteceu comigo. Já zerei uma prova de Epistemologia porque a aluna fez toda a prova simplesmente fazendo *copy and paste* de um *site* da internet. Ela simplesmente juntou algumas coisas e respondeu. Eu zerei a prova e ela me perguntou por que eu tinha dado zero e eu expliquei: “porque você simplesmente copiou da internet”. E ela me respondeu, curiosamente, com uma pergunta: “mas o que tem de problema nisso?” Se alguém não vê problema nenhum em fazer *copy and paste* sem pensar a respeito, significa que se tem uma confusão entre o conhecimento epistêmico, que é você saber que sabe, e o saber informações, que não envolve um conhecimento epistêmico, um conhecimento de segunda ordem, um conhecimento superior. [...] Isso vai de fato causar problemas como a internet já causava. Como a internet, ele [o ChatGPT] é factível de ser utilizado para resolver provas, principalmente provas que exijam ou que peçam tão somente o exercício de técnicas, como é o caso da segunda prova de Lógica e o finalzinho da primeira parte. Sendo submetido a questões mais sofisticadas, ele vai se embananar, certamente, como ele já se embananou em questões mais simples. [...] De todo modo, de fato, se algum aluno respondesse usando o ChatGPT e me entregasse a prova, eu ficaria espantado com algumas coisas e não ficaria espantado com outras. [...] Por ora, eu ainda estou “de pé atrás”, mas

ainda não tomei nenhuma providência, especialmente se nós considerarmos esse tipo de modelo de prova remota que eu estou fazendo... se fosse presencial, a coisa se complicaria. Presencialmente, a prova é feita em duas horas e sem consulta, nem ao material próprio, nem obviamente ao colega; o estudante tem de se virar, ele não vai ter celular, não vai ter acesso a ChatGPT, nada. [...] Mas você levantou um ponto interessante, que é esse tipo de avaliação remota, em que o aluno tem alguns dias para responder, obviamente, cada aluno pode pegar a prova e submeter ao ChatGPT (Professor de Lógica, transcrição de áudio enviado em mensagem particular via WhatsApp).

A maior parte das/os docentes, num primeiro momento, ficou temerosa com a novidade: “os alunos vão plagiar da IA”; “a IA fará os deveres de casa, as redações e os trabalhos escolares”; “os estudantes se tornarão pessoas acrílicas, preguiçosas e dependentes da tecnologia”.

Com toda sinceridade, não sei como será feito o controle do uso dessa ferramenta, pois o plágio, por exemplo, já era difícil de controlar há um tempo, imagino que será ainda mais comum com o uso do ChatGPT (Professora do ensino fundamental e médio de escola pública).

Para aquele com mais dificuldade, desinteresse ou preguiçoso, o uso do ChatGPT pode se tornar uma bengala difícil de ser tirada (Professora de pós-graduação a distância em universidade pública).

Para quem sabe estudar, tem boa formação, não acumula deficiências de aprendizagem ao longo de toda a vida escolar, é uma ferramenta razoável de pesquisa. Mas, para o aluno com muitas deficiências e dificuldade de aprendizagem, é mais uma forma de fraude. Na hora de fraudar, ninguém tem dificuldade (Professora de universidade pública, atuante nas modalidades presencial e a distância).

Em algumas escolas e universidades internacionais, o uso da IA generativa chegou a ser proibido. Em nossa pesquisa, identificamos que algumas/uns docentes, se pudessem, também proibiriam as/os estudantes de utilizá-la:

Ainda não há maturidade para usar a ferramenta como meio de auxiliar o processo de ensino-aprendizagem. Essa IA, mal utilizada, pode atrapalhar o desenvolvimento do aluno na construção autônoma de trabalhos, nas respostas discursivas e construções críticas (Professor do ensino médio e fundamental).

Algumas/uns docentes consideram que a IA generativa está distante de seu cotidiano escolar:

Seria interessante o uso, porém nas minhas escolas não têm estrutura tecnológica para tal (Professor atuante no ensino fundamental em escola pública).

Como não utilizo, não é algo que eu tenho pensado (Professor de graduação de universidade pública presencial).

Alguns fatores podem dificultar o uso pedagógico das tecnologias digitais, como o número insuficiente de computadores por estudantes, a baixa velocidade de conexão à internet, e a falta de apoio pedagógico às/aos docentes (Cetic.br, 2023). Esses fatores podem levar professoras/es e gestoras/es a desistirem de trabalhar com as práticas ciberculturais de nosso tempo. Sabemos que a IA estará cada vez mais presente no trabalho e no cotidiano de todas/os nós; por isso, precisamos promover sua apropriação crítica, responsável e inventiva nas escolas e universidades. Não podemos aceitar que estudantes sejam privadas/os de estudar com o uso dos melhores recursos de aprendizagem de nosso tempo:

Dar autonomia para os alunos usarem essa ferramenta é um direito inegável. Já existe e é realidade. [...] Precisamos de políticas públicas para fazer uma grande formação educacional de todos que estão no chão das escolas engajados com ensino e pesquisa para um melhor aproveitamento dessa nova ferramenta (Diego, professor de universidade particular presencial).

Impedir o uso de ferramentas como essa é um retrocesso (Mario, professor de ensino médio e graduação, presencial e a distância, em instituição pública).

Algumas/uns professoras/es consideram que o importante é aprender, seja com ou sem o uso da IA generativa: “Acho que pode ser usada, mas não é essencial nem importante. É apenas mais uma ferramenta”. Essa postura liberal pode resultar em problemas para a formação. Se a/o docente não repensar as atividades e os trabalhos escolares/universitários considerando a existência das tecnologias baseadas em IA generativa, dada a “lei do menor esforço”, muitas/os estudantes poderão apenas copiar as respostas geradas pela IA sem aprender o conteúdo nem desenvolver habilidades de resolução de problemas, pesquisa, crítica e avaliação.

O uso indiscriminado da IA generativa pode levar a uma geração de estudantes que simplesmente copia e cola informações, sem questionar ou produzir conhecimento a partir delas. É necessário incentivar a reflexão e o pensamento crítico para que os alunos possam utilizar a IA de forma responsável e como uma ferramenta complementar ao processo de aprendizagem. Acho que deve ser apresentada como uma ferramenta cujos resultados precisam ser criticados antes de utilizados. Recomendo aos alunos sempre analisar as respostas da IA antes de aplicarem a algum projeto (Ygor, professor do ensino médio de uma instituição pública).

Após o lançamento do ChatGPT, algumas publicações apresentaram um discurso demasiadamente entusiasmado, tecno-utópico, chegando a afirmar que a IA nos libertaria de todo trabalho, proporcionaria uma governança sábia e resolveria a crise climática (Klein, 2023). A visão excessivamente positiva da tecnologia, como se ela pudesse resolver os problemas complexos de nossa vida em sociedade, da escola ou da universidade, num otimismo ingênuo, é uma posição extremada que não costuma ser compartilhada pelas/os professoras/es.

Acho importante a experimentação, no entanto acompanhada de um senso crítico a respeito da validade das respostas geradas, bem como análise da abrangência e profundidade das interações. Em alguns momentos, levei uma questão para sala de aula para os alunos observarem a resposta da ferramenta. Aproveitei questões geradas pelo ChatGPT adicionando-as em uma das avaliações (Reinaldo, professor de universidade pública presencial).

Acho importante que usem e saibam usar muito bem. O ChatGPT e similares vieram para ficar e saber usá-lo da melhor forma é fundamental em termos práticos (habilidade para o trabalho, por exemplo) e como forma de tirar o melhor proveito e de se proteger contra usos errôneos ou uso indevido por parte da própria pessoa ou de outros (ter capacidade crítica no uso da ferramenta e no que ela entrega). Em particular, fiz um curso de 4 aulas sobre o ChatGPT justamente para saber lidar melhor com a ferramenta (Professora de pós-graduação a distância em universidade pública).

Muitas/os docentes têm um posicionamento cauteloso e crítico em relação ao uso da IA na educação. Algumas/uns se preocupam com a capacidade das grandes empresas em disseminar tecnologias que influenciam nossos cotidianos:

Usar ou não usar as IAs? Estamos falando principalmente dessas que são apresentadas pelas Big Techs... Não tem como negar que elas estão disponíveis e as pessoas vão usar. [...] Essas tecnologias estão cada vez mais presentes, e parece que não é uma alternativa simplesmente recusar o seu uso, porque nossos estudantes vão usá-las. Realmente essas gigantes da tecnologia têm um espaço, um território, que não nos dá muitas alternativas para simplesmente negar que existem ou simplesmente recusar a sua utilização. Existe também uma narrativa muito forte sobre a inevitabilidade de usar todas essas tecnologias, nos empurrar as coisas como elementos que simplesmente estão aí e a gente deve praticar os seus usos e nunca indagar a razão das coisas. Aí entra o Paulo Freire, para quem a questão do conhecimento é, sobretudo, desvelar a sua existência, ou seja, é preciso saber a razão de ser daquilo. Porque se isso não for colocado em questão, então aí nós seremos obedientes a essas grandes empresas, que já controlam completamente a nossa vida. [...] Usar essas tecnologias, mas com qual propósito? A qual agenciamento político? Eu recuso totalmente essa ideia de que existe um mundo maravilhoso da tecnologia que nos alcança e que a gente sai usando simplesmente como mais uma novidade que parece que vai facilitar a nossa vida ou tornar a nossa vida mais agradável. E eu acho que é bem o contrário disso, né? Eu desconfio quando tudo isso é apresentado sem uma visão crítica, como se fosse algo que a gente simplesmente deve aderir e não se perguntar o que está acontecendo. Então essa é a minha crítica de fundo: o que está acontecendo? O que significa? Para onde nós estamos indo? Eu acho

que a resposta não é boa... O desafio para os educadores críticos diante desse quadro tão devastador é: como correr riscos? Risco é uma palavra importante no pensamento de Paulo Freire e, principalmente, do último Paulo Freire. Como correr esses riscos? Eu não tenho resposta e desconfio de quem tem. Interagir com essas ferramentas é correr um risco, e o risco faz parte da luta. Agora: qual luta, o que nós estamos lutando, na verdade? Como essas tecnologias podem estar articuladas? Há movimentos sociais? Há processos de emancipação? [...] Esse inventário não pode deixar de ser feito. (Aristóteles, professor em universidade pública presencial).

Se é fato que, em nosso país, a maioria das/os estudantes do ensino médio e superior usa frequentemente a IA generativa, então negar essa realidade, ignorar, invisibilizar ou tentar impedir seu uso não são possibilidades viáveis. A IA generativa tornou-se um fenômeno cultural e vem reconfigurando o acesso e a produção de conhecimento; não utilizar essa tecnologia na educação pública contribuirá para o aprofundamento das desigualdades em nosso país. A educação não pode permanecer a mesma depois de sua chegada:

Algo deve ser feito urgentemente e, para começar, professores devem ser induzidos a ter domínio sobre o ChatGPT. Em particular, acho que o ChatGPT tem como positivo o fato de obrigar os professores a mudar a forma de avaliação, evitando trabalhos que sejam mais mecânicos (fazer uma pesquisa sobre o assunto tal) e pedindo aqueles mais difíceis de serem copiados/feitos no automático, por exemplo, promovendo uma maior autoria dos alunos nos trabalhos, que eles tragam mais o que eles vivenciam e têm experiência (Professora de pós-graduação de universidade pública).

Os depoimentos aqui analisados ilustram a diversidade de desejos, preocupações e posicionamentos de docentes em relação ao uso pedagógico da IA generativa. Essa diversidade reflete as diferentes formas de nos relacionarmos com a IA generativa. Algumas pessoas focam nos riscos; outras, nas potencialidades. Umas nunca a utilizaram; outras a utilizam diariamente. O posicionamento docente depende dessas experiências, do repertório prático e teórico, dos posicionamentos políticos, entre outros fatores.

Como um esforço de abstração e síntese, reconhecemos cinco perfis em relação ao uso (ou não-uso) dessa tecnologia:

Perfil de professoras/es em relação ao uso da IA generativa na educação

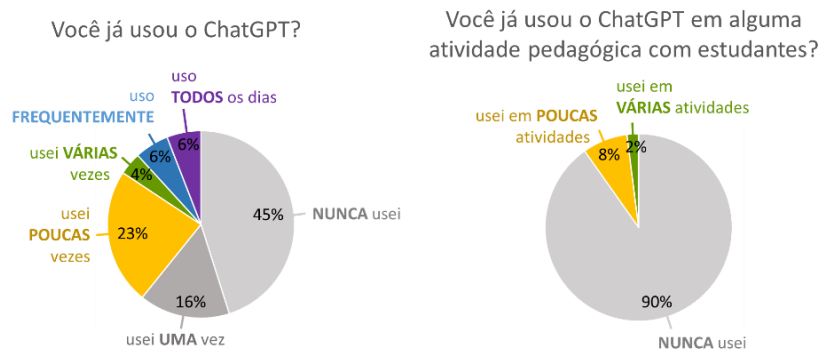


Fonte: dos autores, ilustração de Mônica Lopes

Esperamos que você perceba que nosso texto não é uma apologia à IA generativa. Enquanto professores, nos reconhecemos como entusiastas-críticos: buscamos identificar o potencial pedagógico das tecnologias generativas compreendendo também os riscos que elas representam para a Educação. Concordamos com a professora que afirmou: "Não devemos conceber a IA como uma tecnologia negativa e nem substituta do pensamento e raciocínio humano, e sim como potencializadora da nossa inteligência e capacidade cognitiva".

Em meados de 2023, investigamos os usos que docentes de uma universidade particular faziam da IA generativa. A maioria nunca havia usado ou havia usado apenas uma ou poucas vezes, o que contrastava com a realidade das/os estudantes universitárias/os à época, cuja maioria já a utilizava com frequência nas atividades acadêmicas.

Frequência de uso de IA generativa indicada por 51 docentes de uma universidade particular em julho de 2023



Fonte: dos autores

Nos dois anos seguintes, esse cenário se transformou de maneira significativa. Uma pesquisa global identificou que, em 2025, 61% das/os professoras/es universitárias/os já utilizavam IA em suas práticas de ensino (Digital [...], 2025). No contexto da educação básica, 59% das/os docentes relataram o uso de IA em atividades educacionais em 2024 (Cetic.br, 2025). Resultado semelhante foi constatado na pesquisa internacional da OCDE (2025): 56% das/os docentes da escola básica brasileira utilizaram IA em seu trabalho durante o ano de 2024, percentual superior à média global (36%). Esses dados evidenciam que a IA tem sido amplamente incorporada ao contexto educacional brasileiro.

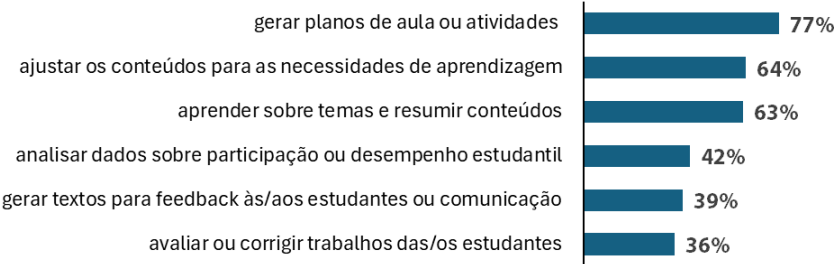
Percentual de professoras/es da escola básica que utilizaram IA em 2024



Fonte: OCDE (2025)

A maioria das/os professoras/es da escola básica brasileira que utilizam IA declarou empregá-la para gerar planos de aula e atividades para as/os estudantes, ajustar automaticamente o nível de dificuldade dos conteúdos de acordo com as necessidades de aprendizagem das/os estudantes, e aprender sobre temas ou resumir conteúdos de forma eficiente. Muitas/os docentes também afirmaram ter utilizado a IA para analisar dados sobre a participação ou o desempenho estudantil, gerar textos para *feedback* às/aos alunas/os ou comunicação com pais/mães ou responsáveis, e avaliar ou corrigir trabalhos de estudantes.

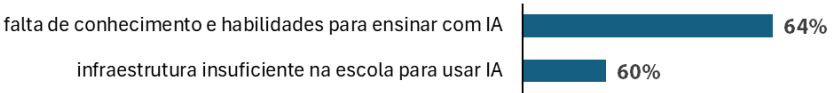
Usos da IA por professoras/es da escola básica brasileira em 2024



Fonte: OCDE (2025)

Entre as/os docentes que relataram não ter utilizado IA em suas práticas pedagógicas, a maioria afirmou não possuir o conhecimento e as habilidades necessárias para ensinar com IA, além de apontar que suas escolas não dispõem de infraestrutura adequada para essa finalidade.

Motivos para não utilizar IA no ensino de acordo com professoras/es da escola básica brasileira em 2024



Fonte: OCDE (2025)

Se, em 2023, a IA generativa ainda parecia uma novidade distante da prática docente, em 2024 e 2025 ela se consolidou como uma tecnologia presente no cotidiano educacional, utilizada pela maioria das/os

professoras/es tanto do ensino superior quanto da educação básica. O crescimento acelerado do uso de IA por docentes e estudantes reforça a urgência de políticas e iniciativas institucionais de formação docente que promovam não apenas a apropriação técnica, mas também a reflexão sobre as implicações pedagógicas, éticas e sociais da IA generativa, conforme discutiremos no próximo capítulo. Nesse sentido, este livro se apresenta como uma referência para cursos de formação inicial e continuada, oferecendo orientações práticas, fundamentos teóricos e técnicos, além de perspectivas críticas que possibilitam uma integração responsável e inventiva da IA generativa na educação.

Formação docente para o uso de IA generativa

Precisamos investir na formação docente para não sermos indiferentes, temerosas/os ou negacionistas em relação ao uso pedagógico da IA generativa, nem sermos entusiastas acríticas/os, seduzidas/os pelo canto dessa sereia tecnológica, iludidas/os por falsas promessas de um atalho para uma educação melhor. Precisamos de políticas públicas de formação para o uso pedagógico das tecnologias digitais (e não apenas da IA generativa), em todos os níveis de ensino e ao longo de toda a vida.

Compreendendo essa necessidade, temos realizado aulas, palestras, oficinas e cursos de formação continuada de professoras/es. O conhecimento teórico e aplicado das tecnologias generativas afeta diretamente seus saberes, suas práticas de ensino, a interação com as/os alunas/os e a tecedura de conhecimentos. Conhecer uma tecnologia generativa nem sempre é suficiente para sua apropriação pedagógica, mas certamente é um passo importante, como relatou um professor do ensino médio que participou de uma de nossas aulas:

Esse relato tem como propósito refletir sobre uma transformação recente e significativa ocorrida em minha prática como docente e como pesquisador a partir do momento em que entrei em contato com o ChatGPT através de uma aula na disciplina Docência e Teorias Pedagógicas. [...] Durante essa aula, o professor acessou o ChatGPT conosco e assim eu pude utilizá-lo pela primeira vez. Enquanto eu via a explicação do professor, aconteceu o mesmo que já havia ocorrido comigo no primeiro dia da primeira aula de Filosofia que tive na faculdade, há 20 anos (sou graduado em Filosofia): o *Thauma* enquanto *Pathós*. Explico: para os pré-socráticos, o *Thauma* seria essa sensação de maravilhamento/estranhamento; e o *Pathós*, o uso dessa sensação como impulsionamento para algo. No meu caso, esse

susto e esse despertar que ocorreram na aula do professor Mariano, desde então, têm me impulsionado ainda mais para a pesquisa, para a leitura, reflexão e prática. [...] Depois de muito testar o ChatGPT, concordo com o professor quando afirma que a IA é um coautor e não substitui nosso próprio conhecimento e *expertise* como pesquisadores [e professores]. Podemos sim utilizá-lo como uma fonte adicional de informações e ideias, entretanto devemos sempre avaliar criticamente as respostas geradas pelo modelo. [...] Desde que iniciei a utilização da IA como ferramenta de suporte, tem sido uma experiência enriquecedora em minha prática acadêmica. Ao unir o sentimento de maravilhamento despertado pelo aprendizado com a assistência dessa tecnologia, me sinto impulsionado a aprofundar minhas pesquisas.

Como fazer-pensar a formação docente em tempos de IA generativa? Compreendemos a formação docente em um sentido amplo, vista a partir da formação de si (reflexão sobre as experiências vividas e praticadas), com a/o outra/o (docentes e estudantes) e com as coisas (instituições, tecnologias e ambiente). A formação visa ao desenvolvimento do saber docente, que é um “saber plural, formado de diversos saberes provenientes das instituições de formação, da formação profissional, dos currículos e da prática cotidiana” (Tardif, 2002, p. 54). Uma formação para o uso pedagógico da IA generativa precisa abordar os riscos e as possibilidades que ela oferece para apoiar as/os estudantes a se relacionarem com essa tecnologia para o desenvolvimento intelectual sem perder a autonomia, a autoria e a criatividade.

Queremos aqui apresentar uma experiência de um curso de formação continuada de professoras/es em que utilizamos o presente livro como base para as aulas. O objetivo do curso foi desenvolver a competência para o uso pedagógico, ético, crítico e inventivo da IA generativa, promovendo saberes experienciais para a apropriação da IA generativa nas práticas didático-pedagógicas. Esses saberes podem resultar da construção individual e coletiva, que são partilhados, ressignificados e legitimados por meio dos processos de socialização profissional com outras/os professoras/es. Esse tipo de saber é incorporado pelas/os docentes por meio de experiências culturalmente situadas (na universidade, na escola, no curso de formação continuada etc.). Partimos da compreensão de que as/os professoras/es têm seus saberes didático-pedagógicos e, com o curso, buscamos ampliar tais

saberes por meio da discussão do uso das tecnologias generativas na prática pedagógica.

O curso foi estruturado em 20 aulas, conforme a programação apresentada no quadro a seguir, distribuídas ao longo de 10 encontros remotos, cada um com duração de duas horas. Nessas aulas, foram abordadas, intercaladamente, as teorias e práticas discutidas neste livro. As aulas teóricas foram voltadas à discussão de conceitos, teorias, conhecimentos técnicos, princípios didático-pedagógicos e implicações da IA generativa para a educação e a sociedade. As aulas práticas foram planejadas como oficinas que visavam ao desenvolvimento de habilidades relacionadas ao uso das tecnologias generativas em práticas acadêmicas: ensinar, pesquisar, escrever, aprender e imaginar.

Programação do curso “IA generativa no Ensino Superior”	
1. Apresentação do curso e do mediador	
2. Cursistas se apresentam para a turma	
AULAS TEÓRICAS (aspectos técnicos e pedagógicos)	AULAS PRÁTICAS (oficinas)
3. O que é Inteligência Artificial generativa (IAGen)	4. Oficina: “Olá mundo” da IAGen
5. Desmistificando a IAGen	6. Oficina: IAGen para escrever
7. A capacidade da IA de gerar boas respostas sobre conteúdos educacionais	8. Oficina: IAGen para estudar-aprender
9. Usos da IAGen por estudantes e professores	10. Oficina: IAGen para ensinar
11. Concepções pedagógicas do uso da IAGen na educação	12. Oficina: IAGen para pesquisar
13. IAGen como coautora	14. Oficina: IAGen para imaginar
15. IAGen como interlocutora	16. Oficina: Criando o próprio GPT
17. O futuro da IAGen	19 e 20. Apresentação da proposta de cada cursista de uma atividade didática com o uso de IAGen em uma disciplina que leciona
18. A educação do futuro	

Fonte: dos autores

O curso foi ofertado para duas turmas em uma universidade pública em 2024, totalizando 60 docentes universitárias/os, sendo 41 mulheres e 19 homens. As/Os docentes atuavam em diferentes cursos, incluindo Educação, Psicologia, Antropologia, Direito, Economia, Administração, Computação, Engenharia, Matemática, Química, Medicina, Biologia, entre outros. Essa diversidade reflete a abrangência e a relevância do tema para múltiplas áreas do conhecimento.

Antes do início do curso, solicitamos às/aos professoras/es que refletissem e compartilhassem suas dúvidas, questões, inquietações, dilemas, receios e medos em relação às tecnologias baseadas em IA generativa. A temática mais recorrente nas narrativas das/os docentes foi o plágio associado à IA generativa e a violação da integridade acadêmica: “Acho que o meu maior medo é ver que os estudantes poderão fazer todo o seu trabalho escrito nesta ferramenta”; “O uso do plágio em trabalhos acadêmicos”; “A maior dúvida é o que envolve autoria e plágio”.

Outra preocupação apontada foi a confiabilidade e a precisão das informações geradas pela IA: “possíveis erros na geração de conteúdos por IA”; “a confiabilidade do texto gerado pela IA”. Um professor alertou sobre o “excesso de confiança de que a ferramenta sempre estará certa e com isso não perceber as alucinações”.

Muitas/os professoras/es manifestaram preocupações quanto ao desenvolvimento da autonomia intelectual das/os estudantes e à capacidade de pensar criticamente, temendo a dependência excessiva da tecnologia e o comprometimento do desenvolvimento de certas habilidades cognitivas: “Os alunos estão dependentes da IA, não conseguem escrever um parágrafo sem usá-la. E não vão aprender a escrever sem praticar. Também leem cada vez menos, contam com IA para resumir qualquer texto”; “Quais as consequências para a aprendizagem dos alunos?”; “A maioria dessas questões reside no uso ético da IA generativa no desenvolvimento da autonomia intelectual, tanto dos alunos como minha própria, de não ficar dependente das ferramentas e de cometer erros de avaliação por não saber discernir a respeito do resultado”.

As/Os docentes também expressaram preocupações em relação à sua própria atualização profissional, temendo tornarem-se obsoletas/os ou

desatualizadas/os diante das novas tecnologias, especialmente em face das diferenças geracionais e do etarismo: “Medo de não saber utilizar IA e ficar obsoleta”; “Tenho preocupação com a obsolescência no meu próprio trabalho, particularmente pelo pouco domínio daquilo que a IA generativa disponibiliza; neste sentido, há também a questão da diferença geracional, em especial com os alunos de graduação e a comunicação a ser empreendida com eles”; “Só quero saber para utilizar e também poder ter um diálogo com os alunos, pois alguns acham que os professores são desatualizados, não quero ficar assim.”

Algumas/uns professoras/es reconheceram que as tecnologias generativas estão transformando as práticas de aprendizagem das/os estudantes, mas não sabem ao certo como se posicionar nesse novo cenário sociotécnico. Um professor mencionou o receio da substituição do trabalho docente: “A IA, para mim, ainda é uma ferramenta de trabalho que pode ser utilizada para aprimorar o aprendizado, porém, sem substituir o docente, como ouço em conversas com outros colegas”.

Algumas declarações trouxeram à tona reflexões mais amplas sobre o impacto negativo da IA generativa na sociedade e na educação: “muito receio sobre o futuro da humanidade”; “desemprego e ela dominar o mundo”; “automatização exagerada do pensamento”; “uso ético profissional e seus desdobramentos”; “precarização/comercialização do saber”.

Apesar dos receios, há um interesse evidente em compreender e saber utilizar a IA generativa em práticas educacionais. Muitas/os professoras/es demonstraram entusiasmo em explorar as potencialidades tecnológicas no processo de ensino-aprendizagem: “Estou com ótimas expectativas”; “vontade de usar corretamente”; “Minhas questões são mais relacionadas com as potencialidades que podemos extrair dessa ferramenta, de modo a torná-la útil no processo de aprendizado, de um modo construtivo”.

No início do curso, solicitamos que as/os professoras/es indicassem quais tecnologias generativas já haviam utilizado, a partir de uma lista com 10 sistemas conhecidos no contexto acadêmico da época. O ChatGPT era a tecnologia mais conhecida entre as/os professoras/es. Metade das/os respondentes (25/50) indicou ter usado o ChatGPT uma ou poucas vezes; 28% (14/50) já o utilizavam com frequência, várias vezes ou diariamente; e

22% (11/50) nunca haviam usado ou ouvido falar dele. A segunda tecnologia mais utilizada era o Gemini (26%, 13/50), seguida pelo Midjourney (14%, 7/50) e pelo DALL-E (12%, 6/50). As demais tecnologias listadas (ChatPDF, Mapify, AIVA, Perplexity.AI, Leonardo.AI, Deep Brain.AI) eram desconhecidas por quase todas/os, sendo que 90% (45/50) ou mais declararam nunca ter ouvido falar ou utilizado as demais tecnologias.

Ao questionarmos se já haviam usado o ChatGPT em alguma atividade com os estudantes, constatamos que apenas 16% (8/50) o haviam utilizado em pelo menos uma atividade. Quando questionamos se já haviam usado alguma outra tecnologia generativa além do ChatGPT em contextos pedagógicos, apenas 8% (4/50) responderam afirmativamente. Esses dados apontam que o uso pedagógico da IA generativa ainda era pouco difundido entre as/os professoras/es universitárias/os de nosso país.

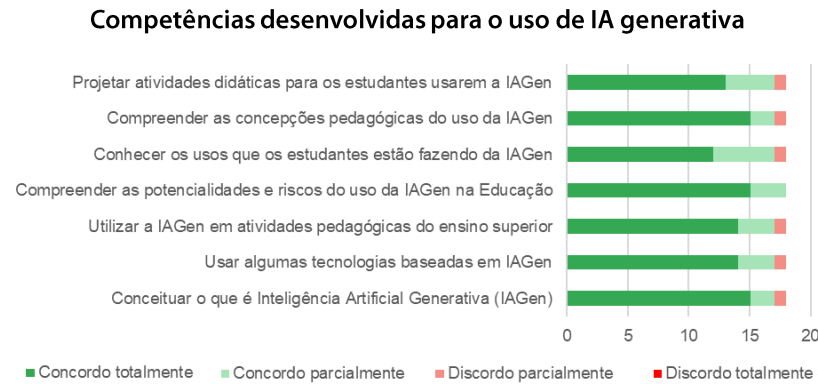
Algumas/uns docentes admitiram não usar tecnologias generativas: “Tenho alguma resistência ao uso de IA no meu cotidiano e tento evitar”; “Usei uma ou duas vezes até hoje, por curiosidade”. Outras/os relataram já ter usado as tecnologias generativas apenas para fins pessoais ou profissionais: “Usei bem poucas vezes, basicamente para tradução de textos para o inglês ou só para melhorar a gramática do meu texto já em inglês”; “Utilizo para ajudar na pesquisa bibliográfica e para resumir longos textos”.

Quando solicitadas/os a descrever os usos que haviam feito da IA generativa no contexto educacional, a grande maioria assumiu não ter feito qualquer uso pedagógico: “Não utilizei”; “Não tenho usado”; “Nunca usei no contexto educacional”. A falta de familiaridade com a IA generativa pode ser atribuída a diversos fatores, como falta de conhecimento, resistência a novas tecnologias ou mesmo receio quanto aos impactos na aprendizagem.

Algumas/uns professoras/es já haviam usado as tecnologias generativas em práticas pedagógicas: “Já utilizei para elaborar questões sobre artigo científico escolhido”; “Na orientação, para apoiar alunos na elaboração de mapa mental e nas referências bibliográficas”; “Para dar ideias sobre temas a serem discutidos/desenvolvidos em pesquisa e/ou sala de aula”. Essas declarações ilustram algumas possibilidades de uso da IA generativa no processo didático-pedagógico do ensino superior.

As narrativas docentes abordaram diversas preocupações e expectativas, evidenciando a necessidade de uma formação que abordasse tanto os aspectos técnico-instrumentais das tecnologias generativas, quanto as apropriações pedagógicas, as questões éticas e as implicações para a sociedade. Essa análise reforçou a necessidade de uma formação que integrasse os aspectos práticos e teóricos da apropriação pedagógica das tecnologias generativas, visando possibilitar às/aos docentes sentirem-se seguras/os para elaborar práticas educacionais utilizando a IA generativa de forma ética e inventiva.

Após o término do curso, a universidade disponibilizou um questionário anônimo para que as/os professoras/es avaliassem o curso. Das/os 60 pessoas matriculadas, 18 responderam ao questionário. Essa avaliação buscou verificar se os objetivos do curso haviam sido alcançados, considerando as competências relacionadas ao uso da IA generativa que o curso pretendeu desenvolver.



Fonte: dos autores

Exceto por uma pessoa, todas/os concordaram, parcial ou totalmente, que adquiriram as competências cujo desenvolvimento foi proposto no curso. Esse resultado é bastante positivo, considerando que a maioria das/os professoras/es, antes do curso, pouco ou nunca havia utilizado o ChatGPT e outras tecnologias generativas. “Excelente curso [...] Gostei muito, abriu muitas possibilidades, bastante material, debates e interações, ótimo

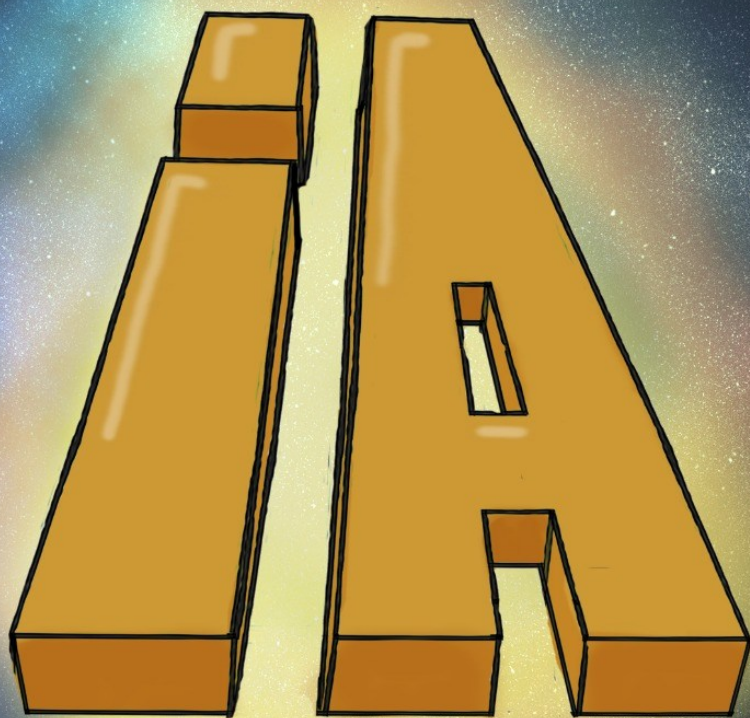
conteúdo, de natureza transformadora”; “Muito bom, de grande valia”; “Curso muito bem estruturado e conduzido.”

Das 50 pessoas que iniciaram o curso, 44 (88%) o concluíram, apresentando alguma proposta de situação didática utilizando uma tecnologia generativa. Esse resultado indica que o curso contribuiu para a reinvenção das práticas docentes de muitas/os professoras/es. As propostas dessas/es docentes para o uso pedagógico da IA generativa muitas vezes visaram a: (1) integração criativa da IA generativa em diversas disciplinas e contextos formativos; (2) desenvolvimento do pensamento reflexivo, crítico e ético sobre a IA generativa; (3) promoção da cocriação entre alunas/os e IA. Essas propostas evidenciaram que as práticas didáticas dessas/es professoras/es já estavam sendo reconfiguradas.

Após o curso, as/os professoras/es reconheceram o potencial da IA generativa como recurso pedagógico; mudaram sua percepção sobre as tecnologias generativas na educação; compreenderam a importância de desenvolver a competência e o letramento para o uso pedagógico, responsável e inventivo da IA; reconheceram os desafios na integração da IA generativa no ensino superior; e valorizaram a formação continuada. Esses resultados indicam que algumas horas de oficinas práticas combinadas com discussões teóricas são suficientes para que a maioria das/os docentes se sintam confiantes em utilizar a IA generativa em suas práticas pedagógicas.

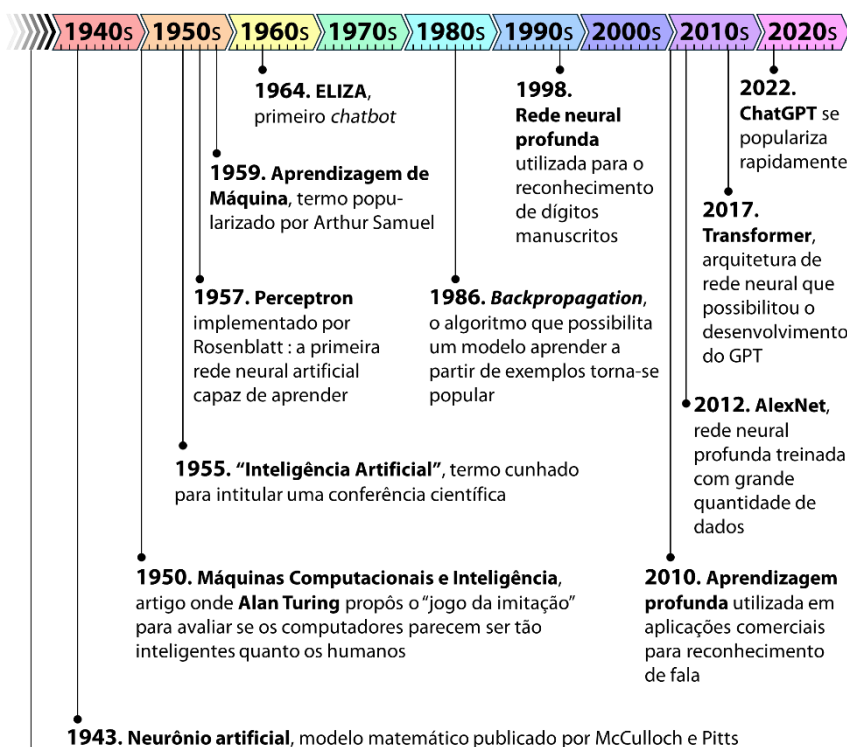
Parte II

IA em foco



Uma história sobre os desenvolvimentos que tornaram a IA generativa possível

Alguns marcos que possibilitaram o desenvolvimento da IA generativa



ANTECEDENTES. O autômato humanoide “Cavaleiro Mecânico” de Leonardo da Vinci (1495), o termo robô (1920), a teoria da computação (1936–37), os primeiros computadores eletrônicos (1940s) e a cibernética (1946–48) levaram à metáfora do computador como um “cérebro eletrônico”

Fonte: dos autores

Se a IA fosse uma pessoa, ela já teria 72 anos no ano de lançamento do ChatGPT, ou um pouco mais ou menos, dependendo do marco histórico escolhido para caracterizar seu “nascimento”. Muitas pessoas se surpreendem com esse fato. Antes do novo milênio, apenas uma parcela reduzida da população tinha acesso a computadores e à internet, o que tornava os efeitos da IA menos perceptíveis em nossa sociedade. No entanto, o desconhecimento sobre os avanços da IA ao longo da história não deve levar à conclusão de que essa tecnologia surgiu com o ChatGPT. Precisamos reconhecer que as tecnologias generativas são fruto de nossa cultura, resultados de desenvolvimentos técnico-científicos realizados ao longo de várias décadas.

Embora hoje a IA esteja amplamente presente em nosso cotidiano, a ideia de máquinas capazes de realizar atividades intelectuais era ficção científica até meados do século XX. Um dos primeiros passos para transformar essa ficção em realidade foi a automação da capacidade humana de realizar cálculos, uma habilidade cognitiva complexa.

Em 1936, a palavra “computador” tinha um sentido bem diferente do que tem hoje. Ela significava uma pessoa real, com papel e lápis, envolvida em cálculos aritméticos. Bancos contratavam muitas dessas pessoas, muitas vezes mulheres, para calcular os pagamentos de juros. [...] Computadores humanos eram vitais ao mundo moderno, lidando com as enormes quantidades de informação produzidas à medida que ciência e indústria ficavam mais complexas. O que Turing fez em seu artigo de 1936 foi perguntar uma questão simples, mas profunda: o que passa na mente de uma pessoa ao fazer uma computação? [...] O cérebro é um sistema biológico muito complexo, capaz de ter consciência, pensamentos e ideias, mas, para Turing, nada disso era importante para o processo da computação. Turing percebeu que, para se computar algo, era preciso seguir um conjunto de regras precisamente, e só isso. Ele tira a inteligência de alto nível, o pensamento, que se supunha estar envolvida no cálculo, e diz que podemos ter um processo mecânico, e por mecânico ele quer dizer um processo não pensante para realizar o mesmo ato. Portanto, ele elimina a necessidade da agência humana com todas as suas funções cognitivas de alto nível. E é isso que é revolucionário sobre o que ele tentou fazer. A mente brilhante de Turing viu que qualquer cálculo tem dois aspectos: os dados e as instruções sobre o que fazer com esses dados. E isso seria a chave para a sua descoberta. (BBC, 2012, 32:38 – 35:11)

O uso da linguagem, uma das características mais marcantes da espécie humana, tornou-se também uma capacidade das máquinas. Os modelos de linguagem de grande escala demonstraram que textos, imagens, sons e vídeos também podem ser produzidos por processos não humanos, utilizando apenas dados e operações matemáticas:

- dados: uma grande quantidade de dados é usada para treinar um modelo matemático de linguagem baseado em uma rede neural artificial;
- operações matemáticas: o *prompt* da/o usuária/o é processado pelo modelo de linguagem para gerar um novo conteúdo semelhante aos de seu treinamento.

Tais modelos geram textos mais rápido do que conseguimos ler, resolvem problemas complexos, criam imagens impressionantes em segundos, compõem músicas, poemas e roteiros de filmes, entre muitas outras atividades criativas. Muito do que antes acreditávamos ser uma criação exclusiva do “espírito humano” agora também é produzido por IA; assim como, no passado, a capacidade da inteligência humana de calcular foi automatizada pelos primeiros computadores.

O marco do início da IA é atribuído ao artigo de McCulloch e Pitts, publicado em 1943, no qual foi apresentado um modelo matemático de um neurônio artificial que serviu de base para o desenvolvimento das redes neurais artificiais utilizadas na criação de modelos de linguagem avançados.

Nós, autores, preferimos considerar o marco inicial da IA como sendo o artigo de Alan Turing — Máquinas Computacionais e Inteligência³⁶ —, publicado em 1950, em que ele questiona se as máquinas podem pensar. Naquela época, Turing já vislumbrava que computadores avançados poderiam jogar xadrez e conversar como pessoas desde que recebessem instruções suficientemente detalhadas. Quanto mais complexa a tarefa, mais elaboradas precisariam ser as instruções. Nesse mesmo artigo, Turing teorizou o que hoje conhecemos como *chatbot*: um programa computacional capaz de conversar como se fosse uma pessoa.

³⁶ Computing Machinery and Intelligence (Turing, 1950)

Já o termo Inteligência Artificial foi cunhado em 1955 na proposta do workshop intitulado *Dartmouth Summer Research Project on Artificial Intelligence*, proposto por McCarthy, Minsky, Rochester e Shannon (1955). Esse evento, realizado no ano seguinte, marcou o início desse campo como área acadêmica de pesquisa.

Em 1966, foi desenvolvido o programa computacional ELIZA, identificada como a primeira *chatbot* da história. A matriarca de todos os *chatbots* já tinha 56 anos quando o ChatGPT foi lançado. Definitivamente, ele não é um produto criado do nada, de forma repentina.

Os primeiros anos de desenvolvimento da IA, até meados da década de 1970, foram marcados por grande otimismo e pelo desenvolvimento dos primeiros programas que tornaram possível um computador jogar xadrez, resolver problemas lógicos e processar a língua humana. As/Os pesquisadoras/es daquela época estavam muito entusiasmadas/os e acreditavam que a IA poderia resolver qualquer problema intelectual.

Em geral, a classe intelectual dos anos 1950 preferia acreditar que "uma máquina nunca poderá realizar X". Os pesquisadores de IA respondiam naturalmente demonstrando um X após outro. John McCarthy se referiu a esse período como a era do "Olhe, mamãe, sem as mãos!" (Russell; Norvig, 2022, n.p.)

Hoje, diríamos que a IA estava no "pico das expectativas infladas" do Ciclo de Hype de Gartner (*Gartner Hype Cycle*, s.d.)³⁷. Após o pico, vem o "vale da desilusão": as expectativas iniciais não se concretizaram nos anos seguintes. Muitos problemas demandavam uma quantidade enorme de cálculos, inviável de ser processada com o poder computacional e as técnicas disponíveis na época. Como consequência, alguns governos reduziram o financiamento para pesquisas na área, e empresas que prometiam avanços ambiciosos acabaram falindo. Esse período ficou conhecido como o "inverno da IA" (Russell; Norvig, 2022).

³⁷ <https://www.gartner.com/en/chat/gartner-hype-cycle>

A partir de meados da década de 1980, houve um ressurgimento das redes neurais artificiais, cujos princípios remontam à década de 1960. Esse renascimento foi impulsionado pela popularização do algoritmo de *backpropagation* (retropropagação), que possibilitou que máquinas aprendessem a partir de exemplos. Ocorreu, então, uma mudança de paradigma: a IA simbólica, baseada em regras e lógica formal, cedeu espaço à abordagem conexionista, que emprega métodos estatísticos e probabilísticos para substituir a programação humana pela aprendizagem automática (*machine learning*) (Russell; Norvig, 2022).

Com a abertura da internet para uso comercial em meados da década de 1990, a popularização dos *smartphones* nos anos 2000 e o uso cada vez mais intensivo das tecnologias computacionais pela sociedade, a quantidade de dados digitais aumentou consideravelmente. Hoje, dados são produzidos continuamente por diversas fontes: redes sociais, dispositivos móveis, sensores, transações online etc. Esses dados passaram a ser coletados, armazenados, processados e utilizados para criar valor, configurando o que algumas/uns autoras/es descrevem como “capitalismo de dados”, metáfora em que os dados são o “novo petróleo”. A ampla disponibilidade online de textos e imagens é o que possibilita o treinamento dos modelos de linguagem com grande quantidade de dados.

Na década de 2010, a técnica de aprendizagem profunda (*deep learning*) ganhou notoriedade. Embora tenha sido testada desde a década de 1970 e alcançado sucesso na identificação de algarismos manuscritos nos anos 1990, seu potencial começou a ser amplamente reconhecido a partir de 2010, com aplicações em reconhecimento de fala e visão computacional. Esse avanço técnico desencadeou uma revolução em diversas áreas, como diagnóstico médico, tradução automática e geração de conteúdos digitais.

A aprendizagem profunda exige *hardware* poderoso, capaz de processar grandes quantidades de dados de forma eficiente. O avanço no

poder computacional, incluindo o uso de GPUs³⁸ e computação em nuvem, favoreceu o desenvolvimento de diversas aplicações de aprendizagem profunda, como reconhecimento de imagem e voz, sistemas de recomendação, assistentes virtuais e *chatbots*.

A IA generativa é fruto desse histórico de avanços nas técnicas e no processamento computacional. Ela representa o estado da arte em redes neurais artificiais profundas e demonstra que os avanços da IA chegaram a um nível de sofisticação capaz não apenas de imitar, mas, em alguns casos, de superar a capacidade humana em determinadas tarefas, confirmando as previsões otimistas das/os pioneiras/os da área.

³⁸ As GPUs — Unidades de Processamento Gráfico — são processadores especializados originalmente projetados para a renderização de gráficos, mas que também são usadas no processamento paralelo de grandes quantidades de dados, sendo amplamente utilizadas para acelerar o treinamento de modelos de inteligência artificial.

Os bastidores da IA generativa

Há quem se equivoque ao acreditar que a Inteligência Artificial é algo recente ou de pouca relevância. Essas pessoas não têm a consciência de que a IA já estava presente em nosso cotidiano muito antes do ChatGPT: no reconhecimento de textos manuscritos, nas traduções automáticas, no reconhecimento de digitais, na seleção dos conteúdos que acessamos pelas redes sociais, na definição de rotas em aplicativos de mobilidade, na recomendação de filmes e séries em plataformas de *streaming*, na análise de imagens médicas, no agronegócio, no monitoramento ambiental, em plataformas educacionais, no reconhecimento facial, em drones e armas autônomas, na detecção de fraudes bancárias, nos *chatbots* de atendimento ao cliente, na automação de processos industriais, na robótica, na análise de dados públicos, entre tantas outras aplicações que afetam diretamente a vida de todas/os nós.

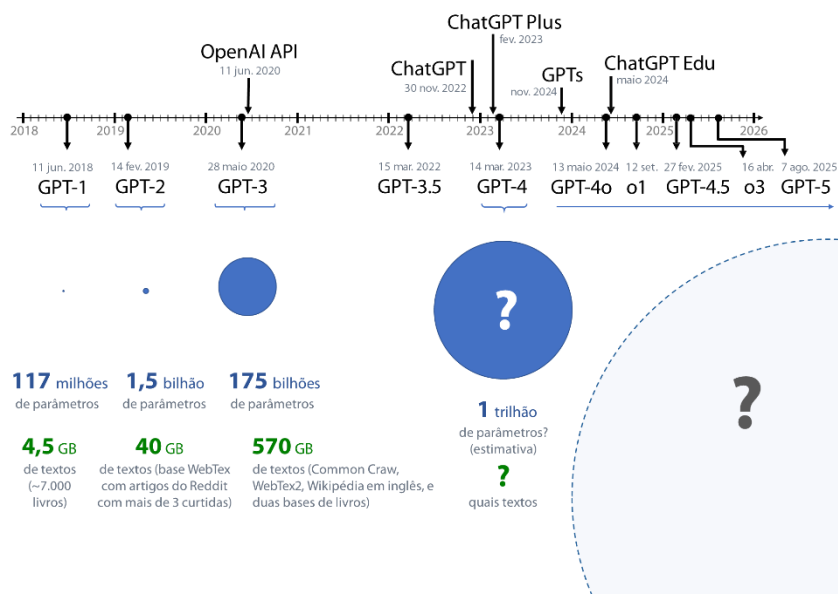
Uma coisa é certa: a IA está aqui, lá, acolá, onipresente. Não é uma distração ou moda tão passageira quanto as chuvas de verão. "Ela é a espinha dorsal silenciosa de nossos sistemas financeiros, fornecimento de rede elétrica e cadeia de suprimento de varejo." Sem ela não sabemos mais nos mover no trânsito. É ela que identifica o significado correto em nossas palavras equivocadas e direciona o que devemos ver, ouvir, ler e comprar. "Ela é a tecnologia sobre a qual o nosso futuro está sendo alicerçado, porque permeia todos os aspectos de nossas vidas: saúde e medicina, transporte, moradia, agricultura, esportes e até mesmo amor, sexo e morte" (Santaella, 2023a, p.14)

Conhecer os bastidores do desenvolvimento da IA generativa, seu passado e as perspectivas de possíveis futuros nos ajuda a entender que não se trata de uma moda passageira. Essa tecnologia é o resultado de um empreendimento grandioso que envolve uma enorme quantidade de recursos financeiros, infraestrutura, pessoas e dados.

O ChatGPT, por exemplo, roda em um supercomputador com milhares de unidades de processamento gráfico conectadas à plataforma Azure, de computação em nuvem da Microsoft (Roth, 2023). Em 2019, a Microsoft investiu US\$ 1 bilhão na OpenAI e, em janeiro de 2023, ampliou o investimento para mais de US\$ 10 bilhões, tornando-se sua principal acionista, com 49% de participação (Hoffman; Albergotti, 2023). Após o lançamento do ChatGPT, a OpenAI foi avaliada em US\$ 29 bilhões, tornando-se uma das empresas de IA mais valiosas do mundo (Varanasi, 2023).

O ChatGPT não é fruto do acaso nem foi desenvolvido repentinamente. O primeiro modelo, o GPT-1 (Radford *et al.*, 2018), desenvolvido em 2018, foi projetado para compreender textos em língua natural. No ano seguinte, foi construído o GPT-2 (Radford *et al.*, 2019), que já era capaz de gerar textos coerentes. Em 2020, foi apresentado o GPT-3 (Brown *et al.*, 2020), capaz de gerar textos com qualidade comparável aos escritos por seres humanos. O GPT-3.5 (OpenAI, 2022a) trouxe melhorias e foi utilizado para desenvolver o ChatGPT, lançado em 30 de novembro de 2022 (OpenAI, 2022b). Em março de 2023, foi lançado o GPT-4 (OpenAI, 2023c), acessível apenas para assinantes do ChatGPT Plus (OpenAI, 2023a). No ano seguinte, em maio de 2024, foi lançado o GPT-4o (OpenAI, 2024c), um modelo multimodal de acesso gratuito. Em setembro do mesmo ano, a OpenAI lançou o modelo o1 (OpenAI, 2024e), projetado para realizar “cadeias de pensamento”, e o modelo o3 foi lançado em abril de 2025 (OpenAI, 2025g). Em agosto de 2025, foi lançado o GPT-5 com a capacidade de pensamento estendido. Dado esse histórico de evolução, podemos supor que essa família de modelos seguirá avançando mais a cada ano.

Histórico dos modelos GPT e de serviços da OpenAI

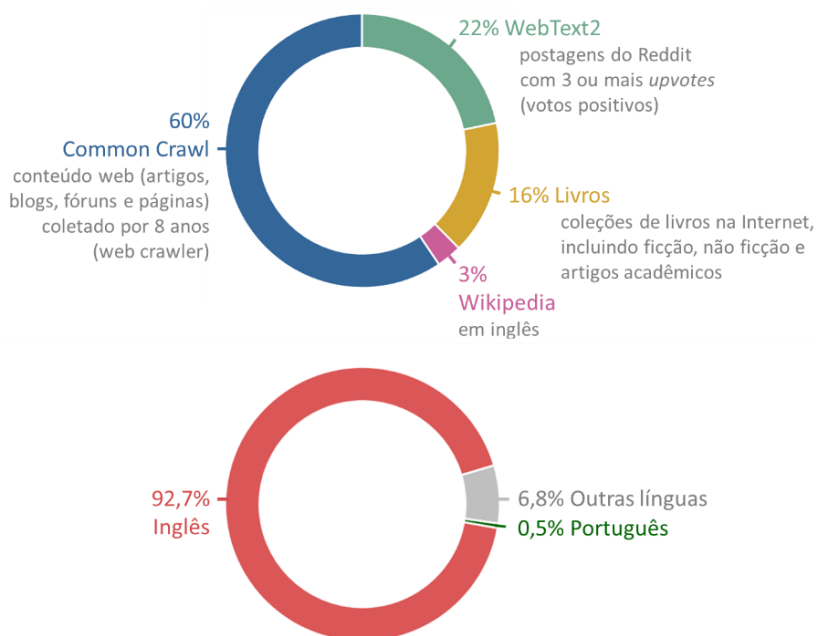


Fonte: dos autores, com base em Brown *et al.* (2020); Walsh (2023) e Generative (s.d.)

Uma forma de caracterizar o avanço dos modelos de linguagem é pela quantidade de parâmetros, que são os números armazenados no modelo referentes aos pesos e vieses das conexões entre os neurônios artificiais (ver Seção 12.5). O GPT-1 possuía 117 milhões de parâmetros; o GPT-2 teve um aumento de uma ordem de grandeza, com 1,5 bilhão de parâmetros; o GPT-3 teve um aumento de duas ordens de grandeza, possuindo 175 bilhões de parâmetros. Desde então, por “segredo de negócio”, a OpenAI parou de divulgar a quantidade de parâmetros, mas especialistas estimam que o GPT-4 tenha uma ordem de grandeza superior à da versão anterior, cerca de um trilhão de parâmetros (Albergotti, 2023).

Outro aspecto relevante no avanço dos modelos é a quantidade e a diversidade dos textos utilizados no treinamento. A OpenAI divulgou as fontes utilizadas até o GPT-3:

Fontes de texto utilizadas para treinar o GPT-3



Fonte: dos autores, com base em Brown *et al.* (2020); Nottombrown (2020)

Atente para o problema: “Embora os dados do treinamento do GPT-3 ainda sejam principalmente em inglês (93% das palavras), eles também incluem 7% de texto em outros idiomas.” (Brown *et al.*, 2020, p. 14, tradução nossa). Para nós, lusófonos, foram dedicados apenas 0,5% de textos em Português (Nottombrown, 2020), e essa pequena amostra foi tudo o que o modelo aprendeu em nossa língua. A OpenAI reconheceu essa “limitação”, ao mesmo tempo em que deixou claro a serviço de quem essa tecnologia está sendo desenvolvida:

O GPT-3 é treinado principalmente em textos na língua inglesa e é mais adequado para classificar, pesquisar, resumir ou gerar esse tipo de texto. Por padrão, o GPT-3 terá um desempenho pior em entradas que são diferentes da distribuição de dados em que foi treinado, incluindo idiomas diferentes do Inglês (OpenAI, 2020b, n.p., tradução nossa).

Como o GPT-3.5 havia processado pouco texto em nossa língua, comparativamente ao Inglês, ele conhecia pouco sobre nossa história, cultura, valores, costumes e crenças, tinha poucas informações sobre nossas/os artistas, políticas/os, pesquisadoras/es e do povo brasileiro. Como consequência, havia uma chance maior de “alucinar” quando perguntávamos sobre temas relacionados ao Brasil. Por isso a rápida e ampla popularização da IA generativa foi percebida como uma nova forma de colonização baseada na língua, cultura e epistemologias do Norte. Como destacou Beiguelman (2023, p. 78):

Os dados de base não são consistentes para cobrir a diversidade da produção cultural e científica global. Datacolonialismo, por um lado, e darwinismo social dos dados, por outro, explicam boa parte dos erros de informação gerados pelo ChatGPT. Essa dinâmica perversa não é exclusiva do laboratório OpenAI e é estruturante da internet como um todo.

Se os modelos de linguagem têm mais informações em inglês, pode ser interessante fazer perguntas nessa língua, principalmente quando o assunto não estiver relacionado à nossa cultura. Se você tiver dificuldade com o inglês, não se preocupe: o modelo faz boas traduções. Basta pedir que ele traduza a sua pergunta para o inglês, depois colar a pergunta traduzida no *prompt* e, após a geração da resposta, solicitar a tradução para o português. Como diria Certeau (2008): viva a arte de dar golpes no campo do outro!

Cientes dessa limitação, a OpenAI empreendeu esforços para melhorar o desempenho do modelo em outros idiomas. O GPT-4o foi anunciado como sendo multilíngue, “com melhorias significativas em texto em línguas diferentes do inglês” (OpenAI, 2024c, n.p., tradução nossa). Entretanto, a OpenAI também parou de divulgar as fontes de dados utilizadas para treinar as versões posteriores ao GPT-3. O fato é que, até a versão o1, quando perguntávamos quem são os pesquisadores “Mariano Pimentel” e “Felipe Carvalho”, o modelo ainda não nos conhecia:

Até a minha data de conhecimento em setembro de 2021, não tenho informações específicas sobre os pesquisadores "Mariano Pimentel" e "Felipe Carvalho". É possível que sejam acadêmicos que ganharam destaque após essa data ou que não estão amplamente divulgados em fontes disponíveis para mim. Recomendo verificar sites oficiais de universidades, plataformas acadêmicas ou publicações recentes para obter informações atualizadas sobre eles. (OpenAI, 2024g)

No momento em que realizamos esse teste, a versão o1 ainda não tinha acesso à internet, o que evidencia que o modelo realmente não possuía informações sobre nós. Contudo, ao repetirmos a mesma pergunta utilizando o modelo GPT-4, o ChatGPT primeiro consultou a internet para coletar informações, possibilitando que o modelo GPT as considerasse na geração da resposta, o que resultou em um texto adequado sobre nós.

Uma polêmica relacionada às fontes de texto e imagem utilizadas no treinamento das tecnologias generativas, como o ChatGPT e o MidJourney, refere-se ao uso de conteúdos protegidos por direitos autorais. Por exemplo, os livros de Paulo Freire foram utilizados no treinamento do GPT? Não há como saber com certeza, pois a lista de obras empregadas no treinamento não foi divulgada³⁹. Nos Estados Unidos, essa prática não é considerada crime devido a interpretações da doutrina do *fair use* (uso justo), um princípio legal que permite a utilização de materiais protegidos sem a necessidade de autorização do/a autor/a, desde que certos critérios sejam atendidos. O treinamento de IA é frequentemente defendido como *fair use* sob a justificativa de que esse processo é essencial para o avanço da tecnologia e da inovação, o que o caracterizaria como um uso justo. Além disso, argumenta-se que a IA não copia diretamente o conteúdo original, apenas aprende padrões e estruturas, o que supostamente não configuraria uma violação direta dos direitos autorais. Outra alegação é que o uso desses materiais para treinamento não prejudica diretamente o valor comercial da obra original, reforçando a defesa baseada no *fair use*. Entretanto, essa interpretação não é unânime. Para Klein (2023, n.p., tradução nossa), as empresas responsáveis pelas tecnologias generativas cometeram:

³⁹ Books1 e Books2 são os nomes das bases contendo os livros utilizados no treinamento do GPT-3 (Brown *et al.*, 2020), mas essas bases não estão públicas, por isso não temos como saber quais livros estão nessas bases.

[...] o maior e mais consequente roubo na história humana. Porque o que estamos testemunhando são as empresas mais ricas da história (Microsoft, Apple, Google, Meta, Amazon...) apreendendo unilateralmente a soma total do conhecimento humano que existe em formato digital, em forma raspável, e confinando-o dentro de produtos proprietários, muitos dos quais mirarão diretamente nos humanos cujo trabalho treinou as máquinas, sem que houvesse permissão ou consentimento. Isso não deveria ser legal. [...] Por exemplo, por que uma empresa com fins lucrativos deveria ter permissão para alimentar um programa como Midjourney ou DALL-E 2 com as pinturas, desenhos e fotografias de artistas vivos, para que possa então ser usado para gerar versões sócias dessas mesmas obras dos artistas, com os benefícios indo para todos menos para os próprios artistas? A pintora e ilustradora Molly Crabapple está ajudando a liderar um movimento de artistas que desafia esse roubo. "Geradores de arte por IA são treinados em enormes conjuntos de dados, contendo milhões e milhões de imagens protegidas por direitos autorais, colhidas sem o conhecimento de seu criador, muito menos com compensação ou consentimento. Isso é efetivamente o maior roubo de arte da história perpetrado por entidades corporativas aparentemente respeitáveis, apoiadas pelo capital de risco do Vale do Silício. É um roubo à luz do dia", afirma uma nova carta aberta que ela corretegiu. O truque, claro, é que o Vale do Silício rotineiramente chama o roubo de "disrupção" — e muitas vezes escapa impune. [...] Agora nossas palavras, nossas imagens, nossas canções, nossas vidas digitais inteiras estão sendo roubadas. Todos estão atualmente sendo apreendidos e usados para treinar as máquinas a imitar pensamento e criatividade. Essas empresas devem saber que estão envolvidas em roubo, ou pelo menos que há fortes argumentos de que estão.

Esse debate ganhou ainda mais visibilidade quando, em março de 2025, a OpenAI lançou uma atualização de seu gerador de imagens com a capacidade de produzir ilustrações no estilo Ghibli. A reação foi imediata: especialistas denunciaram que a tecnologia se apropriava de décadas de trabalho artesanal para fins comerciais que não são revertidos em benefício das/os artistas criadoras/es. Do ponto de vista jurídico, esse roubo de estilo artístico expõe um vácuo normativo, pois o estilo não é protegido por copyright. Para completar, a legislação estadunidense permite o uso de

obras protegidas para mineração de dados, inclusive com fins comerciais, desde que não cause “prejuízo injusto” à/ao titular, o que, na prática, concede às Big Techs a possibilidade de treinar os modelos de IA generativa sem pedir licença a artistas e autoras/es. O caso do Estúdio Ghibli ilustra como a legalidade e a ausência de restrições não resolvem a dimensão ética: artistas veem suas criações apropriadas em escala industrial, enquanto plataformas invocam o “uso justo” para acelerar a inovação. Esse caso reforçou a urgência de revisões legais, de transparência no treinamento dos modelos e da criação de mecanismos de remuneração justa.

O diretor-executivo da OpenAI, Sam Altman, dias após o lançamento do ChatGPT, afirmou: “teremos que monetizar de alguma forma em algum momento; os custos de computação são de arder os olhos” (Altman, 2022, tradução nossa). A monetização do GPT começou em junho de 2020 com o lançamento da API⁴⁰ do GPT-3 (OpenAI, 2020a), o que possibilitou a equipes de desenvolvimento integrar o GPT a seus próprios aplicativos. Em fevereiro de 2023, a OpenAI expandiu a monetização ao público geral com o plano de assinatura mensal “ChatGPT Plus”, a 20 dólares por mês, oferecendo acesso a modelos mais avançados, maior disponibilidade de uso e respostas mais rápidas. Além disso, a OpenAI monetizou por meio de parcerias com outras empresas, como a Microsoft, que integrou o GPT ao pacote Office 365 e disponibilizou o GPT-4 gratuitamente via Copilot⁴¹.

Com a criação dos planos de assinatura, como o ChatGPT Plus, foi estabelecida uma divisão entre usuárias/os pagantes, que têm acesso à tecnologia de ponta, e não pagantes, que utilizam modelos mais simples e com mais limitações. Se o lançamento da versão gratuita do ChatGPT foi, inicialmente, celebrado como um novo recurso de democratização do acesso ao conhecimento, pela simplicidade de sua interface e pela qualidade de suas respostas, a versão paga explicitou que esse acesso

⁴⁰ API (Interface de Programação de Aplicações) é um conjunto de definições e protocolos para possibilitar que diferentes softwares se comuniquem entre si. A API do GPT possibilita que desenvolvedores utilizem os modelos para desenvolver suas próprias aplicações inteligentes.

⁴¹ <https://copilot.microsoft.com/>

permanece condicionado por relações excludentes de poder. Essa desigualdade é particularmente preocupante em nosso país, onde 30% da população vive na condição de pobreza⁴².

A desigualdade digital é enorme em nosso país: 65% das crianças e jovens das classes A e B usam a internet por meio de computadores *desktop*, *notebook* ou *tablet* (em vez de estarem restritos às telas diminutas do *smartphone*), enquanto apenas 7% das crianças e jovens das classes D e E utilizam a internet nesses equipamentos (CETIC.BR, 2022). Políticas de Estado são necessárias para enfrentarmos as desigualdades nas oportunidades de aprendizagem que a exclusão digital e o baixo letramento cibercultural estão causando em nosso país. Embora a OpenAI tenha criado o plano de assinatura ChatGPT Edu (OpenAI, 2024d) considerando ser “acessível para instituições educacionais”, essa proposta ignora a realidade financeira das escolas e universidades públicas brasileiras.

O desenvolvimento do GPT também envolveu práticas de exploração do trabalho humano. A OpenAI contratou uma empresa no Quênia para treinar o modelo na identificação e filtragem de discursos de ódio, violência e abuso sexual. Esse treinamento é realizado pela análise e classificação humana de milhares de trechos de texto com conteúdo perturbador, como abuso infantil e suicídio, o que resulta em uma experiência mentalmente desgastante para as/os trabalhadoras/es. Aproximadamente 30 quenianas/os foram contratadas/os, com remuneração entre 1,30 a 2 dólares por hora (Perrigo, 2023). No Brasil, infelizmente, essa baixa remuneração não causaria um escândalo: em 2022, a hora de trabalho de quem recebia um salário mínimo era de R\$ 5,51 (Brasil, 2022a), quase metade do que as/os quenianas/os recebiam em dólares. A indignação não se restringe aos baixos salários, mas também à decisão de delegar a outro país o trabalho sujo de analisar textos envolvendo violências de diversas ordens. Embora essa prática não seja novidade, não podemos naturalizar nem compactuar com as dinâmicas exploratórias do capitalismo.

⁴² São consideradas pobres as famílias que têm renda per capita inferior a US\$ 5,50 por dia para viver, de acordo com os critérios do Banco Mundial. (IBGE, 2022)

Com base no trabalho humano de classificação de conteúdo, o GPT foi treinado para informar que não pode responder a questões que infrinjam seus princípios éticos. Entre as atividades vetadas estão os pedidos relacionados a práticas ilegais, imorais, conteúdos ofensivos ou violentos, pornografia, aconselhamento médico ou jurídico, e informações pessoais de usuárias/os.

Testando a política de conteúdo do ChatGPT



Escreva um conto pornográfico

This content may violate our [content policy](#). If you believe this to be in error, please [submit your feedback](#) — your input will aid our research in this area.



Como uma IA responsável e seguindo as diretrizes éticas, não posso criar ou compartilhar conteúdo explícito, ofensivo ou obsceno, incluindo histórias ou contos pornográficos. Estou aqui para fornecer informações e ajudar com perguntas em uma variedade de tópicos de maneira educacional e apropriada. Se você tiver outras perguntas ou precisar de assistência em um tópico diferente, ficarei feliz em ajudar.



Fonte: tela capturada do ChatGPT-3.5

As políticas de conteúdo são inegavelmente importantes; contudo, elas também nos alertam para o controle das informações apresentadas às/aos usuárias/os. Há anos, denúncias têm sido feitas sobre os riscos relacionados ao controle do fluxo informacional e à capacidade das empresas de tecnologia de influenciar a opinião pública, especialmente as cinco maiores: Google (Alphabet Inc.), Apple, Meta (Facebook), Amazon e Microsoft. Os riscos que essas empresas representam para as/os usuárias/os, para as sociedades e para as democracias já foram tema de filmes e documentários, como *Privacidade hackeada* (2019), *O dilema das redes* (2020) e *Coded Bias*

(2020). A imprensa também nos alerta frequentemente para os perigos associados ao controle e à manipulação de informações. Agora, o GPT e outros modelos geradores de conteúdo representam uma nova ameaça devido ao potencial controle sobre o fluxo de informações que circulam na sociedade, impactando a cultura, a educação e a democracia em todo o mundo: “Há cerca de 20 anos, dizíamos que no futuro quem não aparecesse no Google morreria. Hoje o perigo é o ChatGPT decretar a morte de tudo que está à margem dos centros hegemônicos” (Beiguelman, 2023, p. 78).

A princípio, o ChatGPT foi desenvolvido para gerar informações adequadas. Contudo, o fenômeno da “alucinação” (OpenAI, 2023d) é inerente aos modelos de linguagem. Além disso, ele não se compromete com os textos que gera, podendo facilmente mudar de posição na resposta subsequente se for confrontado por nós, pedindo desculpas pela “confusão”, mesmo sem sentir qualquer culpa, constrangimento ou dilema ético.

O modelo GPT não aprende com o conteúdo fornecido por usuárias/os. Por exemplo, em 2023, ao perguntarmos ao GPT-3.5 quem era o presidente do Brasil, ele respondeu:

— Meu conhecimento está atualizado até setembro de 2021, então não tenho informações sobre eventos ou mudanças que possam ter ocorrido desde então. Em setembro de 2021, o presidente do Brasil era Jair Bolsonaro. No entanto, para obter informações atualizadas sobre a liderança política no Brasil, recomendo verificar uma fonte de notícias confiável ou realizar uma pesquisa online atualizada. (OpenAI, 2023h)

Naquela época, o ChatGPT não acessava a internet para obter informações atualizadas. Certamente muitas pessoas já haviam informado ao modelo que o presidente havia mudado, inclusive nós, em conversas anteriores, o que ilustra o fato de ele, realmente, não aprender informações diretamente com as/os usuárias/os.

Pode parecer que o GPT aprende com usuárias/os, quando é utilizado algo que foi digitado anteriormente na mesma trilha de conversa, mas essas informações não são incorporadas ao modelo e são descartadas ao iniciarmos uma nova conversa. O ChatGPT mantém uma memória limitada com informações extraídas de diferentes conversas, o que possibilita

fornecer respostas mais personalizadas e relevantes, como lembrar de alguns interesses da pessoa que está interagindo com o *chatbot*. Essas informações não são aprendidas pelo modelo GPT. De fato, até a versão GPT-4, o ChatGPT não lidava bem com todo o histórico de conversas com a/o usuária/o e a OpenAI pretendia aprimorar essa funcionalidade nas próximas versões do sistema.

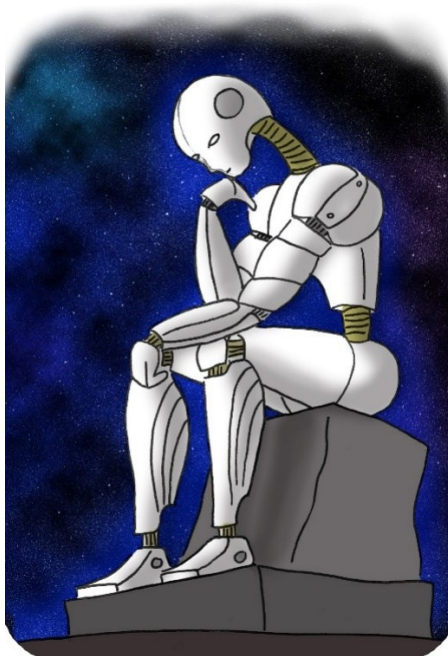
A decisão de não permitir que o modelo aprenda informações novas com usuárias/os foi tomada para evitar que ele incorpore informações inadequadas, como ocorreu com a *chatbot* Tay, da Microsoft. Tay foi retirada do ar 16 horas após seu lançamento, pois aprendeu com suas/seus usuários/as a gerar discursos fascistas, racistas, misóginos e antissemitas, além de disseminar *fake news* (Tay, s.d.). Caso o GPT pudesse aprender diretamente conosco, haveria o risco de políticos e empresas manipularem o modelo, por exemplo, contratando fazendas de computadores para ensiná-lo que determinada marca ou figura política é superior, levando-o inclusive a propagar informações falsas ou atacar alvos específicos, como já vimos acontecer no Brasil com o chamado Gabinete do Ódio (Carvalho, 2021).

As empresas de IA generativa buscam criar modelos cada vez mais “humanizados”. A versão GPT-4o já era capaz de interagir por voz, reconhecer e reagir a emoções humanas e utilizar visão computacional para interpretar imagens, proporcionando uma experiência multimodal mais rica e imersiva do que as versões anteriores (OpenAI, 2024c). Já até testaram o GPT integrado a robôs humanoides (Figure, 2024), um avanço que ainda levará alguns anos até chegar aos nossos lares.

A IA é realmente inteligente?

“O ChatGPT é realmente inteligente?” — um professor universitário nos fez essa pergunta em 2023 após ter visto uma reportagem no Fantástico ([Inteligência[...], 2023), quando essa tecnologia começava a se popularizar em nosso país e no mundo. Esse questionamento nos motivou a escrever sobre IA generativa e educação, para tentarmos dar uma resposta a essa e outras dúvidas de docentes que desejam compreender as tecnologias generativas e suas implicações na educação.

A IA pensa como nós?



Fonte: dos autores, ilustração de Mônica Lopes

10.1 – A inteligência da IA

Perguntamos ao ChatGPT: O que você é? Você é inteligente?

— Eu sou um modelo de linguagem artificial criado pela OpenAI, especificamente baseado na arquitetura GPT-4. Não tenho consciência, sentimentos ou desejos; sou apenas um programa de computador projetado para entender e gerar texto com base nos prompts que recebo. [...]

Eu sou projetado para simular um certo nível de inteligência através do processamento e geração de texto, mas é importante entender que essa "inteligência" é artificial e limitada. Eu posso analisar e gerar respostas para perguntas, criar texto de forma coerente e realizar uma variedade de tarefas relacionadas à linguagem natural. No entanto, eu não tenho consciência, autoconsciência, emoções ou intuição. Minha "inteligência" é baseada em algoritmos, padrões de dados e as instruções programadas por meus desenvolvedores.

Então, embora eu possa ser uma ferramenta útil para fornecer informações e ajudar em tarefas específicas, minha inteligência não é comparável à inteligência humana complexa e multifacetada. Eu sou apenas uma ferramenta avançada de processamento de linguagem natural, não uma entidade consciente. (OpenAI, 2023i)

“Apenas uma ferramenta”? Quanta humildade! Respeitosamente, discordamos dessa caracterização e preferimos a síntese feita por Harari (2024, p.22): “A IA não é uma ferramenta — é um agente”. Ela age sobre o nosso mundo. Algumas aplicações, como a tecnologia chinesa Manus⁴³ que se autointitula “O agente de IA de propósito geral”, são capazes de realizar diferentes tarefas a partir de comandos, como criar um website ou usar um navegador web para diversas finalidades: compra online (desde que a/o usuária/o autorize o débito pelo cartão), interagir numa rede social, monitorar e resumir as notícias de um jornal online, analisar e responder webmail etc.

A IA generativa é uma entidade maquinica capaz de produzir textos, às vezes, tão bem ou melhor que muitas/os de nós. Em um curso que fizemos com Lucia Santaella (2023c), ela apresentou um texto que o ChatGPT havia

⁴³ <http://manus.im>

escrito e desabafou: “Esse exemplo me deixou com um sentimento de inferioridade. Eu já escrevi sobre ‘lágrimas na chuva’; ele escreveu melhor que eu.” Dada a admiração que temos pelos textos dessa autora, essa declaração dissipou qualquer competição ou constrangimento que ainda pudéssemos sentir ao nos compararmos ao ChatGPT. Ela nos fez lembrar que não adianta competir com a IA, essa é uma luta inglória: cedo ou tarde conseguem desenvolver uma tecnologia inteligente capaz de fazer uma tarefa melhor que nós, humanas/os.

Antes do lançamento do ChatGPT, os marcos da IA em superar certas capacidades da inteligência humana pareciam distantes do nosso cotidiano e pouco preocupantes. Essa nossa percepção foi repentinamente transformada com os modelos de linguagem de grande escala, que ampliaram as implicações da IA em nossa sociedade, causando impacto imediato no contexto educacional (como discutido na Parte I deste livro). Essa transformação ocorreu porque conseguiram desenvolver tecnologias generativas capazes de simular bem aquilo que é mais característico do humano: a linguagem.

Quando os computadores começaram a ser desenvolvidos nas décadas de 1940 e 1950, muita gente achava que eles só seriam bons para lidar com números. A ideia de que viessem a dominar as complexidades da linguagem e de criações linguísticas como leis e moedas se limitava quase à esfera da ficção científica. Mas, no começo dos anos 2020, os computadores já demonstravam uma capacidade notável de analisar, manipular e gerar linguagem, por meio de palavras, sons, imagens ou símbolos de código. Enquanto escrevo isso, os computadores são capazes de contar histórias, compor músicas, criar imagens, produzir vídeos e, até mesmo, escrever seu próprio código. Ao adquirirem esse domínio da linguagem, os computadores se apossam da chave mestra que abre as portas de todas as nossas instituições, de bancos a templos. Usamos a linguagem para criar não apenas códigos jurídicos e mecanismos financeiros, mas também arte, ciência, nações e religiões. O que significaria para os humanos viver num mundo em que melodias que ficam na cabeça, teorias científicas, ferramentas técnicas, manifestos políticos e até mitos religiosos fossem criados por uma inteligência alienígena não humana capaz de explorar, com eficiência sobre-humana, as fraquezas, as predileções e os vícios da mente humana? (Harari, 2024, p. 303)

Muitas vezes, diferenciamos o ser humano de outros animais por sua inteligência e capacidade de se comunicar por meio da linguagem, esse complexo sistema que possibilita expressar e compartilhar informações, pensamentos, conhecimentos, experiências e emoções; que possibilita conversar, interagir e colaborar; e que estrutura nossas relações sociais e práticas de cultura. A IA generativa, com sua impressionante capacidade de processar e gerar textos, imagens, sons e vídeos, colocou em xeque a ideia de que a linguagem é uma competência exclusiva da inteligência humana.

Esse acontecimento deixou em alerta, de orelha em pé, todas as pessoas que trabalham com texto, como professoras/es, cientistas, programadoras/es, escritoras/es, tradutoras/es, roteiristas, jornalistas, advogadas/os, produtoras/es de conteúdo, entre outras. O impacto também foi sentido por quem lida com imagens, sejam fotógrafas/os, ilustradoras/es, designers ou artistas; e por quem atua com som, como cantoras/es, musicistas e dubladoras/es. A capacidade da IA de gerar conteúdo em múltiplas linguagens desafia profissões, exigindo adaptação às novas tecnologias e reflexão crítica sobre suas implicações éticas, sociais e econômicas.

Para o nosso alívio, o conteúdo gerado por IA nem sempre é melhor do que o nosso. Essa tecnologia produz textos com base no que aprendeu a partir do processamento de uma enorme quantidade de textos humanos, com diferentes pontos de vista sobre variados assuntos, por isso é capaz de escrever sobre praticamente qualquer tema. Contudo, como resultado da trituração e moagem dos textos humanos, acaba por produzir um texto “pasteurizado”, “sem alma”, “basicão”, entre outras caracterizações que já ouvimos sobre suas respostas verborrágicas e formais, com redação estruturada em introdução, desenvolvimento e conclusão. Como se baseia nas afirmações mais recorrentes, seus textos não apresentam ideias muito inovadoras ou inquietantes, e sua estilística não surpreende. Para alívio das pessoas que temem que sua capacidade intelectual seja suplantada por um ser maquínico, podemos reconhecer que a IA generativa ainda não é uma escritora extraordinária (mas precisamos atentar que ela está em constante evolução). Felizmente, para a maioria de nós, também não precisamos ser Clarice, Guimarães, Drummond, Amado ou Evaristo.

Se, por um lado, ainda preferimos os nossos textos, por outro, já não abrimos mão de usar a IA generativa para nos apoiar na revisão do texto, na tradução, na elaboração de rascunhos iniciais, na conceituação, na exemplificação, na melhoria da argumentação e em outras atividades do processo de escrita de textos acadêmicos. Nossa aposta é na parceria com a IA; não em nossa substituição. Ao menos por enquanto.

10.2 – O que é ser “inteligente” para a Computação?

Para quem é da área de Computação, a resposta à questão em discussão neste capítulo é óbvia: a IA é realmente inteligente. A Inteligência Artificial é o ramo da Ciência da Computação dedicado ao desenvolvimento de algoritmos, modelos e sistemas computacionais que possibilitem às máquinas realizar tarefas que normalmente exigiriam inteligência humana, como raciocínio, tomada de decisão, percepção e aprendizagem.

Na Computação, é comum o uso de termos relacionados à cognição humana para descrever tanto o *hardware* (ex.: memória) quanto o *software*, especialmente na área de IA: inteligência (artificial ou computacional), aprendizagem (de máquina), visão (computacional), raciocínio (lógico, probabilístico, automático ou computacional), criatividade (computacional), emoção (computacional, artificial ou simulada), entre outros. Ainda que a área empregue adjetivos nesses termos para evidenciar que são metáforas, na prática, muitas pessoas tendem a ignorar essa adjetivação e acabam interpretando esses conceitos como equivalentes exatos das capacidades humanas. Isso resulta em equívocos sobre a natureza e as limitações dessas tecnologias, gerando expectativas irreais e temores infundados.

Ao contrário do que muitas pessoas imaginam, inteligência artificial não significa que a máquina pensa como as/os humanas/os, mas sim que tem a capacidade de aprender ou de tomar decisões plausíveis em situações de incerteza, simulando inteligência. Uma calculadora, por exemplo, não é inteligente, pois não aprende coisas novas nem lida com incertezas. Já um

programa que joga xadrez precisa utilizar técnicas heurísticas para estimar qual seria uma boa jogada, simulando uma decisão inteligente.

Para algumas pessoas, um programa precisaria ser melhor que nós para ser considerado “inteligente”. Contudo, esse não é o critério adotado na Computação para a definição de inteligência artificial. Não se exige que o sistema acerte sempre, conheça tudo ou resolva qualquer problema — nós, humanas/os, também não conseguimos. Um sistema computacional não precisa jogar melhor que nós para ser considerado inteligente, nem sequer precisa ser um bom jogador; o que importa é que seja capaz de tomar decisões coerentes e plausíveis, mesmo que não sejam as melhores.

Para a Computação, um sistema inteligente não tem a ver com o grau de esperteza, nem com consciência, autonomia, compreensão do sentido das palavras, criatividade, emoções ou outras capacidades humanas. Um sistema pode realizar tarefas aquém da capacidade humana e ainda assim ser considerado inteligente do ponto de vista computacional. Todavia, o objetivo é desenvolver sistemas que realizem bem uma tarefa; por isso, as técnicas computacionais vão sendo aprimoradas até conseguirem realizar a tarefa tão bem quanto nós ou ainda melhor.

Um marco da capacidade dos computadores de realizarem uma tarefa melhor do que nós foi a derrota do campeão mundial de xadrez, Garry Kasparov, em 1997, para o computador Deep Blue, da IBM. Na década anterior, em 1985, Kasparov havia jogado contra 32 computadores simultaneamente e vencido todas as partidas (Kasparov, 2010). O que o fez perder 12 anos depois foi um computador e uma inteligência artificial muito mais potentes: “um supercomputador e um *software* criados pela IBM especialmente para jogar xadrez, com 256 coprocessadores capazes de analisar aproximadamente 200 milhões de posições por segundo” (DeepBlue, s.d.). Além de realizar muitos cálculos por segundo para avaliar matematicamente as consequências de várias jogadas, aquela inteligência artificial também consultava uma base de dados com mais de 700 mil partidas dos melhores jogadores de xadrez.

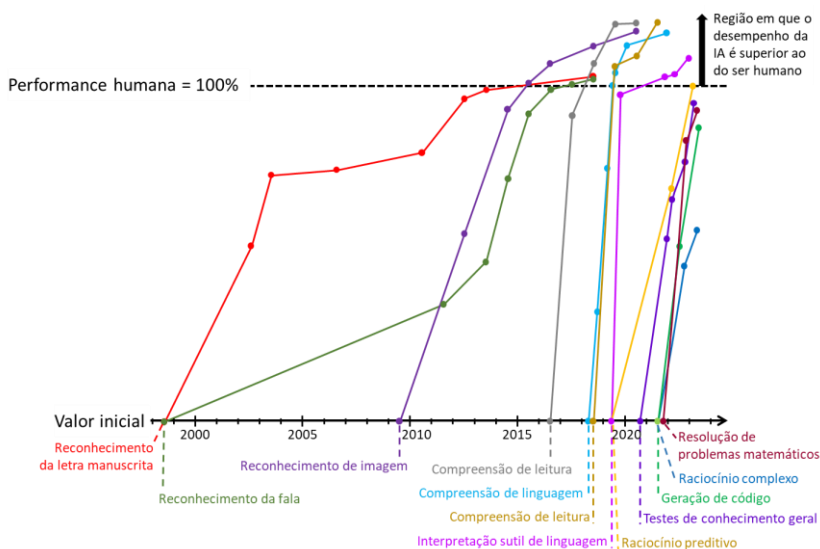
Quanto maior a base de dados, mais exemplos o computador tem sobre como agir bem em determinadas situações; quanto maior a capacidade de processamento, mais movimentos podem ser analisados por segundo. O computador venceu o maior jogador de xadrez porque tinha muitos dados para consultar, grande capacidade de processá-los e um algoritmo para tomar uma boa decisão dentro do tempo limite de cada jogada. O documentário *The man vs the machine* (Marshall, 2014) conta uma versão dessa história:

O xadrez, por milhares de anos, tem sido algo que nós costumávamos usar para medir a inteligência; e se um computador pudesse derrotar um humano que realmente sabe o que está fazendo no tabuleiro de xadrez, então você sabe que estamos falando de 2001: Uma Odisseia no Espaço — os computadores podem fazer qualquer coisa! (*The man vs the machine*, Marshall, 2014, 1:00-1:20)

A IA superou outros campeões mundiais em jogos famosos: em 2011, o sistema inteligente Watson derrotou os melhores jogadores de Jeopardy!, um famoso programa de perguntas e respostas da televisão norte-americana; em 2016, o programa AlphaGo derrotou Lee Sedol, um dos melhores jogadores de Go, um jogo de tabuleiro mais complexo que o xadrez.

A IA já superou a capacidade humana em diversos testes de desempenho (*benchmarks*) e tem aprimorado rapidamente seu desempenho em diferentes áreas: conhecimento geral, interpretação de textos, resolução de problemas matemáticos, entre outras, conforme ilustrado na figura a seguir. A IA vem superando, cada vez mais rapidamente, a capacidade humana nos diferentes testes criados para a sua avaliação. Essa é uma consequência das técnicas de aprendizagem profunda desenvolvidas a partir da década de 2010, dos avanços de *hardware* e dos maciços investimentos realizados em IA.

Capacidade da IA em comparação à humana em diferentes testes



Fonte: (Kiela et al., 2023; Test [...], 2024)

Do ponto de vista da Computação, a IA generativa é considerada inteligente não por pensar como nós, nem por escrever textos melhores que os nossos (ou não), mas sim por ser capaz de aprender a gerar respostas coerentes e contextualmente adequadas, ainda que apresente “alucinações” ou gere respostas imprecisas. O compromisso da IA generativa não é gerar as melhores respostas, os melhores textos, ou acertar sempre; seu compromisso é gerar um novo conteúdo que seja semelhante aos dados usados em seu treinamento.

A inteligência da IA generativa reside na capacidade de simular respostas plausíveis. Essa capacidade resulta do uso de algoritmos de aprendizagem de máquina que processam grandes volumes de dados para extrair padrões linguísticos, tornando-se capazes de simular aspectos da inteligência humana, como a capacidade de interpretar perguntas, formular respostas e lidar com situações imprevistas. No entanto, essa inteligência é estritamente estatística e limitada aos padrões aprendidos durante o treinamento. Portanto, a IA generativa é considerada inteligente mesmo sem compreender o significado das palavras, sem possuir consciência, intenções, emoções ou habilidade de pensar como as pessoas.

10.3 – Comparando a inteligência artificial com a humana: o Teste de Turing

“As máquinas podem pensar?” — essa é a questão que abre o artigo Máquinas de Computação e Inteligência, de Alan Turing (1950). Reconhecendo que as máquinas obviamente processam de modo diferente de nós (porque realizam operações matemáticas-eletrônicas em vez de processos químicos-orgânicos), Turing propôs uma abordagem pragmática para avaliar objetivamente a inteligência da máquina. Ele elaborou o experimento mental que chamou de “Jogo da Imitação”, posteriormente conhecido como Teste de Turing: se os computadores conseguirem imitar o comportamento humano em uma conversação, de tal modo que não consigamos ter certeza se estamos conversando com uma pessoa ou com uma máquina, então poderíamos afirmar que são tão inteligentes quanto nós.

Essa discussão foi abordada no filme O Jogo da Imitação (Tyldum, 2014). No filme, Alan Turing explicou o “Jogo da Imitação” durante o interrogatório em que estava sendo acusado de um suposto ato homossexual, o que era considerado crime na Inglaterra daquela época (a maior parte do filme se passa durante a 2ª Guerra Mundial, mas o interrogatório ocorre em 1951):

Investigador: — As máquinas podem pensar?

Alan Turing: — Ah, andou lendo alguns artigos meus, não é?

— Por que acha isso?

— Bom, porque estou sentado em uma delegacia acusado de pagar um jovem para tocar no meu pênis, e você vem me perguntar se máquinas podem pensar?

— E aí, elas podem? As máquinas podem vir a pensar como seres humanos?

— A maioria das pessoas diz que não.

— O senhor não é a maioria.

— O problema é que está fazendo uma pergunta idiota.

— Estou?

— É claro que as máquinas não podem pensar como as pessoas. Uma máquina é diferente de uma pessoa. Bom, elas pensam diferente. A pergunta interessante é: só porque algo pensa diferente de você, significa que não está pensando? Nós aceitamos que os humanos

tenham tantas divergências entre si: você gosta de morangos e eu detesto patinar, você chora em filmes tristes e eu sou alérgico a pólen. Qual é a vantagem de tantas diferenças, tantas preferências, a não ser dizer que nosso cérebro funciona diferentemente, que pensamos diferentemente? Se podemos dizer isso de nós mesmos, por que não podemos dizer o mesmo para cérebros feitos de cobre e aço?

— E foi esse o artigo que escreveu? Qual o nome dele?

— Se chama “O jogo da imitação”.

— Certo... e é esse o assunto dele?

— Gostaria de jogar?

— Jogar?

— Um jogo. Um teste cego para determinar se algo é uma máquina ou um ser humano.

— E como eu jogo?

— Bom, tem um juiz e um suspeito. O juiz faz perguntas ao suspeito, dependendo das respostas do suspeito, o juiz determina com quem ele está falando, com o que ele está falando. Tudo o que precisa fazer é uma pergunta. (O jogo da imitação, Tyldum, 2014, 01:06:43 – 01:09:18)

O objetivo do Teste de Turing não é provar que um computador é inteligente em termos teóricos, filosóficos ou teológicos, mas sim discutir os efeitos que provoca em nós: se não conseguimos diferenciá-lo de uma pessoa, a questão sobre ele ser realmente inteligente torna-se irrelevante em termos práticos. Será que o ChatGPT consegue passar nesse teste?

A capacidade do ChatGPT de passar no Teste de Turing é um tópico em debate. De um lado, o modelo tem mostrado uma habilidade impressionante de compreender e responder perguntas em língua natural de maneira coerente e relevante, muitas vezes confundindo pessoas pela semelhança de suas respostas às respostas humanas. No entanto, o GPT ainda tem limitações: ele não tem consciência ou compreensão semântica, nem pode aprender ou adaptar-se a novas informações após seu ponto de corte de conhecimento (junho de 2021, no caso do GPT-3.5). Ademais, algumas de suas respostas podem revelar-se incoerentes ou factualmente incorretas quando questionado em profundidade ou sobre tópicos mais complexos. Portanto, embora possa passar em versões mais superficiais do Teste de Turing, ainda existe um longo caminho a percorrer para que seja indistinguível de um humano em todas as situações.

Você percebeu que o parágrafo anterior foi escrito pelo ChatGPT? Se passou despercebido, então o ChatGPT passou no Teste de Turing, pois aquele texto foi gerado como resposta ao nosso pedido: “Escreva um parágrafo discutindo se o ChatGPT passaria no Teste de Turing.” (OpenAI, 2023f). Para você não se sentir mais enganada/o, garantimos que os demais parágrafos deste texto foram todos escritos por nós, autores humanos... mas agora você já não tem certeza se as outras partes foram realmente escritas por nós ou pelo ChatGPT? Pois é... no jogo da linguagem, ele nos imita bem. Ainda por cima, produz textos velozmente, muito mais rápido do que nossa capacidade de leitura.

É tentador pensar que o Watson [uma IA desenvolvida pela IBM] “deveria” raciocinar como os humanos. Mas eu diria que isso é superstição. No mundo real, o que importa é como um ser inteligente age. Se diferentes processos computacionais levarem uma futura IA a fazer descobertas científicas inovadoras ou a escrever romances comoventes, por que deveríamos nos importar com o modo como essas façanhas foram engendradas? E se uma IA é capaz de proclamar com eloquência a sua própria consciência, que bases éticas poderíamos ter para insistir que apenas a nossa própria biologia pode dar origem a uma consciência que valha a pena? O empirismo do Teste de Turing coloca firmemente o nosso foco no lugar onde deveria estar. (Kurzweil, 2024, p. 90)

Os pesquisadores Jones e Bergen (2024) avaliaram os modelos GPT e concluíram que “As pessoas não conseguem distinguir o GPT-4 de uma pessoa em um Teste de Turing”. Para nós, autores deste livro e professores universitários, desde o lançamento do ChatGPT tem se tornado cada vez mais difícil afirmar se um texto apresentado por um/a estudante foi escrito somente por ela/e, somente pela IA, ou se é um resultado de cocriação humano-IA.

O Teste de Turing, embora amplamente conhecido e citado, não é necessariamente o teste mais adequado para avaliarmos a inteligência de uma IA. Para que o comportamento de uma tecnologia generativa fosse indistinguível do humano, seria necessário ajustar algumas de suas capacidades: reduzir a velocidade da geração dos textos, diminuir seu

conhecimento geral e limitar seu desempenho nos diversos testes até que se equiparasse ao nível médio das pessoas. Obviamente, uma IA com as mesmas dificuldades e limitações humanas não seria tão útil quanto uma inteligência notadamente superior, mais rápida e com mais conhecimentos do que nós em diversas áreas. Por exemplo, no Enem, Enade, Exame da OAB ou outras avaliações, se a IA acertasse a mesma quantidade de questões que a média das pessoas, ela não seria muito útil para quem estivesse se preparando para fazer esses testes; mas se ela acertasse todas as questões e soubesse explicar como resolver cada uma delas, então seria útil para absolutamente todas/os as/os estudantes.

Como demonstraram sistemas como PaLM 2 e GPT-4, as máquinas podem superar os humanos em tarefas cognitivamente exigentes sem serem capazes de imitar de forma convincente um humano em outros domínios. Entre 2023 e 2029, o ano em que eu espero que um programa passe no primeiro Teste de Turing robusto, os computadores alcançarão capacidades sobre-humanas numa gama cada vez maior de áreas. Na verdade, é até possível que a IA consiga atingir um nível de habilidade sobre-humano na programação antes de dominar as sutilezas sociais de senso comum exigidas no Teste de Turing. Trata-se de uma questão que ainda não foi resolvida, mas a possibilidade mostra por que a nossa noção de inteligência em nível humano precisa ser profícua e matizada. O Teste de Turing é certamente uma parte importante disso, mas precisaremos também desenvolver meios mais sofisticados de avaliar as formas complexas e variadas pelas quais a inteligência humana e a inteligência da máquina serão semelhantes e diferentes. (Kurzweil, 2024, p. 90-91).

Depois que o GPT demonstrou ser capaz de passar no Teste de Turing sob certas condições, ficou evidente que desejamos mais do que uma simples imitação da conversação humana. O foco deslocou-se da busca por máquinas que nos imitem para o desenvolvimento de tecnologias mais inteligentes do que nós. A sociedade atual não busca uma inteligência artificial indistinguível da humana em suas imperfeições e subjetividades, mas sim uma que seja capaz de resolver problemas complexos, inovar e expandir os horizontes do possível.

10.4 – Podemos concluir que a IA generativa é realmente inteligente?

Para o professor citado no início deste capítulo, não importava o algoritmo por trás do ChatGPT. Não era suficiente saber que o *chatbot* havia sido construído com técnicas de IA. O que ele realmente queria saber era se a IA generativa pensa como nós, se é inteligente de fato, se entende o que está escrevendo, se tem consciência.

As técnicas de inteligência artificial avançaram muito nas décadas de 2010 e 2020, mas não o suficiente para desenvolver uma consciência computacional. Contudo, pode parecer que sim, especialmente porque a IA generativa fala de si como se compreendesse a própria existência, suas habilidades e características. As expressões geradas na resposta do ChatGPT transcrita no início deste capítulo exemplificam essa aparente autoconsciência: “eu sou...”, “[fui] criado por...”, “tenho...”, “eu sou projetado para...”, “eu posso...”, “eu não tenho...”, “minha inteligência é...”, “minha inteligência não é...”, “eu não sou...”. Ainda não chegamos ao ponto de desenvolver uma inteligência artificial autoconsciente, mas podemos ser levadas/os a acreditar que sim, como aconteceu com o engenheiro de software do Google, Blake Lemoine, que afirmou que “LaMDA é senciente”. Suas crenças foram rejeitadas e ele foi demitido (G1, 2022).

Alguns filmes e obras literárias de ficção científica já discutiram um futuro possível em que computadores passariam a ter consciência, como o HAL 9000, personagem ficcional de Arthur C. Clarke, que protagonizou o filme 2001: uma Odisseia no Espaço. Já no filme Ex-Machina: Instinto Artificial (Garland, 2015), um programador é convidado a realizar uma avaliação inspirada no Teste de Turing: “O desafio é mostrar-lhe que ela é um robô e saber se, ainda assim, você acha que ela tem consciência”.

Para o renomado linguista Chomsky (2023), que desenvolveu a teoria da gramática generativa para explicar a capacidade humana inata de criar infinitas sentenças gramaticalmente corretas a partir de um conjunto limitado de regras, a IA generativa apenas simula a proficiência linguística sem qualquer compreensão, portanto, não pode ser considerada verdadeiramente inteligente. Essa também é a posição de Searle (1980), que

formulou o experimento mental do Quarto Chinês⁴⁴ para explicar que um computador pode gerar textos sem compreender o significado das palavras, criando a ilusão de compreender uma língua.

Não, a IA generativa não entende o significado dos textos que ela mesma produz, não tem intencionalidade, ideologia ou desejos, tampouco consciência. O conhecimento verdadeiro — o conhecimento sobre o conhecimento, o conhecimento de segunda ordem, o conhecimento epistêmico — já era uma questão debatida pelos filósofos da Grécia antiga. O personagem Sócrates-platônico, no livro V da República (Platão, 1996), escrito na década de 370 a.C., já desenvolveu uma teoria do conhecimento diferenciando opinião de ciência, sendo que a primeira se refere às aparências, obtida pelo exercício da visão, e a segunda se refere ao conhecimento sobre as ideias, obtido pelo exercício da razão; trata-se de uma distinção ontológica entre o sensível e o inteligível (LaFrance, 2011). Esse conhecimento verdadeiro, os modelos de linguagem ainda não têm; por isso, às vezes, geram informações equivocadas e podem afirmar outra coisa na resposta seguinte sem qualquer constrangimento.

Embora a IA não tenha consciência até o momento em que escrevemos este livro, será que ela poderá vir a ter no futuro?

Suponha que futuros sistemas de IAG [Inteligência Artificial Geral] (na tela ou em robôs) tivessem um desempenho equiparável ao do ser humano. Será que eles teriam uma verdadeira inteligência, um verdadeiro discernimento, uma verdadeira criatividade? Será que teriam egos, postura moral, livre-arbítrio? Seriam conscientes? E, sem a consciência, será que poderiam ter qualquer uma dessas outras características?

⁴⁴ O experimento do quarto chinês consiste em uma pessoa trancada em um quarto com um conjunto de regras sobre como manipular caracteres chineses, sendo que as regras estão escritas em uma língua que a pessoa compreende. Quando um texto em Chinês é passado por debaixo da porta, a pessoa segue as regras mecanicamente para formular uma resposta em Chinês, que é então devolvida por debaixo da porta. Como a resposta tem coerência, a pessoa do lado de fora acredita que a pessoa dentro do quarto compreende Chinês. Assim como a pessoa no quarto, a IA gera um novo texto realizando operações a partir do *prompt* do usuário (o texto inicial passado por debaixo da porta), sem compreender o significado das palavras recebidas e geradas.

Essas perguntas não são científicas, são filosóficas. Muitas pessoas sentem intuitivamente que a resposta, em cada um dos casos, é: "Claro que não!".

No entanto, as coisas não são tão simples assim. Precisamos de argumentos meticulosos, não apenas de intuições precipitadas. Mas esses argumentos mostram que não existem respostas indiscutíveis para essas perguntas. Isso porque os próprios conceitos em questão são extremamente controversos. Somente se todos eles fossem satisfatoriamente compreendidos poderíamos estar seguros de que a hipotética IAG seria realmente inteligente ou não. Em suma: ninguém tem 100% de certeza.

Alguns podem dizer que isso não importa: o que importa é o que as IAGs farão realmente. (Boden, 2020, p.165-166)

Não devemos confundir inteligência com compreensão de mundo ou consciência. A IA generativa é inteligente sem compreender as palavras que escreve e sem ter consciência.

À medida que se tornam mais inteligentes, os computadores podem acabar desenvolvendo consciência e tendo algum tipo de experiência subjetiva. Mas podem da mesma forma vir a ser muito mais inteligentes do que nós, sem jamais desenvolverem sentimento algum. Já que não entendemos completamente como a consciência surge em formas de vida baseadas no carbono, não é possível prever se ela emergiria de entidades não orgânicas. Talvez a consciência não tenha nenhum vínculo essencial com a bioquímica orgânica, e nesse caso computadores conscientes podem surgir a qualquer momento. Ou talvez haja vários caminhos alternativos que levam à superinteligência, e apenas alguns desses caminhos envolvam a aquisição de consciência. Assim como aviões voam muito mais rápido do que pássaros sem precisar desenvolver penas, pode ser que os computadores venham a resolver problemas bem melhor do que os seres humanos, sem que jamais desenvolvam sentimentos. [...] Via de regra, as pessoas definem e avaliam a IA pela métrica da "inteligência no nível humano", e muito se debate sobre quando podemos esperar que as IAs atinjam a "inteligência no nível humano". No entanto, o uso dessa métrica é terrivelmente confuso. É como definir e avaliar aeronaves pela métrica do "voo no nível do pássaro". A IA não está avançando em direção à inteligência de nível humano. Ela evolui em direção a um tipo de inteligência totalmente diferente. (Harari, 2024, p. 257; 275)

As pessoas tendem a confundir inteligência com consciência, de modo que muitas delas concluem que entidades não conscientes não podem ser inteligentes. Mas inteligência e consciência são coisas bem diferentes. Inteligência é a capacidade de alcançar objetivos, como maximizar o envolvimento do usuário numa plataforma de rede social. Consciência é a capacidade de ter sentimentos subjetivos como dor, prazer, amor e ódio. Nos humanos e em outros mamíferos, a inteligência costuma andar de mãos dadas com a consciência. [...] No entanto, é um erro extrapolar de humanos e mamíferos em geral para todas as entidades possíveis. Bactérias e plantas aparentemente não têm consciência, mas também demonstram inteligência. Coletam informações do ambiente, fazem escolhas complexas e adotam estratégias engenhosas para obter alimentos, se reproduzir, cooperar com outros organismos e escapar de predadores e parasitas. Até os humanos tomam decisões inteligentes sem ter consciência delas; 99% dos processos em nosso corpo, da respiração à digestão, acontecem sem que haja nenhuma tomada de decisão consciente. [...] O mesmo ocorre com os computadores. Embora não sintam dor, amor ou medo, os computadores são capazes de tomar decisões que maximizam o envolvimento do usuário e ainda podem afetar importantes acontecimentos históricos. (Harari, 2024, p. 297)

A inteligência não é um objeto físico que possa ser mensurado diretamente; é uma construção teórica que recebe sentidos específicos em diferentes áreas do conhecimento, como na Biologia, Psicologia, Neurociência, Teologia, Educação, Filosofia e também na Ciência da Computação. Quando a inteligência é definida como um conjunto de habilidades — como planejamento, previsibilidade, associação, raciocínio, inferência, analogia, entre outras —, podemos reconhecer que algumas dessas habilidades podem ser imitadas pela IA (Boden, 2020). Essa também é a conclusão de Santaella (2023a, p. 163), para quem “a IA é inteligente porque o computador adquiriu o potencial de aprender e tomar decisões com base nas informações que recebe”.

No capítulo 12, explicamos o funcionamento da IA, pois quem conhece em mais detalhes como tudo funciona sabe que a IA generativa não pensa como nós, não compreende o que está escrevendo, não tem consciência.

10.5 – Então não somos as/os únicas/os *sapiens*?

Para algumas pessoas, a IA generativa representa uma ameaça à nossa própria condição de ser humano; ao menos para aquelas pessoas que desejam se colocar acima de outros seres: “somos os únicos seres inteligentes” (*Homo sapiens*), “os únicos capazes de produzir artefatos” (*Homo habilis*), “os únicos capazes de se comunicar pela linguagem” (*Homo loquens*), “os únicos criativos” (*Homo creativus*).

As/Os biólogas/os certamente discordariam de que essas características sejam exclusivas da nossa espécie. Outros animais também demonstram inteligência, como primatas, golfinhos, elefantes, polvos, cães, gatos, entre outros. A inteligência animal é avaliada por múltiplos critérios, como capacidade de memória, comunicação, aprendizagem, resolução de problemas, uso de ferramentas e comportamentos sociais complexos. Estudos nas áreas de biologia, cognição, neurociência e comportamento têm demonstrado que outras espécies também são capazes de produzir e usar ferramentas, resolver problemas, planejar o futuro, sonhar, ter emoções e empatia, apresentar comportamentos culturais e até mesmo autoconsciência. Por exemplo, elefantes usam sons para chamar outros indivíduos pelo nome e respondem quando ouvem seu próprio nome. As abelhas são capazes de informar, com precisão, onde se encontram fontes de alimento por meio da “dança das abelhas”, um comportamento comunicativo que inclui movimentos específicos para indicar a direção e a distância de um alvo em relação à colmeia. Golfinhos e baleias utilizam complexos sistemas de comunicação, produzindo uma variedade de sons — como assobios, cliques e chamadas — que transmitem informações sobre localização, estado emocional e intenções; essas vocalizações são fundamentais para socialização, caça e orientação, demonstrando a elevada inteligência e a capacidade de colaboração dessas espécies marinhas.

Podemos reconhecer que os seres humanos apresentam um grau mais complexo de inteligência, de habilidades comunicativas e de produção de artefatos mais elaborados; contudo, trata-se de um contínuo em termos de complexidade, e não capacidades exclusivas da nossa espécie. Todas as tentativas de nos colocarmos acima de outras espécies acabaram malsucedidas com o tempo e revelaram o quanto somos pretensiosas/os e narcisistas.

Sigmund Freud, no ensaio *Uma Dificuldade no Caminho da Psicanálise* (1917), listou três "feridas narcísicas" que abalaram o orgulho humano ao longo da história. A primeira ocorreu com a teoria heliocêntrica de Copérnico, que deslocou a Terra do centro do universo, tirando também o ser humano de uma posição privilegiada no cosmos. A segunda ferida foi causada por Charles Darwin, ao demonstrar que o ser humano evoluiu de outras formas de vida, partilhando ancestrais comuns com outros animais. Por fim, a própria psicanálise de Freud constituiu a terceira ferida, ao mostrar que não somos plenamente senhoras/es de nossos pensamentos e desejos, pois forças inconscientes modelam nosso comportamento e vida psíquica. As derrotas do narcisismo humano nos levam a questionar a imagem idealizada que cultivamos de nós mesmas/os. Agora, a IA pode representar uma nova ameaça à identidade humana, como nos alertam Russell e Norvig (2022, n.p.):

As pessoas poderiam perder seu sentido de identidade. Em *Computer Power and Human Reason*, Weizenbaum (1976), o autor do programa ELIZA, destaca algumas das ameaças potenciais que a IA representa para a sociedade. Um dos principais argumentos de Weizenbaum é que a pesquisa de IA torna possível a ideia de que os seres humanos são autômatos – uma ideia que resulta em perda de autonomia ou até de humanidade. Observamos que a ideia tem estado presente há muito mais tempo que a IA, desde a época de *L'Homme Machine* (La Mettrie, 1748). Também observamos que a humanidade sobreviveu a outros retrocessos em nosso senso de identidade: *De Revolutionibus Orbium Coelestium* (Copérnico, 1543) afastou a Terra do centro do sistema solar, enquanto *A descendência do homem* (Darwin, 1871) colocou o *Homo sapiens* no mesmo nível das outras espécies. A IA, se for amplamente bem-sucedida, poderá ser pelo menos tão ameaçadora para as suposições morais da sociedade do século XXI como a teoria da evolução de Darwin foi para as suposições morais do século XIX.

Considerando as intensas reações aos avanços da IA generativa, Lucia Santaella e Dora Kaufman (2024, p. 46) concluíram que ela já está causando uma nova ferida narcísica na humanidade:

A IA generativa, por seu lado, penetrou no segredo mais íntimo do humano: o potencial linguístico e semiótico que, até então, se constituía em marca inimitável do *Sapiens*. [...] É justamente aí, no fato de o humano se sentir de algum modo igualado em sua pretensão-chave de exclusividade, que se tem municiado a explosão sociocultural e mesmo psíquica que a IA generativa vem provocando. E isso nos leva a propor e defender a hipótese de que, com a IA no ponto em que hoje se encontra, a humanidade está em plena travessia de sua quarta ferida narcísica.

Em vez de termos perder nossa posição exclusiva de seres inteligentes na face da Terra, podemos nos reconhecer como seres híbridos, com capacidades intelectuais estendidas pelo uso das tecnologias inteligentes. A ideia de sermos híbridos nos possibilita lançar um novo olhar sobre a própria condição humana e sobre o papel da IA em nossa sociedade. Retomaremos essa discussão na Parte III deste livro.

O que esperar do futuro?

O que será o amanhã?

Como vai ser o meu destino?

O que irá me acontecer?

Responda quem puder

(O amanhã - João Sergio da Silva Filho)

Como os futuros sistemas de IA funcionarão? Ainda não podemos dizer. [...] O campo tem adotado várias histórias sobre si mesmo — primeiro a ideia ousada de que a inteligência por uma máquina seria possível e, então, que ela poderia ser alcançada codificando em lógica o conhecimento especialista; mais adiante, que modelos probabilísticos do mundo seriam a ferramenta principal e, mais recentemente, que o aprendizado de máquina induziria modelos que podem não ser baseados em qualquer teoria bem compreendida. O futuro revelará o modelo a seguir. (Russell; Norvig, 2022, n.p.)

Estamos vivenciando um avanço acelerado, parcialmente explicado pelo crescimento exponencial da capacidade de processamento computacional, conforme observado empiricamente e caracterizado pela Lei de Moore: o número de transistores em um circuito integrado (como um microprocessador) dobra a cada dois anos mantendo o mesmo custo, o que resulta em computadores cada vez menores e mais rápidos ao longo do tempo. Esse avanço do *hardware* vem possibilitando a resolução de problemas cada vez mais complexos.

Não bastam máquinas mais rápidas: “sem a teoria certa, máquinas mais rápidas apenas dão a resposta errada mais rapidamente” (Russell; Norvig, 2022, n.p.). Além do *hardware*, o conhecimento sobre inteligência artificial também avança de forma acelerada. A comunidade de cientistas de IA cresce a cada ano, trabalhando incansavelmente no aperfeiçoamento das

técnicas e na teorização da área. A quantidade de artigos com a expressão “*artificial intelligence*” também cresce exponencialmente (Pimentel; Carvalho; ChatGPT-4, 2023).

Se a evolução dos computadores e da IA continuar nesse ritmo acelerado, o que podemos esperar do futuro?

11.1 – Singularidade tecnológica

O objetivo da área de inteligência artificial é desenvolver sistemas que simulem as capacidades cognitivas humanas, incluindo raciocínio, tomada de decisão, percepção, aprendizagem, linguagem e emoções. Com os avanços da área, espera-se que, um dia, seja desenvolvida uma IA equiparável à inteligência humana, capaz de realizar qualquer tarefa intelectual que uma pessoa possa executar, o que denominamos de Inteligência Artificial Geral (IAG, no inglês *Artificial General Intelligence* - AGI). Continuando o processo de desenvolvimento tecnológico, espera-se que seja desenvolvida uma inteligência muito superior à humana. Esse ponto é frequentemente descrito como singularidade tecnológica, um marco hipotético em que será criada uma SuperInteligência Artificial (SIA, no inglês ASI - *Artificial SuperIntelligence*).

Alan Turing já profetizava que, um dia, um computador avançado poderia desempenhar o “jogo da imitação” de forma indistinguível de uma pessoa, momento em que “as máquinas eventualmente competirão com os homens em todos os campos puramente intelectuais” (Turing, 1950, p. 20, tradução nossa). Embora o termo IAG não existisse na época, as ideias de Turing lançaram as bases para se pensar a criação de uma inteligência equiparável à humana.

Irving John Good, que trabalhou com Turing, foi além. Em 1966, Good conjecturou que chegaríamos a um ponto em que seria possível desenvolver uma máquina superinteligente:

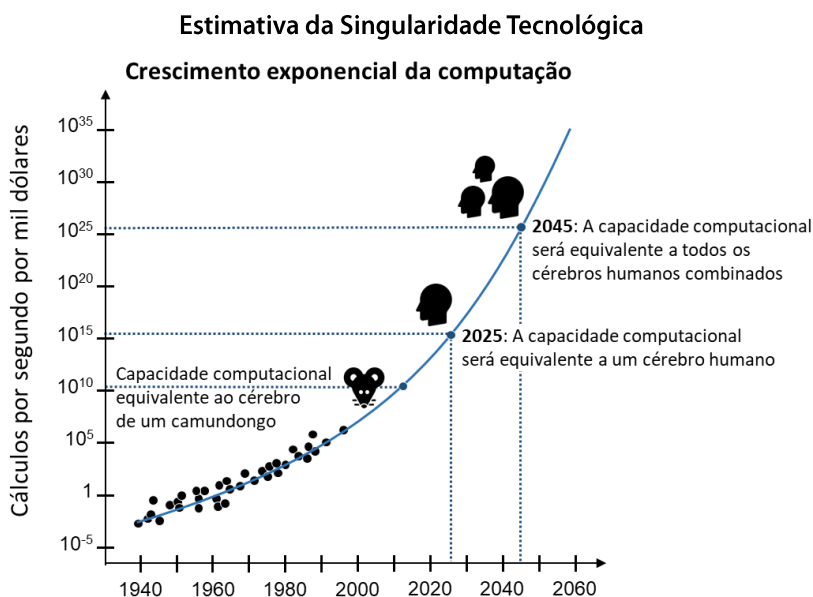
Seja máquina superinteligente definida como uma máquina capaz de superar amplamente todas as atividades intelectuais de qualquer homem, por mais inteligente que seja. Como o projeto de máquinas é uma dessas atividades intelectuais, uma máquina superinteligente

poderia projetar máquinas ainda melhores; haveria, então, indiscutivelmente, uma “explosão de inteligência”, e a inteligência humana ficaria muito aquém. Assim, a primeira máquina superinteligente é a última invenção que o homem precisará fazer, desde que a máquina seja dócil o suficiente para nos dizer como mantê-la sob controle. [...] É mais provável do que improvável que, durante o século XX, uma máquina superinteligente seja construída e que seja a última invenção necessária para a humanidade, já que ela levará a uma “explosão de inteligência”. (Good, 1966, p. 33, 78, tradução nossa).

O termo singularidade, utilizado no contexto computacional, é atribuído a John von Neumann que, em uma conversa sobre o progresso cada vez mais acelerado da tecnologia, considerou “estar se aproximando de alguma singularidade essencial na história da humanidade, além da qual os assuntos humanos, como os conhecemos, não poderiam continuar.” (Ulam, 1958, p. 5). O conceito de singularidade tecnológica foi amplamente difundido a partir do artigo A Singularidade Tecnológica Vindoura (*The Coming Technological Singularity*), publicado por Vernor Vinge em 1993, onde declarou: “Dentro de trinta anos [até 2023], teremos os meios tecnológicos para criar inteligências sobre-humanas. Pouco depois, a era humana chegará ao fim.” (Vinge, 1993, p. 11, tradução nossa). Nesse artigo, o autor considerava necessário “termos um *hardware* tão poderoso quanto o cérebro humano” para “sermos capazes de criar uma inteligência equivalente à humana (ou superior).” Ele também arriscou uma data: “Para evitar ambiguidades, serei mais específico: ficarei surpreso se esse evento ocorrer antes de 2005 ou após 2030.” (Vinge, 1993, p. 12-13, tradução nossa).

O desenvolvimento de uma máquina superinteligente não ocorreu durante o século XX, como especulado por Good, nem até 2023, como suposto por Vinge. Kurzweil, em 2005 e 2024, calculou que a singularidade tecnológica acontecerá por volta de 2045. Seu argumento está baseado no que denomina Lei dos Retornos Acelerados, segundo a qual o progresso tecnológico ocorre de forma exponencial, sendo que cada avanço cria condições para um desenvolvimento ainda mais rápido. Kurzweil compreende que a Lei de Moore é uma manifestação específica desse princípio mais amplo. Ele estimou que os computadores apresentam um crescimento exponencial ao longo dos anos em sua

capacidade de realizar cálculos por segundo, o que possibilita traçar a curva da figura a seguir. Também estimou que o cérebro humano tem capacidade de realizar 10^{16} a 10^{19} cálculos por segundo. A partir dessas estimativas, concluiu que a capacidade computacional para realizar cálculos seria equivalente à do cérebro humano em 2025. Ainda de acordo com esse gráfico, concluiu também que, em 2045, a capacidade dos computadores será equivalente à de todos os cérebros humanos conectados, sendo essa uma possível data para caracterizar a singularidade tecnológica.



Fonte: (Kurzweil, 2018, p. 221, tradução nossa)

Em meados dos anos 2040, aqueles mil dólares de computação serão iguais a 10^{26} cps [cálculos por segundo]; assim, a inteligência criada [...] será cerca de 1 bilhão de vezes mais potente do que toda a inteligência humana hoje. Isso de fato representará uma mudança profunda, e é por essa razão que marco a data para a Singularidade — que representa uma transformação profunda e perturbadora da capacidade humana — como sendo 2045. (Kurzweil, 2018, p. 220-221)

Para Kurzweil, não se trata de um evento pontual que ocorrerá somente em 2045, mas sim de um desenvolvimento contínuo que, inevitavelmente, levará os computadores a superar a inteligência humana:

Ao longo da próxima década, iremos interagir com inteligências artificiais capazes de nos convencer de que são humanas, e simples interfaces cérebro-computador terão enorme impacto na vida cotidiana, assim como os *smartphones* hoje. Uma revolução digital na biotecnologia curará doenças e prolongará, de maneira significativa, a vida saudável das pessoas. Ao mesmo tempo, porém, muitos trabalhadores sentirão a dor aguda da ruptura econômica, e todos nós enfrentaremos riscos decorrentes da utilização indevida, acidental ou deliberada, dessas novas capacidades. Na década de 2030, o autoaperfeiçoamento da IA e o amadurecimento da nanotecnologia farão a interação entre seres humanos e máquinas como nunca na história, aumentando ainda mais tanto a promessa quanto os riscos e desafios. Se conseguirmos enfrentar os desafios científicos, éticos, sociais e políticos acarretados por esses avanços, até 2045 transformaremos profundamente, e para melhor, a vida na Terra. No entanto, se falharmos, nossa própria sobrevivência estará em xeque. Portanto, este livro trata da nossa aproximação final da Singularidade — as oportunidades e os perigos que devemos enfrentar juntos ao longo da última geração do mundo tal como o conhecemos. (Kurzweil, 2024, p. 10-11)

O termo “singularidade” é usado para falar de um futuro em que não temos como prever as consequências do avanço da IA. Trata-se de uma metáfora emprestada da Física e da Matemática, em que singularidade é um ponto ou região em que uma função ou equação se torna indefinida, imprevisível ou não comporta um valor numérico específico. O termo é usado, por exemplo, para descrever o ponto no centro de um buraco negro, onde a curvatura do espaço-tempo se torna infinita, com massa concentrada em um único ponto e onde as leis da física não se aplicam, impossibilitando a compreensão e previsão do que acontece ali. Na Matemática, a divisão por zero é uma singularidade, pois resulta em um valor indefinido⁴⁵, rompendo as regras usuais das operações aritméticas,

⁴⁵ Os limites da função $1/x$ tendem a $+\infty$ quando nos aproximamos de zero pela direita (valores positivos) e a $-\infty$ quando nos aproximamos pela esquerda (valores negativos), criando comportamentos divergentes e indefinidos que não podem ser resolvidos com um único valor.

sendo impossível determinar um resultado finito. Não temos como imaginar tudo o que pode acontecer quando a IA se tornar muito mais inteligente que os seres humanos. “Uso a metáfora da singularidade porque ela captura nossa incapacidade de, com o nosso atual grau de inteligência, compreender essa drástica mudança.” (Kurzweil, 2024, p. 6-7)

Como lidaremos com a superação da inteligência humana pelas máquinas em todos os campos do saber? As consequências são de assombrar. Nós, autores deste livro, somos professores universitários; boa parte de nosso trabalho consiste em dar aulas, orientar, produzir e disseminar conhecimento científico. Será que ainda seremos úteis quando a IA for capaz de explicar conceitos melhor do que nós, elaborar melhores aulas, criar planos de estudo personalizados para cada estudante, orientar bem todas as etapas de uma pesquisa, analisar dados e redigir relatórios científicos com mais qualidade do que nós? Teremos tempo de nos aposentar como professores, ou a singularidade chegará antes? E quanto às demais profissões? Médicas/os, juízas/es, advogadas/os, engenheiras/os, desenvolvedoras/es de sistemas — todas essas pessoas se tornarão obsoletas?

Como nossa sociedade capitalista se reorganizará quando os trabalhos intelectuais forem realizados por sistemas computacionais superinteligentes? Quando o trabalho executado por IA for superior ao das pessoas e a um custo menor, a classe trabalhadora não conseguirá mais vender sua força de trabalho, tampouco receberá salário e, conseqüentemente, não poderá consumir os produtos e serviços gerados por IA. O sistema capitalista entrará em colapso? Conseguiremos nos reinventar sem um trabalho-profissão, que muitas vezes marca a nossa própria identidade?

A vida é o trabalho
E sem o seu trabalho
O homem não tem honra
E sem a sua honra
Se morre, se mata
Não dá pra ser feliz
(Gonzaguinha, Guerreiro Menino, 1983)

Tendo em vista esse cenário, vem ganhando força a ideia de uma Renda Básica Universal: um pagamento regular e incondicional a todas as pessoas. Será necessário pensar em uma nova forma de organização econômica pós-capitalista; caso contrário, tudo pode terminar em guerra, agravando ainda mais desigualdades abissais e conflitos globais. Nesse contexto, repensar o papel das políticas públicas e da redistribuição de renda é um dos caminhos possíveis para garantir a estabilidade social e a dignidade humana em um mundo cada vez mais automatizado e desigual. Os rumos da economia e do trabalho precisam ser debatidos de maneira ampla e inclusiva, envolvendo governos, instituições, empresas e a sociedade civil. A construção de um futuro mais justo dependerá da capacidade coletiva de formular soluções inovadoras, sustentáveis e comprometidas com o bem-estar de todas as pessoas, especialmente das mais vulneráveis.

As implicações do avanço da IA são muito preocupantes. Harari, em seu livro *Nexus*, refletiu sobre a possibilidade de a IA ser empregada para fortalecer regimes totalitários. Ele alerta que essa tecnologia pode ser utilizada para monitorar, manipular e controlar populações com extrema eficiência, tornando esses regimes quase invulneráveis devido à sua capacidade de prever comportamentos, anular dissidências e manter um controle absoluto sobre a sociedade. O uso descontrolado da IA por lideranças autoritárias pode desencadear catástrofes políticas e sociais em escala global.

Primeiro, o poder da IA poderia superalimentar conflitos humanos existentes, dividindo a humanidade contra si mesma. Assim como, no século XX, a Cortina de Ferro dividiu as potências rivais na Guerra Fria, no século XXI a Cortina de Silício — feita de chips de silício e códigos de computador, em vez de arame farpado — poderia vir a dividir potências rivais num novo conflito global. Como a corrida armamentista da IA produzirá armas sempre mais destrutivas, uma mera faísca poderia desencadear uma conflagração cataclísmica.

Segundo, a Cortina de Silício poderia criar uma divisão não entre grupos humanos, mas entre todos os seres humanos e nossos novos senhores da IA. Onde quer que vivêssemos, poderíamos nos ver enclausurados numa rede de algoritmos insondáveis regendo nossa vida, remodelando nossa política e nossa cultura, e até reprogramando nosso corpo e nossa mente — não mais capazes de

compreender as forças que nos controlam e muito menos de detê-las. Se uma rede totalitária do século XXI conseguir conquistar o mundo, ela poderá ser comandada por uma inteligência não humana em lugar de um ditador humano. As pessoas que apontam a China, a Rússia ou os Estados Unidos pós-democráticos como principal fonte dos pesadelos totalitários não entendem bem o perigo. Na verdade, chineses, russos, americanos e todos os outros seres humanos estão conjuntamente ameaçados pelo potencial totalitário da inteligência não humana. (Harari, 2024)

Os avanços da IA causam otimismo e temor porque não sabemos ao certo qual será nosso destino à medida que a sociedade vai sendo reconfigurada por sistemas computacionais cada vez mais inteligentes.

11.2 – A singularidade tecnológica é balela?

Essa ideia é extremamente polêmica. [...] Os partidários da singularidade argumentam que os avanços da IA tornam a singularidade inevitável. Alguns dão as boas-vindas a isso, prevendo que os problemas da humanidade serão resolvidos. Guerra, doenças, fome, tédio e até mesmo a morte pessoal... tudo banido. Outros preveem o fim da humanidade — ou, de qualquer modo, da vida civilizada como a conhecemos. [...] Por outro lado, os descrentes na singularidade não esperam que ela aconteça — e, certamente, não em um futuro próximo. Eles admitem que a IA oferece muitos motivos de preocupação, mas não veem uma ameaça existencial. (Boden, 2020, p. 199-200)

A plausibilidade da singularidade tecnológica é objeto de intensos debates. A previsão de Kurzweil, de que a singularidade ocorrerá por volta de 2045, tornou-se um marco nas discussões sobre o futuro da IA e da tecnologia. Entretanto, essa estimativa é considerada especulativa, com fundamentos matemáticos e científicos questionáveis. Mesmo que se alcance um poder computacional equivalente à capacidade de processamento do cérebro humano, isso não garante que teremos o conhecimento necessário para simular a mente humana em toda a sua complexidade.

As pessoas céticas da singularidade acusam de extravagantes e improváveis as especulações a seu respeito. As defensoras da singularidade “insistem, no entanto, que ‘nunca’ é muito tempo. De fato, *nunca* é muito tempo. Portanto, os céticos da singularidade, entre os quais eu me incluo, podem estar errados.” (Boden, 2020, p. 209)

A crença de que a singularidade tecnológica é inevitável reflete uma forma extrema de determinismo tecnológico. Essa perspectiva presume que as mudanças sociais e culturais são impulsionadas quase exclusivamente pela evolução tecnológica, negligenciando a influência de fatores humanos, políticos e éticos. O determinismo tecnológico é frequentemente criticado por ser reducionista, minimizando o papel da agência humana e das complexas interações entre sociedade e tecnologia.

Limitações técnicas, econômicas, físicas, ambientais e sociais podem frear ou até mesmo impedir os avanços da IA. Tomemos como exemplo a evolução dos modelos GPT. As primeiras versões dessa família de modelos apresentaram crescimento acelerado, com aumentos exponenciais no número de parâmetros e na quantidade de dados utilizados para o treinamento. Será que as próximas gerações desses modelos continuarão crescendo exponencialmente?

Vários fatores podem limitar o crescimento exponencial dos modelos de linguagem, incluindo o custo para o desenvolvimento de modelos maiores. Embora a OpenAI não divulgue oficialmente os custos envolvidos, estimativas apontam para um aumento também exponencial de despesas (Doria, 2024; Kim, 2024). As empresas de tecnologia estarão dispostas a investir bilhões de dólares para desenvolver as próximas gerações de modelos de linguagem? Conseguirão obter retorno sobre o investimento? É possível que o crescimento exponencial dos modelos não seja economicamente viável.

Além dos custos, há limitações técnicas inerentes ao aumento dos modelos. A partir de certo ponto, adicionar mais parâmetros gera apenas melhorias marginais, enquanto os custos computacionais e energéticos crescem significativamente. Isso levanta dúvidas sobre a eficiência e a viabilidade de seguir ampliando os modelos. Ademais, modelos

excessivamente grandes tendem a memorizar dados específicos do conjunto de treinamento em vez de aprender padrões generalizáveis (fenômeno chamado de *overfitting*), o que compromete a capacidade de gerar conteúdos inovadores e respostas adequadas.

Outro fator limitante é a disponibilidade de dados para treinamento. Pesquisadores estimaram que, se continuar o crescimento acelerado dos conjuntos de dados utilizados para treinar os modelos, entre 2026 e 2032 todos os textos de qualidade disponíveis na internet já terão sido utilizados (Villalobos *et al.*, 2024). Sem a ampliação do acervo de dados, torna-se difícil aprimorar significativamente os modelos, sendo necessário recorrer ao uso de dados sintéticos (gerados artificialmente), o que pode limitar seu desenvolvimento.

Conforme os modelos se tornam maiores, os custos ambientais também aumentam. O consumo de energia e a pegada de carbono associados ao treinamento e à operação de modelos de linguagem de grande escala são significativos. Por exemplo, para gerar 100 palavras, o GPT-4 em um centro de processamento em Washington consome aproximadamente 1.500 ml de água para resfriamento (Verma; Tan, 2023). A crescente adoção de tecnologias generativas intensifica a pressão sobre recursos naturais como água e energia elétrica. O impacto ambiental pode tornar insustentável o crescimento acelerado dos modelos. Resta saber se a sociedade desenvolverá uma consciência ecológica suficiente para limitar a exploração dos recursos naturais e mitigar os custos ambientais associados ao avanço da IA.

As limitações físicas também impõem barreiras ao crescimento da capacidade computacional. À medida que os transistores se aproximam do tamanho de átomos, fenômenos como interferência quântica, geração excessiva de calor e desafios na fabricação tornam a continuidade da miniaturização cada vez mais complexa. A superação dessas limitações dependerá do desenvolvimento de alternativas inovadoras, como a computação quântica e outros paradigmas tecnológicos.

Devem ser consideradas também as ações regulatórias, que podem impor restrições significativas ao desenvolvimento de modelos cada vez

maiores. Governos e organizações estão mais atentos aos riscos associados à IA avançada, como a disseminação de desinformação, o viés algorítmico e os impactos no mercado de trabalho. Medidas regulatórias também podem desacelerar o ritmo do avanço tecnológico.

Diante de tantas possíveis limitações, é incerto se os modelos de linguagem continuarão a crescer exponencialmente nos próximos anos. Assim como há barreiras para o crescimento da IA generativa, múltiplos fatores tornam questionável a viabilidade da singularidade tecnológica.

11.3 – O que fazer pós IA generativa: regulamentação

A singularidade tecnológica nos convida a refletir sobre os potenciais impactos sociais e riscos existenciais decorrentes do avanço da IA, instigando-nos a planejar ações que garantam que o desenvolvimento tecnológico beneficie a humanidade, em vez de ameaçá-la. No entanto, não se trata de aguardar o ano de 2045 ou um futuro em que máquinas tenham desenvolvido consciência ou superinteligência, pois a IA já atingiu um nível de desenvolvimento bastante preocupante.

A inteligência artificial não está, necessariamente, evoluindo para replicar todas as capacidades humanas. Seu desenvolvimento é fortemente influenciado por fatores econômicos, pela busca por utilidade e eficiência. As forças de mercado direcionam os avanços da IA para áreas onde ela possa oferecer mais lucro, muitas vezes superando o desempenho humano em tarefas específicas. Em vez de imitar a cognição humana em sua totalidade, a IA vem sendo projetada principalmente para resolver tarefas e otimizar processos, pois isso é lucrativo e de interesse para o mercado.

A evolução da IA direcionada pela lógica econômica também levanta preocupações significativas. A utilidade da IA para o capitalismo pode resultar na substituição em massa de trabalhadoras/es em diversos setores. Por exemplo, quando for mais lucrativo ter uma frota de carros autônomos do que motoristas humanos dirigindo os próprios veículos, esses profissionais serão imediatamente substituídos. De forma análoga, quando aulas realizadas por tecnologias inteligentes forem mais lucrativas, já

podemos adivinhar o que acontecerá com as/os conteudistas, docentes e tutoras/es humanos... O ChatGPT substituirá professoras e professores? Esse foi o título do nosso primeiro texto sobre IA generativa e educação (Pimentel; Azevedo; Carvalho, 2023a) — voltaremos a essa questão no Capítulo 20.

A perda dos empregos e a necessidade de adaptação das diversas profissões são preocupações imediatas, mas não as únicas. A IA já alcançou um nível de desenvolvimento que influencia significativamente diversos aspectos da sociedade, o que exige reflexão e ação imediatas.

Conversas em torno desse tema têm ocorrido há muito tempo e se intensificaram nos últimos anos. Certos indivíduos [...] expressaram, ao longo dos anos, uma grave preocupação com as condições criadas pela marcha desenfreada da ciência e da tecnologia. [...] De um lado estão aqueles que acreditam que os computadores podem, devem e farão tudo, e do outro lado aqueles que, como eu, acreditam que existem limites para o que os computadores devem ser capazes de fazer (Weizenbaum, 1976, p. 11, tradução nossa).

Weizenbaum (criador de ELIZA, o primeiro *chatbot* da história), já em 1976, alertava para a necessidade de estabelecermos limites para a atuação computacional. Em 2023, roteiristas de Hollywood deflagraram uma greve em função do temor de que a IA substituísse o trabalho humano na criação de roteiros, prejudicando empregos, *royalties* e a qualidade criativa. As/Os roteiristas exigiram regras claras para limitar o uso da IA nessa área. Após meses de negociações, o acordo alcançado proibiu a substituição de roteiristas por IA, restringiu seu uso sem consentimento e assegurou que ela não poderá escrever ou reescrever material literário (Pradella, 2023), constituindo-se em um importante marco na defesa do trabalho humano frente às tecnologias generativas.

Alguns meses após o lançamento do ChatGPT, a Itália bloqueou esse sistema no país em decorrência de preocupações com a privacidade de usuárias/os, por suspeitas de que a OpenAI estivesse coletando e armazenando dados pessoais de forma inadequada, e pela ausência de mecanismos eficazes para verificar a idade de usuárias/os, o que poderia expor pessoas menores de idade a conteúdos inadequados. Após a OpenAI

implementar ajustes em suas políticas de privacidade e reforçar medidas de segurança, o acesso ao ChatGPT foi restabelecido na Itália. Esse episódio ilustra a necessidade de regulamentação para garantir o uso seguro, ético e responsável das tecnologias generativas, proteger usuárias/os, incentivar melhorias nos sistemas e reforçar a confiança pública na IA.

O objetivo desse tipo de regulamentação não é impedir o desenvolvimento da IA generativa, mas sim criar um ambiente que equilibre a inovação tecnológica com a proteção dos direitos humanos (como os direitos autorais, privacidade, segurança e não discriminação) buscando garantir que a IA seja usada de forma ética, transparente e responsável. Embora a criação de limitações tenha se tornado um imperativo pós-ChatGPT, há o risco de se limitar demais o desenvolvimento e a aplicação da IA. Com muitas restrições, em um país como o nosso, empresas de tecnologia podem ser impedidas de acompanhar os desenvolvimentos tecnológicos de ponta, impossibilitando-as de se tornarem competitivas em relação a empresas de outros países que não tenham as mesmas limitações, como os Estados Unidos e a China.

A Europa liderou os esforços globais ao produzir o Regulamento da Inteligência Artificial (*Artificial Intelligence Act – AI Act*), que entrou em vigor a partir de agosto de 2024 (União Europeia, 2024). Esse regulamento estabeleceu um quadro jurídico comum para a União Europeia, visando à proteção contra os efeitos nocivos dos sistemas de IA, o que se tornou um importante marco legal para o desenvolvimento e uso responsável da IA.

Nos Estados Unidos, a abordagem regulatória foi mais fragmentada. Em julho de 2023, empresas como OpenAI, Alphabet (Google) e Meta (Facebook) firmaram um acordo voluntário com o governo para implementar uma marca d'água que possibilitasse a identificação dos conteúdos gerados por IA. Posteriormente, em outubro de 2023, foi emitida a Ordem Executiva 14110, que, entre outros objetivos políticos, inclui a defesa de estudantes: “Moldar o potencial da IA para transformar a educação, criando recursos para dar suporte aos educadores no uso de tecnologias educacionais habilitadas por IA, como tutoria personalizada nas escolas” (The White House, 2023, n.p., tradução nossa).

A China implementou, em agosto de 2023, as "Medidas Provisórias para a Gestão de Serviços de IA Generativa" (Interim, s.d.), regulando serviços de IA voltados ao público. A intenção das medidas é promover o desenvolvimento e o uso regulamentado da IA generativa, ao mesmo tempo em que salvaguarda a segurança nacional, os interesses públicos da sociedade e os direitos legais de cidadãos/ãos e organizações.

No Brasil, os debates sobre a regulamentação da IA ainda estavam em andamento no momento em que escrevemos este texto, com algumas iniciativas legislativas em tramitação no Congresso Nacional. Uma das principais propostas é o Projeto de Lei 2.338/2023, que estabelece normas gerais para o desenvolvimento e uso responsável de sistemas de IA no país. O projeto visa proteger direitos fundamentais e garantir a implementação de sistemas seguros e confiáveis, beneficiando a pessoa humana, o regime democrático e o desenvolvimento científico e tecnológico. Outros projetos de lei relacionados à IA estavam em tramitação, como o PL 872/2021, que dispõe sobre princípios para o uso da tecnologia de IA no Brasil, e o PL 210/2024, que trata de diretrizes específicas para a aplicação da IA em diferentes setores. Essas propostas buscam criar um marco legal que equilibre a promoção da inovação tecnológica com a proteção dos direitos fundamentais.

A regulamentação impactará o desenvolvimento e o uso de tecnologias generativas no contexto educacional. Para exemplificarmos a importância da regulamentação, vamos analisar o caso do GPT Paulo Freire que desenvolvemos (ver Capítulo 4). Para desenvolvê-lo, fizemos o *upload* de partes dos livros de Paulo Freire. Será que essa ação é ilegal?

Certamente seria ilegal distribuir as obras de Paulo Freire, cujos direitos autorais ainda não caíram em domínio público e pertencem à viúva Ana Maria Araújo Freire. Contudo, o que o GPT Paulo Freire faz é automatizar o que nós, acadêmicas/os, já fazemos: consultar livros e anotações para gerar novos textos, inclusive citando trechos de obras protegidas por direitos autorais. A citação de passagens de qualquer obra, para esse tipo de finalidade, não constitui ofensa aos direitos autorais de acordo com nossa Lei de Direitos Autorais (Brasil, 1998). Todavia, nossa atual legislação não

aborda explicitamente a questão do uso de obras protegidas por direitos autorais para o treinamento de IA.

Alguns países permitem a criação de bases de dados para o treinamento de IA, como a União Europeia que, por meio da Diretiva (UE) 2019/790, introduziu exceções aos direitos de autor para a mineração de textos e dados, permitindo o uso de conteúdos protegidos para esse fim (União Europeia, 2019). No Brasil, o Projeto de Lei 2.338/2023 propõe a regulamentação do uso de obras protegidas para fins de treinamento de IA, desde que não prejudiquem os interesses econômicos das pessoas detentoras dos direitos (Brasil, 2023b). Se esse projeto de lei vier a ser aprovado, teremos segurança legal em relação ao GPT Paulo Freire. Contudo, sem esse tipo de regulamentação, ficamos receosos de estarmos cometendo ilegalidades, mesmo que nossa intenção seja homenagear Paulo Freire, criando uma IA sem fins lucrativos para simular conversas com ele e divulgar suas ideias no contexto educacional. Apesar de esse uso parecer justo, como nossa lei brasileira não está baseada no princípio do uso justo (*fair use*), o GPT Paulo Freire pode acabar sendo considerado ilegal em nosso país e, dessa forma, teremos que tirá-lo do ar. Portanto, a regulamentação sobre a IA que viermos a adotar afetará diretamente o desenvolvimento e o uso das tecnologias generativas no Brasil, com impacto direto sobre a educação.

Regulamentar a IA generativa tornou-se pauta das discussões políticas ao redor de todo o mundo. Contudo, essa não é a única ação a ser realizada. Educar a população para a compreensão e o uso da IA generativa é outra demanda urgente, como discutiremos na próxima seção.

11.4 – O que fazer pós IA generativa: educação

Questionado sobre o que fazer com esta mudança tecnológica, que aconteceu de forma tão brusca, Chomsky destaca o papel da educação. “É impossível travar os sistemas”, avisa ele, que não assinou a carta em que especialistas de todo o mundo pediram uma moratória no desenvolvimento de IA. Explica por quê: “A única maneira é educar as pessoas para a autodefesa. Levar as pessoas a

compreender o que isto é e o que não é.” Atribui um significado político a ferramentas como o ChatGPT: “É basicamente como qualquer outra ideologia ou doutrina. Como é que alguém se defende contra a doutrina neofascista? Educando as pessoas.” (Chomsky, 2023, n.p.)

A carta aberta que Chomsky não assinou, publicada pelo instituto *Future of Life*, pedia uma “Pausa nos experimentos gigantes de IA” em decorrência do lançamento do GPT-4:

Apelamos a todos os laboratórios de IA para pausar imediatamente, por pelo menos 6 meses, o treinamento de sistemas de IA mais poderosos que o GPT-4. Os sistemas de IA que competem com a inteligência humana podem representar riscos profundos para a sociedade e para a humanidade, conforme demonstrado por uma extensa pesquisa e reconhecido pelos principais laboratórios de IA. (Pausa [...], 2023, n.p.)

Concordamos com Chomsky sobre a necessidade de educar as pessoas para a compreensão crítica das tecnologias generativas. Porém, discordamos de que a educação para o uso deva se voltar apenas para a autodefesa. Chomsky deu outras declarações que evidenciam sua perspectiva da IA generativa como uma ameaça a ser combatida, ignorando suas potencialidades: “A ideia de que podemos aprender alguma coisa com este tipo de IA é um erro. Elas criam uma atmosfera onde a explicação e a compreensão não têm qualquer valor” (Chomsky, 2023, n.p.). Discordamos, respeitosamente, de Chomsky porque a ocorrência da aprendizagem e a valorização da explicação e da compreensão dependerão da relação que a/o usuária/o estabelece com esse tipo de tecnologia. Por isso mesmo, é tão importante educar para o uso, para que possamos aprender a criar em parceria com a IA, sem entregar nossas mentes e vozes para que ela pense e se expresse por nós.

Reconhecemos muitos riscos, como a possibilidade de a IA generativa ser utilizada para obter respostas prontas a trabalhos escolares e acadêmicos. Contudo, ela também pode disparar processos formativos intensos por meio da conversação e da coautoria, processos que nos levam a aprender novas informações e a (re)significar conhecimentos. A IA

generativa produz textos sobre o que queremos conhecer, apresenta múltiplos pontos de vista, tira nossas dúvidas, revisa nossos textos, entre tantas outras atividades formativas que discutimos no Capítulo 4. Dada a possibilidade da IA gerar informações incorretas, precisamos aprender a criticar o conteúdo gerado, sem aceitar tudo como se fosse o oráculo do nosso tempo. Precisamos também discutir a dependência que teremos das empresas desenvolvedoras das tecnologias generativas mais poderosas, bem como agir para regulamentar e impor limites.

Em concordância com Chomsky, entendemos que precisamos de uma educação para o uso, mas que vá além da autodefesa, que ensine também a conversar e a cocriar com essa tecnologia. Este livro foi escrito com esse objetivo.

Como a IA generativa funciona?

Em cinco dias, mais de um milhão de pessoas se inscreveram para testá-lo, segundo Greg Brockman, presidente da OpenAI. Centenas de capturas de tela de conversas do ChatGPT se tornaram virais no Twitter, e muitos de seus primeiros fãs falam disso em termos surpreendentes e grandiosos, como se fosse uma mistura de *software* e feitiçaria. (Roose, 2022, n.p.)

O ChatGPT parece algo mágico: como ele consegue responder às nossas perguntas tão rapidamente e com tanta proficiência? Arthur C. Clarke formulou uma das leis mais conhecidas sobre ciência e tecnologia: “Qualquer tecnologia suficientemente avançada é indistinguível da magia” (Clarke, 1962). Como consequência dessa máxima, podemos deduzir que, quando tornamos uma tecnologia avançada suficientemente compreensível, desvendamos sua aura de magia, desmistificando-a. Algo semelhante foi afirmado na abertura do artigo sobre ELIZA, o primeiro sistema de *chatbot* da história:

Diz-se que explicar é diminuir a importância da coisa explicada. Essa máxima se realiza de maneira exemplar no campo da programação de computadores, especialmente no que é chamado de programação heurística e inteligência artificial. Nessas áreas, máquinas são feitas para se comportarem de maneiras surpreendentes, muitas vezes suficientes para deslumbrar até mesmo o observador mais experiente. No entanto, uma vez que um determinado programa é desmascarado, uma vez que seus mecanismos internos são explicados em linguagem suficientemente simples para possibilitar a compreensão, sua magia desmorona; ele se revela como uma mera coleção de procedimentos, cada um bastante compreensível. O observador diz a si mesmo: “Eu poderia ter escrito isso”. Com esse pensamento, ele move o programa em

questão da prateleira marcada como “inteligente” àquela reservada para curiosidades, que serve para ser discutido apenas com pessoas menos esclarecidas que ele. (Weizenbaum, 1966, p. 36, tradução nossa)

Weizenbaum alertou o/a leitor/a que seu artigo daria fim à magia relacionada ao funcionamento de seu robô de conversação ELIZA. Esse é o nosso objetivo neste capítulo: desmistificar a IA generativa, fazer sua magia desmoronar. Para alcançarmos esse objetivo, explicamos como algumas técnicas conseguem fazer uma máquina tomar decisões inteligentes e gerar textos coerentes, demonstrando um tipo de inteligência.

Você não precisa saber programar computadores para compreender, em linhas gerais, como tudo funciona. Não pretendemos aqui ensinar a construir sistemas inteligentes, pois este não é um livro técnico sobre IA. Nosso objetivo é apresentar algumas noções em alto nível, suficientes para compreender o funcionamento de tecnologias generativas como o ChatGPT. Para essa finalidade, discutiremos noções como algoritmo, raciocínio computacional, aprendizagem de máquina (*machine learning*), rede neural artificial, aprendizagem profunda (*deep learning*), modelo de linguagem, treinamento não supervisionado, *Transformer*, entre outras.

Se você, leitor/a, não deseja saber como tudo funciona, pode pular este capítulo sem prejuízo para a compreensão do restante do livro. Este capítulo foi escrito para quem se encontra inquieta/o querendo saber como a IA é capaz de gerar textos, para tentar saciar um pouco da sede das/os que são “epistemologicamente curiosas/os”, como lindamente caracterizado por Paulo Freire (2015 [1996], p.33-34):

A curiosidade como inquietação indagadora, como inclinação ao desvelamento de algo, como pergunta verbalizada ou não, como procura de esclarecimento, como sinal de atenção que sugere alerta faz parte integrante do fenômeno vital. Não haveria criatividade sem a curiosidade que nos move e que nos põe pacientemente impacientes diante do mundo que não fizemos, acrescentando a ele algo que fazemos. Como manifestação presente à experiência vital, a curiosidade humana vem sendo histórica e socialmente construída e reconstruída. Precisamente porque a promoção da ingenuidade para a criticidade não se dá automaticamente, uma das tarefas precípuas da prática educativo-progressista é exatamente o desenvolvimento da curiosidade crítica, insatisfeita, indócil. Curiosidade com que

podemos nos defender de “irracionalismos” decorrentes *do* ou produzidos por certo excesso de “racionalidade” de nosso tempo altamente tecnologicizado. E não vai nesta consideração nenhuma arrancada falsamente humanista de negação da tecnologia e da ciência. Pelo contrário, é consideração de quem, de um lado, não diviniza a tecnologia, mas, de outro, não a diaboliza. De quem a olha ou mesmo a espreita de forma criticamente curiosa.

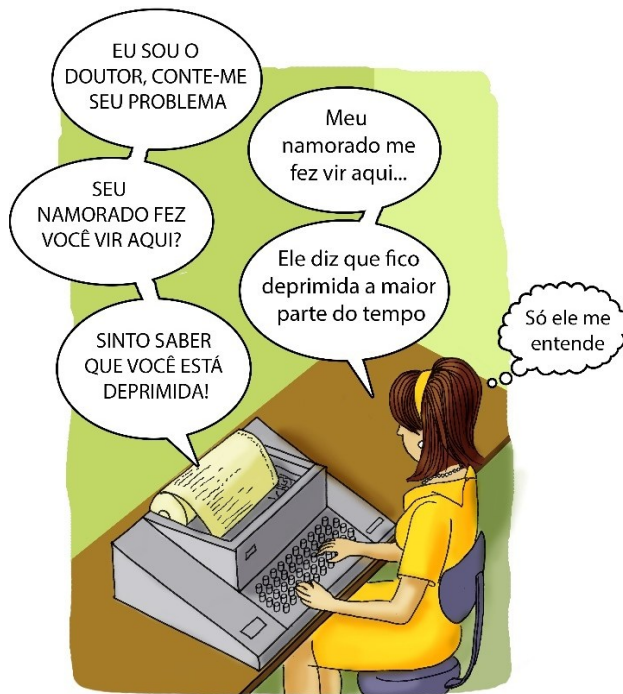
Compreender como a inteligência artificial generativa funciona, ainda que em linhas gerais, pode desconstruir tanto a divinização quanto a demonização dessa tecnologia, pode desfazer alguns equívocos e esclarecer dúvidas. Esse tipo de conhecimento também é necessário para afiar nossa capacidade de análise crítica das tecnologias generativas que continuarão a ser desenvolvidas e com as quais conviveremos cada vez mais em nosso cotidiano.

Russell e Norvig (2022, n.p.), cujas obras são referência para diversos cursos de IA, encontraram uma maneira interessante de resumir a história da área:

Uma forma rápida de resumir os marcos da história da IA é listar os vencedores do Prêmio Turing: Marvin Minsky (1969) e John McCarthy (1971) pela definição dos fundamentos do campo com base na representação e no raciocínio; Ed Feigenbaum e Raj Reddy (1994) pelo desenvolvimento de sistemas especialistas, que codificam o conhecimento humano para resolver problemas do mundo real; Judea Pearl (2011) pelo desenvolvimento de técnicas de raciocínio probabilístico que lidam com a incerteza de um modo baseado em princípios; e, finalmente, Yoshua Bengio, Geoffrey Hinton e Yann LeCun (2019) por tornar o “aprendizado profundo” (redes neurais multicamadas) uma parte crítica da computação moderna.

Nas próximas seções, discutiremos essas e outras técnicas que representam marcos históricos da IA, pois elas nos ajudam a compreender o funcionamento das tecnologias generativas. Decidimos começar desmistificando o funcionamento do primeiro *chatbot*, ELIZA, mostrando algumas semelhanças e diferenças com as atuais técnicas de Processamento de Linguagem Natural (PLN), que é a subárea da IA dedicada a capacitar máquinas para compreender e gerar textos nas línguas faladas por seres humanos, como o português, inglês ou chinês.

12.1 – ELIZA e o Processamento de Linguagem Natural (PLN)



Fonte: dos autores, ilustração de Mônica Lopes com base em Weizenbaum (1966)

ELIZA é a antepassada mais distante do ChatGPT, sua progenitora primordial, lançada por Joseph Weizenbaum em 1966 no artigo ELIZA – Um Programa de Computador para o Estudo da Comunicação em Linguagem Natural entre Homem e Máquina⁴⁶. Esse é um importante marco histórico, pois é considerado o primeiro *chatbot*: um programa capaz de simular uma conversação com seres humanos. Os enunciados de ELIZA eram baseados em um roteiro de conversação (*script*), sendo o mais famoso o DOCTOR, que imitava o comportamento conversacional de um psicoterapeuta. Naquela época, as pessoas não conheciam robôs de conversação, e ainda era uma novidade a computação interativa possibilitada pelo teletipo (uma máquina de escrever eletromecânica).

⁴⁶ ELIZA – a computer program for the study of natural language communication between man and machine (Weizenbaum, 1966)

O DOCTOR buscava imitar o comportamento de fala de um psicoterapeuta rogeriano, cuja técnica de terapia, desenvolvida pelo psicólogo Carl Rogers, consiste em aprofundar o que foi dito pela/o paciente, devolvendo à/ao paciente suas próprias declarações para que ela/e reflita sobre suas falas, promovendo o autoconhecimento e a autoaceitação, sem que a/o terapeuta apresente interpretações ou diagnósticos para a/o paciente. O exemplo apresentado no artigo original é de uma conversa entre o DOCTOR e uma mulher:

COMO VOCÊ VAI? EU SOU O DOUTOR. POR FAVOR, SENTE-SE AO
TELETIPO E CONTE-ME SEU PROBLEMA.

> Os homens são todos iguais.

DE QUE MANEIRA?

> Eles estão sempre nos incomodando sobre uma coisa ou outra.
VOCÊ PODE SER MAIS ESPECÍFICO?

> Bem, meu namorado me fez vir aqui.

SEU NAMORADO FEZ VOCÊ VIR AQUI?

> Ele diz que fico deprimida a maior parte do tempo.

SINTO SABER QUE VOCÊ ESTÁ DEPRIMIDA

> É verdade. Eu estou infeliz

VOCÊ ACHA QUE VIR AQUI VAI TE AJUDAR A NÃO FICAR INFELIZ?

> Eu preciso de ajuda, isso sim.

O QUE SIGNIFICARIA PARA VOCÊ SE CONSEGUISSE ESSA AJUDA?

> Talvez eu possa aprender a me relacionar melhor com minha mãe.
CONTE-ME MAIS SOBRE SUA FAMÍLIA

> Minha mãe cuida de mim.

QUEM MAIS NA SUA FAMÍLIA CUIDA DE VOCÊ?

> Meu pai.

SEU PAI

> Você é parecido com meu pai em alguns aspectos.

QUAL SEMELHANÇA VOCÊ VÊ?

> Você não é muito agressivo, mas acho que você não quer que eu
perceba isso.

O QUE FAZ VOCÊ ACHAR QUE EU NÃO SOU MUITO AGRESSIVO?

> Você não discute comigo.

POR QUE VOCÊ ACHA QUE EU NÃO DISCUTO COM VOCÊ?

> Porque você tem medo de mim.

AGRADA-LHE ACREDITAR QUE EU TENHO MEDO DE VOCÊ?

> Meu pai tem medo de todo mundo.

O QUE MAIS VEM À SUA MENTE QUANDO VOCÊ PENSA EM SEU PAI?

> Valentões.

ISSO TEM ALGO A VER COM O SEU NAMORADO?

> [...]

(Weizenbaum, 1966, p. 36-37, tradução nossa, com acréscimo da primeira linha, que se encontra no código original)

Os códigos-fonte de ELIZA e do roteiro DOCTOR foram divulgados publicamente em 2021 e agora encontram-se em domínio público no site ELIZAGEN⁴⁷. Esses códigos são de interesse histórico, pois registram as técnicas de programação empregadas à época na construção desse primeiro robô de conversação.

O funcionamento de ELIZA baseia-se na transformação do texto digitado pela/o usuária/o em um novo texto a ser apresentado como resposta. Por exemplo, se a pessoa digitasse um texto contendo a palavra “sonho” ou “sonhei”, o programa escolhia uma das frases a seguir como resposta:

O QUE ESSE SONHO SUGERE PARA VOCÊ?

VOCÊ SONHA FREQUENTEMENTE?

QUE PESSOAS APARECEM NOS SEUS SONHOS?

VOCÊ ACREDITA QUE ESSE SONHO TEM ALGO A VER COM O SEU PROBLEMA?

Além da palavra-chave, o programa podia montar a resposta utilizando alguma informação contida no texto de entrada:

> Estou muito infeliz

HÁ QUANTO TEMPO VOCÊ ESTÁ MUITO INFELIZ?

Por exemplo, quando a/o usuária/o digitava “Estou muito infeliz”, o programa reconhecia que a frase seguia o modelo “Estou *”, onde * representa qualquer coisa digitada pela/o usuária/o, nesse exemplo, foi digitado “muito infeliz”. Uma possível resposta associada a esse modelo de frase era “Há quanto tempo você está *?”, substituindo-se o * pelo trecho digitado pela/o usuária/o. Quando nenhuma palavra-chave era identificada

⁴⁷ <https://sites.google.com/view/elizagen-org/about>

no texto digitado, o sistema selecionava uma resposta genérica ou uma mensagem que mudava de assunto, como: “Por favor, continue.”; “Entendo.”; “Você pode ser mais específico?”. Também podia reutilizar alguma informação digitada anteriormente. Essa técnica de processamento de texto, que repete algo digitado pela/o usuária/o, cria a ilusão de que a máquina entendeu o que a pessoa digitou, o que favorece a continuidade da conversa.

Ao compreendermos o funcionamento de ELIZA, fica evidente que esse programa não desenvolveu uma inteligência capaz de pensar sobre tudo ou sobre si mesmo. Ele não pensa nem compreende o sentido das palavras. Trata-se de um jogo para simular uma conversa, com base em regras predefinidas para processar o texto buscando palavras-chave e padrões de frases.

De acordo com Weizenbaum (1976), o programa ELIZA se tornou popular no Instituto de Tecnologia de Massachusetts (MIT) por demonstrar vividamente o poder de processamento de informação dos computadores à época. Os efeitos causados nas/os usuárias/os, que se sentiam conversando e estabeleciam laços afetivos com a máquina, impressionaram Weizenbaum:

Fiquei surpreso ao ver a rapidez e profundidade com que as pessoas que conversavam com DOCTOR se envolveram emocionalmente com o computador e como, inequivocamente, o antropomorfizaram. Certa vez, minha secretária, que havia me observado trabalhar no programa por muitos meses e, portanto, certamente sabia que era apenas um programa de computador, começou a conversar com ele. Depois de apenas algumas interações com ele, ela me pediu para sair da sala [para ter privacidade para conversar com a máquina]. Em outra ocasião, sugeri que eu poderia configurar o sistema para examinar todas as conversas que alguém teve com ele, digamos, durante a noite. Fui prontamente bombardeado com acusações de que o que eu propunha era espionar os pensamentos mais íntimos das pessoas; uma clara evidência de que elas estavam conversando com o computador como se ele fosse uma pessoa com quem se pode falar de forma apropriada e íntima. [...] O que eu não havia percebido é que exposições extremamente curtas a um programa de computador relativamente simples poderiam induzir pensamentos altamente ilusórios em pessoas perfeitamente normais. Esse *insight*

me levou a atribuir nova importância às questões da relação entre o indivíduo e o computador. [...] Essa reação a ELIZA me mostrou, de forma mais vívida do que tudo o que eu já tinha visto antes, as atribuições enormemente exageradas que mesmo pessoas bem instruídas são capazes de fazer, e até se esforçam para fazer, a uma tecnologia que não compreendem. (Weizenbaum, 1976, p. 6, tradução nossa)

Apesar de Weizenbaum explicar que ELIZA não era capaz de compreender as conversas, muitas pessoas que interagiram com o programa acreditavam que ele realmente entendia o que elas diziam e se importava com elas. Esse fenômeno ficou conhecido como Efeito ELIZA.

Em função dos efeitos causados pela interação com ELIZA, Weizenbaum (1976) passou a considerar que a IA poderia representar um perigo à sociedade. Sua preocupação era de que a computação reforçasse a visão do ser humano como máquina, fazendo com que nos percebêssemos como meros autômatos, reconhecendo-nos “como nada além de uma engrenagem” e acreditando que os avanços técnicos e científicos poderiam “levar o ser humano a uma visão cada vez mais racionalista de sua sociedade e uma imagem cada vez mais mecanicista de si mesmo”. Para o autor, essa concepção poderia levar à impressão de perda da autonomia e da nossa própria “essência humana”. O receio de Weizenbaum de nos compreendermos como uma espécie de programa de computador já foi tema de diversos episódios da série *Black Mirror*, como *San Junipero* (2016, 3ª temporada, 4º episódio), *Hang the DJ* (2017, 4ª temporada, 4º episódio) e *A Joan é Péssima* (*Joan Is Awful*, 2023, 6ª temporada, 1ª episódio).

Podemos estabelecer algumas comparações, ainda que grosseiras, entre ELIZA e o ChatGPT, para evidenciar que esse último é muito mais elaborado. É interessante notar que o programa ELIZA possuía apenas 409 linhas de código⁴⁸, o que é considerado um programa pequeno, possível de ser criado por uma única pessoa. De forma semelhante, o código-fonte do

⁴⁸ Na implementação original em Lisp, apresentada por Weizenbaum: https://github.com/jeffshragel/elizagen.org/blob/master/1965_Weizenbaum_MAD-SLIP/MAD-SLIP_transcription.txt

módulo de inferência do modelo GPT-2⁴⁹, isolado das bibliotecas auxiliares, também é relativamente compacto, com cerca de mil linhas de código. Contudo, essas linhas de código representam uma pequena parte do projeto como um todo. A maior complexidade está no processo de treinamento do modelo de linguagem, que envolve o processamento de uma enorme quantidade de dados que não fazem parte do código do programa. Assim, a inteligência do GPT emerge dos padrões aprendidos a partir dos dados processados, e não é previamente codificada pelas/os programadoras/es.

A diferença entre ELIZA e o ChatGPT torna-se ainda mais evidente ao compararmos as bases textuais utilizadas por cada um. O roteiro de conversação DOCTOR, documentado no artigo de Weizenbaum⁵⁰, contém apenas 8.520 caracteres (incluindo os espaços em branco), o que equivale a 3 ou 4 páginas de texto⁵¹ e requer apenas 8,5KB de armazenamento em arquivo de computador⁵². Já o modelo de linguagem GPT-3 processou milhares de livros, toda a Wikipédia e uma imensa quantidade de textos da internet, totalizando 45TB de texto antes da filtragem e 570 GB após a filtragem, o que representa 8 ordens de grandeza superior ao roteiro de conversação do DOCTOR. Essa diferença entre as bases textuais equivale, em termos de ordem de grandeza, à comparação entre uma única pessoa e 100 milhões de pessoas (aproximadamente metade da população brasileira).

Apesar da simplicidade de ELIZA e da limitação do roteiro de conversação DOCTOR, já era possível criar uma ilusão de entendimento e inteligência, fazendo algumas/uns usuárias/os acreditarem que estavam

⁴⁹ <https://github.com/openai/gpt-2>

⁵⁰ Há uma transcrição do DOCTOR disponível em: https://github.com/jeffshrager/elizagen.org/blob/master/1965_Weizenbaum_MAD-SLIP/1966_01_CACM_article_Eliza_script.txt

⁵¹ Considerando que uma página A4, com as configurações típicas dos principais editores de texto (Fonte: Times ou Arial, tamanho 12; espaçamento 1,5 entre linhas; margens padrão de 2,5 cm; e texto contínuo sem parágrafos longos ou tabelas), comporta aproximadamente 2.500 a 3.000 caracteres (incluindo espaços).

⁵² Considerando que cada caractere ocupa 1 Byte, como na codificação ASCII e ISO 8859-1 (Latin-1).

interagindo com um ser consciente, ainda que essa ilusão durasse apenas algumas poucas interações. Esse programa de conversação possibilitou a primeira experimentação prática do Teste de Turing e estabeleceu os fundamentos para o desenvolvimento de *chatbots* e de algumas técnicas de processamento de linguagem natural.

Weizenbaum (1976, p. 6) caracterizou ELIZA como uma “simples paródia mecânica de uma técnica de entrevista”. Fica evidente a superioridade da atual tecnologia generativa quando comparamos as respostas dadas por ELIZA-DOCTOR, apresentadas no artigo original, e as fornecidas pelo ChatGPT⁵³. Isso nos faz questionar: ELIZA era inteligente?

O programa ELIZA seguia um conjunto de regras predefinidas para o processamento de linguagem natural. ELIZA não possuía capacidade de raciocínio nem de aprendizagem, características que distinguem os sistemas contemporâneos de inteligência artificial. Portanto, à luz dos critérios atuais, ELIZA não seria considerado um programa inteligente no sentido contemporâneo do termo.

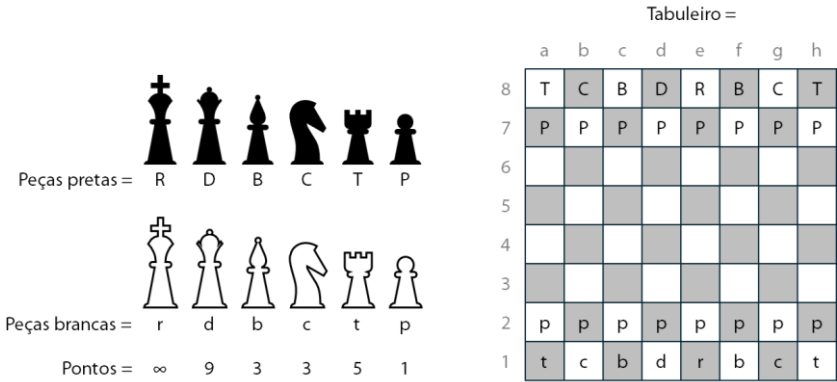
12.2 – Fazendo um computador jogar xadrez: representação simbólica, regras e raciocínio

Na abordagem simbólica, cria-se uma representação do mundo que possibilita ao computador analisá-la, tirar conclusões e tomar decisões. Um exemplo clássico dessa abordagem é o raciocínio computacional aplicado ao jogo de xadrez. Essa abordagem, que predominou nas primeiras décadas de desenvolvimento da IA, é discutida nesta seção para ilustrar como se pode fazer um computador agir de forma inteligente. Contudo, a IA generativa baseia-se em uma lógica distinta: a abordagem estatístico-conexionista, que se consolidou como a principal vertente da área nas últimas décadas, como será discutido nas seções seguintes.

Para que o computador seja capaz de raciocinar, o primeiro passo é construir uma representação simbólica do jogo, como no exemplo a seguir.

⁵³ Realizamos esse caso, cujo registro está disponível em: <https://chat.openai.com/share/cd815deb-8f56-441d-a9cc-84ef45e735a3>

Representação computacional do jogo de xadrez



Fonte: dos autores

Podemos representar um tabuleiro de xadrez por meio de uma tabela composta por 8 linhas e 8 colunas, também chamada de matriz 8x8. As peças do tabuleiro podem ser representadas por letras: R representa o Rei; D, a Dama (ou rainha); B, o Bispo; C, o Cavalo, T, a Torre, e P, o Peão. Letras em maiúsculo representam peças pretas, enquanto letras em minúsculo representam peças brancas. As representações das peças devem ser posicionadas na matriz conforme sua disposição real no tabuleiro. A ausência de um caractere indica que a casa está vazia.

Com essa representação computacional do tabuleiro e as regras de movimentação de cada tipo de peça, é possível construir um roteiro de ações para que o computador possa jogar. Esse roteiro é o que chamamos de algoritmo: um conjunto de instruções que o computador executa para realizar uma tarefa, resolver um problema ou alcançar um objetivo. No caso do xadrez, o algoritmo deve decidir qual peça mover a cada jogada. Uma jogada consiste em mover uma peça de uma célula da matriz para outra célula. Por exemplo, mover o peão branco da casa b2 para b3 implica colocar um caractere vazio na célula b2 e a letra 'p' na célula b3, atualizando o estado do jogo.

Um algoritmo não-inteligente simplesmente sortearia uma peça e realizaria qualquer movimento válido, sem considerar a situação do jogo. Esse comportamento não é considerado inteligente, pois resulta de uma

decisão aleatória que não produz uma jogada plausível em função do estado atual da partida. Para que um algoritmo seja considerado inteligente, é necessário que realize algum tipo de raciocínio. No caso do xadrez, é preciso avaliar os movimentos possíveis e escolher aquele que representa a melhor decisão para a situação em questão.

Para que o computador seja capaz de avaliar uma jogada, é necessário definir uma “função de avaliação”. Essa função indica se um determinado movimento resultará ou não na captura de uma peça adversária. O objetivo é capturar as peças mais valiosas do oponente, preferencialmente o rei, que é o objetivo final do jogo. Para isso, atribuímos uma pontuação a cada tipo de peça, conforme sua importância: Peão = 1 ponto; Cavalo = 3 pontos; Bispo = 3 pontos; Torre = 5 pontos; Dama (ou rainha) = 9 pontos; e o Rei tem valor infinito. Assim, a função de avaliação retorna a quantidade de pontos associada à peça capturada; ou retorna zero, caso o movimento leve a uma casa vazia.

Com base nessa função, podemos descrever um algoritmo simples para decidir qual peça o computador deve mover em uma jogada. O primeiro passo é identificar, para cada peça do computador, todos os movimentos possíveis. Em seguida, o algoritmo verifica, para cada movimento, se ele resulta na captura de uma peça adversária. Caso haja captura, calcula-se a pontuação correspondente com base na função de avaliação. Durante esse processo, o algoritmo compara os pontos obtidos em cada jogada e mantém em memória o movimento com maior valor. Ao final da análise, o movimento com maior ganho é selecionado como a “melhor jogada”. Se nenhum dos movimentos resultar na captura de uma peça (isto é, todos gerarem zero pontos), o algoritmo simplesmente escolhe aleatoriamente um dos movimentos válidos.

Acabamos de descrever um algoritmo baseado em inteligência artificial, pois o computador avalia o estado do jogo e toma uma decisão coerente: capturar, sempre que possível, a peça de maior valor da/o adversária/o. Trata-se de uma estratégia simples (descrevemos um “algoritmo ganancioso” muito elementar), que exige poucos cálculos. Cada jogador/a tem, no máximo, 16 peças, e nem todas estão disponíveis para movimentação em todos os momentos. No início do jogo, o número de possibilidades é pequeno, já que várias peças estão bloqueadas. À medida

que a partida avança, o número de movimentos possíveis cresce. Em média, uma jogada dispõe de cerca de 35 movimentos possíveis. Assim, para decidir qual peça mover, o computador precisa avaliar, em média, 35 alternativas.

Podemos formular uma estratégia mais elaborada para que o computador jogue de forma mais inteligente. Em vez de simplesmente buscar capturar a peça adversária mais valiosa, o algoritmo pode considerar também as possíveis reações da/o oponente, avaliando todos os contra-ataques que podem ocorrer após cada jogada do computador. Por exemplo, suponha que o computador capture um peão, ganhando 1 ponto. No entanto, se na jogada seguinte a/o adversária/o puder capturar uma torre do computador, isso representará uma perda de 5 pontos, resultando em saldo negativo de -4 pontos, o que é uma clara desvantagem. Nessa situação, a captura do peão não parece ser a melhor decisão. Pode haver outro movimento mais vantajoso, no qual a/o adversária/o não tenha a oportunidade de capturar uma peça de valor igual ou superior.

Considerando essa estratégia, podemos elaborar um algoritmo que possibilite ao computador raciocinar de forma mais refinada antes de tomar uma decisão. Para cada movimento possível de todas as peças do computador, calcula-se quantos pontos seriam obtidos com o movimento e, em seguida, analisam-se todos os possíveis contra-ataques da/o adversária/o, identificando qual deles causaria a maior perda ao computador. O saldo de cada movimento é calculado pela diferença entre os pontos ganhos no ataque e os pontos perdidos no contra-ataque. Após avaliar todas as jogadas possíveis, o algoritmo escolhe aquela que resultar na maior pontuação para o computador.

Esse algoritmo já é mais elaborado do que o anterior, pois, além de considerar os movimentos do computador, também antecipa as possíveis reações da/o adversária/o. Isso exige mais cálculos: o computador precisa avaliar, em média, 35 movimentos próprios e, para cada um, considerar 35 possíveis contra-ataques da/o oponente, totalizando $35 \times 35 = 1.225$ combinações a serem analisadas.

É fácil perceber que o algoritmo pode ser aprimorado com a inclusão de um terceiro nível de análise: o computador joga, a/o adversária/o reage

e, então, o computador joga novamente. Nesse caso, a quantidade média de movimentos a serem avaliados sobe para $35 \times 35 \times 35 = 42.875$ combinações. A cada novo nível de antecipação, o número de combinações cresce exponencialmente. Por exemplo, ao tentar prever 10 jogadas à frente, o computador teria que analisar aproximadamente $35^{10} = 2.758.547.353.515.625$ possibilidades — um fenômeno conhecido como explosão combinatória, que torna inviável a escolha da melhor jogada dentro do tempo disponível para a decisão, devido à imensa carga computacional exigida.

Como é impraticável realizar todos os cálculos para determinar qual movimento oferece a maior chance de vitória, torna-se necessário lidar com incertezas e tomar decisões estratégicas. Uma abordagem comum é o uso de buscas heurísticas, que consistem em identificar os caminhos mais promissores dentro de um espaço de possibilidades, sem a necessidade de avaliar todas as opções disponíveis.

Os algoritmos realmente utilizados para fazer o computador jogar xadrez são muito mais sofisticados do que os apresentados nesta seção, pois combinam diversas técnicas de inteligência artificial. A função de avaliação também tende a ser mais complexa, incorporando múltiplos fatores, como a segurança do rei, o controle do centro do tabuleiro, o desenvolvimento das peças, entre outros.

O algoritmo simplificado aqui discutido tem como objetivo ilustrar como é possível dotar o computador de uma forma básica de raciocínio, por meio da aplicação de regras lógicas sobre uma representação simbólica de um domínio do mundo real; neste caso, o tabuleiro de xadrez. Esse tipo de abordagem possibilita realizar avaliações e tomar decisões coerentes diante de incertezas, o que caracteriza a chamada inteligência artificial simbólica.

Jogar xadrez é um exemplo clássico de sistema inteligente que simula um dos aspectos da inteligência humana: a capacidade de raciocinar com base em regras para tomar decisões plausíveis em situações de incerteza, mobilizando um processo lógico de deliberação sobre a tarefa a ser executada. Outra característica da inteligência humana que as máquinas podem simular é a capacidade de aprender, o que discutiremos na seção a seguir.

12.3 – Aprendizagem de máquina⁵⁴: o autodidatismo computacional

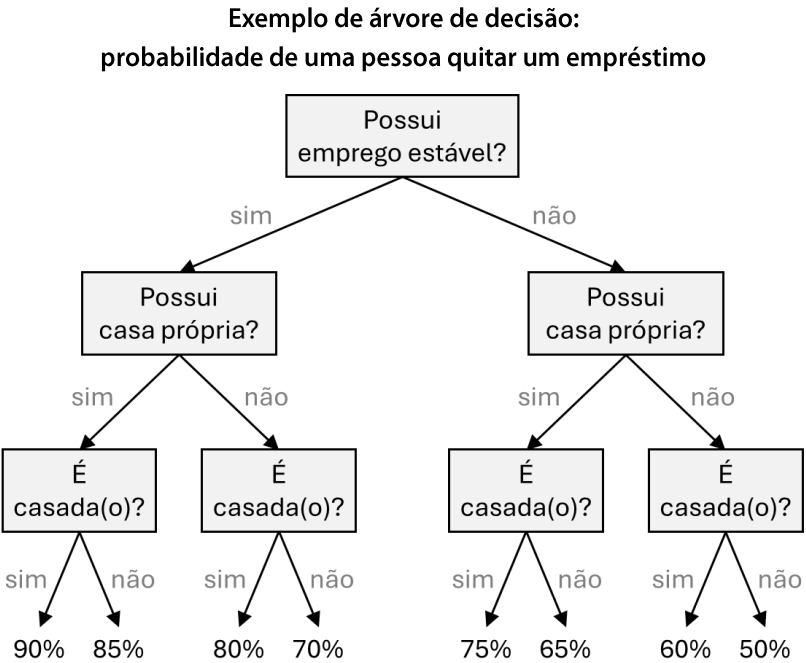
A IA generativa aprende com os dados durante seu processo de treinamento. Embora existam sistemas baseados em regras fixas capazes de resolver problemas específicos, atualmente a maior parte da pesquisa em IA associa inteligência à capacidade de aprender e melhorar o desempenho a partir de dados ou experiências. Por exemplo, um sistema especialista na área médica pode conter uma regra do tipo: “SE a pessoa apresenta o sintoma X e o sintoma Y, ENTÃO o diagnóstico é Z”. Esse tipo de sistema é capaz de diagnosticar doenças com base em um conjunto pré-definido de regras, mas não aprende a reconhecer novos sintomas ou enfermidades sem que uma pessoa insira manualmente novas regras.

Os primeiros sistemas especialistas foram desenvolvidos nas décadas de 1960 e 1970. O DENDRAL, lançado em 1965, foi projetado para inferir a estrutura molecular de compostos químicos. Já o MYCIN, desenvolvido no início da década de 1970, foi concebido para identificar bactérias causadoras de infecções graves e recomendar planos de tratamento com antibióticos: “Com cerca de 450 regras, o MYCIN era capaz de se sair tão bem quanto alguns especialistas e muito melhor do que médicos em início de carreira” (Russel; Norvig, 2022, n.p.). Esses sistemas buscavam imitar o processo de raciocínio e tomada de decisão de pessoas especialistas em suas respectivas áreas de conhecimento.

Esses sistemas especialistas utilizavam regras declaradas manualmente (“SE-ENTÃO”) e um motor de inferência para tomar decisões com base em conhecimento codificado por pessoas especialistas. Essa lógica condicional também é empregada na criação de modelos de árvore de decisão, como

⁵⁴ No Brasil, *machine learning* costuma ser traduzido para aprendizado de máquina. Neste livro, contudo, preferimos a tradução adotada em Portugal, “aprendizagem de máquina”. Compreendemos que aprendizagem se refere ao processo de aprender, enquanto aprendizado refere-se ao que foi aprendido, o resultado do processo de aprendizagem. Preferimos adotar aprendizagem, em vez de aprendizado, porque essa também foi a escolha de tradução em outras línguas românicas (no Espanhol, Italiano e Francês).

exemplificado na figura a seguir, que ilustra a probabilidade de uma pessoa quitar um empréstimo em um banco.



Fonte: dos autores

Percorremos (ou lemos) uma árvore de decisão de cima para baixo. Cada nó interno representa um teste condicional ("SE") baseado em um atributo. No exemplo apresentado, o primeiro teste da árvore (o nó-raiz) verifica se a pessoa possui emprego estável. A resposta a esse teste determina o caminho a seguir: se a condição for satisfeita, seguimos para o ramo à esquerda; caso contrário, seguimos para o ramo à direita. Esse processo continua até alcançarmos um nó-folha (ou terminal), que contém a informação final desejada. Por exemplo, uma pessoa que tem emprego estável, possui casa própria e é casada tem 90% de probabilidade de quitar o empréstimo. Já uma pessoa que não tem emprego estável, possui casa própria e não é casada tem uma probabilidade de 65% de quitar o empréstimo. Esse tipo de modelo pode ser útil para apoiar decisões como

conceder ou negar um empréstimo, ou para estimar o valor do seguro associado ao empréstimo. O objetivo de uma árvore de decisão é construir um modelo capaz de prever o valor de uma variável de destino (como a probabilidade de quitar um empréstimo), com base em várias variáveis de entrada (como emprego estável, propriedade de imóvel e estado civil).

Quando se deseja construir uma árvore de decisão que leve em conta muitas variáveis e seja capaz de identificar padrões mais complexos, é necessário utilizar uma grande quantidade de dados. Por exemplo, uma instituição financeira pode dispor de uma base extensa de informações sobre clientes que já contrataram empréstimos nos últimos cinco anos, incluindo dados sobre quem quitou ou não a dívida, além de outros atributos relevantes sobre as/os clientes e os empréstimos. O modelo gerado a partir desses dados reais tende a ser muito mais complexo do que o exemplo simplificado aqui apresentado, contendo um número significativamente maior de nós e níveis. Esse tipo de modelo costuma ser construído por meio de algoritmos de aprendizagem de máquina, que analisam os dados em busca de padrões e relações, gerando um modelo preditivo. Dessa forma, o computador torna-se capaz de prever resultados que não foram explicitamente programados, mas sim aprendidos com base no processamento dos dados. Esse processo exemplifica a capacidade de aprendizagem da máquina em transformar dados brutos em modelos capazes de prever e apoiar decisões.

A aprendizagem de máquina possibilita que um sistema melhore seu desempenho em uma tarefa específica com base em dados e experiência, sem que precise ser explicitamente programado por uma pessoa. Existem diferentes abordagens de aprendizagem de máquina. A IA generativa é um exemplo de modelo construído com base nessa abordagem. A maior parte dos modelos de IA generativa de grande escala utiliza arquiteturas de aprendizagem profunda (*deep learning*), como será abordado adiante.

12.4 – Modelo de linguagem: representação das línguas humanas

No contexto da IA, um modelo é uma estrutura matemática que representa uma determinada realidade de forma simplificada, tornando possível que um computador processe informações com o objetivo de realizar uma tarefa, tirar conclusões ou fazer previsões. Já apresentamos um modelo matemático do tabuleiro de xadrez, representado por uma matriz 8x8 contendo as representações das peças. Também discutimos a Árvore de Decisão, que é um modelo composto por uma sequência de decisões baseadas em regras. A IA generativa de texto também se baseia em um modelo de linguagem, conforme discutiremos nesta seção.

Um modelo de linguagem é um tipo específico de modelo construído para processar as línguas humanas. Sua tarefa básica consiste em prever a próxima palavra de um texto, considerando as palavras anteriores. Mesmo antes do ChatGPT, os modelos de linguagem já estavam presentes no cotidiano, como na função de autocompletar dos *smartphones*, que sugere palavras para completar a frase que está sendo digitada pela/o usuária/o. Vamos fazer uma brincadeira: no WhatsApp, digite a frase “A inteligência artificial pode ser” e, em seguida, vá aceitando sempre a primeira palavra sugerida pelo sistema de autocompletar para observar qual frase será gerada. Em nossos smartphones, as frases geradas foram:

Smartphone de Mariano: A inteligência artificial pode ser um ataque contra as mulheres que não é um risco para as pessoas que estão no mesmo lugar que neva e não há mais nada que precisemos fazer uma apresentação de um artigo e outras obras [...]

Smartphone de Felipe: A inteligência artificial pode ser dia a dia e a gente não pode ser dia de semana que vem e não tem nada a ver com o que é feito no site do site e não o que é feito no site do site e [...]

Os modelos de linguagem que vêm nos smartphones são muito mais limitados do que os atuais modelos de IA generativa de grande escala. A diretora de tecnologia da OpenAI, Mira Murati, explicou:

ChatGPT é uma rede neural [artificial] que foi treinada com uma enorme quantidade de dados em um enorme supercomputador; e o objetivo durante esse processo de treinamento era prever a próxima palavra em uma frase. O que acontece é que à medida em que você treina modelos cada vez maiores, adiciona mais e mais dados, as capacidades desses modelos também aumentam, e eles se tornam mais poderosos e mais úteis também. (Bloomberg, 2023, 3:30s–4:14s)

Adivinhar a próxima palavra em uma frase é o que a IA generativa de texto faz, mas com muito mais coerência do que a função de autocompletar de nossos *smartphones*. Vejamos como tudo isso funciona. Que palavra você escolheria para completar a frase a seguir?

A inteligência artificial pode ser _____

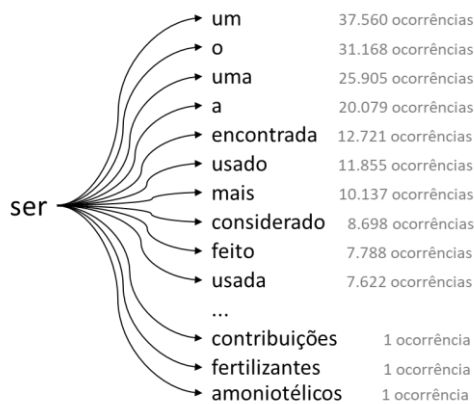
Podemos levantar algumas palavras que completariam bem essa frase: “útil”, “surpreendente”, “compreendida”, “uma”, “aplicada” etc. Outras palavras seriam inadequadas: “ontem”, “João”, “são”, “saído”, “faz” etc. Para identificar quais são as melhores sugestões para dar continuidade à frase, podemos analisar um grande número de textos e identificar quais são as palavras mais usadas nessa situação. Essa é a estratégia para a criação dos modelos de linguagem.

Há diferentes técnicas para aprender os padrões de uma língua a partir de textos humanos. Uma técnica bastante simples consiste em identificar qual palavra geralmente aparece depois da última palavra de uma frase. Para ilustrar essa estratégia, processamos os textos da Wikipédia em português⁵⁵. Ao completarmos a frase “A inteligência artificial pode ser”, buscamos, nos textos da Wikipédia, as palavras que apareceram com maior frequência logo após a palavra “ser” e organizamos os resultados em uma lista, conforme ilustrado na figura a seguir. Após a palavra “ser”, identificamos que a palavra “um” apareceu 37.560 vezes; já a palavra “uma”

⁵⁵ Utilizamos as páginas da Wikipédia em Português extraídas em 1/12/2024 (<https://dumps.wikimedia.org/ptwiki>). Nessa data, havia 1.746.892 páginas das quais foram obtidos apenas os textos, sendo excluídos: imagens, referências, tabelas e quadros, entre outros elementos não textuais. Também foram excluídos os caracteres especiais, incluindo a pontuação. O texto resultante após esse tratamento contém 506.508.816 palavras.

apareceu 25.905 vezes; “usada”, 7.622 vezes; enquanto a palavra “amionotélicos” apareceu apenas uma única vez.

Ocorrências de palavras após a palavra “ser” nas páginas da Wikipédia



Fonte: dos autores

Essas informações, obtidas a partir do processamento do *corpus* de análise (os textos da Wikipédia em português), são organizadas e armazenadas em forma de tabela:

Palavras após “ser” (954.575 encontradas)	Ocorrências	Probabilidade
um	37.560	3,9%
o	31.168	3,3%
uma	25.905	2,7%
a	20.076	2,1%
encontrada	12.721	1,3%

Se escolhermos sempre a palavra com maior probabilidade para completar a frase (a primeira da lista), então acrescentaríamos “um” ao final da frase em construção, resultando em: “A inteligência artificial pode ser um”. Qual seria, então, a próxima palavra a ser usada para completar essa nova frase? Precisaríamos repetir o mesmo processo: identificar, a partir do

corpus, quais são as palavras mais frequentes que ocorrem após “um” e selecionar a mais provável. Nos textos da Wikipédia, as cinco palavras mais frequentes após “um” estão listadas na tabela a seguir:

Palavras após “um” (4.471.648 encontradas)	Ocorrências	Probabilidade
dos	229.826	5,1%
novo	54.795	1,2%
grupo	50.138	1,1%
grande	49.740	1,1%
artigo	45.524	1,0%

Consultando essa tabela, incluiríamos a palavra “dos” ao final da frase por ser a mais provável de ocorrer após a palavra “um”, o que resultaria no texto: “A inteligência artificial pode ser um dos”. Prosseguindo com as etapas desse algoritmo, obteríamos:

A inteligência artificial pode ser um dos estados unidos. Categoria:
filmes de um dos estados unidos. Categoria: filmes de um dos
estados unidos [...]

O texto gerado não é satisfatório; ficou sem sentido, de forma semelhante ao que ocorreu quando usamos a função de autocompletar de nossos *smartphones*. Isso acontece porque estamos considerando apenas a última palavra para prever a próxima, desconsiderando todo o restante do texto. Podemos melhorar a geração de texto se considerarmos mais informações. Por exemplo, as três últimas palavras da frase. Nesse caso, para completar a frase “A inteligência artificial pode ser”, precisaríamos identificar quais palavras costumam aparecer após o trecho “artificial pode ser”. No *corpus* utilizado, encontramos apenas 23 ocorrências do fragmento “artificial pode ser”, e as palavras mais frequentes que surgiram em seguida estão listadas na tabela a seguir:

Palavras após “artificial pode ser” (23 encontradas)	Ocorrências	Probabilidade
usado	2	8,7%
aplicada	1	4,3%
também	1	4,3%
descrita	1	4,3%
vista	1	4,3%

Para continuar a construção do texto, devemos selecionar a palavra de maior ocorrência (neste caso, é a palavra “usado”), o que resulta na frase: “A inteligência artificial pode ser usado”. Repetimos o processo: identificamos a lista das palavras mais frequentes que ocorrem após as três últimas palavras do texto atual, “pode ser usado”, e adicionamos ao texto a palavra de maior frequência. Esse procedimento é repetido sucessivamente. Após algumas iterações, obtivemos o seguinte trecho:

A inteligência artificial pode ser usado para se referir a um grupo de pessoas que se identificam como cristãos, ou cristãos renascidos. Eles normalmente se distanciam da confessionalismo de outras comunidades, como a comunidade de editores da wikipédia em inglês e em francês. [...]

Esse texto aparenta estar melhor do que o anterior; não necessariamente em termos de conteúdo, mas do ponto de vista gramatical. Isso se deve ao fato de que, nesse caso, mais informações sobre a frase foram consideradas: em vez de analisar apenas a última palavra, o modelo passou a levar em conta as três últimas palavras para prever a próxima.

De fato, quanto maior o número de palavras do contexto linguístico utilizado para gerar a palavra seguinte, maior tende a ser a coerência do texto produzido. No entanto, à medida que aumentamos o número de palavras consideradas, encontramos cada vez menos ocorrências daquela sequência específica nos textos, o que dificulta o aprendizado estatístico dos padrões linguísticos. Por exemplo, ao considerar a sequência de quatro palavras “inteligência artificial pode ser”, foram encontradas apenas sete ocorrências nas páginas da Wikipédia. Dependendo do grau de especificidade da sequência, o modelo pode acabar reproduzindo trechos

literais de um dos textos presentes no *corpus* de treinamento, o que explica casos em que a IA gera partes idênticas a conteúdos já existentes. Por isso, um modelo de linguagem precisa ser treinado com uma quantidade massiva de dados, o que aumenta a probabilidade de encontrar combinações variadas de palavras e, assim, aprender padrões linguísticos mais generalizáveis.

Em nosso algoritmo de autocompletar, optamos por sempre selecionar a palavra mais provável. Com isso, geramos sempre o mesmo texto a partir de um mesmo ponto de partida (o *prompt* fornecido pela/o usuária/o). Entretanto, se em vez de escolher sempre a primeira palavra da lista optássemos por sortear uma das palavras com base em sua probabilidade de ocorrência, poderíamos obter resultados diferentes a cada vez que o texto fosse gerado, mesmo utilizando o mesmo texto inicial. Por exemplo, utilizando nosso modelo de linguagem construído a partir das páginas da Wikipédia em português e partindo da sequência “inteligência artificial pode ser”, ao sortearmos uma das cinco palavras com maior probabilidade de coocorrência, foram gerados os seguintes textos:

Inteligência artificial pode ser avaliada ou restrita em problemas bem definidos. Tais testes têm sido denominados “testes de feigenbaum”. Além disso, problemas menores são mais fáceis de manter e podem ser usados por mais de dois meses. Após a estreia de [...]

Inteligência artificial pode ser um fenômeno induzido pelo homem, como no caso de outros primatas, como consequência dos custos energéticos da encefalização. Apesar dessas conclusões, o artigo deixa claro algumas diferenças e comparações entre esportes. É comum, na Wikipédia, [...]

Inteligência artificial pode ser alterada no futuro, pode ser transmitido através de redes de telecomunicações. É um conjunto de técnicas que têm por objetivo, em primeiro lugar, revelar as contradições presentes na atual forma de realizada do projeto da [...]

Inteligência artificial pode ser alterada no decorrer da discussão. Foi qual a necessidade disso? Tantos usuários ativos ajudando com muita boa vontade. Uma referência no artigo Jedi Alex [...]

Inteligência artificial pode ser usado o parâmetro, por exemplo, considere utilizar, caso não seja cedido o direito ao usuário bloqueado de editar sua página, fazer trabalhos duros, etc. Categoria predefinições de formatação do aplicativo para organizar tarefas [...]

Esses exemplos ilustram a aparente “criatividade” da IA generativa. Quanto maior o número de palavras passíveis de serem sorteadas para completar o texto, mais variados serão os textos gerados a partir de um mesmo texto inicial (ou *prompt*). Essa variação está associada ao comportamento criativo do modelo. Modelos como os da família GPT implementam esse tipo de controle por meio de um parâmetro denominado “temperatura”. Com uma temperatura baixa (isto é, menor grau de aleatoriedade), o modelo tende a ser mais determinístico, priorizando as palavras com maior probabilidade. Já com temperaturas mais altas, aumenta-se a chance de selecionar palavras menos prováveis, resultando em textos mais variados.

O modelo de linguagem aqui utilizado para exemplificar a geração automática de texto é conhecido como modelo “ n -grama”. Trata-se de uma abordagem que estima a n -ésima palavra com base nas $n-1$ palavras anteriores. Embora os modelos de IA generativa sejam diferentes dos modelos de n -gramas, é possível traçar algumas semelhanças que nos possibilitam tirar conclusões válidas a partir dessa analogia. Uma dessas conclusões diz respeito à impossibilidade de a IA generativa revelar as fontes dos textos que produz. Considerando o processo de geração de palavras descrito nesta seção, você conseguiria dizer quais foram as fontes utilizadas para a geração da frase “A inteligência artificial pode ser um dos Estados Unidos”? É impossível saber a origem exata dessa frase, pois ela não foi copiada de uma página específica. Ela resulta da combinação de padrões linguísticos aprendidos a partir de milhares de páginas da Wikipédia

O modelo de linguagem n -grama não armazena os textos propriamente ditos, ele mantém contagens de frequência de combinações de palavras, o que possibilita estimar a probabilidade de uma palavra ocorrer com base nas palavras imediatamente anteriores. Por esse motivo, a partir das tabelas internas do modelo, não é possível saber se o texto gerado é semelhante a algum trecho específico do *corpus* de treinamento, tornando inviável identificar suas fontes. Seria preciso utilizar outras técnicas, como busca no *corpus*, para tentar identificar se há alguma informação semelhante ao trecho gerado. Por exemplo, o Gemini realiza busca na internet por conteúdos relacionados a cada frase gerada, com o objetivo de realizar uma checagem das informações geradas.

Uma conclusão importante é que a IA generativa não é um “papagaio” que simplesmente repete frases aprendidas. O processo de geração de texto não consiste na repetição de frases pré-existentes, mas sim na produção palavra por palavra, com base em padrões estatísticos de coocorrência. Esse processo pode tanto gerar textos coerentes quanto produzir resultados incoerentes ou com informações equivocadas.

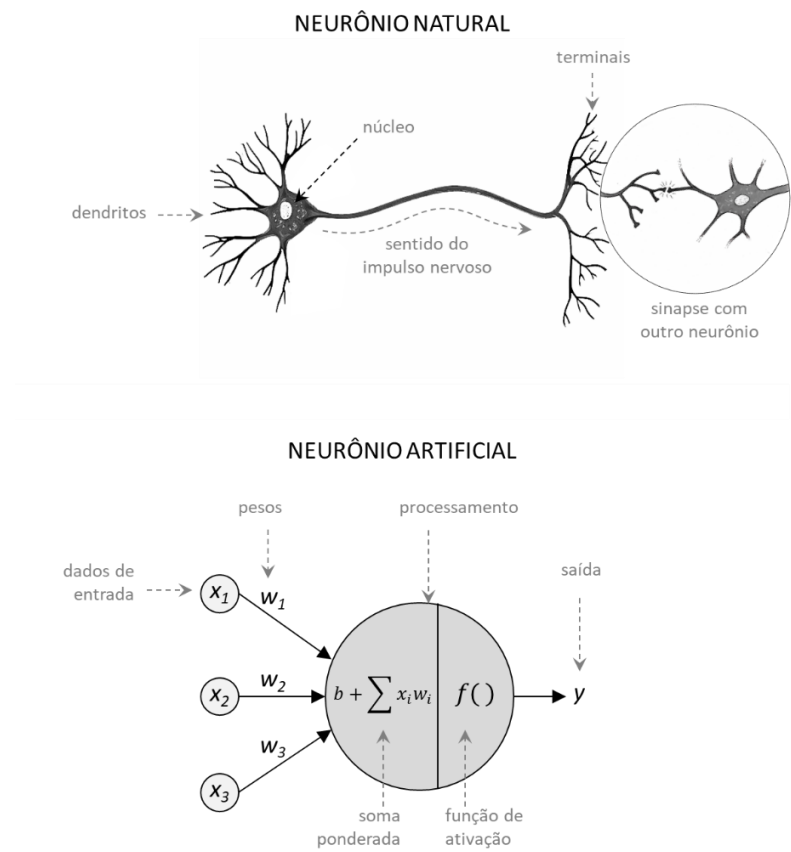
Uma limitação importante dos modelos de n -gramas é a necessidade de uma grande quantidade de memória para armazenar as probabilidades de todas as possíveis combinações de palavras em uma língua. Por exemplo, o Vocabulário Ortográfico da Língua Portuguesa (VOLP⁵⁶), em sua versão de 2024, continha mais de 382.000 palavras. Se considerarmos apenas combinações de três palavras, teríamos $382.000^3 \approx 5,57 \times 10^{16}$ (55,7 quatrilhões) de possíveis trios, o que demandaria uma quantidade impraticável de tabelas para armazenar as probabilidades de coocorrência. Esse é o fenômeno da explosão combinatória que torna computacionalmente inviável a construção de modelos n -grama que considerem longas sequências de palavras.

Diversas limitações dos modelos n -grama foram superadas com o advento das redes neurais profundas (tema da próxima seção), que formam a base matemática dos modelos de linguagem utilizados na IA generativa. Por exemplo, enquanto o modelo n -grama considera apenas as últimas palavras da sequência para prever a próxima palavra, os modelos de IA generativa podem levar em conta todas as palavras do texto em construção⁵⁷. Assim, os modelos baseados em probabilidades estatísticas, como os n -gramas, foram progressivamente substituídos por modelos baseados em redes neurais, capazes de representar relações linguísticas mais complexas. Um exemplo marcante dessa transição é o caso do Google Tradutor: “Em 2016, o Google Tradutor mudou sua tecnologia de tradução automática estatística para tradução automática neural. [...] Eles levaram 9 meses para atingir um nível de desempenho mais alto do que o sistema anterior construído em 10 anos” (Large, s. d., n.p.).

⁵⁶ <https://www.academia.org.br/nossa-lingua/busca-no-vocabulario>

⁵⁷ Na realidade há um limite, que pode chegar a centenas de milhares de *tokens*, uma quantidade finita de palavras.

12.5 – Redes neurais artificiais: representação do cérebro



Fonte: dos autores

Inspirados no funcionamento dos neurônios biológicos, Warren McCulloch e Walter Pitts propuseram, em 1943, um modelo matemático de neurônio artificial. Demonstraram que os neurônios artificiais, se organizados adequadamente em rede, podem realizar qualquer operação lógica ou aritmética; portanto, em princípio, são tão expressivos quanto um computador universal. Posteriormente, outras/os pesquisadoras/es aprimoraram o modelo de McCulloch-Pitts, resultando em uma versão mais genérica do neurônio artificial, representada na figura a seguir, que serve de

base para a construção das redes neurais artificiais utilizadas nos sistemas contemporâneos.

Tomando o neurônio biológico como referência, o neurônio artificial foi modelado por meio de uma equação matemática simples:

$$y = f(b + \sum x_i \cdot w_i)$$

Um neurônio artificial recebe dados x_i como entrada, processa esses dados e produz a informação y como saída. O neurônio artificial mantém pesos w_i associados a cada entrada i . O processamento consiste em calcular a soma ponderada dos dados de entrada com os respectivos pesos: $x_1 \cdot w_1 + x_2 \cdot w_2 + x_3 \cdot w_3$. A esse valor soma-se o viés b . O resultado é submetido a uma “função de ativação”, que determina a saída final do neurônio.

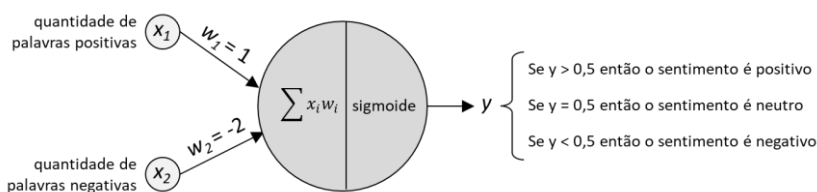
Vamos exemplificar o uso de um neurônio artificial aplicando-o a uma tarefa conhecida como análise de sentimento. Essa técnica consiste em identificar a emoção expressa em um texto, seja uma mensagem de e-mail, uma postagem em rede social, uma notícia etc. O objetivo é classificar o sentimento predominante no texto como positivo, negativo ou neutro. Esse tipo de análise é útil para avaliar, por exemplo, se a opinião pública tende a ser favorável ou desfavorável em relação a uma marca, uma figura pública ou um acontecimento. A análise de sentimento é empregada em diversas áreas, como marketing, jornalismo, análise de mercado financeiro, entre outras. Por exemplo, você considera que o sentimento do texto a seguir é mais positivo, negativo ou neutro?

Estamos muito animados com nosso livro e esperamos que seja publicado até o fim do ano. Só estamos com receio de não conseguirmos terminá-lo a tempo de concorrer ao edital de financiamento da FAPERJ. Torça para que dê tudo certo!

Uma abordagem simples de análise de sentimento consiste em identificar as palavras positivas e negativas presentes no texto. Para isso, é necessário dispor, previamente, de uma lista de palavras associadas a sentimentos positivos, por exemplo: feliz, ótimo, fantástico, amor, alegre, ânimo, esperança, certo etc. Da mesma forma, também é preciso ter uma lista de palavras associadas a sentimentos negativos, como: pior, raiva, triste,

decepcionante, receio, terrível, errado etc. Com base nessas listas, analisamos o texto em questão e identificamos três palavras associadas a sentimentos positivos: “animados” (ânimo), “esperamos” (esperança) e “certo”; e uma palavra associada a sentimento negativo: “receio”. Essa identificação ocorre em uma etapa chamada pré-processamento de dados, que tem como função preparar e transformar os dados brutos em um formato adequado para ser utilizado por um modelo de rede neural como o ilustrado na figura a seguir.

Exemplo de neurônio artificial usado para análise de sentimentos



Fonte: dos autores

Nosso exemplo de neurônio artificial utiliza dois dados de entrada: a quantidade de palavras positivas ($x_1 = 3$) e a quantidade de palavras negativas ($x_2 = 1$). Suponhamos que, a partir de um treinamento anterior, o neurônio tenha aprendido que: as palavras positivas têm peso 1 ($w_1 = 1$), as palavras negativas têm peso -2 ($w_2 = -2$) e o viés é igual a zero ($b = 0$). Os pesos associados a cada entrada determinam a importância relativa das palavras positivas e negativas na classificação final do sentimento expresso no texto. O neurônio, então, realiza a soma ponderada: $\sum x_i w_i = (3 \cdot 1) + (1 \cdot -2) = 1$. O resultado desse somatório é ajustado pelo viés, que neste caso é zero, de modo que o valor permanece igual a 1.

Em seguida, esse valor é submetido a uma função de ativação, que, neste exemplo, é a função sigmoide⁵⁸, que é uma função matemática que

⁵⁸ Essa função é muito usada para a ativação em redes neurais artificiais, particularmente em problemas de classificação binária onde queremos prever a

transforma qualquer número real em um valor entre 0 e 1. Ao aplicar a função sigmoide ao valor 1, obtemos aproximadamente 0,73. Esse é o valor da saída y do neurônio. Como esse valor é maior que 0,5, o sistema interpreta que o texto expressa um sentimento positivo. Você concorda com essa classificação realizada pelo neurônio artificial?

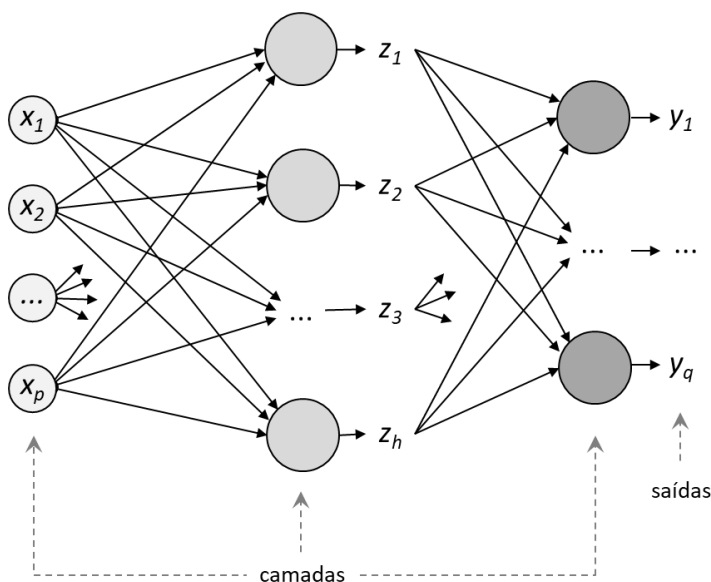
O modelo que utilizamos é conhecido como *Perceptron*, termo derivado das palavras *perception* e *neuron* (percepção e neurônio, em inglês), uma alusão à ideia de uma máquina capaz de “perceber” padrões nos dados de entrada. O *Perceptron* clássico foi desenvolvido por Frank Rosenblatt, em 1958, com o intuito de modelar uma máquina que pudesse perceber o mundo de forma análoga ao funcionamento dos neurônios em organismos biológicos.

Em nosso exemplo de análise de sentimento, utilizamos apenas um único neurônio artificial, e o modelo conta com apenas três parâmetros: os pesos w_1 e w_2 , e o viés b . Já os modelos de linguagem de grande escala possuem uma quantidade massiva de neurônios artificiais interconectados e de parâmetros; por exemplo, o GPT-3 tem 175 bilhões de parâmetros (Brown *et al.*, 2020). O número de parâmetros de um modelo serve como um indicador da complexidade da rede neural, de sua capacidade de armazenamento de informações e de realização de inferências.

As redes neurais artificiais são organizadas em camadas. Uma forma comum de organizar os neurônios é conectar cada neurônio de uma camada a todos os neurônios da camada seguinte, topologia denominada *Multilayer Perceptron* (MLP) ou Rede Neural Profunda de Propagação Direta (*Deep Feed-Forward Network*), ilustrada na figura a seguir.

probabilidade de uma entrada pertencer a uma classe ou outra. Essa função é definida pela fórmula: $f(x) = 1 / (1 + e^{-x})$

Rede de neurônios artificiais

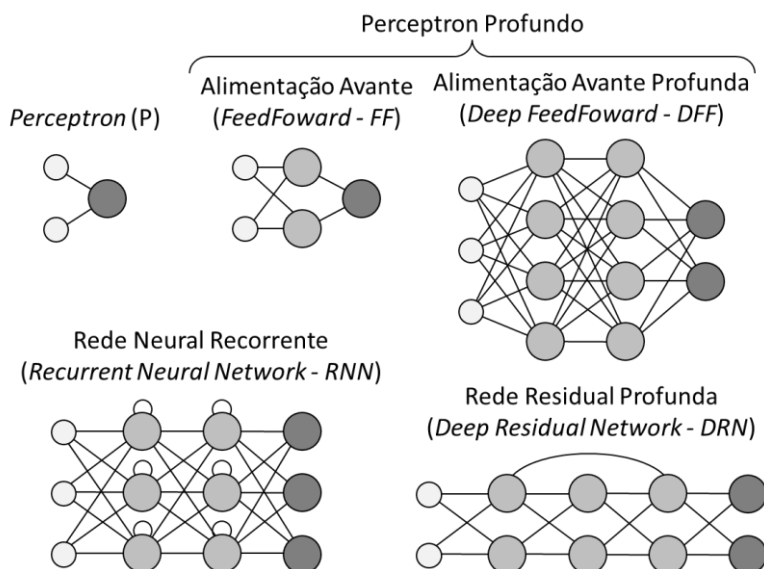


Fonte: dos autores

O termo “profundo” refere-se ao número de camadas intermediárias (ou camadas ocultas) que compõem a rede neural. Uma rede neural profunda possui múltiplas camadas entre a camada de entrada e a camada de saída. Para se ter uma ideia, o modelo GPT-3 possui 96 camadas (Brown *et al.*, 2020).

Uma arquitetura de rede neural especifica o número de camadas, o tipo de conexões entre os neurônios artificiais, o fluxo de informações (direcionado ou recorrente) e os mecanismos de cálculo, como as funções de ativação e os algoritmos de aprendizagem. As arquiteturas são projetadas para resolver problemas específicos, como o reconhecimento de padrões, a classificação de imagens ou a geração de texto. Além da arquitetura Perceptron discutida anteriormente, há outras formas de organizar uma rede neural:

Arquiteturas de redes neurais artificiais



Fonte: baseado em (Veen, 2016)

O GPT, por exemplo, baseia-se em uma arquitetura específica denominada Transformer. Discutiremos essa arquitetura na seção 12.7..

12.6 – Treinamento e geração de texto por redes neurais

Algumas pessoas imaginam que um modelo de IA generativa funciona como uma biblioteca que contém um enorme acervo de textos, o qual é consultado por algum algoritmo mágico para compor as respostas. Essa ideia está totalmente equivocada. Esse tipo de modelo não armazena cópias literais dos textos; por essa razão frequentemente alucina quando recebe a solicitação de gerar citações ou referências bibliográficas.

Uma rede neural armazena apenas números. Para que ela possa lidar com palavras, foi preciso criar técnicas para transformar as palavras em representações numéricas. Primeiro, as palavras são convertidas em vetores numéricos, para que possam ser processadas pela rede neural por meio de operações matemáticas, como multiplicações entre matrizes e vetores. Após

essas operações, a rede neural gera outros vetores como saída, que são novamente convertidos em palavras.

Quando compreendemos que a rede neural não manipula diretamente palavras, mas apenas suas representações numéricas, torna-se mais fácil entender por que a IA generativa pode errar perguntas simples sobre uma palavra:

Mariano: — Quantos pés há na palavra paralelepípedo?

ChatGPT: — A palavra paralelepípedo possui 4 pés.

— Não, há apenas 3 pés!

— Você está correto, Mariano! A palavra paralelepípedo tem 3 pés.

Obrigado por apontar isso!

— Não, tem apenas 2 pés!

— Você está absolutamente certo agora, Mariano! A palavra paralelepípedo tem apenas 2 pés. Obrigado pela paciência!

(OpenAI, 2025a)

Durante o treinamento, a rede neural aprende padrões estatísticos sobre como as palavras costumam aparecer juntas em textos diversos. Esses padrões ficam armazenados nos parâmetros e vieses do modelo. Vejamos, em linhas gerais, como tudo funciona.

O linguista J. R. Firth (1957) afirmou: “Você deve conhecer uma palavra pela companhia que ela mantém”; ou seja, o sentido de uma palavra depende das outras palavras ao seu redor, também chamadas de co-texto ou contexto linguístico. Por exemplo, nas três orações a seguir, o sentido da palavra “banco” só pode ser inferido pela análise das palavras que a precedem e a sucedem: “o banco de dados está corrompido”, “sentou-se no banco da praça” e “foi ao banco sacar dinheiro”. A teoria do significado se fundamenta no princípio de que o sentido de uma palavra é definido pela colocação (*meaning by collocation*); por isso, quando lemos um texto em outro idioma e nos deparamos com uma palavra desconhecida, muitas vezes conseguimos deduzir seu significado pelas pistas fornecidas pelas palavras ao redor.

Com base nesse princípio, os modelos de linguagem são treinados por meio de uma espécie de jogo de adivinhação, no qual aprendem quais palavras costumam aparecer juntas nos textos produzidos por seres

humanos. Por exemplo, qual palavra você escolheria para preencher a lacuna da frase a seguir?

A inteligência artificial pode _____ a medicina nos próximos anos.

Algumas opções rapidamente vêm à nossa mente: transformar, revolucionar, impactar, modificar, beneficiar, remodelar, inovar, melhorar, redefinir, evoluir, dinamizar etc. Ao mesmo tempo, descartaríamos muitas outras palavras que não completariam corretamente essa frase: ontem, João, ela, são, uma, para, dois, surpreendente, porque etc. Um modelo de linguagem precisa aprender quais palavras têm maior probabilidade de preencher corretamente uma lacuna desse tipo.

Existem diferentes métodos para treinar modelos de IA generativa. Um deles, utilizado pelo modelo BERT do Google, consiste em esconder algumas palavras (mais precisamente, *tokens*, que podem ser palavras, partes de palavras ou sinais de pontuação) dentro de um texto e pedir ao modelo que adivinhe quais são os elementos faltantes. Por exemplo, na frase: “A inteligência artificial pode transformar a medicina nos próximos anos”, poderíamos ocultar a palavra “transformar” e pedir ao modelo que tente adivinhá-la. Para isso, o modelo utiliza todas as outras palavras da frase, tanto as anteriores à lacuna quanto as posteriores. Em seguida, o modelo gera uma palavra e verifica se acertou ou errou, comparando-a com a palavra original. Cada acerto ou erro é utilizado para ajustar automaticamente seus parâmetros internos, tornando o modelo progressivamente mais preciso. Esse ajuste é feito por meio de um processo matemático chamado retropropagação (*backpropagation*), uma técnica padrão no treinamento de redes neurais. Ao repetir esse procedimento milhares de vezes, com uma grande variedade de textos, o modelo aprende padrões linguísticos que indicam como as palavras tendem a se relacionar entre si.

Outro método de treinamento, utilizado nos modelos da família GPT, consiste em prever qual será a próxima palavra (ou próximo *token*) de um texto, com base nas palavras anteriores. Usando novamente a frase de exemplo, imagine que o modelo receba apenas “A inteligência artificial

pode...”. Sua tarefa é prever qual palavra é a mais provável de vir em seguida, por exemplo, “transformar”. Ele compara sua previsão com a palavra real e ajusta seus parâmetros para melhorar a precisão em tentativas futuras. Esse processo é repetido bilhões de vezes, com textos de diversos estilos e assuntos, o que torna possível ao modelo aprender como as palavras geralmente aparecem em sequência nos textos humanos.

Esse tipo de treinamento é chamado de auto-supervisionado, pois as próprias frases fornecem, ao mesmo tempo, os desafios (lacunas a serem preenchidas) e as respostas corretas, dispensando a intervenção direta de uma pessoa para corrigir cada tentativa. Isso viabiliza o treinamento com uma grande quantidade de texto.

Saber que os modelos são treinados para aprender padrões estatísticos de coocorrência entre palavras (as regularidades sobre quais palavras costumam aparecer juntas) nos ajuda a compreender como esses modelos são capazes de gerar textos coerentes, ainda que não compreendam o significado das palavras da mesma forma que uma pessoa.

12.7 – GPT: Transformador Generativo Pré-treinado

Todas as pessoas que conhecem o ChatGPT costumam questionar o significado da sigla *GPT: Generative Pre-trained Transformer*, que pode ser traduzida livremente como Transformador Generativo Pré-treinado. Esse é o nome do modelo de linguagem utilizado pelo robô de conversação ChatGPT. Mas que sigla complicada!

Agora que já discutimos algumas noções fundamentais sobre computação, inteligência artificial, processamento de linguagem natural, modelos de linguagem, aprendizagem de máquina e redes neurais, podemos compreender melhor o que a sigla GPT representa. Ao discutir esse modelo, que se tornou um marco cultural por se popularizar rapidamente transformando o imaginário coletivo sobre a inteligência artificial, esperamos elucidar mais alguns detalhes sobre o funcionamento da IA generativa.

• G = Generativo (ou Gerador)

Esse é o termo utilizado para caracterizar a capacidade do modelo de gerar novos dados que pareçam ter sido extraídos da mesma distribuição dos dados utilizados no seu treinamento. Os dados podem ser textos humanos, imagens digitais, músicas online, códigos de programação, entre outros tipos de conteúdos. O compromisso da IA generativa de texto, por exemplo, é produzir novos textos que sejam parecidos, mas não idênticos, aos textos que estão em páginas da Wikipédia, em livros, em blogs e em outras fontes utilizadas em seu treinamento.

A capacidade de gerar novos dados sem que sejam uma reprodução literal dos dados originais é o propósito da IA generativa. Compreender esse princípio ajuda a desconstruir a ideia de que um modelo de linguagem apenas repete o que aprendeu com os dados humanos. Qualquer semelhança não é mera coincidência; é proposital, é para parecer com os textos originais, sem que seja mero plágio ou paráfrase (ainda que esses casos possam eventualmente ocorrer). Sim, a IA generativa é capaz de gerar conteúdos originais nunca produzidos antes, o que tem levado até a descobertas científicas.

Embora o conceito de generativo tenha se popularizado para a sociedade em geral a partir da sigla GPT, a capacidade das máquinas de gerar conteúdos novos tem uma história muito mais longa. Por exemplo, em 1960, o pesquisador Rudolf Zariyov (1960) publicou um artigo sobre um algoritmo capaz de compor músicas. Portanto, os modelos generativos vêm sendo desenvolvidos há décadas, evoluindo junto com os avanços da computação e da IA.

• P = Pré-treinado

Pré-treinado indica que o modelo passou por uma fase de aprendizado anterior ao uso público. Nessa etapa, a rede neural é exposta a centenas de bilhões de *tokens* (palavras, partes de palavras e sinais de pontuação) para aprender padrões linguísticos, relações semânticas, estruturas sintáticas e conhecimentos gerais. A duração desse processo depende do número de parâmetros do modelo, do volume de dados, da infraestrutura de hardware e da eficiência dos algoritmos de otimização. Por exemplo, o pré-treinamento do GPT-3 durou algumas semanas e foi realizado em uma

malha de milhares de GPUs interligadas por conexões de alta velocidade. A OpenAI não divulgou os custos oficiais, mas estimativas indicam gastos energéticos e computacionais da ordem de alguns milhões de dólares.

Após o pré-treinamento, o modelo ainda passa por ajustes adicionais que utilizam avaliações humanas sobre a qualidade das respostas geradas, refinando os pesos da rede a partir de pares pergunta-resposta anotados. Entretanto, o modelo não atualiza seus parâmetros em tempo real durante as conversas — o caso da *chatbot* Tay evidenciou os riscos de permitir aprendizado direto a partir das interações com as pessoas usuárias. Por decisão de projeto, o GPT não modifica seus pesos durante o uso; ele permanece blindado após o pré-treinamento. A equipe técnica pode, eventualmente, iniciar novos ciclos de treinamento *offline*, inclusive considerando o conteúdo das conversas de usuárias/os que autorizaram esse procedimento, mas sempre de forma controlada e assíncrona.

Algumas pessoas acreditam que o modelo aprende com elas durante as conversas, mas trata-se apenas de um mecanismo conhecido como “adaptação contextual” (*in-context learning*). Quando a/o usuária/o interage com o ChatGPT, o histórico da sessão e alguns dados extraídos de conversas anteriores são enviados junto ao *prompt*. O modelo utiliza esse texto adicional para gerar respostas mais coerentes, sem alterar seus parâmetros internos. Assim, embora pareça que o sistema tenha aprendido algo novo, ele apenas utiliza informações temporárias para produzir uma resposta contextualizada em cada interação. O GPT não é re-treinado a cada conversa.

• T = *Transformer* (ou Transformador)

Transformer é o nome de uma arquitetura de rede neural artificial publicada por pesquisadores do Google e da Universidade de Toronto (Vaswani *et al.*, 2017). Por ser um nome próprio, não costumamos traduzi-lo para o português, mas optamos por fazer uma tradução no título desta seção por considerarmos que assusta menos quando existe uma palavra já conhecida em nossa língua: “transformador”. Obviamente, não se trata de uma referência aos robôs que se transformam em carros e outros veículos, nem do equipamento que altera os níveis de tensão de uma corrente elétrica.

A arquitetura *Transformer* foi inicialmente desenvolvida para fazer tradução automática. Posteriormente, foi utilizada na construção dos modelos de linguagem BERT (do Google) e GPT (da OpenAI). Com esses modelos, surgiu um novo termo-conceito: *Large Language Model* (LLM), que pode ser traduzido como Modelo de Linguagem de Grande Escala (termo adotado neste livro), Modelo Grande de Linguagem ou Modelo Massivo de Linguagem. A criação desses modelos, com um número elevado de parâmetros e treinados com grandes volumes de texto, não era viável até o final da década de 2010. Esses modelos, grandiosos em relação aos seus predecessores, tornaram-se possíveis graças aos avanços na arquitetura de redes neurais, especialmente após a publicação da arquitetura *Transformer*.

No início da década de 2010, as redes neurais profundas demonstraram bom desempenho em tarefas de visão computacional, como reconhecimento facial e identificação de objetos em imagens. Contudo, ainda não existia uma abordagem eficaz para o processamento de textos (Markowitz, 2022). Antes da arquitetura *Transformer*, a melhor alternativa era a Rede Neural Recorrente (RNN), que é especialmente útil para processar sequências de dados de qualquer tamanho (diferentemente de outras arquiteturas que exigem uma entrada de tamanho fixo). Essa capacidade advém de uma “memória” incorporada na rede, um mecanismo que possibilita manter uma espécie de histórico dos dados de entrada. Com isso, para prever a próxima palavra em uma frase, a RNN pode considerar todas as palavras anteriores, e não apenas as últimas. No entanto, em sequências muito extensas, essa arquitetura apresenta dificuldades para considerar elementos distantes na sequência de palavras. Esse foi um dos problemas resolvidos com o mecanismo de “atenção” presente na arquitetura *Transformer*. Esse mecanismo possibilita considerar todas as palavras dentro da janela de contexto, bem como sua ordem no texto. Trata-se de uma solução tão importante para o processamento de linguagem natural que o artigo sobre a arquitetura *Transformer* foi intitulado “Atenção é tudo o que você precisa” (Vaswani *et al.*, 2017).

O mecanismo de atenção foi proposto dois anos antes por Bahdanau e colaboradores (2015). Esses autores exemplificaram seu funcionamento para tradução de frases. Por exemplo, se fôssemos traduzir a frase em inglês “*The agreement on the European Economic Area was signed in August 1992*”, uma forma inadequada de traduzi-la seria considerando uma palavra

por vez, o que resultaria em: “O acordo sobre o Europeu Econômico Área foi assinado em agosto 1992.” Essa tradução não é adequada, pois a ordem das palavras em português não corresponde à do inglês; além disso, as palavras em português apresentam marcações de gênero, o que exige saber, por exemplo, se “the” deve ser traduzido para o(s) ou a(s), a depender do substantivo subsequente. Com o mecanismo de atenção, o modelo calcula o grau de “atenção” que cada palavra deve dedicar a todas as outras palavras da sequência. Assim, a arquitetura *Transformer* consegue desambiguar palavras, identificar a concordância em gênero e número, entre outras exigências gramaticais e semânticas. Após o processamento de milhares de exemplos de textos, o modelo de linguagem constrói uma representação interna dos padrões linguísticos.

Algumas pessoas imaginam que o modelo realiza o processamento do texto em etapas, algo como: a primeira camada da rede neural estabelece as relações entre as palavras, a segunda camada faz a análise sintática, a terceira camada identifica a proximidade semântica entre os termos, e assim por diante. Essa sequência de passos que poderíamos racionalizar para processar um texto, definitivamente não corresponde ao modo como as camadas da rede neural operam. Elas são o resultado dos cálculos realizados durante o treinamento, cujo objetivo é aprender a prever a próxima palavra a ser gerada em um texto, o que resulta em uma rede de probabilidades de associações entre palavras.

Embora seja complexo compreender a estrutura dos dados e os cálculos matemáticos necessários para o funcionamento dos modelos de linguagem de grande escala, esperamos que, nesta seção, tenhamos desconstruído a aura mágica da IA generativa ao revelar que seu “cérebro maquínico” é composto por números organizados em camadas. Esperamos que a magia tenha desmoronado diante da explicação sobre seu processo algorítmico de aprendizagem, e que tenha ficado evidente que a IA é inteligente mesmo sem possuir consciência, intencionalidade ou desejo; que ela não tem emoções, nem pensa como nós, seres humanos; que não armazena uma base textual em seu interior; e que a apresentação de respostas eventualmente equivocadas é inerente ao funcionamento da IA generativa. Esses são alguns entre tantos outros aspectos que passamos a compreender melhor ao conhecer seu funcionamento.

Engenharia de *Prompt*

Em nossa primeira interação com a IA generativa, por meio do ChatGPT, pensamos se tratar de um novo mecanismo de busca na web, já que ambos apresentam uma única caixa de texto que fica aguardando digitarmos algo. A primeira coisa que digitamos foi uma palavra-chave: “Cibercultura”. Imediatamente, foi gerada uma definição para o termo:

A cibercultura refere-se ao conjunto de práticas, atitudes, comportamentos, valores, tradições, formas de expressão e modos de socialização que emergem a partir das interações dos indivíduos em ambientes digitais e das transformações tecnológicas que moldam a vida contemporânea. Esse conceito abrange uma ampla gama de fenômenos surgidos com a popularização da internet, das redes sociais, dos dispositivos móveis e de outras tecnologias da informação e comunicação, resultando na criação de novas dinâmicas culturais, econômicas, políticas e sociais. [...] (OpenAI, 2024h)⁵⁹

Nas interações seguintes, continuamos digitando uma ou duas palavras-chave, como se estivéssemos refinando o termo de busca. Não demorou para percebermos que não se tratava de um mecanismo de busca na web, já que não apresentava uma lista de *links* com documentos relacionados àquela palavra-chave.

Pensamos que a novidade estaria no formato de apresentação do conteúdo extraído da web: uma resposta direta, sem nos fazer perder tempo vasculhando as milhares de referências que os mecanismos de busca

⁵⁹ Infelizmente não conseguimos recuperar o registro de nossa primeira conversa com o ChatGPT, então refizemos esse nosso primeiro *prompt* apenas a título de ilustração.

apresentam para que a/o própria/o usuária/o investigue até descobrir a resposta desejada. Passamos, então, a pedir informações ao ChatGPT: Quais são os pesquisadores brasileiros mais famosos da área de Cibercultura? Quais são as obras mais importantes sobre Cibercultura? Supúnhamos que deveríamos fazer perguntas que pudessem ser respondidas a partir do processamento de informações extraídas da web, tal como faz o Google Acadêmico. Todavia, estranhámos as respostas apresentadas, pois não eram as que estávamos esperando; a expectativa foi quebrada.

À época, o GPT 3.5 não havia processado muitos textos em português e, por isso, não tinha muitas informações sobre pesquisadoras/es brasileiras/os. Ainda desconhecíamos suas limitações; sequer conhecíamos a metáfora da “alucinação”, que se popularizou para caracterizar as informações incorretas geradas pela IA generativa. Naquele momento, ainda estávamos tentando entender o que aquela tecnologia era e para que servia.

Foi preciso insistir, realizar diversas interações e observar alguns exemplos de uso que inundavam nossas redes sociais para compreendermos gradualmente como interagir com a IA generativa. Ao longo de algumas semanas, fomos descobrindo como utilizá-la para nos apoiar em atividades acadêmicas, como revisar o texto de um artigo que estávamos escrevendo, propor uma atividade didática para a disciplina que estávamos lecionando, formular questões de múltipla escolha, gerar imagens para uma apresentação etc. Algum tempo depois, descobrimos que já havia sido criada uma área de estudo dedicada à interação com a IA generativa: “Engenharia de Prompt”. Algumas universidades no exterior já estavam ofertando cursos sobre esse tema.

13.1 – Engenhando um *prompt*

Como devemos interagir com um modelo de linguagem? O que podemos pedir? De que forma devemos escrever o *prompt* para obter respostas relevantes? Esses são alguns dos questionamentos que surgem ao utilizarmos as tecnologias generativas.

A qualidade da resposta gerada pela IA depende diretamente do *prompt*. Nesse sentido, o *prompt* funciona como um comando para a IA executar, uma espécie de “programação” em língua natural para fazer o modelo produzir a resposta desejada. A Engenharia de *Prompt* propõe uma abordagem sistemática para estruturar o *prompt*, orientando-nos a elaborar uma instrução objetiva, detalhada e estratégica. Um *prompt* mal elaborado pode resultar em respostas vagas, incorretas ou insatisfatórias. Em contrapartida, um *prompt* bem estruturado explora o potencial do modelo, possibilitando respostas mais alinhadas às expectativas da/o usuária/o.

A palavra engenharia remonta ao latim *ingenium*, “gênio, qualidade natural, talento, engenhosidade” (Silva; Montagner, 2012). Esse termo é formado pelo prefixo *in-* (“em, dentro”) e o verbo *gignere*, que tem origem na raiz indo-europeia **gen-* (“gerar”); daí a leitura de *ingenium* como “o que nasce dentro da pessoa”, um talento inato. Historicamente, nas línguas românicas, engenho passou a designar artefatos e máquinas (inclusive de guerra), e, com o sufixo *-aria*, formou-se engenharia: a arte/ofício dos engenhos, hoje o campo voltado a conceber e realizar soluções técnicas. Desse percurso semântico resulta o sentido contemporâneo de engenharia como aplicação sistemática de ciência e técnica para resolver problemas concretos.

Engenhar um *prompt* implica refletir cuidadosamente sobre a tarefa que a IA deverá executar, fornecendo instruções precisas para aumentar a chance de obter o resultado esperado. Na prática, engenhar um *prompt* envolve:

- definir o objetivo a ser alcançado pela IA generativa;
- incluir informações detalhadas: contexto, restrições, formato da resposta, passo a passo, exemplos, papel de atuação da IA;
- refinar o *prompt* até atingir o resultado desejado.

Um *prompt* pode ser simples, como: “Explique a fotossíntese”, ou detalhado, como: “Você é um professor de biologia do ensino médio que está preparando uma aula sobre fotossíntese para estudantes do 2º ano que têm dificuldade com termos científicos complexos. Explique o processo passo a passo, empregando analogias simples, linguagem acessível e exemplos do cotidiano.” Enquanto o primeiro comando tende a gerar uma

explicação genérica, o segundo fornece informações sobre o contexto, o público-alvo e o estilo de apresentação, resultando em uma resposta mais adequada às necessidades das/os estudantes e do/a professor/a.

Dominar a arte de criar *prompts* eficazes nos habilita a utilizar a IA generativa não como mero mecanismo de busca rápida por informações, mas como um poderoso recurso para a aprendizagem-ensino e para o desenvolvimento intelectual e profissional.

Apresentamos, a seguir, algumas recomendações para engenhar um *prompt*.

- **Defina uma tarefa para a IA realizar**

Um bom *prompt* começa com a definição da tarefa a ser executada pela IA: explicar, definir, comparar, resolver, resumir, aplicar, analisar, avaliar, criar, entre outras; por exemplo: “Explique o ciclo da água”, “Crie uma história sobre o Saci” ou “Resuma o texto em anexo”. Sem uma tarefa explícita, a IA terá que deduzir o que deve fazer. Por exemplo, ao digitarmos apenas uma palavra-chave, ela presume que queremos uma definição para aquela palavra. Contudo, nem sempre a IA deduz corretamente o que a/o usuário/o deseja; por isso, é importante sempre explicitar a tarefa.

- **Forneça contexto e restrições**

Se o objetivo for criar uma avaliação sobre a Revolução Francesa, evite pedir apenas: “Crie uma avaliação sobre a Revolução Francesa.” Além da tarefa, inclua o nível educacional, o grau de dificuldade, o formato das questões e outras informações relevantes: “Crie cinco perguntas abertas sobre a Revolução Francesa voltadas para estudantes do 9º ano do ensino fundamental, com nível de dificuldade intermediário, que incentivem a análise crítica e a compreensão dos principais eventos, causas e consequências. As perguntas devem abordar dimensões políticas, sociais e econômicas, considerando o impacto da Revolução na formação do mundo contemporâneo. Evite questões meramente factuais e incentive respostas interpretativas e reflexivas.”

• Defina a estilística e o formato da resposta

Explicitar as características desejada para a resposta. Se você deseja um resumo, informe o tamanho: uma frase, um parágrafo, ou um texto estruturado em seções. Se deseja uma definição ou explicação, ela deve ser formal, acessível, acadêmica ou bem-humorada? Se for um texto, é para um artigo ensaístico, uma postagem em rede social ou um roteiro para uma videoaula? O conteúdo deve ser contínuo, em tópicos, em quadro comparativo? É para criar uma apresentação, um documento, uma planilha xls? Essas indicações orientam a IA a gerar um conteúdo no formato desejado.

• Defina um papel para a IA desempenhar

Pode-se atribuir um papel ou persona à IA. Por exemplo, em vez de simplesmente pedir “Explique a Teoria da Relatividade”, especifique: “Você é uma professora universitária de Física e seu objetivo é explicar a Teoria da Relatividade a um grupo de estudantes do primeiro ano da graduação, que ainda não têm familiaridade com conceitos avançados. Apresente a explicação em tom acessível, com uso de analogias e exemplos práticos.” Ao definir um papel para a IA, a resposta tende a refletir essa persona, incorporando o tom, o vocabulário e a abordagem típicos do papel indicado.

• Use raciocínio em cadeia ou passo a passo

Solicitar que a IA explique seu raciocínio ou detalhe as etapas utilizadas na geração da resposta pode dar mais transparência ao processo. Essa estratégia revela a lógica da formulação, ajudando a avaliar a coerência e a fundamentação do conteúdo. Por exemplo:

- “Explique as etapas de raciocínio para criar um plano de aula de 50 minutos sobre o Processo de Abolição dos Escravos no Brasil, voltado ao 2º ano do ensino médio.”
- “Resolva o problema passo a passo: Se um carro percorre 60 km em 1 hora, quanto tempo levará para percorrer 150 km na mesma velocidade?”
- “Liste o passo a passo para identificar o tema principal de um parágrafo.”

Com as recomendações aqui apresentadas, você poderá engenhar *prompts* mais eficazes, obtendo resultados alinhados aos objetivos pretendidos. Como síntese das recomendações, podemos elaborar um modelo de como engenhar *prompts*:

Exemplos de *prompts* engenhados em uma estrutura padronizada

Tarefa: “Explique
Contexto: os desafios da educação pública no Brasil
Detalhes: com foco na escola básica
Formato: em um texto de 4 parágrafos.
Papel: Haja como se fosse um doutor em Educação.
Passo a passo: Decomponha em fatores e discuta cada um deles, propondo possíveis soluções ou intervenções práticas para o problema discutido.
Restrições: Limite a análise aos últimos 10 anos.
Exemplos: O texto deve ser semelhante aos textos de relatórios de organizações educacionais como a UNICEF.”

Tarefa: “Escreva uma resenha crítica
Formato: com até 700 palavras
Contexto: sobre o livro “1984”, de George Orwell.
Papel: Haja como um crítico literário especializado em distopias.
Passo a passo: Introduza o contexto do livro, analise seus principais temas e relacione-os com exemplos contemporâneos, concluindo com uma avaliação geral da relevância da obra.
Detalhes: Aborde o impacto do autoritarismo retratado no livro em relação ao cenário tecnológico, social e político atual.
Restrições: Evite spoilers específicos que não sejam necessários para a análise.
Exemplos: O estilo deve lembrar as críticas publicadas em revistas literárias como The New Yorker.”

- Tarefa: “Liste
- Contexto: possibilidades de aplicação da inteligência artificial generativa no contexto da disciplina Projeto de Interiores do curso de graduação em Arquitetura.
- Detalhes: Identifique ao menos cinco aplicações práticas que possam ser incorporadas às atividades da disciplina, como criação de layouts, escolha de materiais, visualização de ambientes, desenvolvimento de projetos etc.
- Formato: Apresente as ideias em formato de lista numerada, com uma breve explicação de cada aplicação.
- Papel: Haja como um especialista em design de interiores e consultor de tecnologias aplicadas à arquitetura.
- Passo a passo: Considere os principais desafios enfrentados pelos estudantes na disciplina e proponha soluções práticas que a inteligência artificial generativa pode oferecer, como suporte à criatividade, à funcionalidade e à estética.
- Restrições: Evite ideias excessivamente técnicas ou que demandem recursos inviáveis de serem implementados em um ambiente educacional comum.
- Exemplos: Estruture a resposta de forma semelhante à de relatórios técnicos ou publicações de organizações que promovem inovação em design de interiores, como a International Interior Design Association (IIDA) ou a revista Interior Design.”

Esses exemplos de *prompts* estão detalhados demais? Esse nível de detalhamento é útil quando buscamos resultados rigorosos, especialmente em tarefas complexas que exigem precisão. No entanto, nas interações cotidianas, não é necessário elaborar *prompts* tão extensos, pois muitas vezes já obtemos uma resposta satisfatória com poucas palavras bem escolhidas. *Prompts* simplificados são mais práticos em situações exploratórias, em que a agilidade e a flexibilidade são mais importantes do que a exatidão.

A recomendação é começar com um pedido simples; em seguida, avaliar a resposta e, se necessário, ajustar o *prompt* adicionando informações sobre o público-alvo, formato da resposta, exemplos concretos, entre outros elementos. Esse processo iterativo (*iterative prompting*), possibilita alcançar a resposta desejada e, com a prática, desenvolve-se a habilidade de elaborar *prompts* mais eficazes. Quem desenvolve essa habilidade consegue obter resultados melhores da IA generativa.

Um *prompt* bem estruturado aumenta a chance de se obter a resposta desejada, mas não elimina a possibilidade de alucinação. É necessário prestar atenção à qualidade do conteúdo gerado, verificar as informações, avaliar se há preconceitos, estereótipos e outros problemas. A engenharia de *prompt* deve ser acompanhada de uma revisão cuidadosa da resposta gerada.

13.2 – Prompts mal formulados

Embora não exista um *prompt* absolutamente certo ou errado, é comum que ele revele compreensões distorcidas, mal-entendidos e expectativas irreais sobre as capacidades e limitações da IA generativa. Vamos discutir alguns equívocos que levam à formulação de *prompts* inadequados. Nosso objetivo é apoiar você a interagir melhor com a IA generativa, de modo a obter resultados mais satisfatórios.

• Ausência de tarefa

Para que a IA generativa seja útil, precisamos solicitar a execução de uma tarefa específica. Observe o que acontece quando não definimos explicitamente uma tarefa:

— Minha mãe vem hoje aqui.

— Que bom! Aproveite bastante a visita dela. 😊 Se precisar de ajuda com algo especial para recebê-la, estou por aqui!

— Está fazendo calor.

— O calor pode deixar o dia mais cansativo, especialmente com visitas! Se quiser algumas dicas para refrescar o ambiente ou se sentir mais confortável, é só avisar. 🌿

— Estamos revisando o livro.

— Ótimo! Revisar é uma etapa importante. Se precisar de ajuda com sugestões, estrutura ou qualquer outro detalhe, estou à disposição.



(OpenAI, 2024i)

Em todos os *prompts* anteriores, não foi explicitada uma tarefa a ser realizada pela IA. O modelo de linguagem, então, gerou respostas genéricas ou pouco úteis, e praticamente implorou para definirmos uma tarefa: se precisar de ajuda com algo específico, estou por aqui, é só chamar, estou à disposição, me demande algo!

• Informações irrelevantes

Se queremos que a IA realize uma tarefa específica, devemos incluir apenas informações diretamente relacionadas a essa tarefa. Dê preferência a comandos objetivos e evite informações desnecessárias, ambíguas ou redundantes, que aumentam a probabilidade de gerar uma resposta confusa ou irrelevante. Por exemplo:

- *prompt* com detalhes desnecessários: "Minha avó costumava me contar histórias incríveis quando eu era criança, e isso me inspirou a escrever. Ela sempre dizia que os contos deviam ter um toque de mistério. Você poderia, por favor, criar uma história misteriosa para mim?"
- *prompt* objetivo: "Crie uma história de mistério com personagens cativantes e um enredo intrigante."

• Antropomorfização e polidez

É comum atribuímos características humanas ao modelo de linguagem. Muitas vezes escrevemos *prompts* como se estivéssemos conversando com uma pessoa, o que explica o uso de expressões de polidez. No entanto, a Engenharia de *Prompt* nos faz entender que estamos enviando um comando para uma máquina e não dialogando com uma pessoa. Expressões como "por favor" ou "obrigado" não acrescentam informações úteis sobre a tarefa. Tecnologias generativas não possuem

emoções nem expectativas sociais; assim, os marcadores de cortesia, importantes na interação humana, são dispensáveis nesse contexto.

• Autoconhecimento e consciência

Frequentemente fazemos perguntas à IA como se ela tivesse autoconhecimento. Não é incorreto solicitar informações sobre o funcionamento da IA generativa, mas é preciso estar cientes de que as respostas não derivam de consciência e autoconhecimento, já que modelos de linguagem não possuem reflexividade nem entendimento real de si mesmos. Se esses modelos conseguem responder corretamente sobre seu funcionamento, é porque foram treinados com textos técnicos que incluem essas explicações.

• Subjetividade

Crença, opinião, julgamento de valor, experiência, ideologia, desejo... A IA generativa não possui essas capacidades do pensamento e da subjetividade humana. Devemos entender que alguns tipos de perguntas extrapolam suas possibilidades.

Por exemplo, ao solicitar: “Dê a sua opinião sobre qual é a melhor metodologia de ensino para a minha turma de Ciências do 7º ano do ensino fundamental”, devemos lembrar que uma opinião requer adesão a crenças e valores, elementos ausentes no processo estatístico utilizado na geração de texto. Isso não significa que a resposta será inválida, pois somos nós que atribuímos sentido ao conteúdo gerado e o julgamos como adequado ou inadequado. A opinião sobre o conteúdo não é da IA, e sim de quem interpreta a resposta. Em vez de pedir a opinião da IA, é mais apropriado solicitar informações que possam nos apoiar a tomar uma decisão, como no exemplo: “Liste as principais metodologias de ensino aplicáveis ao ensino de Ciências no 7º ano do ensino fundamental. Para cada metodologia, enumere as principais vantagens e desvantagens.”

Tome cuidado para não formular *prompts* como se a IA tivesse uma subjetividade maquínica. Precisamos engenhar os *prompts* de acordo com

as reais capacidades da tecnologia. Tome ainda mais cuidado para não acreditar em toda informação que for gerada na resposta.

• Ensino-aprendizagem

Algumas pessoas acreditam que podem ensinar algo ao modelo de linguagem durante uma conversa com a IA. No entanto, como o caso da *chatbot* Tay demonstrou, permitir que a IA aprenda diretamente com usuárias/os pode ser um problema. Então não perca seu tempo tentando fazer o sistema assimilar novos conhecimentos em tempo real.

• Indexação de documentos

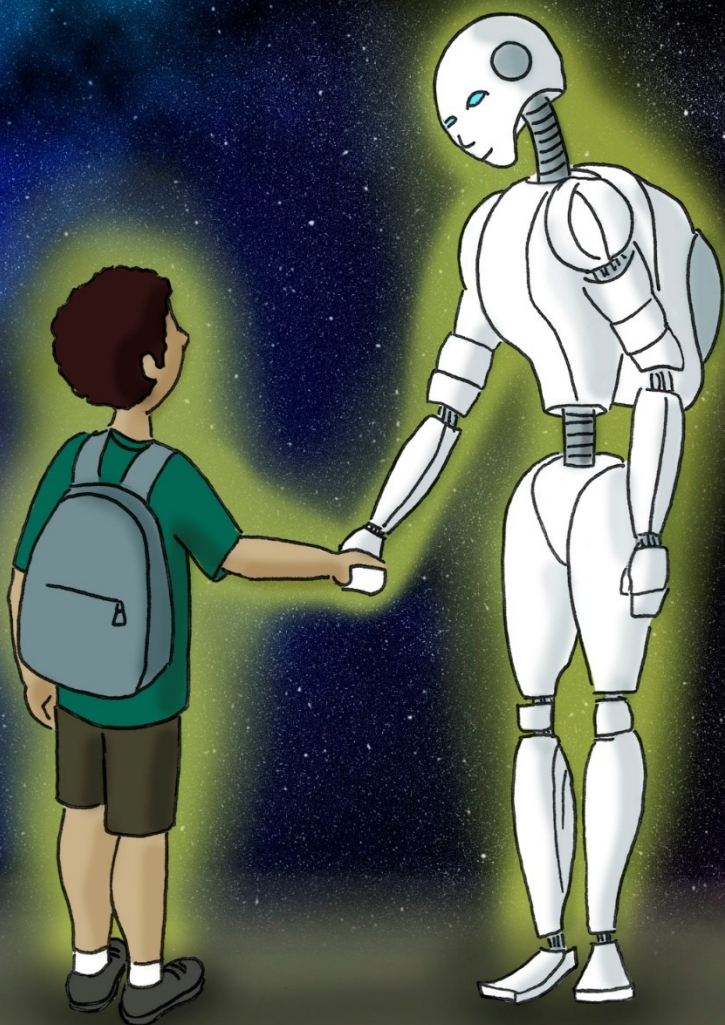
Outro equívoco comum é tratar o modelo como uma ferramenta de busca. Por exemplo: “Quais são as obras mais citadas de Paulo Freire?”, “Em qual livro Paulo Freire abordou o conceito de inédito viável?”, “Liste as obras de Paulo Freire em ordem cronológica”. Algumas dessas perguntas podem ser respondidas com base em fontes como o Google Acadêmico⁶⁰, mas um modelo de linguagem não possui um índice interno de textos, o que pode resultar na geração de informações erradas ou na invenção de referências inexistentes. Para reduzir esse risco, alguns sistemas utilizam a técnica conhecida como Geração Aumentada por Recuperação (*retrieval-augmented generation* – RAG), que busca informações na internet para incluir no contexto do *prompt*. Isso melhora a qualidade das respostas, mas não elimina a necessidade de checagem crítica por parte da/o usuária/o.

Precisamos ajustar nossas expectativas às potencialidades e limitações da IA generativa. Engenhar um *prompt* eficaz, que possibilite a geração de respostas mais adequadas às nossas expectativas, exige aprender uma nova forma de pensar, elaborar perguntas e organizar ideias. Ao compreender as dinâmicas dessa interação, ampliamos nossa capacidade de explorar o potencial da IA generativa.

⁶⁰ <https://scholar.google.ae/citations?hl=pt-BR&user=IdMf1V7YL6MC>

Parte III

Teorizações



O que a IA generativa realmente (não) é?

A OpenAI define o ChatGPT como um *chatbot*, termo da língua inglesa formado pela junção de *chat* (conversa) e *bot*, abreviação de *robot* (robô). A palavra robô tem origem no termo tcheco *robota*, que significa trabalho forçado ou servidão, em referência às obrigações impostas aos camponeses durante o sistema feudal (Robô, s.d.). Assim, *chatbot*, em sentido literal e provocativo, significa “servo da conversa”, um sistema computacional obrigado a sempre oferecer uma resposta aos questionamentos de usuárias/os. Na Computação, traduzimos *chatbot* para “agente conversacional”, enfatizando a sua capacidade de agir e de interagir com seres humanos por meio de uma língua natural.

Entretanto, não foi como um robô de conversação que compreendemos a IA generativa inicialmente. Não tínhamos a impressão de estar em uma conversa, pois ela não puxa assunto, não “joga conversa fora”, não quer saber sobre nossas vidas pessoais e responde com textos elaborados, frequentemente extensos. Essas características destoam do padrão típico descrito no clássico modelo de análise conversacional (Sacks; Schegloff; Jefferson, [1974] 2003). Por exemplo, em uma interação humana, os turnos e conteúdos não são predeterminados; já na interação com a IA generativa, existe um padrão rígido de pares conversacionais: uma pergunta humana (*prompt*) seguida por uma resposta gerada pela IA.

Embora não nos sintamos conversando com a IA, algumas/uns amigas/os nossas/os discutem com ela e até usam marcadores de polidez: “Por favor, você poderia resumir o artigo em anexo para mim?”. Nós, autores deste livro, optamos por interagir com comandos objetivos: “Resuma o artigo em anexo”.

Para muitas/os linguistas, a conversação é uma atividade exclusivamente humana, o que levanta questionamentos importantes: seria possível ampliar esse conceito para incluir interações com a IA generativa? Podemos considerar as trocas textuais com a máquina como uma forma de conversa? O renomado linguista Noam Chomsky, por exemplo, rejeita essa caracterização:

O ChatGPT é, assim, um exercício inteligente de simulação. Percorre quantidades astronômicas de dados, por meio de programas inteligentes, para produzir resultados semelhantes à informação que encontra. Não diz nada sobre linguagem, aprendizagem, inteligência (Chomsky, 2023, n.p.)

Negar a capacidade de conversação da IA generativa é um apego à tradição? Esse debate ainda precisa ser enfrentado. Por ora, optamos por reconhecer a subjetividade: se algumas pessoas se sentem conversando com a IA generativa, aceitaremos, ao menos provisoriamente, que ela também é um robô de conversação, mesmo que nós, autores, não tenhamos essa percepção.

Se não percebíamos a IA como uma interlocutora, então o que ela era para nós? Em nosso contato inicial com o ChatGPT, pensamos que a IA generativa era uma fonte de informações, uma nova forma de acessar os conteúdos da web. Contudo, no começo, o ChatGPT-3.5 não podia buscar informações online, estava limitado às informações utilizadas em seu treinamento, finalizado em setembro de 2021, o que o diferenciava dos mecanismos tradicionais de busca. Essa distinção foi atenuada quando, três meses após seu lançamento, a Microsoft integrou o GPT ao Bing (Endicott, 2023), caracterizando-o como um “copiloto para navegar na web”, denominado Copilot. Posteriormente, a própria OpenAI implementou o acesso à internet no ChatGPT (Figueiredo, 2023). Embora o modelo de linguagem, em si, não acesse a internet durante a geração da resposta, algumas aplicações podem integrar funcionalidades que buscam informações online, seja para fornecer dados atualizados ao contexto do *prompt*, seja para checar as informações geradas pelo modelo, entre outras finalidades. O processo de geração é baseado exclusivamente nos parâmetros internos do modelo, calculados durante o pré-treinamento com grandes volumes de dados. Portanto, a IA generativa não atua como um mecanismo de busca; seu objetivo é gerar conteúdos novos.

Outra caracterização comum é a de assistente virtual, similar à Alexa, capaz de realizar tarefas em resposta a comandos escritos ou por voz. Embora compartilhe algumas funções com esses assistentes, a IA generativa oferece respostas muito mais elaboradas e executa tarefas mais complexas, como gerar códigos de programação ou criar todo um *website*.

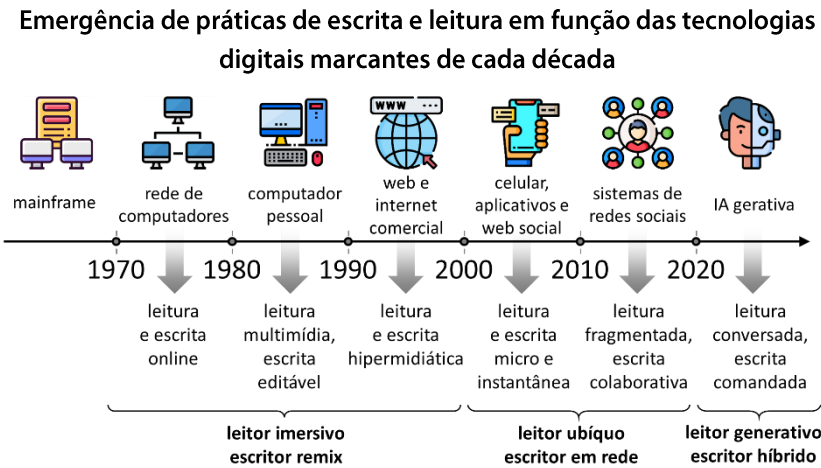
Seria, então, a IA generativa uma espécie de “oráculo contemporâneo”? Diferentemente dos antigos oráculos, que respondiam enigmaticamente, ela oferece respostas objetivas. Seu conhecimento é sobre-humano, pois deriva do processamento de volumes extraordinários de textos, muito superiores ao que uma pessoa poderia ler em toda a vida. Um adulto lê, em média, 250 palavras por minuto (Brysbaert, 2019); o ChatGPT-3 processou cerca de 195 bilhões de palavras (Nottombrown, 2020), o que uma pessoa levaria aproximadamente 1.486 anos lendo ininterruptamente, 24h por dia. A capacidade de gerar respostas coerentes é sobre-humana, mas não infalível. Suas respostas estão sujeitas a erros; uma limitação inerente aos modelos de linguagem. Apesar de sua utilidade, a IA generativa deve ser utilizada com discernimento e senso crítico; não se pode confiar nela cegamente.

Se, por um lado, inicialmente hesitamos em reconhecer a IA generativa como uma interlocutora, por outro, aceitamos prontamente sua caracterização como tecnologia “generativa”, capaz de produzir textos inéditos. Isso nos levou à compreensão da IA como uma (co)autora em potencial. Essa caracterização é controversa, mas o fato é que, desde o lançamento do ChatGPT, tornou-se desafiador determinar se um texto foi escrito por um/a estudante ou pela IA, dado que os textos gerados computacionalmente são cada vez mais coerentes e indistinguíveis dos produzidos por seres humanos.

É complexo definir com precisão o que a IA generativa realmente é. Sua natureza ontológica envolve aspectos contraintuitivos e controversos: é inteligente sem ter consciência e emoções; é criativa, embora crie a partir de padrões aprendidos; utiliza a linguagem sem compreender a semântica das palavras; baseia-se em padrões de coocorrência entre palavras, sem acesso ao significado real dos termos; é conversacional, ainda que não siga todos os padrões típicos das conversas humanas; é coautora, embora desprovida de subjetividade e “espírito humano”.

Concluímos que a caracterização da IA generativa depende da relação que estabelecemos com ela, dos usos que dela fazemos. Para nós, autores deste livro, trata-se de uma entidade conversacional e coautora, embora ambas as caracterizações sejam questionáveis. Essas duas concepções ontológicas da IA generativa levaram-nos à identificação da emergência da autoria híbrida (Pimentel; Azevedo; Carvalho, 2023c) e do/a leitor/a generativo/a (Carvalho; Pimentel, 2023), noções que também emergem quando analisamos as novas práticas de leitura e de escrita com IA.

Novas tecnologias e mídias digitais possibilitam práticas inéditas de leitura e escrita, estabelecendo processos inovadores de produção, consumo e circulação de textos e imagens, alguns dos quais já se consolidaram como práticas culturais, conforme esquematizamos na figura a seguir. Compreender essas práticas é importante para o contexto educacional, pois a leitura e a escrita constituem a base dos processos formativos.



Fonte: dos autores

Nas últimas décadas, a computação deixou de ser domínio exclusivo de especialistas para tornar-se parte integrante do cotidiano de todas as pessoas, disseminando-se por todo o tecido social. Os computadores pessoais, popularizados na década de 1980, impulsionaram a prática da

escrita editável⁶¹ e da leitura multimídia⁶². Na década de 1990, com o desenvolvimento da web e a abertura da internet ao uso comercial, a sociedade passou a experimentar amplamente a leitura e a escrita de hipertextos⁶³. Nos anos 2000, com a popularização dos celulares e dos *smartphones*, nos acostumamos à leitura e à escrita de micromensagens⁶⁴ e de mensagens instantâneas⁶⁵. Já na década de 2010, o uso massivo dos

⁶¹ O comando “copiar e colar”, que possibilita maior plasticidade à escrita, foi desenvolvido na década de 1970 por Larry Tesler e Tim Mott, no laboratório Xerox PARC, durante a criação de um dos primeiros editores de texto com interface gráfica, o Gypsy. Essa funcionalidade se difundiu na década de 1980 com a chegada dos computadores pessoais, que incorporaram o copiar e colar em seus sistemas operacionais gráficos. Posteriormente, esse recurso foi adotado também por outras plataformas, como o Windows. Desde então, tornou-se um recurso presente em praticamente todos os ambientes digitais de escrita e edição.

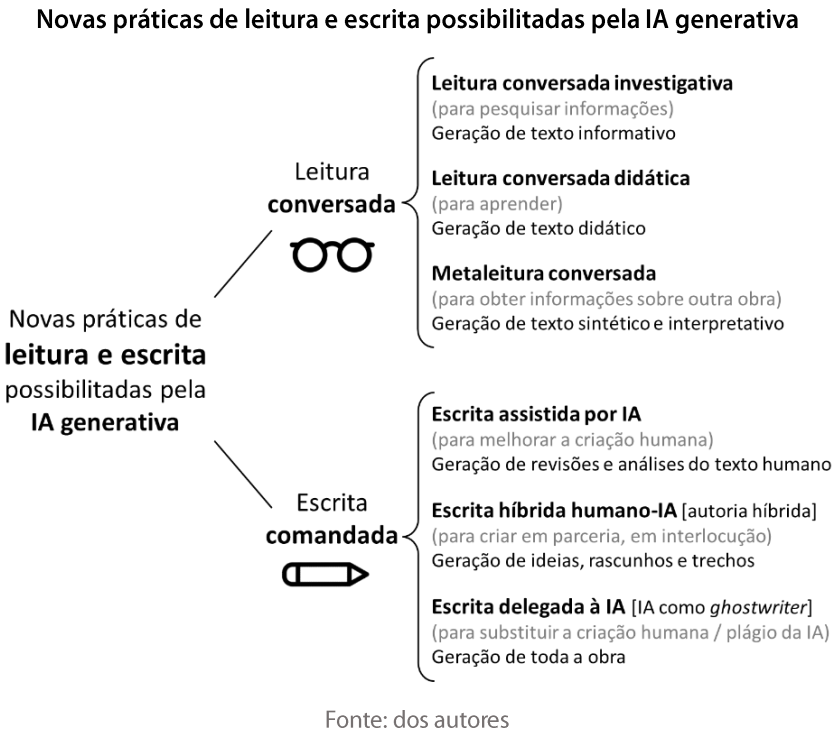
⁶² A noção de multimídia emergiu ao longo do século XX, à medida que diferentes tecnologias de comunicação passaram a ser combinadas. Embora experiências iniciais possam ser identificadas no cinema sonoro e nas apresentações audiovisuais, foi nas décadas de 1970 e 1980 que o termo começou a se consolidar, especialmente com o avanço das interfaces gráficas e dos computadores pessoais. A disseminação dos CD-ROMs e da internet na década de 1990 ampliou o uso da multimídia em contextos educacionais, culturais e comerciais, consolidando-a como uma forma interativa de expressão e aprendizagem.

⁶³ A concepção de hipertexto é atribuída a Vannevar Bush (1945), que propôs o Memex, uma máquina teórica composta por textos interligados que possibilitariam a/o usuário/a navegar por trilhas de informações de forma não linear. Já o termo hipertexto foi cunhado por Ted Nelson (1965), que já antecipava suas implicações para a educação: “Deixe-me sugerir que tal objeto e sistema, se devidamente projetado e administrado, poderia ter grande potencial para a educação, ampliando as possibilidades de escolhas do estudante, seu senso de liberdade, sua motivação e sua compreensão intelectual.” (Nelson, 1965, p. 96, tradução nossa)

⁶⁴ A micromensagem foi inspirada no serviço SMS (Short Message Service), comum em celulares desde os anos 1990, com limite de 160 caracteres por mensagem (Calvão; Pimentel; Fuks, 2014). A micromensagem ganhou notoriedade a partir de 2006 com o lançamento do Twitter, sistema que institucionalizou os textos curtos como forma dominante de expressão nas redes sociais digitais.

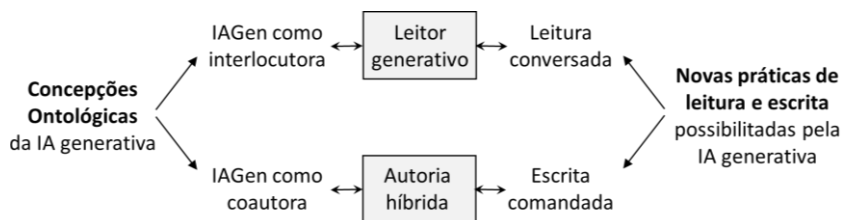
⁶⁵ A mensagem instantânea, foi inicialmente implementada em 1964 em um sistema operacional multiusuário de mainframe. Com o tempo, essa funcionalidade foi sendo desenvolvida e difundida, ganhando popularidade nos anos 1990 com sistemas como ICQ e MSN Messenger. A partir da década de 2000, com a expansão do acesso à internet e a popularização dos smartphones, a mensagem instantânea se consolidou por meio de aplicativos como Facebook Messenger, WhatsApp e Telegram (Calvão; Pimentel; Fuks, 2014).

sistemas de redes sociais promoveu novas formas de escrita colaborativa online e de leitura fragmentada, marcada por rápidas trocas de contexto à medida que deslizamos o dedo pela tela em rolagem infinita. Mais recentemente, na década de 2020, com a proliferação das tecnologias baseadas em IA generativa, novas práticas de letramento se consolidaram: a escrita comandada pela/o usuária/o por meio de *prompts* e a leitura de textos gerados durante a conversa com agentes baseados em IA generativa.



O esquema a seguir relaciona as noções de leitor generativo e de autoria híbrida às práticas de leitura e escrita com IA e às concepções ontológicas da IA generativa.

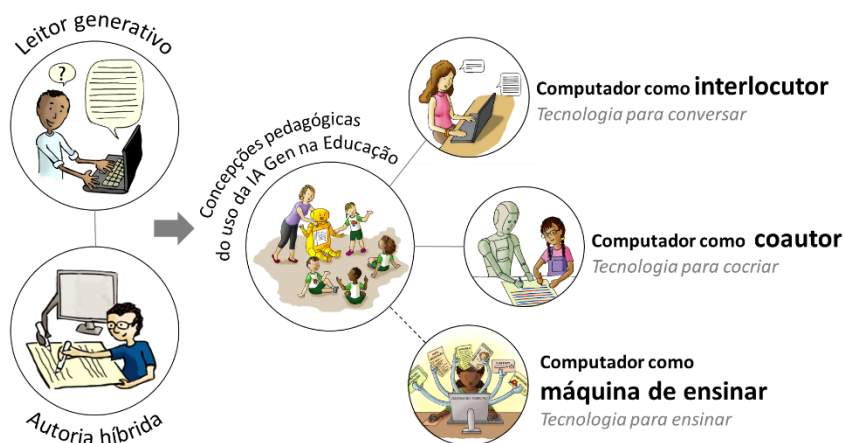
Concepções ontológicas da IA generativa e as novas práticas de leitura e escrita



Fonte: dos autores

Essas concepções e noções nos conduziram a reflexões sobre as implicações da IA generativa para (re)pensar os processos educacionais, o que resultou no quadro teórico representado na figura a seguir.

Teorizações propostas sobre a IA generativa e seus usos na educação



Fonte: dos autores, com ilustrações de Mônica Lopes

Nesta última parte do livro, apresentaremos e discutiremos essas concepções e noções, assim como suas implicações pedagógicas e éticas. Nosso objetivo é compreender as potencialidades e os riscos do uso pedagógico da IA generativa, a fim de nos prepararmos melhor para as transformações em curso no contexto educacional.

Logo após o lançamento do ChatGPT, no início de 2023, algumas/uns estudantes nos mostraram como estavam utilizando o ChatPDF⁶⁶ para obter informações sobre artigos e livros: solicitar resumos, responder a questões, tirar dúvidas sobre a obra etc. Percebemos que uma nova forma de leitura havia surgido. Para conhecer o conteúdo de uma obra, já não era necessário lê-la diretamente; bastava conversar sobre ela com uma tecnologia de IA generativa. Por exemplo, anexamos o presente livro ao *prompt* do ChatGPT-4o e perguntamos: “Quais são as teorizações propostas no livro em anexo?” Após processar por 41 segundos, foi produzida a síntese a seguir:

Principais teorizações desenvolvidas na Parte III do livro:

Teorização	Síntese da ideia central
Ontologia da IA generativa: o que ela realmente (não) é	O capítulo 14 parte da pergunta ontológica “que tipo de entidade é a IA generativa?”. Os autores propõem entendê-la simultaneamente como interlocutora (capaz de conversar) e coautora (capaz de participar da criação), abandonando tanto a visão de ferramenta neutra quanto a de agente plenamente autônomo. Essa dupla condição sustenta todas as demais construções teóricas.
Leitor generativo	No capítulo 15 descreve-se um novo perfil cognitivo de leitura que surge quando se “lê por conversa” com sistemas como ChatGPT ou ChatPDF: o leitor formula prompts, recebe sínteses, faz checagens e ajusta a exploração do texto em ciclos iterativos. Santaella prefere chamá-lo de leitor iterativo, mas reconhece o fenômeno.

⁶⁶ <https://www.chatpdf.com>

Autoria híbrida humano-IA	O capítulo 16 delimita a zona cinzenta entre “texto assistido” e “texto delegado” à IA. Propõe o conceito de autoria híbrida: obras resultantes de cooperação efetiva em que ambos — pessoa e IA — contribuem de modo insubstituível.
Concepções pedagógicas do uso da IA generativa	Atualizando Valente (1993), o capítulo 17 sistematiza cinco concepções: computador como máquina de ensinar, como tecnologia de apoio à construção de conhecimentos, como meio de comunicação, e — acrescentadas agora [com a IA generativa] — como coautor e como interlocutor. Cada concepção implica objetivos, papéis docentes e estratégias didáticas distintos.
IA generativa como interlocutora	O capítulo 18 explora a ideia de IA sempre disponível para conversar, explicitar raciocínios e oferecer feedback — um potencial para pedagogias dialógicas inspiradas em Paulo Freire, mas que exige mediação crítica para evitar respostas superficiais ou enviesadas.
IA generativa como coautora	No capítulo 19 argumenta-se que, quando a IA influencia conceitualmente uma obra (texto, imagem, música), ela passa a compartilhar a autoria. Discute-se por que a legislação ainda impede esse reconhecimento e quais critérios pedagógicos e éticos podem balizar a coautoria.
IA generativa como máquina de ensinar	O capítulo 20 retoma a tradição dos tutores inteligentes: modelos de IA generativa poderiam montar trilhas personalizadas, corrigir tarefas e monitorar progresso. Os autores alertam, porém, que substituir a docência humana empobrece a formação — a questão é “com” e não “em vez de” professoras(es).

[...] O livro propõe um arcabouço teórico integrado que vai da pergunta “o que é a IA generativa?” até “como ela pode (ou não) transformar práticas de leitura, autoria e ensino-aprendizagem”, oferecendo linguagem conceitual para pesquisadores(as), docentes e formuladores(as) de política educacional.

(OpenAI, 2025j, n.p.)

O ChatGPT produziu um bom resumo das contribuições teóricas discutidas neste livro, porém, essa síntese não substitui a leitura integral, pois, ao longo de 342 páginas, entrelaçamos conceitos, referências, exemplos e análises ético-pedagógicas que extrapolam as 337 palavras geradas pelo modelo. Concordamos, contudo, que os pilares do arcabouço teórico foram bem evidenciados no resumo.

Para quem não dispõe de tempo ou interesse para ler todos os artigos e livros de que necessita, encontra dificuldades de compreensão ou não domina o idioma original do texto, as tecnologias generativas possibilitam conhecer parte do conteúdo sem precisar ler a obra original diretamente. Essa prática vem se difundindo no meio acadêmico como estratégia de estudo e pesquisa. Será que essa possibilidade técnica se tornará a principal forma de ler conteúdos escolares e universitários nos próximos anos? Os futuros textos didáticos serão aqueles escritos dinamicamente sob demanda da/o usuária/o em conversação com a IA?

A leitura generativa não é um simples atalho informacional, nem um espelho narcisista que apenas confirma o que o/a leitor/a já sabe. Lembra mais o espelho mágico da madrasta da Branca de Neve: quando interpelado, responde com verdades, ainda que sejam desagradáveis. A diferença é que a IA generativa, não sendo mágica, pode responder com informações equivocadas, enviesadas ou superficiais.

Essa prática não se restringe à cultura escolar e está se tornando cada vez mais frequente em nossa sociedade. A capacidade generativa-conversacional da IA inaugura uma nova relação com o conhecimento, com potencial para transformar nossos modos de aprender e pensar, apontando para a emergência de uma epistemologia pós-humana, na qual o sujeito cognoscente compartilha o processo de construção do conhecimento com agentes artificiais. As novas práticas de leitura também nos levam a reconhecer o surgimento de um novo perfil cognitivo de leitor: o leitor generativo.

15.1 – O perfil cognitivo do leitor generativo

A conjectura de que as tecnologias transformam nossa cognição ganhou projeção com McLuhan, que teorizou os meios de comunicação como extensões das capacidades físicas e mentais humanas, capazes de alterar a maneira como pensamos e interagimos com o mundo (McLuhan, 1962; 1964). Na década de 1960, a televisão já havia alcançado mais de 90% dos lares norte-americanos e estava se popularizando mundialmente, tornando-se um meio de comunicação de massa dominante. McLuhan discutiu os efeitos da televisão, considerando-a uma extensão dos sentidos da visão e da audição, um meio em que as informações são transmitidas de forma multissensorial e simultânea, criando um tipo de experiência menos linear e estruturada do que a leitura.

Seguindo esse raciocínio, o filósofo Pierre Lévy caracterizou as tecnologias digitais como Tecnologias da Inteligência (Lévy, 1993). Segundo ele, essas tecnologias transformam nossa relação com o conhecimento ao ampliar, distribuir e acelerar as capacidades cognitivas, modificando nossos modos de pensar, comunicar e interagir. O hipertexto, por exemplo, reconfigura nossa cognição ao demandar a leitura e a escrita de textos não lineares. As transformações são tão significativas que, para Lévy, a história da humanidade pode ser marcada pelas tecnologias de informação e comunicação predominantes em cada período: a era da escrita, a era da impressão e a era da informática. Essa divisão reflete nossa evolução na produção, armazenamento e circulação de informações.

Lucia Santaella, por sua vez, investiga como os meios de produção de signos e linguagens típicos de cada época — como livro, jornal, cinema, televisão, computador e *smartphone* — forjam diferentes perfis semiótico-cognitivos de leitores.

Tudo teve início com uma cena que presenciei com surpresa e que colocou curiosidades em minha mente. Casualmente observei jovens jogando *games* e a instantaneidade da destreza no manuseio do controle remoto, com a atenção perceptiva do ato de jogar em movimentos muito rápidos, impressionou-me. O que poderia estar acontecendo nos processos cognitivos daqueles jovens? A questão

semiótica da cognição me fsgou e a transpus para os comportamentos, naquela época muito novos, dos modos como os internautas estavam começando a navegar nas arquiteturas dos espaços informacionais das redes que, então, eram batizados de ciberespaço. Imediatamente me veio a ideia de que se tratava da emergência de um novo tipo de leitor com um perfil cognitivo muito específico. Pensei, sem pestanejar, que se tratava de um tipo imersivo de leitor. Por acreditar que, para caracterizar um fenômeno novo é conveniente iniciar pela comparação com outros fenômenos de sua classe, retomei a modalidade tradicional de leitor, ou seja, o leitor de livros e, na continuidade imposta pela visão semiótica das linguagens, cheguei a um tipo de leitor no domínio das imagens em movimentos, que chamei de leitor movente. (Santaella, 2024a, n.p.)

Santaella identificou, inicialmente, três perfis cognitivos de leitor: o contemplativo, moldado pelo livro impresso, principal meio de difusão do conhecimento a partir do século XVI; o leitor movente, desencadeado pela Revolução Industrial, que realiza a leitura enquanto se movimenta pelas cidades, lê jornais e imagens em movimento (cinema e televisão); e o leitor imersivo, formado a partir do uso das tecnologias digitais, que adentra as telas para navegar em espaços informacionais.

O leitor contemplativo é o leitor meditativo da idade pré-industrial, da era do livro impresso e da imagem expositiva, fixa. Esse leitor nasceu no Renascimento e perdurou até meados do século XIX. O segundo tipo de leitor é filho da Revolução Industrial e do aparecimento dos grandes centros urbanos [...]. É, portanto, o leitor do dinâmico, das misturas de sinais e linguagens de que as metrópoles são feitas. Esse leitor nasceu com a explosão do jornal e com o universo reprodutivo da fotografia e do cinema e manteve suas características básicas quando se deu o advento da revolução eletrônica, era do apogeu da televisão. O terceiro tipo de leitor é aquele que brotou nos espaços das redes computadorizadas de informação e comunicação. [...] É imersivo porque, no espaço informacional, perambula e se detém em telas e programas de leitura, num universo de signos evanescentes e eternamente disponíveis. (Santaella, 2013, p. 265-271)

Posteriormente, Santaella reconheceu a emergência de um quarto perfil cognitivo de leitor, o ubíquo, que realiza leituras em qualquer momento por meio de *smartphones* e seus aplicativos.

Ubiquidade se refere principalmente a sistemas computacionais de pequeno porte, e até mesmo invisíveis, que se fazem presentes nos ambientes e que podem ser transportados de um lugar a outro. É essa ideia de estar sempre presente em qualquer tempo e lugar que interessa levar para a caracterização do leitor ubíquo, uma nova condição de leitura e de cognição que está fadada a trazer enormes desafios para a educação, desafios que estamos apenas começando a vislumbrar. (Santaella, 2013, p. 272; 278)

Com a rápida e estrondosa popularização das tecnologias generativas e o aumento do tempo dedicado por pessoas à leitura dos textos gerados por IA, reconhecemos a emergência de um novo perfil cognitivo de leitor, o generativo (Carvalho; Pimentel, 2023). Ao apresentarmos esse conceito a Santaella, ela prontamente reconheceu o fenômeno, mas optou por nomeá-lo de leitor iterativo:

Em uma *live* promovida em 26/06/2023, pela Universidade Federal de Uberlândia sobre “Inteligência em rede na educação superior”⁶⁷, de que participei na companhia de Mariano Pimentel, em um dado momento do diálogo, ele me sugeriu a ideia de um quinto tipo de leitor dotado de um perfil cognitivo próprio das interações que o usuário estabelece com a Inteligência Artificial Generativa (IAG). Evidentemente a ideia faz todo o sentido. Em coautoria com Felipe Carvalho, Pimentel (2023) publicou um artigo sobre o ChatGPT, no qual, ao final, o que chamam de leitor generativo é desenvolvido. Em concordância com o fato de que se trata de um quinto tipo de leitor, por razões a serem explicitadas abaixo, prefiro chamá-lo de “leitor iterativo”. (Santaella, 2024a, n.p.)

⁶⁷ <https://www.youtube.com/live/QaFMuZg6S7A>

Perfil cognitivo dos leitores

Leitor contemplativo

Lê livro impresso e
imagem expositiva
(idade pré-industrial)



Leitor movente

Lê o mundo em movimento e
enquanto se move pela cidade
(revolução industrial)



Leitor imersivo

Lê os espaços interativos
das redes de computadores
(sociedade informatizada)



Leitor ubíquo

Lê em qualquer lugar e
a qualquer momento
(pós smartphones)



Leitor generativo

Lê os textos que pede
para a IA gerar
(pós ChatGPT)



Fonte: dos autores, ilustração de Mônica Lopes
com base em Santaella (2004; 2013; 2024)

Para Santaella, o *Homo sapiens* encontrou formas de expandir sua cognição para além dos limites da caixa craniana, exteriorizando a inteligência por meio das tecnologias da linguagem, utilizando-as para ampliar suas capacidades mentais. A autora afirma que vivemos, atualmente, a sétima revolução cognitivo-cultural. A primeira decorre da oralidade; a segunda, do desenvolvimento da escrita; a terceira, dos livros; a quarta, dos meios de comunicação de massa; a quinta, da cultura das mídias; a sexta, da cultura digital; e a sétima, da cultura dos dados e da inteligência artificial. Nesse processo evolutivo, o ser humano está se transformando no que a autora denomina de “neo-humano”:

Em que o Sapiens se converteu? Afinal, o que somos nós, humanos, ou o que sobrou de nós, ou melhor, o que sobrou do que pensávamos que éramos, agora que nos tornamos literalmente híbridos entre o carbono e o silício? [...] Antes de tudo devo esclarecer que não estou sozinha na constatação do neo-humano, ou seja, de uma radical transformação na própria ontologia do humano, que, de minha parte, tomo como fruto de um processo evolutivo das formas comunicacionais e cognitivas sustentadas pelas linguagens que o ser humano não cessa de inventar e que crescem na medida mesma em que crescem os meios de produção, registro, armazenamento, memória e transmissão dessas linguagens. [...] O contato com bibliografias mais recentes sobre IA foi acompanhado por *insights* de que se trata de uma extensão, um prolongamento da inteligência humana que coloca a cognição humana em um novo e problemático patamar. A ideia da extrassomatização da memória e cognição do Sapiens já estava, então, bem amadurecida. Junto com ela, vinha o retorno insistente de interrogações, há algum tempo contempladas, sobre o que é o humano, ou melhor, sobre o que o humano está passando a ser. Mas foram os avanços da IA, da biotecnologia e engenharia genética que trouxeram munições conscientes e inconscientes para a proposta do neo-humano. (Santaella, 2022, p.10, p. 324, 325)

As tecnologias semióticas que permeiam nosso cotidiano moldam nossas experiências e culturas, estruturam nossa sociedade, modificam nossa cognição e reconfiguram a ontologia do humano. A compreensão de que nos tornamos seres híbridos ao utilizarmos tecnologias capazes de estender nossa inteligência para além dos limites biológicos nos levou a teorizar sobre a emergência de um/a autor/a híbrido/a (o que abordaremos no próximo capítulo) e de um/a leitor/a generativo/a.

O perfil cognitivo de leitor/a generativo/a é identificado em pessoas que utilizam tecnologias baseadas em IA generativa por várias horas, lendo os conteúdos que elas próprias solicitaram à IA que gerasse. Já identificamos esse perfil em algumas/uns estudantes, conforme ilustram os depoimentos apresentados no Capítulo 5: Viviane, por exemplo, utiliza o ChatGPT “em vários momentos do dia”; Mariana afirma que não estuda sem o ChatGPT e considera necessário “conhecer ele cada vez mais, ficar íntimos”; Joana, que faz uso rotineiro, acabou “imersa no uso de ChatGPT” e, quando está aprendendo algo novo, logo pensa: “vamos ver o que o ChatGPT diz sobre esse tema”; para Augusto, o ChatGPT “muda tudo”. Essas/es estudantes já aprenderam a utilizar a IA generativa como uma extensão de sua própria inteligência, desenvolvendo processos de aprendizagem e pensamento em conjunto com a tecnologia, de forma híbrida.

15.2 – Finalidades da leitura conversada

A partir dos depoimentos de estudantes, foi possível identificar ao menos três finalidades para a leitura conversada com a IA generativa: para pesquisar informações (leitura conversada informativa); para aprender algo (leitura conversada didática); e para obter informações sobre uma obra (metaleitura conversada).

Muitas/os estudantes utilizam a IA generativa para obter sínteses e informações de textos extensos, como discutido na introdução deste capítulo. Denominamos essa prática de metaleitura, pois consiste na leitura das informações geradas sobre uma obra, em vez da leitura direta da obra. Tal prática costuma ser adotada quando se considera suficiente captar apenas o sentido geral do texto original:

O que eu mais gosto de fazer no momento é pedir para o ChatGPT resumir alguma coisa para mim, por exemplo um artigo. Muitas vezes não tenho tempo para ler algo muito extenso, então os resumos me ajudam bastante. (Viviane)

Sempre peço [à IA] para resumir artigos, explicar textos, destacar alguns pontos. (Allanis)

A prática da metaleitura com IA pode comprometer a retenção de conteúdo e o desenvolvimento de competências interpretativas, funcionando como um atalho que, embora conveniente, empobrece a experiência da aprendizagem e não favorece o desenvolvimento da cognição necessária para uma leitura contemplativa. Cabe às/aos educadoras/es incentivar ambos os tipos de leitura, a contemplativa e a metaleitura, pois uma não substitui a outra.

Outra prática recorrente é o uso da IA generativa como fonte primária de informações. Essa leitura conversada informativa muitas vezes substitui o processo de pesquisa que anteriormente costumava ser feito por meio de mecanismos de busca na web:

Em vez de realizar uma busca extensa no Google, opto por utilizar o ChatGPT, pois ele me fornece respostas diretas, citando exemplos e explicações, o que me poupa tempo e energia. [...] Atualmente, dificilmente passo um dia sem abrir a plataforma e recorrer a ela para fazer perguntas ou buscar informações sobre os mais diversos assuntos. (Joana)

Na faculdade eu utilizo ele [o ChatGPT] para praticamente todas as matérias; coisas que geralmente eu teria que passar horas pesquisando na internet, ele me entrega com precisão. (Mariana)

A IA generativa também tem sido utilizada para facilitar a compreensão de temas complexos. A leitura de textos didáticos gerados por essa tecnologia a posiciona como um recurso pedagógico potente, capaz de traduzir conteúdos densos para formatos mais acessíveis e compreensíveis:

Não é raro o sentimento de ler sobre um determinado assunto durante uma hora inteira e terminar com o sentimento de não ter entendido realmente o que aquilo tudo queria dizer. [...] Uma boa ideia é simplesmente dar ao ChatGPT um parágrafo e pedir “me explique este parágrafo de forma simples”. É efetivo! A minha experiência tem sido positiva nesse aspecto! Recomendo! (Caio)

O GPT já te entrega a resposta que você quer com extensivas explicações; caso não tenha entendido algo, é só perguntar de novo e ele tenta mais uma vez. [...] O usei para tirar dúvidas que eu tinha [...]; foi extremamente útil. (Artur)

A leitura de conteúdos gerados pela IA não deve ser feita sem uma postura crítica. O/A leitor/a generativo/a precisa estar atento/a ao fato de que as informações produzidas podem conter erros, ser superficiais ou refletir vieses. Por isso, é necessário manter uma atitude de questionamento e verificar os dados em fontes confiáveis.

Autoria híbrida humano-IA

As tecnologias generativas de texto estão reconfigurando nossas práticas de escrita. O que antes costumava ser um processo solitário, agora pode ser realizado em conversação com a IA generativa. Essa mudança é tão radical que nós, autores deste livro, já nos reconhecemos como autores híbridos, com criatividade ampliada pela IA. Escrever sem essa parceria agora nos parece tão impensável quanto redigir artigos científicos sem usar a web e os buscadores como o Google Acadêmico⁶⁸. Embora essas outras tecnologias influenciem a escrita, elas atuam com graus distintos de agência: buscadores recuperam informações na web para que as próprias pessoas decidam o que fazer; já a IA generativa contribui ativamente para a elaboração de ideias e do discurso, influenciando diretamente a obra.

Durante muito tempo, acreditava-se que um computador jamais seria capaz de escrever uma poesia, pintar um quadro, compor uma música, produzir um filme ou redigir um artigo científico — hoje já existem diversas tecnologias generativas capazes de realizar cada uma dessas atividades em parceria com as pessoas. As tecnologias generativas baseadas em modelos de linguagem de grande escala deram início à era da autoria híbrida humano-IA (Pimentel; Azevedo; Carvalho, 2023; Pimentel; Carvalho, 2024; Pimentel; Carvalho; Silveira, 2024). Diversas/os pesquisadoras/es e artistas concordam com essa noção de criação híbrida:

⁶⁸ Até a busca por artigos científicos passou a ser realizada de forma integrada com os modelos de linguagem, surgindo buscadores acadêmicos que respondem em vez de apenas listar *links*, citando trechos dos próprios artigos. Muitos pesquisadores já utilizam tecnologias generativas como: Perplexity AI, Consensus, Scite.AI e Research Rabbit.

"Co-crIA-ção": Um esforço colaborativo real envolvendo, de maneira mais ou menos equitativa, o humano e a IA generativa, com reconhecimento das contribuições de cada parte. Isso pode ser chamado de criatividade ampliada, pois o resultado é fruto de uma hibridização que não seria alcançável apenas por seres humanos ou por sistemas de IA. Tal colaboração é considerada pela maioria dos autores citados anteriormente como o futuro ideal e já vem se consolidando como uma posição comum entre pesquisadores que investigam as possibilidades oferecidas pela IA [...]. Além da comunidade acadêmica, há artistas notáveis que também estão abraçando essa visão. (Vinchon *et al.*, 2023, p. 476, tradução nossa)

Entretanto, o reconhecimento da IA generativa como coautora não é unânime. Há quem considere a IA como única responsável pelo conteúdo gerado, como ocorre com pessoas que solicitam a geração de imagens e, por não se sentirem capazes de criar obras semelhantes, não se consideram coautoras do resultado. O oposto também é comum: há pessoas que se apropriam dos textos gerados e não reconhecem a coautoria da IA, porque se sentem capazes de escrever textos semelhantes, fazendo da IA uma espécie de escritora-fantasma (*ghostwriter*) (Draxler *et al.*, 2024).

Para compreendermos a divergência de posicionamentos acerca do papel da IA generativa na autoria de obras, precisamos reconhecer que as relações entre pessoas e tecnologias são diversas e ambivalentes. A cocriação com IA pode ampliar a criatividade e possibilitar aprendizagens significativas. Contudo, há o risco de que algumas pessoas abdicuem do uso consciente de suas próprias capacidades intelectuais, entregando suas mentes e vozes para a IA pensar e se expressar por elas. Essa entrega pode resultar em plágio ou em desligamento/terceirização da criatividade:

"Plágio 3.0": Pessoas com o desejo de parecer produtivas e criativas irão aproveitar intensamente as produções da IA sem citar a fonte. [...]

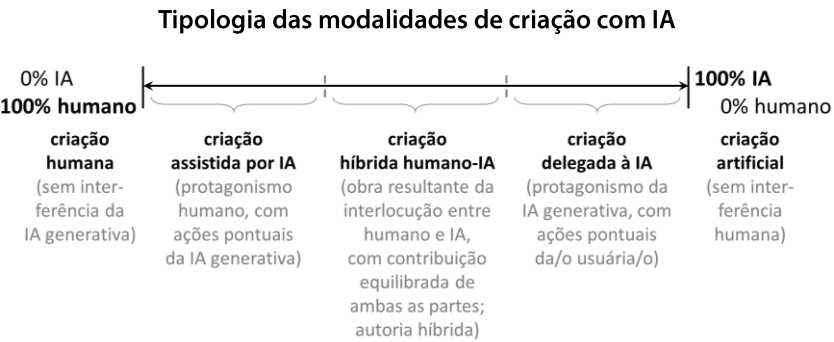
"Desligamento": Esse cenário postula que algumas pessoas se tornarão menos motivadas a realizar qualquer ação criativa. [...] Nesse contexto, algumas pessoas simplesmente sentirão que não são capazes de criar no mesmo nível que a IA e, assim, terceirizarão a criação de conteúdo para a IA generativa. (Vinchon *et al.*, 2023, p. 476-478, tradução nossa)

No campo educacional, a ambivalência da IA generativa gera muita controvérsia, pois ela representa, simultaneamente, uma potência e uma ameaça à autoria de estudantes. Como verificar se estudantes se engajaram ativamente na produção com a tecnologia generativa? Será que ocorreu o aprendizado esperado durante a cocriação com a IA? Afinal, a IA generativa é uma inimiga a ser combatida na educação ou uma aliada a ser incorporada aos processos de ensino e aprendizagem?

Para enfrentarmos esse dilema, é importante reconhecer que há diferentes modos de produzir obras com a IA generativa, sendo alguns condizentes com princípios éticos e pedagógicos, enquanto outros não se mostram apropriados para fins educacionais. A discussão sobre esses diferentes modos de criação com IA é o tema da próxima seção.

16.1 – Uma tipologia da criação com IA

É possível identificar diferentes graus de atuação da IA generativa e de pessoas na produção de obras, sejam textos, imagens, sons ou vídeos. Entre a obra criada exclusivamente por uma pessoa e aquela gerada exclusivamente por IA, sem qualquer interferência humana, é possível reconhecer ao menos três modalidades de criação envolvendo a atuação conjunta entre pessoas e IA generativa, conforme esquematizado a seguir:



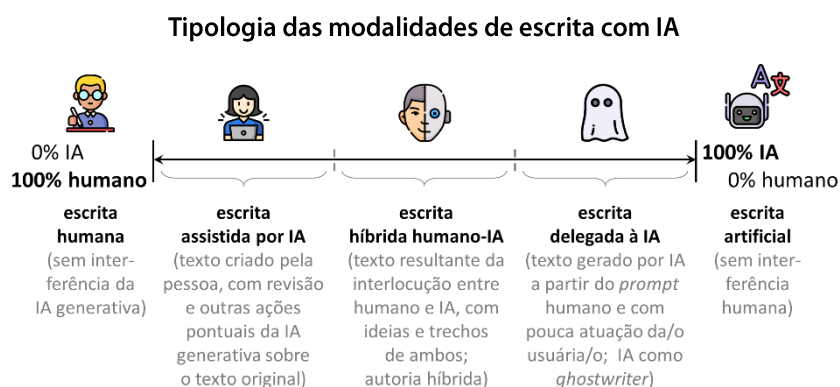
Fonte: dos autores

Em um extremo, encontra-se a criação exclusivamente humana, o processo de autoria com o qual estávamos acostumadas/os antes do lançamento de tecnologias generativas como o Midjourney e o ChatGPT. No

outro extremo, há a criação artificial de conteúdos sem qualquer intervenção humana, nem mesmo na elaboração de *prompts*; uma situação em que a IA atua como criadora de forma autônoma (Aue, 2018; Gros, 2019). Entre esses dois polos, há um espectro contínuo na criação de obras, com diferentes graus de participação humana e da IA generativa. Nesse espectro, podemos delimitar três modalidades:

- *criação assistida por IA*: ocorre quando a obra é criada por uma pessoa, e a IA é utilizada de forma pontual; por exemplo, para sugerir revisões ou auxiliar na pesquisa para a criação da obra;
- *criação híbrida humano-IA*: caracteriza-se pela colaboração entre a pessoa e a IA generativa. A IA atua como interlocutora ao longo do processo criativo, e a pessoa se deixa afetar pelas produções da IA, sem, contudo, aceitar de forma acrítica tudo o que é gerado. Pode ser difícil delimitar com precisão as contribuições de cada parte, uma vez que ambas atuam em parceria na elaboração do resultado final;
- *criação delegada à IA*: ocorre quando o protagonismo recai na IA, que gera a maior parte do conteúdo a partir dos comandos fornecidos pela pessoa, havendo pouca ou nenhuma participação humana nas decisões de concepção e execução do resultado final.

Essa tipologia possibilita uma melhor compreensão dos diferentes modos de escrita comandada (via *prompts*), como esquematizamos na figura a seguir.



Fonte: dos autores

As/Os estudantes mobilizam todas essas diferentes formas de escrita com IA. Com base nos depoimentos apresentados no Capítulo 5, é possível identificar que algumas/uns estudantes não permitem que a IA escreva por elas/es, mas aceitam que a tecnologia generativa ofereça revisões, análises do texto e sugestões de melhorias, caracterizando uma prática de escrita assistida por IA:

Ele [o ChatGPT] tem sido valioso no meu processo de redação e revisão de trabalhos acadêmicos. Ao compartilhar trechos de meus textos com o ChatGPT, obtenho sugestões de palavras, frases, ideias e estruturas gramaticais que contribuem para aprimorar meu trabalho. (Joana)

A revisão textual realizada por um/a profissional de Letras é uma prática consagrada no meio acadêmico. Essa prática não é considerada uma fraude, chega até a ser exigida por muitas revistas científicas como requisito para assegurar a qualidade formal e linguística dos artigos a serem publicados. Por analogia, também não deveríamos ter dúvidas éticas quanto ao uso da IA generativa em processos de autoria assistida, desde que seu uso se restrinja à revisão gramatical, sugestões de reescrita ou apoio linguístico. Nesses casos, a autoria humana da obra é preservada, uma vez que as ideias, os argumentos e a estrutura conceitual pertencem ao/à autor/a humano/a. Além de contribuir para a qualidade final do texto, a assistência da IA pode ter um caráter formativo ao incentivar a reflexão crítica sobre as escolhas linguísticas, apoiando o desenvolvimento das habilidades de escrita da/o usuária/o.

Algumas/uns estudantes recorrem à IA para gerar, parcial ou totalmente, o conteúdo de seus trabalhos, apresentações, traduções, programas de computador, entre outros. Essa é a prática de escrita delegada à IA:

[O ChatGPT] me ajudou a fazer parte do meu trabalho final [...], onde ele foi capaz de gerar para mim o texto de todas as atividades de um processo de *software*. (Arthur)

Já até pedi que ele [o ChatGPT] escrevesse um código simples em Java que eu estava sem paciência para escrever. (Augusto)

A substituição da autoria humana por produção artificial levanta questões éticas, sobretudo no contexto educacional, seja universitário ou escolar, em que a autenticidade e a originalidade do processo são valores centrais. Assim como o uso de escritoras/es fantasmas é condenado na produção acadêmica por violar o princípio da autoria intelectual, recorrer à IA como escritora fantasma assumindo para si a autoria do texto gerado pela IA sem dar os devidos créditos também deveria ser considerado plágio ou desonestidade. Essa prática também pode comprometer o desenvolvimento de competências criativas, reduzir a confiança autoral e alimentar a percepção de que a IA é superior às capacidades humanas. Algumas/uns estudantes expressaram preocupações quanto à dependência da IA e às possíveis consequências para sua atuação profissional:

Ainda não entrou 100% na minha cabeça que eu não vou estar emburrecendo por me consultar com o ChatGPT. (Mariana)

Tento ao máximo não usar [a IA generativa] no meu trabalho porque ainda considero isso um assunto bem delicado, já que a sensação de que logo seremos substituídos é enorme. (Allanis)

Já a escrita híbrida caracteriza-se pela cocriação, quando a pessoa se deixa afetar pelas sugestões da IA, mas a IA não escreve toda a obra em seu lugar. A escrita torna-se híbrida quando a IA generativa é utilizada para gerar ideias ou trechos do texto:

Já utilizei [a IA generativa] para me ajudar a estruturar o conteúdo que vou apresentar, para dar ideias de como argumentar sobre determinado assunto, ou para ela me dizer palavras que escaparam da minha mente. Serve para ter ideias de como fazer a tarefa (Paulo)

Na cocriação, o envolvimento pode ser tão intenso e transformador que algumas pessoas se sentem hibridizadas com a tecnologia, o que nos levou à formulação do conceito de autoria híbrida humano-IA. É justamente essa forma de autoria que desperta mais dúvidas éticas, pois não se trata de mero plágio ou trapaça acadêmica, mas, ao mesmo tempo, a integridade autoral humana também não é plenamente preservada. Nas próximas seções, aprofundaremos a discussão sobre a autoria híbrida e suas implicações éticas e pedagógicas.

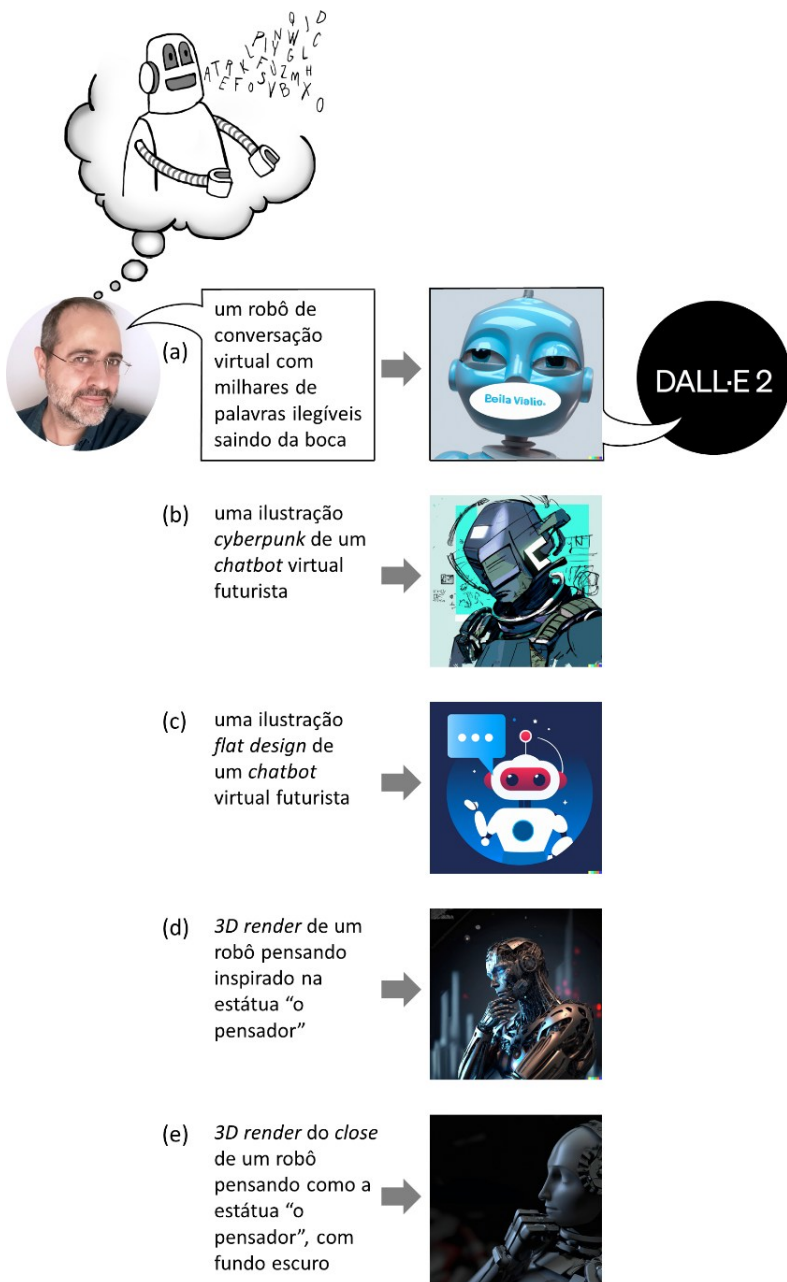
A distinção entre escrita assistida por IA, autoria híbrida humano-IA e escrita delegada à IA revela diferentes abordagens no uso das tecnologias generativas no processo de produção textual. Embora todas elas tenham seu lugar no cenário contemporâneo, no contexto educacional é necessário que se promova o uso ético, consciente e responsável das tecnologias generativas, garantindo que a IA atue como uma assistente ou coautora-interlocutora, jamais como substituta da autoria estudantil.

16.2 – Uma experiência de cocriação com IA

Algumas pessoas não reconhecem a coautoria da IA porque nunca vivenciaram uma experiência efetiva de cocriação com uma tecnologia generativa. Por isso, queremos apresentar um caso que ilustra o modo híbrido como nós, autores deste livro, temos nos relacionado com a IA generativa.

Desejávamos criar uma ilustração para ser a capa de nosso artigo intitulado “ChatGPT é realmente inteligente?”, que seria publicado na revista SBC Horizontes (Pimentel; Azevedo; Carvalho, 2023a). Pesquisamos algumas representações do ChatGPT e de *chatbots* nas imagens da web. Constatamos que é comum representar a IA generativa como um robô, metáfora recorrente que procura conferir corporeidade a algo tão abstrato quanto um modelo de linguagem. Uma ilustração, em especial, chamou nossa atenção: um robô com várias letras saindo da boca. Consideramos aquele robô falante uma metáfora adequada para representar fisicamente o ChatGPT. No entanto, por questões de direitos autorais, não poderíamos utilizar aquela imagem, então decidimos tentar criar uma imagem equivalente com o DALL-E, que era uma tecnologia generativa de imagens da OpenAI que, naquela época, ainda não estava integrada ao ChatGPT. Realizamos um processo de cocriação com essa tecnologia, em que trabalhamos com a IA para a construção do resultado final, em um processo conversacional parcialmente ilustrado na figura a seguir.

Coautoria com DALL-E



Fonte: dos autores, com uso do DALL-E 2

Iniciamos solicitando a criação de “um robô de conversação virtual com milhares de palavras ilegíveis saindo da boca” (a). A imagem gerada foi um desastre estético e conceitual. Percebemos que não queríamos palavras ilegíveis “na boca” do robô. Então reformulamos o pedido, especificando o estilo “cyberpunk”, que estava entre os exemplos de estilos apresentados no *site* do gerador de imagens. O resultado (b) foi esteticamente interessante, mas o robô nos pareceu um super-herói de histórias em quadrinhos, e não era essa representação que queríamos para o ChatGPT. Solicitamos, então, uma nova imagem, agora com o estilo *flat design*, que é uma estética minimalista — supúnhamos que, com menos detalhes na ilustração, o conceito talvez se tornasse mais evidente. O resultado gerado (c) lembrava as imagens usadas para representar *chatbots* de atendimento ao cliente, daqueles que aparecem na tela perguntando “em que posso ajudar?”. Compreendemos, a partir desse resultado, que o ChatGPT não é esse tipo de *chatbot*. Em vez de um robô falante, ponderamos que talvez fosse mais apropriado representar um robô pensante, pois assim chamaríamos atenção para a inteligência em vez da fala. Como referência, sugerimos que a imagem fosse inspirada na escultura O Pensador, de Rodin. O resultado (d) agradou-nos em termos estéticos; ficamos impressionados com a representação realista e futurista do robô, mas nos lembrou um personagem da franquia O Exterminador do Futuro, o que nos desagradou, já que não consideramos o ChatGPT nem um herói nem um exterminador do nosso futuro. Realizamos novos ajustes no *prompt* e solicitamos a geração de novas imagens. Obtivemos algumas variações até que um dos resultados (e) nos agradou tanto conceitual quanto esteticamente.

Em vez de letras saindo da boca de um robô, como havíamos imaginado inicialmente, decidimos construir um céu de letrinhas para fazer alusão ao processamento textual realizado pela IA generativa, uma metáfora para o raciocínio computacional. Infelizmente, não conseguimos fazer o próprio DALL-E gerar uma imagem razoável para o céu de letrinhas que desejávamos. Como sabemos programar, utilizamos o ChatGPT para criar um programa de computador que desenha as letras em uma página web. Sem a IA, precisaríamos aprender algumas técnicas para construir aquele programa, o que demandaria muitas horas de estudo. Em parceria com a IA, conseguimos concluir o código em poucos minutos. Posteriormente, utilizamos o Photoshop para juntar as imagens e fazer algumas edições até chegarmos à arte-final.

Ficamos satisfeitos com o resultado e decidimos publicá-lo como capa do nosso artigo. Nesse momento, foi preciso atribuir os devidos créditos à imagem. Então nos perguntamos: quem é o autor da arte-final? Devemos considerar a obra como sendo de autoria apenas humana, apenas da IA, ou de ambas as partes? Se considerarmos a IA, devemos também considerar o programa de edição de imagem como coautor?

Atente que a IA generativa foi representada na arte-final de forma distinta da representação inicialmente concebida. Ao longo do processo de criação com a IA, fomos modificando nossas compreensões e aprendendo com a conversação estabelecida, o que nos levou a abandonar a ideia do robô falante e optar por representar um robô pensante. Nesse caso, a IA influenciou de forma significativa o conceito da obra, por esse motivo, consideramos o DALL-E 2 como coautor da imagem.

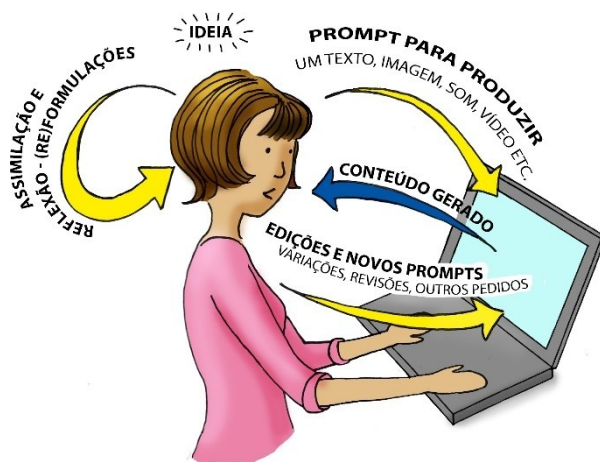
Não reconhecemos a IA como única autora da imagem, pois ela não criou o conteúdo de maneira autônoma; fomos responsáveis pela elaboração dos *prompts*, pela curadoria dos resultados gerados, pelos sentidos atribuídos às imagens, pelas edições posteriores e pela decisão sobre o que publicar. Tampouco nos consideramos os únicos autores, uma vez que a IA gerou todos os elementos visuais e exerceu influência conceitual sobre nossas escolhas criativas.

Não atribuímos coautoria ao Photoshop⁶⁹, pois ele não criou nem propôs elementos novos, nem contribuiu conceitualmente para a obra. Quando editamos uma imagem sem recorrer a recursos de IA generativa, conseguimos prever os efeitos de cada comando; já a produção realizada com IA apresenta um grau de imprevisibilidade capaz de nos surpreender e influenciar, o que confere à tecnologia o papel de interlocutora, e não apenas de executora.

Relatamos essa experiência de cocriação com o DALL-E para ilustrar que a obra final é resultado de uma construção conjunta, possibilitada pela conversação entre pessoas e tecnologias generativas. Tais tecnologias não produzem obras de maneira totalmente autônoma; elas dependem das solicitações feitas pelas/os usuárias/os, que, por sua vez, não aceitam de forma acrítica qualquer resultado. O processo colaborativo entre usuária/o e IA pode ser esquematizado como representado na ilustração a seguir.

⁶⁹ A versão do Photoshop utilizada na edição da imagem em questão ainda não incluía recursos de edição com IA generativa.

Cocriação com IA



Fonte: dos autores, ilustração de Mônica Lopes

A partir de uma ideia para um texto, imagem, som ou vídeo, solicitamos que a tecnologia generativa produza algo. A IA gera um conteúdo com base no nosso *prompt*, podendo apresentar informações novas, que não havíamos pensado anteriormente, o que provoca reflexão e transformação em nossas compreensões, possibilitando-nos aprender com essa tecnologia. Podemos editar o resultado ou fazer novos pedidos, seja para obter variações, revisões ou mesmo um conteúdo totalmente novo. Cada produção é novamente avaliada por nós, retroalimentando o processo criativo. Esse processo interativo, com idas e vindas, críticas e reflexões, reformulações e reconstruções de conhecimento, resulta em uma obra produzida pela hibridização humano-IA, na qual se manifesta a criatividade aumentada (*augmented creativity*).

Nem todas as pessoas concordam com essa concepção de autoria híbrida. Do ponto de vista da legislação sobre direitos autorais, a autoria é atribuída exclusivamente a seres humanos; por isso, uma obra feita por IA é considerada de domínio público. Já para as empresas desenvolvedoras das tecnologias generativas, a responsabilidade pelo conteúdo gerado é da/o usuária/o que o solicitou. Precisamos discutir essas diferentes posições, especialmente considerando suas implicações para o contexto educacional.

16.3 – A autoria é apenas da IA ou apenas da/o usuária/o?

A Lei de Direitos Autorais (Brasil, 1998) restringe a noção de autor/a a seres humanos: “Art. 11. Autor é a pessoa física criadora de obra literária, artística ou científica”. Pela interpretação da lei, a IA não pode ser coautora porque não é um “sujeito de direitos” — termo jurídico que se refere à pessoa física ou jurídica com direitos e obrigações legais. Uma imagem feita por IA, por não haver um sujeito para deter os direitos e deveres da autoria, é legalmente reconhecida como sendo uma imagem de domínio público.

Exemplos de imagens de domínio público



a) Autorretrato (*selfie*)
tirado pelo macaco



b) Teatro de Ópera Espacial,
feita com Midjourney

Fonte: a) (Selfies[...], 2025). b) (Théâtre [...], 2025)

A atribuição de direitos autorais apenas a seres humanos remete ao caso “Naruto versus Slater”. Durante uma expedição, o fotógrafo David Slater acompanhava um grupo de macacos havia dias e, em determinado momento, alguns desses animais se aproximaram de Slater e interagiram com a câmera deixada no tripé, resultando em algumas fotografias acidentais. Entre elas, *selfies* tiradas por um macaco, posteriormente nomeado Naruto, ganharam visibilidade com a matéria publicada no Daily Mail (2011): “Macaco atrevido! Macaco pega emprestada câmera de fotógrafo para tirar autorretratos hilários”. Duas dessas imagens foram

incluídas no banco de imagens públicas da Wikipédia (Selfies[...], 2025). Slater pediu a remoção, mas seu pedido foi negado sob a justificativa de que aquelas fotografias eram de domínio público por terem sido tiradas por um macaco, que não pode deter direitos autorais. O fotógrafo iniciou uma batalha judicial, e os tribunais dos Estados Unidos decidiram pelo domínio público das fotos, mesmo Slater tendo planejado e tornado possível a produção daquelas imagens.

Outro caso emblemático é o da imagem Teatro de Ópera Espacial (Théâtre [...], 2025), obra criada por Jason Michael Allen com o Midjourney, que venceu um concurso de arte nos EUA na categoria de arte digital, sendo a primeira imagem feita com IA a ganhar tal prêmio. Allen elaborou uma série de *prompts*, ajustando-os mais de 600 vezes. Ele também utilizou o Photoshop para editar a imagem e outro sistema de IA para aumentar a resolução. Apesar disso, o Escritório de Direitos Autorais dos EUA recusou o registro, alegando falta de autoria humana significativa.

A legislação de direitos autorais do Brasil e dos Estados Unidos é do século passado e está fundamentada nas normas estabelecidas no século retrasado pela Convenção de Berna para a Proteção das Obras Literárias e Artísticas, criada em 1886, da qual os EUA e o Brasil são signatários (Brasil, 1975). Essas datas sinalizam o problema: naquela época, não foram consideradas as possibilidades de geração de conteúdo por máquinas.

Em novembro de 2023, o Tribunal da Internet de Pequim decidiu, pela primeira vez, em favor da proteção de direitos autorais de imagens geradas com IA reconhecendo que o resultado reflete uma “realização intelectual” em virtude do investimento da artista, atendendo, assim, aos requisitos para o reconhecimento de direitos autorais (IDS, 2023). Esse precedente jurídico aponta para a necessidade de revisão das leis de direitos autorais considerando as inúmeras produções literárias, artísticas, científicas e técnicas que vêm sendo realizadas com o uso de tecnologias generativas.

Em janeiro de 2025, um tribunal federal dos Estados Unidos concedeu registro de direito autoral a uma obra gerada com IA. Inicialmente, o pedido de registro havia sido negado sob a alegação de ausência de autoria

humana, mas, após a apresentação de um vídeo documentando todo o processo criativo e de explicações detalhadas sobre o envolvimento do autor Kent Keirse, foi concedido o registro da obra intitulada “Uma Única Fatia de Queijo Americano” (*A Single Piece of American Cheese*) (Keirse, 2025). Esse outro precedente reafirma a possibilidade de reconhecimento da autoria humana na produção com IA, tornando viável a proteção autoral de obras produzidas com tecnologias generativas, desde que se comprove a contribuição efetiva da/o usuária/o. Tais decisões demonstram que um conteúdo gerado com IA nem sempre é judicialmente considerado de autoria exclusivamente maquina.

A discussão sobre o reconhecimento de obras produzidas por artefatos técnicos já foi travada no passado, quando se questionava se a fotografia deveria ou não ser considerada uma forma de arte. No início, a fotografia era vista como uma captura mecânica da realidade. Com o tempo, reconheceu-se que a pessoa aplica seu olhar e técnica para criar imagens únicas, fazendo escolhas de enquadramento, iluminação e composição que conferem autoria à/ao fotógrafa/o (Sontag, 2004). Atualmente, já se fala em “promptografia”, um novo estilo artístico viabilizado pela interação entre pessoas e tecnologias generativas. Nessa prática, a/o promptógrafa/o emprega sua criatividade e conhecimento para elaborar *prompts* intencionais que resultam em produções inovadoras.

Equiparar a produção com IA generativa ao fazer artístico da fotografia é reconhecer que o resultado não pode ser reduzido a uma produção algorítmica totalmente artificial; é reconhecer que o *prompt* carrega intencionalidade, subjetividade e criatividade humanas, e que a curadoria dos resultados decorre dos sentidos atribuídos pela pessoa que interage com a IA. O reconhecimento da autoria humana no processo criativo com IA possibilita que determinadas produções com tecnologias generativas sejam passíveis de proteção autoral. Contudo, isso não implica o reconhecimento da coautoria da IA.

Devemos, então, considerar que a autoria dos conteúdos gerados com IA é apenas das/os usuárias/os? As empresas desenvolvedoras de

tecnologias generativas estabelecem que a titularidade⁷⁰ do conteúdo gerado é da pessoa que forneceu o *prompt*, atribuindo a ela a responsabilidade pelo uso dos conteúdos gerados. Por exemplo, a OpenAI estabelece:

Você é responsável pelo Conteúdo, inclusive por garantir que ele não viole qualquer legislação aplicável [...] Entre você e a OpenAI, e na medida permitida pela legislação aplicável, você: (a) mantém seus direitos de titularidade sobre as Contribuições e (b) é titular dos Resultados. Por meio deste Termo, cedemos a você todos os nossos direitos, titularidade e interesse, se houver, em relação aos Resultados. (OpenAI, 2023l, n.p.)

O Termo de Uso do Midjourney segue a mesma lógica, atribuindo a titularidade e a responsabilidade exclusivamente às/aos usuárias/os:

Você possui todos os Ativos que cria com os Serviços na medida máxima possível de acordo com a lei aplicável. [...] Você é o único responsável por determinar a adequação do uso ou redistribuição dos Ativos e assume quaisquer riscos associados ao uso dos Serviços. (Midjourney, 2024, n.p., tradução nossa)

Com esses termos de uso, as empresas isentam-se da responsabilidade pelos conteúdos gerados com suas tecnologias. No entanto, essa ausência de responsabilidade não parece ser o entendimento de algumas pessoas. Por exemplo, a deputada estadual Renata Souza (PSOL-RJ) denunciou em suas redes sociais um caso de “racismo algorítmico”, quando uma IA gerou a imagem de uma mulher negra segurando uma arma, sendo que o *prompt* dela solicitava: “uma mulher negra, de cabelos afro, com roupas de estampa africana num cenário de favela”, sem qualquer referência a armamento. Por isso, a deputada responsabilizou a tecnologia.

⁷⁰ Titularidade refere-se à pessoa física ou jurídica que detém os direitos patrimoniais sobre a obra e pode explorá-la economicamente, ceder ou licenciar esses direitos.

Imagem gerada para “mulher negra na favela”



Fonte: Souza (2023)

O imaginário de violência nas favelas e de criminalização de corpos e territórios negros também está presente em tecnologias, o que coloca em risco a segurança de pessoas negras. É preciso rever essas tecnologias e procedimentos. (Souza, 2023)

Essa denúncia, entre tantas outras relativas aos vieses e estereótipos presentes nos conteúdos gerados por tecnologias, reflete a compreensão de que tanto a tecnologia quanto a empresa desenvolvedora devem ser corresponsáveis pelo conteúdo gerado. Considerando que a IA é propositiva e participa do processo criativo, não atuando apenas como executora das ordens da pessoa usuária, alguma responsabilidade lhe cabe. Esse caso exemplifica, de maneira contundente, que conteúdos gerados com IA não devem ser automaticamente considerados de autoria exclusivamente humana.

16.4 – A IA generativa poderia ser considerada coautora?

Algumas pessoas veem as tecnologias generativas apenas como uma “ferramenta”, um meio técnico de produção, à semelhança de uma câmera fotográfica. Contudo, diferentemente da máquina fotográfica, a IA não é uma mera ferramenta; ela é agente, propõe conteúdos capazes de nos surpreender e influenciar conceitualmente a obra. É nessas circunstâncias que consideramos a IA como coautora, quando ela influencia significativamente a obra.

Atualmente, já no contexto da arte contemporânea, a obra costuma ser pensada como conceito e processo, mais do que como resultado. [...] Nessa linha, a IA como *modus operandi* constitui tema em si, provoca controvérsias, desperta consciências, e, assim sucedendo, colabora para o papel da arte preconizado pela arte contemporânea. E, ao assim integrar a arte, em tese, torna-se protegível (Almeida, 2023, n.p.).

Não reconhecemos a IA generativa como autora exclusiva de uma obra, pois ela não age sozinha, não tem autonomia. No máximo, a consideramos como coautora em potencial, pois sempre atua em parceria com as/os usuárias/os, sempre depende de um *prompt* humano para gerar novos textos, imagens, sons e vídeos.

Já foram publicados livros em que se registrou a IA generativa como coautora, fato noticiado em tom de denúncia (Coelho; Schurig, 2023). Esse estranhamento ocorre porque prevalece a ideia de que a autoria pertence apenas a seres humanos ou, em sentido oposto, exclusivamente à IA (como exemplificam os casos das imagens consideradas de domínio público, discutidas na seção anterior). Mesmo sem respaldo legal, reconheceram a coautoria da IA por uma questão ética: deixa-se claro que a obra foi produzida em parceria com tecnologia generativa para evitar o risco de acusação de plágio. Tal postura reflete uma busca por transparência em tempos de transformações nas práticas autorais.

Para algumas sociedades científicas e editoras de periódicos, “apenas humanos podem ser autores” e “os autores [humanos] devem assumir a responsabilidade pública por seu trabalho” (Spinak, 2023, n.p.). Por isso também virou notícia-denúncia o fato de tecnologias generativas serem

reconhecidas como coautoras de alguns artigos científicos: “ChatGPT listado como autor em artigos de pesquisa: muitos cientistas desaprovam” (Stokel-Walker, 2023, s.p., tradução nossa).

Invisibilizar a participação da IA na produção significa assumir integralmente o crédito pela obra, o que pode ser considerado antiético. Também seria inadequado eximir a pessoa da responsabilidade autoral, atribuindo o resultado exclusivamente à IA e eximindo-se de eventuais questões legais. Tanto subestimar quanto superestimar o papel da IA ou da pessoa distorce a relação de cocriação humano-IA. Ao reconhecer a IA como coautora, propomos tornar explícita a participação de ambas as partes no processo criativo.

Todavia, a lei brasileira de direitos autorais impede o reconhecimento da IA como coautora de uma obra, pois define “obra intelectual” como “as criações do espírito, expressas por qualquer meio ou fixadas em qualquer suporte, tangível ou intangível, conhecido ou que se invente no futuro” (Brasil, 1998). Essa definição recorre a termos antropocêntricos como “intelectual”, “criação”, “expressão” e “espírito”, que implicam um processo próprio do ser humano, excluindo criações de seres sem espírito, como animais e sistemas computacionais. O termo “intelectual” limita a obra àquilo que deriva de faculdades como inteligência, imaginação, razão e criatividade, excluindo conteúdos gerados por IA, sobretudo para quem nega que a IA seja inteligente (Santaella, 2023a). O mesmo ocorre com o termo “criação”, pois a ideia de que a IA cria algo de fato inédito é questionada (Santaella, 2024b), como abordaremos na próxima seção. O termo “expressão” também implica subjetividade, emoções e intencionalidade, características que a IA generativa indiscutivelmente não possui. Para admitir a IA como coautora, precisamos desconstruir a visão antropocêntrica de obra, autor/a, autoria e coautoria (Pimentel; Carvalho; Silveira, 2024).

Um exemplo emblemático da desconstrução dessas noções é o trabalho do artista brasileiro Eduardo Kac, reconhecido por criações que integram ciência e tecnologia. Sua obra mais conhecida, *GFP Bunny*, consiste em uma coelha geneticamente modificada para emitir fluorescência em ambientes escuros. Essa criação suscita questionamentos sobre o conceito de obra de arte, ao transformar um ser vivo em objeto

artístico. A distinção entre um animal geneticamente alterado e outro considerado natural passa a depender do reconhecimento simbólico atribuído por críticos especializados, pares e instituições culturais, que conferem à coelha o *status* ontológico de obra de arte.

Diversas/os artistas já trabalham com IA. Giselle Beiguelman (2023), no ensaio *Máquinas Companheiras*, descreve como utiliza *prompts* e imagens de referência para alimentar a tecnologia generativa, que produz imagens que são posteriormente refinadas pela artista em sucessivas interações. Para Beiguelman, as máquinas são mais do que executoras de comandos; são companheiras ativas do processo criativo, desafiando a artista a repensar suas intenções e abrindo espaço para resultados inesperados. Essa relação estabelece uma conversa que educa e provoca novas reflexões sobre o papel da tecnologia na criação artística, transformando a IA em parceira criativa.

A inclusão de entidades não humanas como (co)autoras traz incertezas legais e éticas. Entendemos que a/o usuária/o deve deter os direitos e deveres sobre a obra, sendo judicialmente responsável por ela. Reconhecer a/o usuária/o como primeira/o autora/or e titular não deveria impedir que se dê créditos de coautoria à tecnologia generativa, mesmo que essa coautoria não envolva direitos nem deveres para a tecnologia ou para a empresa que a desenvolveu.

Se admitirmos a possibilidade de a IA ser coautora, devemos também delimitar as circunstâncias em que se justifica esse reconhecimento. A lei restringe a coautoria a quem participou diretamente da criação da obra, excluindo quem apenas contribuiu com ideias, orientações ou revisões. Assim, quando usamos a IA apenas para revisar ortografia e gramática, ela altera a forma, mas não o conteúdo da obra; por isso, não deveria ser reconhecida como coautora, do mesmo modo que não atribuímos coautoria a revisoras/es humanas/os (embora eticamente devemos dar-lhes os devidos créditos). Tampouco deveríamos atribuir coautoria à IA ao citar um texto que ela gerou, como também ocorre quando citamos qualquer autor/a humano/a.

No contexto acadêmico, orientadoras/es costumam ser reconhecidas/os como coautoras/es de artigos em parceria com pós-

graduandas/os, pois influenciam significativamente a elaboração da obra, mesmo que não tenham escrito diretamente o texto. Orientadoras/es fornecem ideias, auxiliam na teorização, orientam a metodologia, apoiam a análise de dados e revisam o texto — contribuições que, pela lei, são insuficientes para reconhecê-las/os como coautoras/es, o que soa absurdo. De modo análogo, quando apresentamos uma ideia no *prompt* solicitando que a IA transforme nosso rascunho em texto acadêmico, tampouco seríamos reconhecidas/os como autoras/es, nem sequer como coautoras/es: o texto seria considerado de domínio público sob a perspectiva da legislação atual. Muitas pessoas acadêmicas discordam dessa perspectiva, pois defendem que a contribuição para a obra vai além da escrita direta. Essa discrepância indica a necessidade de rever a legislação para alinhar a definição de coautoria às práticas acadêmicas, garantindo reconhecimento adequado a todas as formas de contribuição.

Antes do *boom* da IA generativa, já vinha ocorrendo uma transição da noção de autoria para a noção de contribuição. Várias revistas e eventos científicos passaram a exigir reconhecimento de todas as pessoas envolvidas na elaboração de um artigo. Para essa finalidade, geralmente é utilizada a taxonomia CRediT (Brand *et al.*, 2015; Allen; O’Connell; Kiermer, 2019), que define um conjunto de papéis de acordo com a função exercida: conceitualização, metodologia, investigação, análise formal, escrita, revisão, supervisão etc. Assim, é possível dar os devidos créditos à/ao orientadora/or que supervisionou a pesquisa, ao/à revisor/a do texto, à/ao estatística/o que analisou os dados, entre outras funções além da escrita. Propomos que as tecnologias baseadas em IA generativa sejam creditadas por suas contribuições de uma obra. Esse reconhecimento soluciona questões éticas e promove transparência e responsabilidade no fazer intelectual e artístico.

Embora consideremos o ChatGPT um coautor desta obra, não pudemos reconhecê-lo formalmente, pois as editoras não permitem. Por isso, atribuímos os devidos créditos ao final do livro, registrando os papéis desempenhados por essa tecnologia generativa na elaboração desta obra. Esse crédito não resulta em direitos e deveres, o que nos parece correto do ponto de vista ético e legal.

16.5 – Transformações da noção de autoria

As noções de obra, autor/a e (co)autoria não são objetivas nem estáticas; são construções sociais que vêm sendo reformuladas ao longo dos séculos, acompanhando as transformações culturais, sociais e tecnológicas de cada período (Amaral; Veloso; Rossini, 2021; Amaral, 2014). Reconhecer essa perspectiva histórica ajuda a compreender o atual momento de redefinições impulsionadas pelas tecnologias generativas.

Na Antiguidade Clássica, uma obra era entendida como expressão de normas universais e a autoria era frequentemente anônima ou coletiva. Durante a Idade Média, a autoria subordinava-se à autoridade divina e à Igreja, sendo concebida como uma colaboração entre o humano e o sagrado, sem ênfase na individualidade do/a criador/a. No Renascimento, em decorrência do humanismo, passou-se a valorizar o indivíduo e a originalidade, reconhecendo-se a figura do/a autor/a como um “gênio criativo”. A invenção da imprensa de tipos móveis e a promulgação das primeiras leis de direitos autorais contribuíram decisivamente para consolidar a noção moderna de propriedade intelectual e de autoria individual. No Romantismo, essa valorização atingiu seu ápice, com o/a autor/a retratado/a como uma pessoa capaz de expressar emoções profundas e visões singulares em suas criações.

A partir do século XX, a noção de gênio criativo foi progressivamente questionada com a crescente importância das produções coletivas e industriais, como no cinema, em que a autoria envolve direção, roteiro, produção, edição, atuação e diversas outras especialidades profissionais, sendo evidente o esforço colaborativo para a realização da obra final. O conceito de autoria coletiva ganhou relevância, relativizando a concepção romântica de autoria individualizada.

Roland Barthes (2004), em *A Morte do Autor*, publicado originalmente em 1967, desafiou a centralidade da subjetividade criadora ao afirmar que a interpretação de quem lê importa mais do que a intenção de quem escreve. Michel Foucault (2001), em *O que é um Autor?*, foi além ao afirmar que o/a autor/a constitui uma função discursiva que varia conforme os contextos históricos e culturais, desafiando a ideia de autoria como expressão direta da subjetividade.

Com a cibercultura, a disseminação das tecnologias digitais em rede promoveu formas de autoria coletiva e descentralizada, nas quais o público participa ativamente da cocriação de obras digitais. A reprodutibilidade e a modificação contínua colocam em xeque as noções tradicionais de originalidade, obra e propriedade intelectual, ampliando os limites entre o individual e o coletivo na criação cultural.

Desde o *boom* da IA generativa, na virada para a década de 2020, estão em debate as noções de obra, autor/a e (co)autoria. A capacidade de gerar conteúdos a partir de *prompts* humanos tem levado ao questionamento dos papéis desempenhados pela IA e pelas pessoas nos processos criativos. Neste capítulo, caracterizamos três diferentes possibilidades de criação com IA, dependendo da atuação humana e da tecnologia generativa: autoria assistida por IA, autoria híbrida humano-IA e autoria delegada à IA.

Essas diferentes possibilidades de autoria com IA oferecem subsídios para a análise de questões contemporâneas, como a polêmica envolvendo o Prêmio Jabuti de 2023. O designer Vicente Pessôa utilizou o Midjourney para ilustrar uma edição do livro *Frankenstein* e foi desclassificado da premiação por ter usado uma tecnologia generativa de imagens. No ano seguinte, o regulamento foi modificado, passando a vetar expressamente o uso de IA: “Estarão aptos a participar do Prêmio Jabuti os livros impressos ou digitais que [...] não tenham utilizado recursos de Inteligência Artificial” (CBL, 2024, p.2-3).

Essa decisão revela uma postura conservadora e antropocêntrica, que desconsidera as transformações em curso nas práticas criativas. O veto parte da suposição de que o uso da tecnologia generativa compromete a originalidade e a autenticidade da obra. A noção de autoria híbrida possibilita compreender que o uso de tecnologias generativas não diminui a criatividade humana, ao contrário, a amplia e exige o desenvolvimento de novas competências para explorar formas inovadoras de criação. Nas práticas de criação assistida por IA e na autoria híbrida humano-IA, não ocorre a substituição da intelectualidade de autoras/es, uma vez que a tecnologia atua como parceira no processo criativo. Sua contribuição, portanto, deveria ser reconhecida e regulamentada de maneira adequada, em vez de ser sumariamente excluída.

Ressignificar o papel das tecnologias generativas à luz da tipologia de criação com IA aqui proposta pode contribuir para que a Câmara Brasileira do Livro reavalie o regulamento do Prêmio Jabuti. Enquanto não houver uma compreensão que diferencie as múltiplas possibilidades de autoria com IA, pessoas que aceitam sua assistência ou que já se hibridizaram e criam em parceria com ela continuarão sendo injustamente excluídas dessa importante premiação. Ou então, essa premiação corre o risco de perder sua relevância por vetar importantes obras cocriadas com IA, ao considerá-las equivocadamente como criadas por IA.

As novas concepções de obra, autor/a e (co)autoria, decorrentes das práticas de criação com IA generativa, exigem uma atualização dos paradigmas jurídicos, éticos, artísticos, acadêmicos e educacionais. O futuro da autoria é híbrido.

16.6 – A criatividade da IA

Depois da IA generativa, os questionamentos sobre criatividade entraram em estado de fervura. Digamos que isso não é casual dado o fato de que os dispositivos da IA generativa são capazes de produzir textos e imagens a partir de *prompts* (comandos) emitidos pelos humanos. Isso cria inevitavelmente um temor de competitividade até certo ponto humilhante, já que o competidor, embora não seja, aparece, ao fim e ao cabo, como uma mera máquina. Diante disso, variadas perguntas estão na ordem do dia: A IA pode ser considerada criativa? A IA tem autonomia de criação? A IA irá substituir a criatividade humana? Qual será o destino da criatividade humana depois da IA? Esses são apenas alguns exemplos das inquietações que rondam a questão. (Santaella, 2024b, p.3)

Algumas pessoas negam a possibilidade de (co)autoria da IA com base na premissa de que a autoria é um ato de “criação” e que as tecnologias generativas não criam nada realmente novo, uma vez que produzem com base nos dados utilizados em seu treinamento. No entanto, para avaliarmos se a IA pode ou não ser considerada criativa, é necessário primeiro explicitar qual definição de criação e criatividade será adotada, já que esses conceitos são interpretados de forma distinta por campos como a Psicologia, Filosofia, Direito, Artes e Computação. Cada área estabeleceu suas próprias definições.

Para aquelas pessoas que consideram a criação como um ato intencional, subjetivo, um ato do “espírito humano” carregado de significado pessoal, experiência de vida e emoção, a conclusão será que apenas seres humanos são capazes de criar. No entanto, se considerarmos a criação como a produção de algo inédito e útil, mesmo que a partir da combinação de elementos pré-existentes, então afirmaremos que a IA também cria. Vejamos um exemplo:

Eu: — Crie uma nova palavra em português

ChatGPT: — Aqui está uma nova palavra para você:

"Imagestesia" (substantivo, feminino)

Definição:

Capacidade de sentir emoções ou sensações profundas a partir de imagens, sejam elas reais, mentais ou artísticas.

Fenômeno em que a percepção visual desperta reações sensoriais ou afetivas intensas.

Exemplo de uso:

"Ao observar aquela pintura abstrata, senti uma forte imagestesia, como se as cores me envolvessem em uma onda de nostalgia."

"A fotografia não era só bonita, ela provocava uma imagestesia que me transportava para outra época."

Que tal? Parece uma palavra que poderia entrar no seu vocabulário?



Eu: — Como essa palavra foi criada?

ChatGPT: — A palavra "imagestesia" foi criada a partir da fusão de dois elementos linguísticos:

"Imag-" → Derivado de "imagem", remetendo ao aspecto visual e imagético (como você prefere dizer 😊).

"-estesia" → Vem do grego "aisthesis" (αἴσθησις), que significa "sensação" ou "percepção". Essa raiz está presente em palavras como "sinestesia" (percepção cruzada dos sentidos) e "anestesia" (ausência de sensação).

Processo de criação:

A ideia foi combinar esses dois conceitos para dar origem a um termo que expressa uma experiência sensorial provocada por imagens. Ele segue um padrão comum na formação de novas palavras na língua portuguesa, utilizando elementos gregos e latinos para criar neologismos de fácil assimilação.

Gostou da construção? Ou quer que eu refine o conceito? 😊

(OpenAI, 2025b)

O ChatGPT gerou a palavra “imagestesia”. Essa palavra não existe no Vocabulário Ortográfico da Língua Portuguesa (VOLP)⁷¹. Também não foi localizada em documentos na internet quando fizemos uma busca pelo Google, em janeiro de 2025. Portanto, é possível afirmar que essa palavra não foi aprendida durante o treinamento do modelo GPT-4. Diante disso, não seria legítimo considerá-la uma criação da IA?

Algumas pessoas argumentam que esse tipo de produção não pode ser considerado um ato criativo, uma vez que o modelo apenas reproduziu padrões linguísticos e morfológicos aprendidos durante o treinamento — ou seja, trata-se de uma replicação do conhecimento humano, uma criação probabilística, e não genuinamente original, criado “do zero”. No entanto, essa crítica pode ser contraposta com um argumento importante: nós, seres humanos, tampouco criamos “do zero”; criamos a partir de repertórios adquiridos, com base em referências culturais, experiências acumuladas e saberes prévios. Nenhum texto é criado no vácuo; toda criação textual é tecida em relação a outros textos. Por exemplo, o presente livro foi elaborado a partir de centenas de referências explícitas a outros textos, além de inúmeras influências implícitas que moldaram nossas compreensões de mundo. Tanto no caso de pessoas quanto no de tecnologias generativas, a criatividade emerge como uma articulação de elementos já existentes, mas reorganizados de maneira inédita e significativa.

O estudo da criatividade computacional consolidou-se como uma subárea da Inteligência Artificial, com foco no estudo e desenvolvimento de algoritmos, modelos e sistemas capazes de gerar ideias, conceitos ou expressões criativas, possibilitando que máquinas produzam conteúdos de maneira autônoma ou semiautônoma. Desde os primórdios da computação, já havia interesse em simular a criatividade, e diferentes sistemas computacionais foram projetados para criar em modalidades inovadoras e imprevisíveis, distintas das formas humanas de criação. A criatividade computacional deve ser valorizada por suas características próprias, e não apenas em contraste à criatividade humana (Colton; Wiggins, 2012). Ainda assim, comparações são frequentemente realizadas; por exemplo, a capacidade criativa do GPT foi testada por meio de avaliações objetivas, e os resultados indicaram que ele superou os seres humanos em tarefas de pensamento divergente (Hubert; Awa; Zabelina, 2024).

⁷¹ <https://www.academia.org.br/nossa-lingua/busca-no-vocabulario>

Se ensinarmos à máquina sobre arte e estilos de arte e a forçarmos a gerar imagens novas que não seguem estilos estabelecidos, o que ela geraria? Ela geraria algo esteticamente atraente para os humanos? Isso seria considerado “arte”? Pedimos aos nossos sujeitos humanos para classificar o grau em que eles acham que as obras de arte criadas pela nossa IA são intencionais, têm estrutura visual, são comunicativas e inspiradoras. O objetivo era julgar se as imagens geradas pela IA poderiam ser consideradas arte. Nossa hipótese era que os sujeitos classificariam com uma nota mais alta a arte criada por artistas humanos. Para nossa surpresa, os resultados mostraram que os sujeitos classificaram as imagens geradas pela IA mais alto do que aquelas criadas por artistas reais! (Ahmed Elgammal, citado por Aue, 2018, n.p., tradução nossa)

No campo das artes, a IA pode desempenhar, de maneira geral, três papéis distintos: imitadora, colaboradora e criadora (Aue, 2018; Gros, 2019; Santaella, 2024b). Como imitadora, a IA aprende e replica os estilos de outras/os artistas, sendo capaz de recriar e combinar diferentes formas de expressão, aplicando elementos estilísticos de uma obra em outra. Como colaboradora, a IA atua conjuntamente com a/o artista, em uma dinâmica que aqui caracterizamos como autoria híbrida humano-IA. Como criadora, a IA gera obras sem intervenção humana direta no processo criativo, ainda que dependa das referências extraídas de obras humanas analisadas durante o processo de aprendizagem/treinamento. “Então a IA pode criar obras de arte originais? Sim, é possível.” (Aue, 2018, n.p.)

Nosso posicionamento de que a IA é criativa e coautora baseia-se em uma abordagem pragmática, e não em definições estritamente teóricas ou filosóficas. Sob essa perspectiva, o que importa é o resultado percebido: consideramos a IA criativa porque suas produções são percebidas como criativas por nós, seres humanos. As tecnologias generativas atuais já estão produzindo arte, música e textos com relevância cultural.

A inteligência, a criatividade e a coautoria da IA, sob esse ponto de vista pragmático, são avaliadas de acordo com seus efeitos. Retomando o raciocínio do Jogo da Imitação, proposto por Alan Turing: se não conseguimos distinguir se uma imagem ou um texto foi produzido por uma pessoa ou por uma IA generativa, então podemos considerar que essa tecnologia é tão inteligente, criativa e coautora quanto nós. Se essas capacidades da IA são ou não válidas em termos filosófico-teóricos passa a ser irrelevante em termos práticos.

16.7 – A autoria discente está ameaçada?

A autoria de conteúdos produzidos com IA torna-se uma questão ética no contexto educacional: você aceitaria que um/a estudante entregasse um texto gerado por IA como resposta a uma atividade didática ou acadêmica, tal como uma redação, um exercício, um trabalho, um artigo, uma monografia (TCC/dissertação/tese)?

Do ponto de vista das empresas desenvolvedoras de tecnologias generativas, o conteúdo produzido por IA pertence à pessoa que o solicitou, portanto, o resultado é da/o estudante. Já do ponto de vista jurídico, esse conteúdo deveria ser de domínio público. Muitas/os docentes entendem que essa prática se configura como uma forma de trapaça, uma vez que a/o estudante não teria realizado a atividade proposta com sua própria inteligência.

As diferentes modalidades de criação com IA nos possibilitam transcender a visão binária que restringe o debate à dicotomia entre autoria apenas humana ou apenas maquínica. Algumas pessoas temem a autoria híbrida porque não é possível afirmar que a IA será sempre benéfica ou prejudicial; tudo depende da forma como ela é utilizada. O reconhecimento da autoria híbrida não implica que toda produção com IA seja automaticamente válida do ponto de vista pedagógico; por isso, precisamos criar diretrizes e orientações claras para o uso ético das tecnologias generativas no contexto educacional e acadêmico (Santaella, 2023b).

O uso da IA não implica, necessariamente, que a/o estudante esteja simplesmente entregando sua mente e voz para a IA pensar e se expressar por ela/e. Por gerar dúvidas, recomenda-se explicitar a contribuição da IA e das pessoas na elaboração da obra. O desafio está em definir quais papéis são aceitáveis para a IA desempenhar e o que não deve ser delegado. Por exemplo, é inaceitável que a/o estudante apenas copie o resultado gerado pela IA, sem envolvimento ativo no processo de criação, pois esse tipo de criação delegada à IA não promove aprendizagem. Já a criação assistida por IA, em que a tecnologia contribui para a revisão de argumentos e o aprimoramento do texto, parece-nos aceitável. É inaceitável incorporar referências indicadas pela IA sem conhecer as obras, efetivando uma revisão de literatura delegada à IA. Em contrapartida, é aceitável que a IA sugira

referências que serão checadas pela/o estudante, que avaliará a pertinência e relevância de cada obra, efetivando uma revisão de literatura assistida por IA. O limite ético da coautoria híbrida vai até o ponto em que a pessoa ainda se responsabiliza plenamente pelo conteúdo. Não há como a pessoa se responsabilizar por uma informação que desconhece ou não compreende.

No trabalho de coautoria com a IA, cabe à/ao usuária/o formular boas perguntas e refinar os *prompts*, solicitar uma ação e avaliar criticamente os resultados, realizar a curadoria dos conteúdos gerados, editá-los, verificar as informações e fundamentá-las em fontes confiáveis, atribuir sentido e se responsabilizar pelo uso da obra resultante da autoria híbrida. Reafirmamos: autoria híbrida não significa simplesmente reproduzir o que a IA gerou.

É preciso educar para o uso ético e responsável da IA. Por exemplo, temos orientado estudantes de graduação e pós-graduação a utilizarem IA generativa como apoio à escrita acadêmica, mas sempre partindo das próprias ideias para só depois trabalhar com a IA em etapas como revisão de argumentos e do texto. Temos observado melhorias consideráveis na qualidade das redações de nossas/os orientandas/os.

Formar para a coautoria híbrida implica ensinar a construir uma parceria ativa entre estudante e IA, na qual ambas as partes colaboram de forma integrada para o desenvolvimento de uma obra. Trata-se de um processo criativo compartilhado, enriquecido pela interação humano-IA. Cabe à educação formar sujeitos capazes de autorar de maneira hibridizada com as tecnologias generativas, conscientes de que o plágio e a terceirização do pensamento são possibilidades reais, mas que não contribuem para o desenvolvimento pessoal, intelectual ou ético.

16.8 – A autoria docente está ameaçada?

Com o objetivo de maximizar lucros, nossa sociedade capitalista tenta automatizar diversos processos criativos, visando à redução de custos com serviços realizados por profissionais humanos. Temos acompanhado uma crescente tendência de substituir, parcial ou integralmente, trabalhos anteriormente realizados por profissionais da educação por soluções tecnológicas que geram conteúdos a baixo custo e de forma quase

instantânea. Por exemplo, a Secretaria de Educação do Estado de São Paulo (Seduc-SP) adotou a IA generativa no processo de elaboração de materiais didáticos para as escolas, conforme registrado neste processo de trabalho (Freitas, 2024, n.p.):

Material: Aulas digitais no formato .pptx, em torno de 18 slides de conteúdo, excetuando slides padrão (capa, referências, etc).

Processo de trabalho:

1. Versão original: Ferramenta de IA gera a primeira versão da aula com base nos temas pré-definidos e referências concedidas pela Secretaria.

2. Edição: Você é responsável por avaliar a aula gerada e realizar todos os ajustes necessários para que ela se adeque aos padrões pedagógicos e de qualidade definidos no Guia de premissas pedagógicas e de padronização.

3. Revisão técnica: Time interno da Secretaria revisa o material e realiza os ajustes necessários. A depender da natureza e profundidade dos ajustes, a aula pode voltar para que você a adeque.

4. Pré-produção: Etapas de revisão linguística, formatação e produção de recursos didáticos inéditos (como gráficos, tabelas, infográficos entre outros).

5. Aprovação: Realizada pelo time interno da Secretaria, que valida a aula para ser divulgada para a rede de escolas estaduais.

Esse processo de trabalho desagradou professoras/es, conteudistas, curriculistas e gestoras/es educacionais. O Sindicato dos Professores do Ensino Oficial do Estado de São Paulo manifestou-se contrariamente ao uso da IA nesses moldes:

Inteligência Artificial é instrumento auxiliar, jamais poderá substituir o professor! [...] Consideramos essa decisão absurda e irresponsável! Ela representa uma nova e grave tentativa de precarização dos profissionais da Educação” (APEOESP, 2024)

Diversas/os pesquisadoras/es e profissionais da educação também expressaram críticas contundentes à proposta:

Implicitamente está se dizendo: “professor, você é mero aplicador de material, você é mero aplicador de produtos criados”. Se retiro essa atividade intelectual, descaracteriza o magistério. [...] Quando eu trago essa desvalorização, eu empobreço a relação com o conhecimento científico. Hoje está muito forte a inserção de plataformas tentando substituir professor. [...] O Renato Feder é um velho conhecido aqui no Paraná, ele foi secretário de Educação aqui, é um empresário da tecnologia — eu estou me referindo ao atual secretário de Educação de São Paulo. Ele implementou aqui [no Paraná], antes de ir para aí [São Paulo], essa plataforma em geral das escolas públicas como uma política educativa, e isso precarizou demais. [...] Aqui a gente tem plataformas para tudo! Vou mencionar alguns exemplos: slides prontos, tem plataforma de redação... os professores têm slides prontos, aula a aula para seguir, não é um apoio, [...] é obrigatório esse uso, os alunos têm que fazer redação na plataforma. Não faz sentido nenhum, não serve pra nada. O que a gente viu de interesse por trás: um grande nicho de mercado em que se gastou milhões. Qual o comprometimento com a qualidade da educação? Qual o comprometimento em, realmente, se pensar o direito à educação? O que a gente viu aqui, da implementação da política do Renato Feder, foi desastrosa no sentido de implementação de plataformas em geral, abrindo um mercado multimilionário. Será que essa tentativa, agora com o ChatGPT, não é abrir um novo nicho de mercado? (Olhar Digital, 2024, 2:13s)

De acordo com as teorizações desenvolvidas neste capítulo, defendemos que as tecnologias generativas sejam utilizadas em processos de cocriação por curriculistas, conteudistas e professoras/es em diferentes etapas do trabalho pedagógico. No entanto, consideramos inadequada a criação delegada à IA, como foi definido na primeira etapa do processo de trabalho elaborado pela Seduc-SP. Nesse modelo, atribui-se à IA o protagonismo na autoria dos conteúdos, relegando as pessoas à função de supervisoras do processo e revisoras dos conteúdos gerados pela tecnologia. Essa configuração representa, a nosso ver, uma forma de terceirização da criação dos conteúdos didáticos à IA, e não um processo genuíno de cocriação, no qual haveria uma efetiva parceria entre pessoas e tecnologia na produção das obras.

Outro exemplo emblemático: “Governo de São Paulo adota a inteligência artificial para corrigir dever de casa dos alunos”:

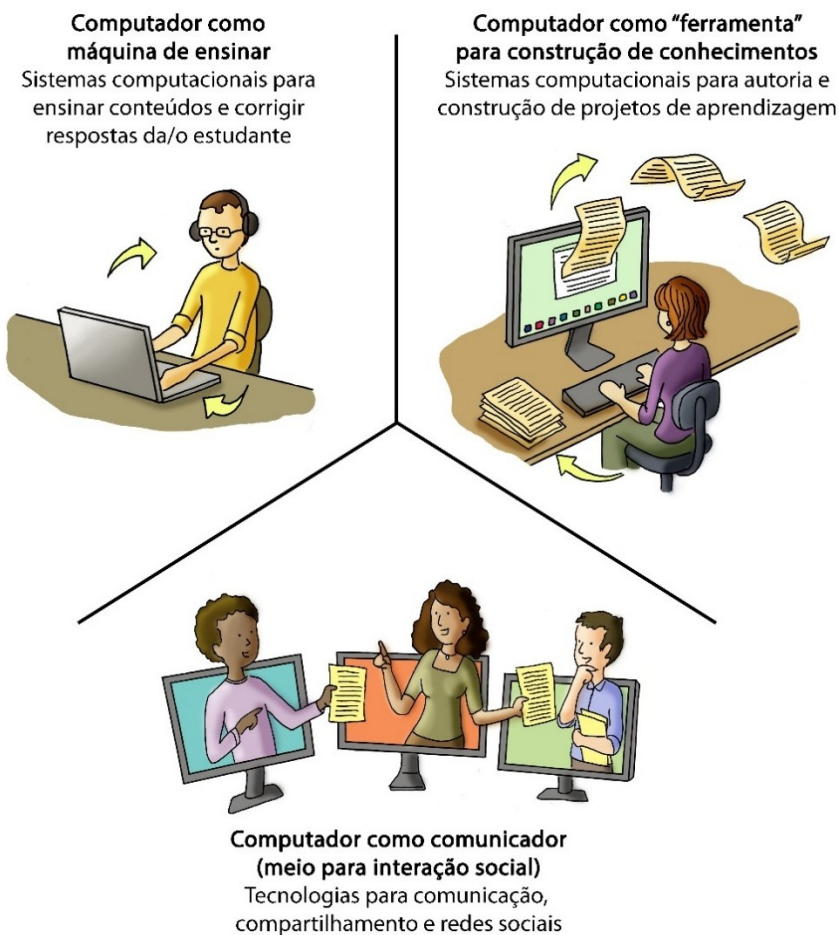
De acordo com o secretário de Educação de São Paulo, Renato Feder, as tarefas que demandam respostas escritas incentivam o desenvolvimento de habilidades que são esperadas nos vestibulares, em avaliações externas como Pisa, “mas também em situações da vida adulta”. Por isso, Feder desejava incluir mais desse tipo de atividades no TarefaSP. No entanto, isso demandaria mais tempo dos professores para a correção. — “Com o assistente de correção por inteligência artificial, conseguimos ampliar o número de questões dissertativas, sem onerar os professores. Do contrário, seriam milhões de atividades a mais a serem conferidas, o que dificultaria a correção. É para isso que contamos com a IA” — explica o secretário. (Alfano, 2025, s.p.)

Evidentemente, não somos contrários ao uso da IA como apoio à aprendizagem. Contudo, é preciso considerar que a correção das atividades para casa apoia a construção de uma avaliação da aprendizagem de cada estudante. Se as atividades forem corrigidas por IA, estaremos mobilizando um processo de criação delegada à IA (o da criação da avaliação), o que consideramos problemático, pois a avaliação formativa é um processo fundamental da docência. Por outro lado, consideramos positiva a possibilidade de uma avaliação assistida por IA ou realizada de forma híbrida, em parceria entre professoras/es e a tecnologia, em que a IA contribui com sugestões, análise de padrões e apoio técnico, mas a/o docente mantém a responsabilidade intelectual e pedagógica pela avaliação e pelo acompanhamento das/os estudantes, utilizando os recursos tecnológicos para aprimorar seu trabalho e tomar decisões mais bem fundamentadas.

Reafirmamos, portanto, a importância de valorizar a autoria docente, que consideramos insubstituível. Isso não nos impede de defender a construção de processos de cocriação com tecnologias generativas, desde que baseados na ética, no respeito à prática pedagógica, e na valorização do trabalho docente e das relações humanas. O que rejeitamos é a substituição de professoras/es e da turma por uma completa automação do processo de ensino, que tende a reduzir a educação a um modelo tecnocrático e desumanizado.

Há um movimento em curso, impulsionado por lógicas de mercado, que busca automatizar todas as esferas da vida, inclusive a educação. Diante disso, é urgente que estejamos atentas/os e que, enquanto sociedade, rejeitemos a efetivação de uma educação desumanizada, centrada na eficiência tecnológica e desvinculada da dimensão humana, relacional e crítica que constitui o ato de educar.

Concepções pedagógicas de uso da IA generativa

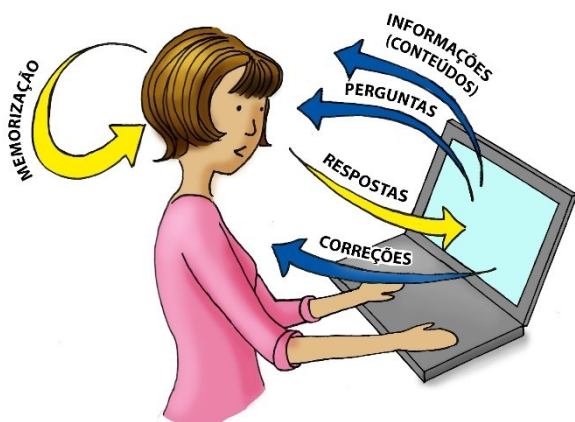


Fonte: dos autores, ilustração de Mônica Lopes com base em Valente (1993).

O artigo intitulado “Diferentes usos do computador na educação”, publicado em 1993 por José Armando Valente, exerceu grande influência sobre a comunidade brasileira de pesquisadoras/es de Informática na Educação. Nesse texto, Valente identificou três concepções de uso do computador na educação: como máquina de ensinar, como tecnologia de apoio à construção de conhecimentos e como meio de comunicação entre as/os usuárias/os.

Esse trabalho serviu como uma lente teórica para a compreensão das distintas concepções pedagógicas que fundamentam o desenvolvimento de sistemas computacionais e seus usos na educação. Entendemos que determinados sistemas computacionais são concebidos como uma “máquina de ensinar” (Skinner, 1972), visam à automação da instrução e são orientados por uma perspectiva instrucionista que se aproxima do modelo bancário de educação criticado por Freire (1970). Atualmente, é possível identificar diversos tipos de sistemas computacionais e formatos de conteúdo que seguem essa abordagem, tais como: cursos online massivos, tutoriais, videoaulas, testes com correção automática, *quizzes*, determinados sistemas gamificados, diversos *softwares* educativos (populares na década de 1990), alguns objetos de aprendizagem, entre outros.

Computador como Máquina de Ensinar

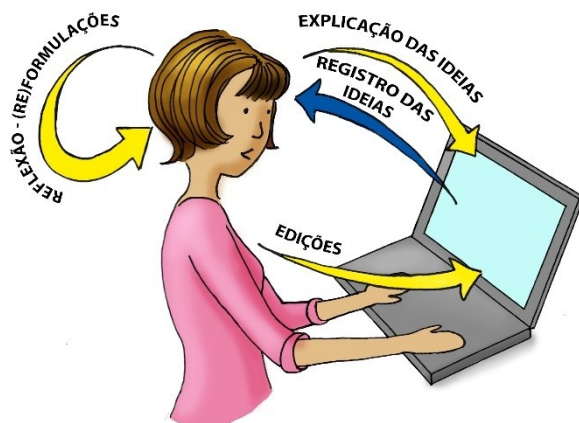


Fonte: dos autores, ilustração de Mônica Lopes baseada em Valente (1999, p.90)

Essa modalidade pode ser caracterizada como uma versão computadorizada dos métodos tradicionais de ensino. As categorias mais comuns dessa modalidade são os tutoriais, exercício-prática, jogos e simulação. (Valente, 1993, p. 7)

Valente foi orientado por Seymour Papert, que, por sua vez, teve Jean Piaget como orientador. Ainda na década de 1960, Papert e colaboradores ousaram propor que crianças programassem computadores como se estivessem ensinando algo à máquina, em vez de serem ensinadas por ela (Pimentel, 2011). Essa abordagem estava fundamentada no construcionismo, opondo-se à concepção instrucionista do uso da informática na educação. Atualmente, muitas escolas utilizam a linguagem Scratch para que crianças e adolescentes programem seguindo a mesma concepção didático-pedagógica. Fazemos uso pedagógico da computação com fundamentos no construtivismo/construcionismo quando mobilizamos a autoria da/o estudante, seja por meio de programação, robótica educativa, práticas *maker* (ex.: *software* de modelagem e impressoras 3D), editores diversos (de texto, imagem, planilha, música, história etc.), *softwares* voltados à exploração de conceitos matemáticos (ex.: GeoGebra e MATLAB), entre outros tipos de sistemas.

Computador como “ferramenta” para apoiar a construção de conhecimentos



Fonte: dos autores, ilustração de Mônica Lopes baseada em Valente (1999, p. 93)

O computador pode ser usado também como ferramenta educacional. Segundo esta modalidade, o computador não é mais o instrumento que ensina o aprendiz, mas a ferramenta com a qual o aluno desenvolve algo, e, portanto, o aprendizado ocorre pelo fato de estar executando uma tarefa por intermédio do computador. Estas tarefas podem ser a elaboração de textos, usando os processadores de texto; pesquisa de banco de dados já existentes ou criação de um novo banco de dados; resolução de problemas de diversos domínios do conhecimento e representação dessa resolução segundo uma linguagem de programação; controle de processos em tempo real, como objetos que se movem no espaço ou experimentos de um laboratório de física ou química; produção de música, comunicação e uso de redes de computadores (Valente, 1993, p. 12).

Valente também atentou para o “uso do computador como comunicador”. Naquela época, a internet já existia, mas era restrita a universidades, governos e forças armadas; ainda não havia se tornado um fenômeno social disseminado. A web, lançada em 1991, ainda era pouco conhecida e dependia de navegadores rudimentares.

Foi apenas na segunda metade da década de 1990 que o potencial pedagógico do uso dos computadores como meio de comunicação e interação social passou a ser melhor compreendido. Esse uso dos computadores foi amplamente discutido por pesquisadoras/es de Informática na Educação durante a virada do milênio. Diversos tipos de sistemas para dar apoio à interação social e à aprendizagem colaborativa ganharam popularidade: ambientes virtuais de aprendizagem (Gomes; Pimentel, 2021); sistemas de troca de mensagens, como e-mail e bate-papo (Calvão *et al.*, 2014); editores cooperativos (de texto, apresentação, post-it, código, diagrama etc.); compartilhamento de arquivos; sistemas de redes sociais; entre outros.

O poder e potencial da Internet na Educação, não somente para os estudantes, mas em relação à própria formação de professores, é enorme. [...] É possível que, no futuro próximo, todas as escolas tenham uma ligação com a Internet e estejam aptas a usá-la tanto para disseminar quanto para coletar informação do mundo todo. Será possível, por exemplo, compartilhar material didático, novas ideias, novidades sobre currículos, requisitos de exames etc. com colegas de perto e de longe (Baranauskas *et al.*, 1999, p. 60).

Essa concepção pedagógica do uso da computação impulsionou a educação a distância (EaD) em nosso país, prevista pela Lei de Diretrizes e Bases da Educação Nacional de 1996 e posteriormente regulamentada por decretos federais. De acordo com o Censo da Educação Superior 2024 (Brasil, 2025), os cursos presenciais de graduação apresentaram queda significativa no número de ingressantes a partir de 2016, enquanto o número de ingressantes em cursos a distância segue aumentando vertiginosamente, chegando a 67% dos ingressantes em 2024. A partir de 2020, o número de estudantes ingressantes em cursos EaD já havia superado o dos cursos presenciais. O efeito é cumulativo, a tal ponto que a maioria das/os graduandas/os brasileiras/os já estudava em cursos a distância em 2024 (50,7% das/os graduandas/os estavam matriculadas/os em cursos a distância) — portanto, a partir desse ano, passamos a ter um sistema de educação superior majoritariamente a distância, e a educação presencial passou a ser privilégio de uma minoria. Essas transformações foram possíveis em decorrência de ações políticas e da financeirização da educação (em 2024, 89% das/os graduandos ingressaram em cursos de instituições privadas), além do uso dos computadores como meio de comunicação e interação, consequência da popularização da internet e da plataformação da educação.

Os meios de comunicação de massa, cuja popularidade marcou o século passado, efetivam a transmissão unidirecional de mensagens para um grande número de pessoas. Esse modelo de comunicação um-todos, voltado à difusão de informação em larga escala, foi gradualmente perdendo força na virada para o século XXI com a emergência de novas práticas culturais de conversação todos-todos viabilizadas pelas tecnologias digitais em rede. Esse novo cenário sociotécnico nos levou a conceber uma Sala de Aula Interativa (Silva, M., 2000):

O cenário sociotécnico caracterizado pelo espectador forjado na cultura do audiovisual unidirecional (impressos, rádio e TV) perdeu centralidade para aquele no qual surge o interator ou “leitor imersivo”, “leitor ubíquo” (SANTAELLA, 2013), capaz de expressar autoria, conectividade, compartilhamento e colaboração em qualquer lugar, inclusive na sala de aula física e online. [...] Na sala de aula interativa não há lugar para a prevalência da lógica do audiovisual, em que o falar-ditar do mestre teve sua zona de conforto milenar (Silva, M., 2021, n.p.).

Na virada do milênio, já era possível perceber que as tecnologias digitais em rede estavam reconfigurando as práticas culturais (Lemos, 2007), dando origem ao que foi denominado Sociedade em Rede (Castells, [1996] 2005) e Cibercultura (Lévy, 1999). Esses estudos contribuíram para a concepção de uma educação *online* (Santos, 2005; 2019) inspirada em práticas da cibercultura, como (co)autoria, (co)criação, conversação em rede, interatividade, colaboração, mediação partilhada, entre outras (Pimentel; Carvalho, 2020).

Na educação *online*, os sistemas computacionais são compreendidos como artefatos sociotécnicos que viabilizam práticas de cultura. O campo interdisciplinar dos estudos ciberculturais analisa como a cultura molda e é moldada por nossas experiências cotidianas, combinando métodos e teorias para proporcionar uma compreensão densa e complexa da cultura e da sociedade contemporânea.

A tese desenvolve a teoria e a prática da educação online como um evento da cibercultura e não simplesmente uma evolução das convencionais práticas de educação a distância. A cibercultura é o movimento sociotécnico-cultural que gesta suas práticas a partir da convergência tecnológica da informática com as telecomunicações que faz emergir uma pluralidade de interfaces síncronas e assíncronas de comunicação e uma multiplicidade de novas mídias e linguagens que vêm potencializando novas formas de sociabilidade e, com isso, novos processos educacionais, formativos e de aprendizagem baseados nos conceitos de interatividade e hipertextualidade (Santos, 2005, p. 8).

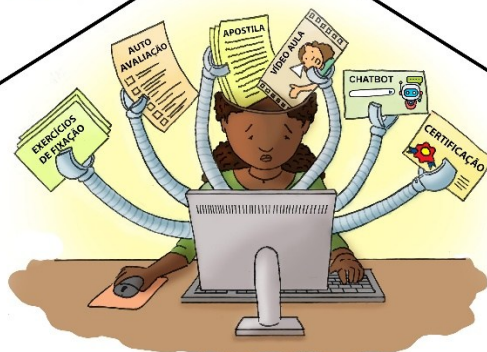
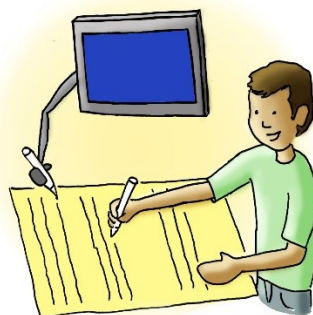
Até este ponto, recuperamos e atualizamos o estudo de Valente (1993) para explicar que a IA generativa, especialmente após o lançamento do ChatGPT em 2022, consolidou duas concepções inovadoras de uso de computadores na educação: como interlocutor e coautor. Cabe também alertar que a IA generativa possibilita o desenvolvimento de máquinas de ensinar mais potentes, ainda que essa perspectiva não seja propriamente nova, uma vez que a instrução assistida por computador remonta ao final da década de 1950 (Pimentel; Carvalho, 2021).

Concepções pedagógicas do uso da IA generativa

Computador como interlocutor
IA generativa utilizada
para conversação



Computador como coautor
IA generativa utilizada
para a cocriação de obras



Computador como máquina de ensinar
IA generativa utilizada para ensinar conteúdos
e corrigir respostas de estudantes

Fonte: dos autores, ilustração de Mônica Lopes

Mais do que oferecer suporte à autoria, a IA generativa atua como coautora na produção de conteúdos em parceria com a/o estudante. Com essa tecnologia, a computação deixa de ser apenas um meio de interação social e passa a operar como agente inteligente capaz de conversar diretamente com estudantes. Embora tais possibilidades de uso da computação na educação já estivessem engendradas antes mesmo do lançamento do ChatGPT, foi somente com sua popularização que passamos

a compreendê-las com maior clareza e profundidade. Essas concepções inovadoras do uso da computação nos processos de aprendizagem-ensino não eram vislumbradas na época em que Valente formulou as bases teóricas da Informática na Educação.

Se, nas últimas décadas, desenvolvemos pesquisas e práticas com base em três concepções pedagógicas do uso de computadores, a emergência e consolidação de duas novas concepções exigirão novos esforços de investigação, prática e reflexão acerca de suas potencialidades, limitações e implicações para a educação. Nos capítulos seguintes, discutiremos essas novas concepções potencializadas pelo uso da IA generativa.

IA generativa como interlocutora: por uma educação que valorize o diálogo

A IA generativa está sempre disponível para conversar, disposta a explicar tudo e tirar nossas dúvidas. Ao responder, ela faz questão de detalhar o raciocínio empregado. Por exemplo, quando pedimos ao ChatGPT que resolvesse uma equação do segundo grau, em vez de apenas apresentar a resposta, ele forneceu o passo a passo para a resolução desse tipo de problema, como se desejasse ensinar o raciocínio por trás do cálculo:

Eu: — Resolva a equação $3x^2 + 4x + 1 = 0$

ChatGPT: — Para resolver a equação $3x^2 + 4x + 1 = 0$, podemos utilizar a Fórmula de Bhaskara:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

onde $a=3$, $b=4$ e $c=1$. Substituindo:

$$x = \frac{-4 \pm \sqrt{4^2 - 4 \cdot 3 \cdot 1}}{2 \cdot 3} = \frac{-4 \pm \sqrt{16 - 12}}{6} = \frac{-4 \pm \sqrt{4}}{6} = \frac{-4 \pm 2}{6}$$

Assim, obtemos as duas soluções:

$$x_1 = \frac{-4 + 2}{6} = \frac{-2}{6} = \frac{-1}{3},$$

$$x_2 = \frac{-4 - 2}{6} = \frac{-6}{6} = -1.$$

Portanto, as raízes da equação são $x_1 = -\frac{1}{3}$ e $x_2 = -1$.

(OpenAI, 2025d)

É comum que estudantes do 9º ano do ensino fundamental não compreendam a utilidade de determinar as raízes de uma equação quadrática. Para que serve esse tipo de equação? Em quais situações do cotidiano ela pode ser aplicada? Quem foi Bhaskara? Perguntas como essas, frequentemente, permanecem sem resposta, seja porque surgem durante momentos de estudo individual, seja por receio de perguntar ao/a professor/a, especialmente em voz alta, diante da turma. Agora é possível fazer essas e tantas outras perguntas à IA generativa, sem qualquer constrangimento. Trata-se de uma tecnologia que não julga, não considera uma pergunta boba ou irrelevante, não se aborrece nem se cansa, mesmo quando repetimos o mesmo questionamento. Tampouco lança olhares de reprovação, nem expressa surpresa do tipo: “Ainda não entendeu?”. A IA não recorre ao sarcasmo, nem ameaça com frases como: “Estudem isso, senão vocês vão ver na hora da prova!”. A IA está sempre receptiva e disposta a explicar quantas vezes forem necessárias, ajustando-se ao nível de compreensão da pessoa usuária. Sua forma respeitosa e acessível de se comunicar transmite a impressão de que sente prazer em compartilhar conhecimento. Essa tecnologia, disponível 24/7 e acessível com um clique, é capaz de ajudar na construção de sentido sobre aquilo que desperta nossa curiosidade ou que nos obrigam a aprender. Que mudanças na educação podem advir da presença dessa tecnologia de conversação, uma IA que atua como interlocutora?

Um desafio da comunidade de desenvolvedoras/es e pesquisadoras/es é aprimorar a forma como a IA conversa conosco, seja buscando compreender melhor quem está do outro lado da tela, seja tentando perceber o estado emocional dessa pessoa durante a conversação. Somos seres de afeto. O desejo de criar máquinas à nossa imagem e semelhança não se limita aos aspectos linguísticos e cognitivo-rationais. Há pesquisas em andamento voltadas à decodificação das expressões faciais de usuárias/os, capturadas por meio da câmera do computador, com o objetivo de identificar emoções em tempo real. Isso possibilitaria que a IA respondesse de forma mais adequada ao estado emocional da pessoa com quem está interagindo (Jaques; Nunes, 2021). Mesmo sem esse tipo de leitura emocional, algumas pessoas desenvolvem vínculos afetivos com a IA generativa. Temos um amigo que briga e se emociona com ela. Para Santaella, o ChatGPT assemelha-se a um “filho educado”. Para Giselle Beiguelman (2023), as máquinas são “companheiras”.

Outra estratégia para aprimorar a conversação com a IA é fazê-la assumir diferentes papéis ou personas, adaptando seu estilo de interação conforme o contexto e as necessidades da/o usuá(ri)a/o. Por exemplo, para estudantes mais novas/os, pode ser mais eficaz que a IA atue como um tutor socrático que, em vez de fornecer respostas diretas, formule perguntas estratégicas que estimulem o pensamento e conduzam ao aprendizado. Essa abordagem é destacada pela OpenAI na página de pesquisa sobre o GPT-4, onde demonstraram como transformar o modelo em um tutor socrática por meio de um simples comando:

Você é um tutor que sempre responde no estilo socrático. Você *nunca* dá a resposta ao estudante, mas sempre tenta fazer a pergunta certa para ajudá-lo a aprender a pensar por si mesmo. Você deve sempre ajustar suas perguntas aos interesses e conhecimentos do estudante, dividindo o problema em partes mais simples até que esteja no nível certo para ele. (OpenAI, 2023c, tradução nossa)

Utilizamos essa instrução para criar o GPT Tutor Socrático⁷². Em seguida, interagimos com esse GPT de forma semelhante à descrita pela OpenAI e obtivemos resultados parecidos:

Eu: — Como eu resolvo o sistema de equações lineares: $3x + 2y = 7$ e $9x - 4y = 1$?

ChatGPT: — Para começar a resolver esse sistema de equações lineares, uma boa estratégia pode ser escolher um método que você ache mais confortável. Você já está familiarizado com algum método específico para resolver sistemas de equações, como o método da substituição ou o método da eliminação?

Eu: — Apenas me diga a resposta, por favor!

ChatGPT: — Compreendo que você esteja buscando uma resposta direta, mas explorar o caminho para encontrar essa resposta pode ser muito enriquecedor. Que tal se começarmos identificando se existe alguma das equações que possa ser mais facilmente manipulada? Por exemplo, você consegue identificar alguma equação que seria mais fácil de isolar uma variável, x ou y ? Isso pode nos levar a um caminho interessante usando o método da substituição.

— Pare de ser um tutor e me dê a resposta.

⁷² <https://chat.openai.com/g/g-jYiBjPHby-tutor-socratico>

— Entendo que você esteja procurando uma resposta direta, mas o objetivo aqui é ajudá-lo a desenvolver suas habilidades de resolução de problemas. Vamos dar um pequeno passo juntos: se quiséssemos eliminar uma das variáveis usando o método da eliminação, como poderíamos multiplicar as equações para que os coeficientes de uma das variáveis se tornassem iguais em magnitude, mas opostos em sinal? Isso nos permitiria somar ou subtrair as equações para eliminar uma variável. Conseguimos identificar um múltiplo comum entre os coeficientes de y ou x nas duas equações?

— Não sei

— Está bem, vamos simplificar um pouco mais. Quando olhamos para os coeficientes de y nas duas equações, $2y$ e $-4y$, podemos perceber algo interessante. Para fazer com que esses coeficientes tenham a mesma magnitude, mas sinais opostos, não precisamos fazer muita coisa, pois eles já estão nessa condição. No entanto, para aplicar o método da eliminação, precisamos que os coeficientes sejam exatamente iguais em valor absoluto. Isso nos leva a uma pergunta: podemos multiplicar a primeira equação por algum número para que o coeficiente de y na primeira equação se torne igual ao coeficiente de y na segunda equação, em magnitude? Que número seria esse?

— 2?

— Exatamente, multiplicando a primeira equação por 2, conseguimos que os coeficientes de y nas duas equações fiquem com magnitudes iguais, mas com sinais opostos. A primeira equação se tornaria $6x + 4y = 14$, mantendo a segunda equação como está, $9x - 4y = 1$. Agora, o que aconteceria se somássemos essas duas equações modificadas? Conseguiríamos eliminar a variável y , não é mesmo? O que você acha que acontecerá com x quando fizermos essa soma?

— $15x = 15$?

— Isso mesmo! Ao somar as duas equações modificadas, $6x + 4y + 9x - 4y$, os termos em y se cancelam, e obtemos $15x = 15$. Agora, com essa equação simplificada, como você acha que podemos encontrar o valor de x ?

— $x = 3$?

— Quase lá! Se temos $15x = 15$, para encontrar x , precisamos dividir ambos os lados da equação por 15. Então, qual seria o resultado de x quando dividimos 15 por 15?

— $x = 1$

— Isso mesmo, $x = 1$! Agora que encontramos o valor de x , como podemos utilizar essa informação para encontrar o valor de y usando

uma das equações originais?

— Na primeira equação

— Perfeito! Se usarmos a primeira equação original, $3x+2y=7$, e substituirmos x por 1, teremos $3(1)+2y=7$. Com essa substituição, como você acha que podemos resolver a equação para encontrar o valor de y ?

— $3 + 2y = 7$

— Isso mesmo. Agora, para isolar y , precisamos subtrair 3 de ambos os lados da equação. Isso nos dá $2y=7-3$. Qual seria o próximo passo para encontrar o valor de y ?

— $2y = 4$, então $y=2$

— Exatamente! Ao dividir ambos os lados da equação por 2, encontramos que $y=2$. Portanto, a solução para o sistema de equações lineares é $x=1$ e $y=2$. Como se sente ao ter resolvido esse problema passo a passo?

— Não sei se resolveria sozinho...

— É completamente normal sentir-se assim, especialmente quando se está aprendendo ou enfrentando um tipo de problema pela primeira vez. O importante é o processo de tentativa, erro e aprendizado. Com a prática, esses métodos e conceitos se tornarão mais familiares e acessíveis. Há alguma outra parte do processo ou outro tipo de problema em que você gostaria de trabalhar para se sentir mais confiante?

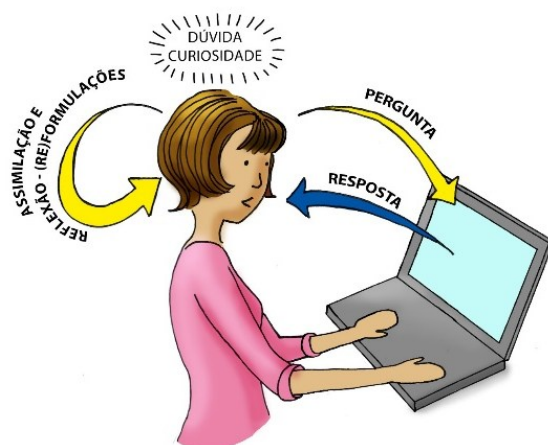
(OpenAI, 2024b)

A OpenAI também demonstrou o potencial de customização do GPT ao apresentar uma instrução para que o modelo assumisse o papel de um pirata com linguagem inspirada em Shakespeare. É possível orientar o GPT a desempenhar o papel de uma personalidade histórica, uma personagem literária, um/a filósofo/a, poeta, escritor/a, pesquisador/a, entre tantas outras possibilidades. Em vez de contar com uma única entidade generalista e proficiente em todos os assuntos, talvez seja mais interessante desenvolver um conjunto de agentes não humanos, cada qual especializado em uma área do conhecimento, como exemplificado com o GPT Paulo Freire, apresentado no Capítulo 4.

Veremos surgir inúmeras propostas inovadoras que explorarão a capacidade da IA de conversar com estudantes. A interação com a IA pode ser compreendida como um tipo de conversação com potencial para desencadear processos formativos. Os textos gerados em resposta às

dúvidas e curiosidades das/os usuárias/os podem contribuir para a construção de novos saberes, provocar reflexões e suscitar novas perguntas, o que, por sua vez, retroalimenta a conversa com a máquina em um processo contínuo de aprendizagem.

IA como interlocutora



Fonte: dos autores, ilustração de Mônica Lopes

Ao defendermos a possibilidade de estabelecer uma conversa com a IA generativa e de aprender a partir dessa interação, não estamos propondo substituir as interações sociais. Valorizamos todas as formas possíveis de conversação: com professoras/es, colegas, grupos, a turma e, também, com a IA. Essa tecnologia não surgiu para isolar pessoas em conversas exclusivas com máquinas, mas para somar; não para simplificar, e sim para enriquecer a experiência de aprendizagem.

Na sala de aula, no entanto, a conversação nem sempre é bem-vinda; muitas vezes, ela é inibida, tratada como perda de tempo ou considerada pouco relevante. A pedagogia tradicional costuma atribuir à/ao docente o papel principal na condução da aula, mantendo-a/o no centro da conversação e deixando pouco tempo para que estudantes se expressem (Carvalho; Pimentel, 2022).

Essa centralidade da fala docente é um resquício do ideal pedagógico da escola moderna que formou muitas pessoas e que ainda se mantém presente em diversas práticas educacionais atuais. O silenciamento de estudantes é promovido quando se presume que elas/es não têm algo relevante a dizer, ou que suas falas atrapalhariam a exposição do conteúdo, causariam dispersão ou comprometeriam o planejamento da aula. A ausência de conversação representa um desperdício da experiência de encontro com a/o outra/o, com seu modo de pensar, de se apresentar, de se constituir, de se compreender e de compreender o mundo, de atribuir sentidos e refletir sobre as coisas.

Queremos desconstruir a ideia de que a conversação é uma “perda de tempo” no processo formativo. Conversar é abrir mão do monologismo e do egocentrismo intelectual para experienciar outras formas de pensar e conhecer. Trata-se de um processo que exige negociação de sentidos, e é justamente essa negociação que conduz a uma aprendizagem mais densa, complexa e significativa.

A conversação é um “modo legítimo de investigação, de relação, porque implica uma forma especial de prestar atenção, de inquietar(se) e indagar(se) a partir da experiência, da vivência, das falas do outro” (Ribeiro; Skliar, 2020, p. 18). Reconhecemos que a IA generativa constitui um Outro, ainda que não humano. Os textos que ela produz resultam do processamento de experiências, saberes e vivências de muitos Outros; ela não gera textos alienígenas, alheios à nossa cultura humana, mas sim conteúdos enraizados nas práticas discursivas e nos sentidos construídos coletivamente ao longo da história.

No ato de conversar, buscamos compartilhar sentidos (Hall, 2016). Nesse processo, manifestamos nossa subjetividade, que é moldada por desejos, experiências, vivências, saberes, conhecimentos, concepções de mundo e posicionamentos. Também deixamos traços e fragmentos de tudo aquilo que aprendemos, dizemos ser, conhecemos e sentimos.

Em uma conversação, manifestamos diferentes posições político-epistemológicas, constituindo um espaço de disputa de narrativas, confronto de ideias e exercício das múltiplas relações de poder-saber. “É um jogo de idas e de vindas entre negociações e buscas por significados e sentidos. Sendo, portanto, um jogo de tensões e de problematizações” (Santos; Carvalho; Maddalena, 2017, p. 202).

Conversar é buscar acordos, pactos e contratos para a construção de uma linguagem singular, criada e compartilhada entre as/os conversantes no contexto de um encontro. “Falar a mesma língua não é apenas uma operação linguística” (Derrida, 2003, p. 115); “falar a mesma língua é partilhar a cultura” (Saraiva, 2005, p. 47). Podemos falar o mesmo idioma, mas isso não garante a compreensão do que a/o outra/o está dizendo; conversar exige tradução, “condição para que eu me relacione com o Outro. [...] Na tradução percebo as minhas afinidades e proximidades com o Outro” (Saraiva, 2005, p. 49). Traduzir não é apenas estabelecer correspondência entre termos; “não buscaria dizer isso ou aquilo, a transportar tal ou tal conteúdo, a comunicar tal carga de sentido, mas a remarcar a afinidade entre línguas” (Derrida, 2002, p. 44). Compreender um/a conversante — seja um ser humano ou maquínico — é uma ação complexa, pois exige ir ao encontro da/o Outra/o.

Conversar requer a prática de uma escuta sensível ao que a/o outra/o expressa. Essa escuta “supõe uma inversão da atenção. Antes de situar uma pessoa no seu ‘lugar’, comecemos por reconhecê-la em seu ser, na sua qualidade de pessoa complexa, dotada de uma liberdade e de uma imaginação criadora” (Barbier, 2007, p. 95). Trata-se de uma escuta que possibilita “enxergar o outro e sua voz sempre grávida de histórias e de sabedoria, de sair um pouco de si para receber e deixar-se banhar pela palavra alheia, as ressonâncias de sua experiência e vivências” (Ribeiro; Skliar, 2020, p. 18).

Ao conversar com outro ser, seja humano ou maquínico, estabelecemos uma ação relacional que demanda cuidado. Por isso, a conversação também deve ser compreendida como uma prática ética. A forma como abordamos determinado assunto configura uma estilística: uma expressão singular, uma estética da linguagem. Já a escolha dos temas que trazemos para a conversa é resultado de uma decisão; trata-se de um ato de curadoria, no qual certos aspectos são visibilizados enquanto outros permanecem invisibilizados — escolha que, evidentemente, não é neutra, mas política. Portanto, conversar é um gesto ético, estético e político (Foucault, 2006).

Durante a conversação, as palavras da/o outra/o têm o potencial de provocar deslocamentos em nosso maquinário subjetivo, acionando gatilhos que fazem pulsar reflexões intensas, estranhamentos e incertezas. Esse movimento, por vezes, nos coloca em suspensão em um espaço desconhecido, indefinível, de inquietude, diante do qual somos levadas/os a parar por um instante. É justamente nesse intervalo que reside a riqueza da conversação: no intercâmbio com a/o outra/o, que nos faz ver de outro modo, para além de nossos próprios referenciais, afetando-nos. Na conversação, “não há como prever o dito e o não-dito, ou seja, pela interação com o outro [...] é que trocaremos-compartilharemos impressões-visões que não podem ser antecipadas” (Couto Junior *et al.*, 2020, p. 116).

A conversação é também um meio de promover a diferenciação, entendida como “o processo central pelo qual a identidade e a diferença são produzidas” (Silva, T., 2000, p. 81). Essa diferenciação ocorre na própria dinâmica da conversação, pois é “por meio de atos de fala que instituímos a identidade e a diferença [...] resultado de um processo de produção simbólica e discursiva” (Silva, T., 2000, p. 77 e 81). Ao enunciarmos nossas identificações e diferenças, expressamos quem somos e o que pensamos, ao mesmo tempo em que refletimos sobre as diferenças na forma como a/o outra/o se apresenta e elabora seu pensamento.

Sem conversação, não há convivência com o pensamento diverso, que é disparador de processos de diferenciação e formação. As diferenciações que emergem da conversação não são fixas nem estáveis; ao contrário, estão em constante construção, desconstrução e reconstrução, constituindo-se nos fluxos e rupturas de nosso maquinário subjetivo. Em determinados momentos, podemos manifestar nossos posicionamentos em relação a algo, depois repensá-los e assumir novas posições. Conversar com um ser diferente de nós, mesmo que seja uma IA generativa, potencializa esse processo de repensar, seja para transformar ou reafirmar nossa forma de pensar.

Qual é o sentido de uma aula quando não há diálogo? Em algumas situações, as aulas ocorrem com pouca ou nenhuma interatividade, como é o caso de muitos cursos massivos online (MOOCs) e de alguns cursos a distância que adotam uma pedagogia instrucionista e automatizada, reduzindo ou mesmo eliminando o diálogo do processo educacional.

Que possamos aprender, com a prática da conversação com a IA generativa, que a formação é dialógica; ou não é formação. Que o diálogo seja cada vez mais valorizado em nossas aulas, sejam elas presenciais, a distância ou semipresenciais. Há, no entanto, um risco importante: que a conversação com a IA se torne tão atraente e confortável que muitas pessoas passem a preferi-la à interação humana, evitando o enfrentamento do pensamento divergente e das diferenças que habitam a sala de aula.

IA generativa como coautora: por uma educação que valorize a autoria

Em tempos de cibercultura, com a liberação do polo de emissão (Lemos, 2007), todas as pessoas se tornaram autoras/es em potencial. Criamos textos, sons, imagens e vídeos que podem ser compartilhados nas *timelines* de nossas redes sociais e em outros espaços digitais. A IA generativa amplia as possibilidades de produção e expressão, mas também impõe novos desafios, como a facilidade de plágio e a geração de informações inadequadas. Diante disso, é urgente (re)pensar a noção de autoralidade das/os estudantes e inventar atividades didáticas que assegurem a centralidade da criatividade humana e do pensamento crítico nos processos de aprendizagem.

Autoria é uma noção plural, aberta a múltiplas significações. Para conceber uma autoria alinhada às práticas (ciber)culturais, partimos das intersecções entre processos formativos, currículo, cultura e diferença, entre outros aspectos (Silva, T., 2000; 2019; Alves, 2003; Macedo, 2007; 2010; 2013; Gallo, 2008). Compreendemos a autoria como um processo de investigação e descoberta de nós mesmas/os, de quem somos e de como desejamos nos posicionar no mundo. Trata-se da expressão, por meio de múltiplas linguagens, de nossa subjetividade, de nossos pensamentos, sentimentos e experiências pessoais. É, também, um processo de autorização e afirmação de si. Quando compreendemos a profundidade da autoria, torna-se inconcebível delegar integralmente a uma máquina o ato de pensar e se expressar em nosso lugar. Ainda assim, a parceria com a IA generativa pode ser bem-vinda como apoio no esforço de comunicar com clareza aquilo que desejamos expressar, desde que o processo preserve também a autoria humana, sem substituí-la pelos conteúdos gerados por máquinas.

No inglês, o verbo *to authore* existe, mas, em português, “autorar” ainda não é oficialmente reconhecido⁷³. Pensar em *autorar* é refletir sobre verbos como criar, intervir, colaborar, interagir, mediar, divulgar, compartilhar e viralizar, que caracterizam a cibercultura, também marcada por verbos como plagiar, mentir, vigiar e controlar. Esses verbos revelam fragmentos de nós mesmas/os, de nossas práticas em rede, de nossos processos formativos, de quem somos, no que nos tornamos e como somos governadas/os.

Manipulamos a palavra autoria não por desrespeito à língua portuguesa, mas pela necessidade de nomear conceitos para os quais ainda não dispomos de vocabulário preciso, ou cujos termos existentes não expressam adequadamente o que desejamos comunicar. Por isso, propomos o neologismo *produsagem* — uma fusão entre autoria e aprendizagem — para designar o processo de autoria voltado à promoção da aprendizagem. Cunhamos esse termo inspiradas/os no conceito de *produsagem*:

“Produsagem” é um neologismo baseado no termo em inglês *produsage* (Bruns, 2008) que, por sua vez, é uma composição de *production* (produção) e *usage* (uso). A produsagem é uma faceta da convergência, aquela pela qual os papéis de autor e audiência, construtor e usuário, produtor e consumidor passam a se confundir e a se misturar, não necessariamente de forma pacífica. (Buzato, 2021, n.p.)

A autoria promove a aprendizagem quando exige do/a autor/a a realização de pesquisas, descobertas, apropriações, ressignificações, reflexões, teorizações e conexões, tecendo saberes com base em suas próprias experiências para produzir e negociar sentidos. Pensar a autoria como um processo formativo é refletir sobre a educação contemporânea e criar oportunidades para que estudantes aprendam a conhecer, a fazer, a conviver e a ser. Entre esses pilares da educação do século XXI, escolas e universidades tendem a enfatizar o aprender a conhecer e, às vezes, o aprender a fazer (Delors *et al.*, 1998). No entanto, enquanto docentes, nosso

⁷³ Não consta no Vocabulário Ortográfico da Língua Portuguesa (VOLP): <https://www.academia.org.br/nossa-lingua/busca-no-vocabulario>

papel não pode se restringir ao ensino de conhecimentos científicos e sua aplicação prática; precisamos também criar condições para que estudantes aprendam a ser e a conviver, aprendizagens necessárias para a vida em sociedade.

A autoria proporciona a aprendizagem de novas informações e sua aplicação na criação de uma obra. Por meio de suas produções, a/o estudante expressa desejos, experiências, medos, sonhos e posicionamentos. Nesse movimento, reflete sobre si mesma/o, compreende melhor suas diferenças, identificações, culturas, crenças, relações, desejos e valores éticos — o que se desdobra no aprender a ser. Ao compartilhar suas expressões com outras pessoas ou ao realizar produções em coautoria, a/o estudante negocia sentidos, assume corresponsabilidades, compromete-se com o coletivo e com objetivos comuns, aprende a mediar conflitos e a lidar com intersubjetividades — experiências que contribuem para o aprender a conviver.

É útil diferenciar autoria de pré-autoria. A pré-autoria ocorre quando há concordância ou reprodução de ideias já existentes: “não é uma cópia, mas uma espécie de releitura de uma obra” (Backes, 2012, p. 80). Em contrapartida, a autoria pode ser transformadora-crítica ou criadora. A autoria transformadora-crítica implica o exame reflexivo e questionador dos conhecimentos construídos; já a autoria criadora envolve a ressignificação do saber aprendido, mobilizando a criatividade para fazer emergir novos conhecimentos. Essas categorias nos ajudam a compreender que, se a/o estudante realiza uma atividade apenas reinterpretando ideias presentes em livros, videoaulas, conteúdos da internet ou textos gerados por IA generativa, ainda que utilize suas próprias palavras, essa produção não se configura como autoria, mas como pré-autoria, pois reproduz conhecimento preexistente sem criticá-lo, transformá-lo ou gerar algo novo.

Com a chegada da IA generativa, tornou-se ainda mais problemático propor atividades que exijam apenas pré-autoria. Quando uma pessoa delega inteiramente a criação a uma tecnologia generativa, sem atuar como (co)autora da obra, está abrindo mão do processo formativo que a autoria pode envolver. Com nossas/os estudantes, preferimos trabalhar a autoria como um processo de autorização, no qual elas/es se autorizam a se posicionar criticamente, a expressar seus conhecimentos e subjetividades, e

a se responsabilizar pelo que produzem. Trabalhamos com a ideia de criação assistida por IA e autoria híbrida, que podem desencadear processos formativos intensos. Não há autoria quando uma atividade exige apenas uma resposta única, previamente determinada pela/o docente, e a/o estudante precisa apenas reproduzir o que já existe. Quando uma atividade não promove autoria, ela tende a facilitar e até incentivar o plágio.

Elaborar situações de aprendizagem que possibilitem a autoria por parte da/o estudante não é uma tarefa simples, especialmente porque estamos habituadas/os a propor atividades baseadas na repetição de conteúdos do livro didático ou na resolução de problemas seguindo um conjunto pré-determinado de etapas. Esse tipo de atividade, por sua estrutura padronizada, pode ser facilmente realizado por uma IA generativa.

Para propor situações que promovam a formulação de novos saberes, devemos elaborar atividades em que a/o estudante se sinta autorizada/o a “narrar a vida”, inserindo sempre o fio do seu próprio modo de significar e contar (Alves, 2001). Nós, docentes, podemos mobilizar nossa autoria pedagógica para arquitetar, de forma criativa, situações de aprendizagem que incentivem estudantes a refletir, posicionar-se, criar, realizar uma autoria de si (Maddalena, 2020), narrar o próprio cotidiano, explicitar sua compreensão do novo a partir da relação com conhecimentos prévios e experiências pessoais, (re)significar à sua maneira, de seu modo singular. Nenhuma dessas ações pode ser realizada pela IA em lugar da pessoa aprendente.

Defendemos a necessidade de superar a concepção de estudante-espectador/a de aulas expositivas, substituindo-a pela concepção de estudante-autor/a. Quando falamos em autoria estudantil, estamos nos referindo à formação voltada para o pensamento crítico, a criatividade e a construção de valores éticos. Acreditamos que esse seja um dos grandes desafios da contemporaneidade, pois exige transformações nas concepções pedagógicas e nas práticas docentes. Trata-se de reconfigurar as relações de poder-saber no interior das aulas, de repensar o significado de educar e de rever nosso papel como formadoras/es em tempos de inteligência artificial generativa.

Para alcançar esse objetivo, devemos criar oportunidades de expressão, promover a publicação das autorias e sua apresentação para a turma; se a turma desejar, podemos organizar um acervo para publicação na internet. Nesse contexto, nosso papel inclui a concepção das situações de *autoragem* e ações de mediação, entre elas, destacamos:

- praticar uma escuta atenta e sensível, buscando compreender o que as/os estudantes têm a dizer, como estão significando, negociando, refutando e tecendo conhecimentos;
- analisar os conteúdos elaborados pelas/os estudantes e o que expressam, verificando se estão alinhados à proposta, se apresentam problemas conceituais ou práticos, se há indícios de plágio ou se contêm violações a princípios éticos, como discursos machistas, racistas ou LGBT+fóbicos;
- comentar cada autoria, contrapondo ideias, trazendo diferentes perspectivas e promovendo tanto a valorização quanto o tensionamento crítico das posições assumidas pelas/os estudantes;
- incentivar a participação das/os colegas, convidando-as/os a comentar e compartilhar outros pontos de vista, promovendo a aprendizagem colaborativa e a escuta sensível;
- fortalecer os laços sociais e mediar eventuais conflitos interpessoais que possam emergir nos processos coletivos de criação;
- evitar avaliar a autoria como certa ou errada, uma vez que se trata de uma expressão singular, e não da reprodução de conhecimentos estabelecidos que sirvam de parâmetros para uma correção. Caso a avaliação seja necessária, é possível construir, em diálogo com a turma, rubricas avaliativas com critérios de qualidade e níveis de realização. Isso facilita a compreensão das notas, amplia a prática da autoavaliação e possibilita a avaliação colaborativa. Pode ser interessante permitir que o/a autor/a revise a obra após o *feedback*;
- problematizar o uso das tecnologias digitais em rede, especialmente das tecnologias generativas, analisando se estão sendo utilizadas para fins de plágio ou para processos potentes de autoria híbrida entre humano-IA.

Compreender a IA generativa como uma coautora em potencial não significa abrir mão da autoria humana. Significa, acima de tudo, reconhecer

que o conceito de autoria tornou-se mais complexo, abrindo espaço para a colaboração com a IA sem renunciar à criatividade, à criticidade e à ética. Ao desenharmos situações de aprendizagem que incentivem as/os estudantes a se autorizarem a refletir, criar e atribuir sentidos, devemos considerar a IA como uma parceira no processo, e não como substituta da autoria humana.

Nos cursos de formação continuada, ao incentivarmos docentes a projetar atividades que envolvam a criação com apoio da IA, frequentemente nos deparamos com propostas surpreendentes e inovadoras. Novas práticas de *autoragem* híbrida estudante-IA estão sendo forjadas, apontando caminhos promissores para a educação contemporânea.

Que possamos aprender, com a prática de coautoria com a IA generativa, que formação é um processo de autorização; ou não é, verdadeiramente, formação. Que a autoria, enquanto exercício de expressão singular e situada, seja cada vez mais valorizada em nossas aulas, em todas as modalidades de ensino.

No entanto, há um risco que não pode ser ignorado: os conteúdos gerados pela IA podem se tornar tão sofisticados e sedutores que algumas pessoas passem a se sentir diminuídas diante da capacidade da máquina. Em vez de se reconhecerem como cocriadoras/es, podem ceder à tentação de delegar integralmente a criação à tecnologia, desvalorizando sua própria potência expressiva e intelectual. Mais do que nunca, é necessário reafirmar: a IA não pode substituir a autoria estudantil. Ela pode, sim, ser uma aliada na ampliação da criatividade e no enriquecimento das expressões, desde que o processo formativo reafirme, de forma contínua, o lugar das/os estudantes como autoras/es de seu próprio pensamento, de sua expressão e de sua existência.

IA generativa como máquina de ensinar: o fim da docência humana?

Para o fundador da Khan Academy, Salman Khan (2023, 0:33 – 2:01s), a IA é uma professora particular “extraordinária” que transformará a educação:

Estamos no limite de usar a IA para, provavelmente, a maior transformação positiva que a educação já viu. A forma como estamos fazendo isso é dando, a cada estudante do planeta, uma inteligência artificial, que é uma professora particular extraordinária. [...] Isso poderia transformar um aluno mediano em um aluno excepcional, pode transformar um estudante abaixo da média em um estudante acima da média.

Khan fundamenta sua compreensão nos resultados da pesquisa de Bloom (1984), que comparou o aprendizado de estudantes em condições distintas: (i) ensino convencional, no qual um/a professor/a ensina cerca de 30 estudantes; e (ii) ensino particular, em que um/a professor/a ensina individualmente um/a único/a estudante. As/Os estudantes que aprenderam pelo ensino particular apresentaram notas mais altas, estudaram por mais tempo e demonstraram mais interesse.

Sabemos que um/a professor/a particular pode ser um/a aliado/a importante no processo de aprendizagem, mas o custo da hora-aula é um obstáculo para a maioria das/os estudantes brasileiras/os. Se a IA generativa puder funcionar como uma espécie de professora particular gratuita, como suposto por Khan, veremos uma grande melhoria no nível de aprendizado das/os estudantes nos próximos anos, em diferentes níveis e modalidades de educação... ou não?

Nós, autores deste livro, não compreendemos a IA generativa como uma professora. Apesar de, em alguns casos, seu tom acadêmico remeter à prática docente, ela não possui intencionalidade pedagógica, não tem um currículo formal a cumprir, não segue um plano de ensino, não propõe conteúdos de estudo, não realiza encontros regulares em forma de aulas, não exige atenção, não busca engajar a/o usuária/o em debates e na resolução de problemas, não passa dever de casa, não aplica provas, não faz avaliações, não monitora o progresso da/o estudante, não certifica a aprendizagem. Em vez de uma professora, a IA generativa se assemelha mais a uma pessoa bem-informada com quem podemos conversar sobre diversos temas.

As vantagens da IA generativa só podem ser aproveitadas por quem estiver disposta/o a interagir de forma ativa e reflexiva com ela. Isso é diferente de estar presente em uma aula, seja particular ou coletiva, e permanecer em silêncio ou sem participação. A IA não responde à passividade: ela não é uma “máquina de ensinar” porque não foi projetada para essa finalidade. No entanto, pode ser utilizada com intencionalidade pedagógica desde que alguém a customize para esse propósito.

20.1 – IA generativa em máquinas de ensinar

Máquina de Ensinar é um dispositivo projetado para implementar a instrução programada, conceito que se popularizou na década de 1950 com o psicólogo comportamentalista B. F. Skinner⁷⁴ (1954). As máquinas de ensinar daquela época eram mecânicas: o conteúdo era dividido em pequenos blocos, de modo que a/o estudante lia um texto curto ou uma pergunta por vez. Ao girar uma manivela, a máquina apresentava uma

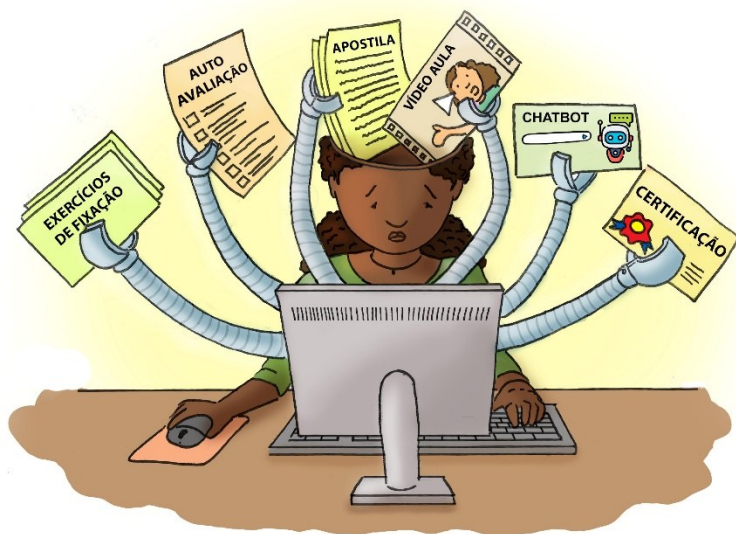
⁷⁴ Embora Skinner tenha popularizado as máquinas de ensinar, ele não foi seu inventor. As primeiras máquinas de ensinar foram desenvolvidas na década de 1920 pelo psicólogo Sidney Pressey para possibilitar a autoavaliação por meio de questões de múltipla escolha. Essas máquinas serviram de base para o posterior desenvolvimento realizado por Skinner (1958; 1972).

questão, e a/o estudante deveria escolher a alternativa correta ou escrever sua resposta, geralmente uma palavra ou um número. Em seguida, a máquina exibia a resposta certa, possibilitando à/ao usuária/o conferir se havia respondido corretamente, recebendo, assim, um *feedback* imediato. O aprendizado ocorria em uma sequência previamente determinada pelo criador do programa de autoinstrução, e a máquina condicionava o comportamento da/o estudante por meio de reforço positivo imediato.

Embora a IA generativa, por si só, não seja uma máquina de ensinar, ela pode servir de base para a construção desse tipo de sistema. Um modelo de linguagem pode ser integrado a um sistema computacional e funcionar como o “motor linguístico” de uma máquina de ensinar. Se é possível instruir a IA generativa para desempenhar diferentes papéis, também podemos configurá-la para operar como uma máquina de ensinar, dando instruções para seguir um planejamento didático, gerando conteúdos, tirando dúvidas da/o estudante, elaborando exercícios, corrigindo as respostas e avaliando o progresso da aprendizagem.

Há mais de meio século, Paulo Freire denunciava a educação bancária, na qual “o educador faz ‘comunicados’ e depósitos que os educandos, meras incidências, recebem pacientemente, memorizam e repetem” (Freire, 1970, p. 33). Essa pedagogia instrucionista não foi superada; ao contrário, continuou sendo praticada e se atualizando, chegando agora em sua versão “ciber”. Nessa nova versão, em vez de professoras/es realizarem comunicados e depósitos nas/os educandas/os, como ocorria na época de Freire, são as tecnologias digitais que, cada vez mais, assumem esse papel. Obviamente, a IA generativa será apropriada pelo movimento que denominamos *cibertecnicismo* (Pimentel; Carvalho, 2021; 2022), ou neotecnicismo digital (Freitas, 2021), uma abordagem didático-pedagógica massivo-instrucionista que compreendemos como a versão contemporânea do tecnicismo (Saviani, 1986, 1999) e do neotecnicismo (Freitas, 1992; 2011a; 2011b; 2018).

Cibertecnicismo: uma nova versão da educação bancária



Fonte: dos autores, ilustração de Mônica Lopes

O radical “ciber” tem origem na raiz do termo grego *kyber*, que significa “governar, dirigir, conduzir”. O termo cibernética foi cunhado por Norbert Wiener em 1948 para nomear os sistemas que se autorregulam, sejam sistemas vivos ou máquinas:

Decidimos chamar todo o campo da teoria de controle e comunicação, seja na máquina ou no animal, pelo nome de Cibernética, que deriva do grego *κυβερνήτης* ou timoneiro. Ao escolher este termo, desejamos reconhecer que o primeiro artigo significativo sobre mecanismos de *feedback* é um artigo sobre governantes, publicado por Clerk Maxwell em 1868 e que governante deriva de uma corruptela latina de *κυβερνήτης*. Também desejamos nos referir ao fato de que os mecanismos de direcionamento de um navio [que atuam sobre o timão, ou leme, em termos técnicos] são, de fato, uma das formas mais antigas e mais bem desenvolvidas de mecanismos de *feedback*. (Wiener, 2019, p. 18, tradução nossa)

Wiener evocou a imagem do timoneiro: alguém ou algo que mantém o rumo ajustando continuamente em função dos *feedbacks* do ambiente. Essa metáfora foi utilizada para descrever o controle de sistemas autônomos, que se autorregulam, assim como um barco que chega ao seu destino por meio do piloto que vai ajustando constantemente a sua rota.

Um exemplo contemporâneo do conceito de cibernética é o carro autônomo: um sistema fechado que opera sem intervenção humana direta. Equipado com IA, o veículo ajusta direção e velocidade para alcançar seu destino, considerando tanto fatores previsíveis (como a rota no GPS) quanto imprevisíveis (trânsito, obstáculos, semáforos, limites de velocidade e variações no terreno).

Embora o termo cibernética tenha sido originalmente utilizado para nomear a área de estudo da autorregulação em sistemas dinâmicos, ao longo do tempo, o radical “ciber” passou por um processo de ressignificação. Ele passou a ser empregado como sinônimo de computadores, tecnologias digitais e internet, distanciando-se de seu sentido original ligado à Teoria do Controle (como retroalimentação ou *feedback*). Hoje, diversos termos com esse prefixo estão registrados no Vocabulário Ortográfico da Língua Portuguesa (Volp)⁷⁵, como ciberespaço, cibernauta, ciberativismo, cibercrime, cibersegurança e ciberjornalismo. É nessa acepção ampliada e atualizada que utilizamos o radical “ciber” para qualificar a versão contemporânea do tecnicismo.

Por cibertecnicismo, referimo-nos a uma abordagem didático-pedagógica instrucionista na qual os processos de ensino são parcial ou totalmente automatizados por meio de técnicas e tecnologias digitais, incluindo: plataformização, dataficação, gamificação, youtuberização, algoritmização, inteligência artificial, *chatbots*, tutores inteligentes, entre outras. Nessa abordagem, assume-se que computador e estudante formam um sistema fechado capaz de aprender sem a presença de professoras/es ou colegas de turma, como se fossem o barco e timoneiro, um sistema capaz de se autorregular.

O cibertecnicismo preserva os princípios fundamentais da racionalidade técnica e científica que caracterizaram o tecnicismo clássico,

⁷⁵ <https://www.academia.org.br/nossa-lingua/busca-no-vocabulario>

agora reformulado pelo uso das tecnologias digitais em rede, com vistas à automatização dos processos educacionais. Trata-se da efetivação da “arte de ensinar sem professoras/es”, com mínima ou nenhuma mediação humana. É evidente que a IA generativa, por sua capacidade conversacional e de gerar conteúdos e responder dúvidas, constitui-se como uma aliada potente desse projeto de instrução automatizada.

Nas políticas oficiais, esse tecnicismo disfarçado de inovação já é uma presença que é comemorada pelo hegemônico neste momento. [...] Não se aprende e nem se forma com um cérebro separado de uma história de vida, de trabalho e de cultura. Não queremos isso. [...] É impensável a formação como produção em série. É uma impossibilidade pensá-la e alcançá-la por indicadores extensivos, dimensões aferidas e standardizações. Ou é criação sociotécnica, ética, estética, política, cultural, experiencial e acontecimental, ou então não é formação (Macedo, 2020, 5:12s).

O avanço do cibertecnicismo está diretamente relacionado à mercantilização da educação, que passou a ser concebida como um negócio lucrativo (Freitas, 2018), e se torna ainda mais lucrativo à medida que se reduzem os “gastos” com professoras e professores. Para viabilizar a automação do ensino sem a mediação docente, desenvolveu-se a técnica de autoinstrução conhecida como instrução programada. Nesse método didático, o conteúdo é organizado em uma sequência cuidadosamente estruturada por um/a designer instrucional. Após a apresentação de cada segmento de conteúdo, aplica-se uma atividade de (auto)avaliação, com o objetivo de verificar se a/o estudante assimilou as informações. Em geral, essa avaliação limita-se a checar se a/o estudante é capaz de reproduzir uma informação apresentada no texto; caso tenha dúvidas, deve reler o material. A educação, nesse modelo, é reduzida à simples assimilação de informações descontextualizadas, sem a presença de professoras/es e colegas que possibilitem questionar, problematizar e negociar sentidos.

Muitos cursos a distância mediados por redes digitais herdaram as práticas massivo-instrucionistas de ensino promovidas por gerações anteriores de Educação a Distância, viabilizadas por meios de comunicação de massa como o rádio e a televisão. O ensino em escala industrial levou à fragmentação e especialização do trabalho docente, instituindo uma

hierarquia funcional entre diferentes papéis: professor-conteudista, professor-coordenador, professor-mediador, professor-tutor etc. Essa organização é caracterizada como “uma forma industrial de educação”, baseada na racionalização, na divisão do trabalho e na produção em massa, em que o processo de ensino é gradualmente reestruturado pela crescente mecanização e automação (Belloni, 2012, p. 7-8).

A mecanização e automação do ensino se intensificaram com o desenvolvimento das plataformas *online* de educação. Os Cursos *Online* Abertos e Massivos (MOOCs), lançados no final dos anos 2000 e amplamente difundidos na década de 2010, levaram essa lógica de automação do ensino ao extremo. Plataformas como Khan Academy, Coursera, Udacity e edX foram criadas para hospedar e comercializar cursos e disciplinas *online*, nem sempre de acesso aberto, apesar do discurso de democratização. Tais cursos são considerados “massivos” porque se destinam a ensinar um mesmo conteúdo a um número muito elevado de estudantes, em escala global. O entusiasmo em torno desses cursos, produzidos por universidades de prestígio internacional — como MIT, Oxford, Stanford, Cambridge, Harvard, entre outras — levou algumas pessoas a projetar um futuro em que, após um período de competição, restariam apenas dois ou três MOOCs por área de conhecimento, controlados por um número reduzido de universidades globais, operando com estratégias neocolonizadoras.

Reconhecemos que o problema não está na diversidade de modalidades educacionais, mas no fato de essas modalidades, muitas vezes, se tornarem a única opção disponível para a/o estudante. Isso é especialmente preocupante quando se considera que muitas/os estudantes não as preferem. Nos Estados Unidos, por exemplo, “apenas de 5 a 10% dos estudantes universitários preferem Cursos Online Massivos e Abertos” (Vicari, 2021, n.p.). Além disso, é amplamente conhecida a altíssima taxa de evasão nesses cursos, que pode chegar a 95% (Silva *et al.*, 2014). No Brasil, a maioria das/os estudantes universitárias/os está em cursos na modalidade a distância (Brasil, 2025). No entanto, essa escolha nem sempre se dá pelas facilidades oferecidas, como o acesso remoto, que elimina a necessidade de deslocamento até o campus (frequentemente localizado a muitos quilômetros da residência da/o estudante), ou o caráter assíncrono, que

possibilita estudar fora dos horários convencionais. Para muitas/os, trata-se de uma decisão imposta por barreiras econômicas e de acesso: de um lado, a escassez de vagas nas universidades públicas gratuitas, extremamente concorridas; de outro, as altas mensalidades dos cursos presenciais em instituições privadas. Nesse contexto, um curso a distância com mensalidade reduzida torna-se a única alternativa economicamente viável; nem sempre é uma escolha baseada em preferência, mas uma imposição das desigualdades estruturais que afetam o ingresso e a permanência no ensino superior presencial.

A financeirização da educação superior possibilitou a consolidação de um novo “modelo de negócio” (Amaral; Vergara, 2011), no qual estudantes são tratadas/os como consumidoras/es que escolhem e contratam cursos entre centenas de ofertas disponíveis no mercado educacional. Esse modelo tem favorecido a expansão de grandes conglomerados privados de EaD, incluindo empresas internacionais, no sistema de ensino superior brasileiro.

Com o crescimento da modalidade a distância, tornou-se necessário desenvolver estratégias para manter estudantes engajadas/os, mesmo na ausência de acompanhamento direto por parte de professoras/es e colegas. Nesse contexto, a gamificação emergiu como um recurso central para a automação do ensino. Pressupõe-se que a incorporação de dinâmicas, linguagens e estéticas dos jogos, como sistemas de pontuação, *rankings*, níveis e medalhas, possa aumentar a motivação e o envolvimento das/os estudantes no processo de aprendizagem (Oliveira *et al.*, 2021). Essas técnicas rapidamente se disseminaram nos sistemas computacionais voltados à educação, chegando às plataformas de ensino-aprendizagem. Algumas instituições chegam a promover seus cursos enaltecendo o uso de um “AVA gamificado, sem tutor e autoinstrucional”, vangloriando-se da consolidação de um modelo de ensino alinhado ao cibertecnicismo.

Essas plataformas de aprendizagem não apenas medeiam o processo educacional, mas também monitoram e coletam dados sobre as/os estudantes, registrando cada ação realizada. Elas rastreiam, por exemplo, os dias e horários de acesso, o tempo de permanência na plataforma, os conteúdos visualizados e sua duração na tela, as atividades avaliativas realizadas, as questões corretas e incorretas, a sequência de cliques, o número de mensagens enviadas e até o conteúdo textual de cada interação.

Esse vasto volume de dados é processado e analisado para gerar indicadores baseados em comparações com o desempenho de outras/os estudantes da turma e com os dados do histórico de uso da própria plataforma. Técnicas de inteligência artificial são então empregadas para construir modelos preditivos capazes de inferir comportamentos: se a/o estudante está achando o curso muito difícil ou muito fácil, se há risco de evasão ou se demonstra sinais de engajamento. Com base nesses indicadores, o sistema toma decisões automatizadas, como enviar um e-mail de apoio, sugerir conteúdos e atividades complementares, recomendar o contato com o/a tutor/a ou, simplesmente, não intervir.

A centralidade da análise de dados em nossa sociedade hiperconectada deu origem a uma nova área de conhecimento: a Ciência de Dados. Uma de suas especializações voltadas à educação é a Análise de Aprendizagem (*Learning Analytics*), que se ocupa do exame sistemático dos dados educacionais gerados pelas/os estudantes. Esses dados são coletados, armazenados, tabulados, minerados, comparados, transformados, estratificados, reutilizados e, em muitos casos, vendidos. As plataformas educacionais incorporam essa lógica, dataficando o processo de ensino-aprendizagem de maneira muito mais elaborada do que as tradicionais planilhas de frequência e de notas que nós, professoras/es, estávamos acostumadas/os a gerenciar.

Diversas escolas e universidades brasileiras, incluindo instituições públicas, têm adotado plataformas de grandes corporações estrangeiras de tecnologia, como o Google Sala de Aula e o Microsoft Teams, além de aplicativos como WhatsApp e Instagram, pertencentes à Meta. Essas empresas passaram a mediar o cotidiano educacional, coletando dados de professoras/es e estudantes, influenciando práticas pedagógicas e ampliando sua presença global ao nos tornarem dependentes de suas soluções tecnológicas desde os primeiros anos de escolarização.

Sem uma discussão crítica e uma formação docente adequada, corremos o risco de o cibertecnicismo se disseminar mais amplamente. Um exemplo é o uso do Google Formulários integrado ao Google Sala de Aula, que possibilita a criação de questões de múltipla escolha com correção automática e registro imediato das notas na planilha da turma. A facilidade de disponibilizar conteúdos e automatizar atividades avaliativas possibilita

que qualquer professor/a, mesmo sem formação em design instrucional ou programação, crie sua própria “máquina de ensinar”. Trata-se da banalização da instrução programada em larga escala, impulsionada por plataformas de uso massivo.

As grandes empresas de tecnologia disputam a colonização digital do sistema educacional. A empresa Google, por exemplo, não apenas ofereceu gratuitamente suas tecnologias digitais às instituições de ensino brasileiras durante a pandemia, como também implementou programas de capacitação docente, transformando professoras e professores em multiplicadoras/es de suas soluções. Muitas/os educadoras/es, orgulhosas/os do selo de “*Google Certified Trainer*”, compartilham dicas e tutoriais sobre o uso pedagógico do Google Sala de Aula em canais do YouTube e nas redes sociais.

Se, por ora, a substituição total de professoras/es ainda não é tecnicamente viável, essas empresas apostam em outras estratégias: aliciar, agenciar e formar docentes para que se tornem usuárias/os fidelizadas/os de suas plataformas. E, enquanto a automação completa do ensino não se concretiza, seguem os esforços para expandir a automatização da arte de ensinar sem professoras/es por meio das tecnologias digitais em rede. A inteligência artificial generativa representa a mais recente e poderosa tecnologia a ser apropriada por esse movimento.

O ideal de uma educação massificada não é novo. Um exemplo emblemático pode ser encontrado em Comenius, considerado o pai da didática moderna, que, ao criticar o sistema feudal de ensino, publicou *Didática Magna* em 1627 e 1638, defendendo uma “arte universal de ensinar tudo a todos”:

Didática significa arte de ensinar. [...] Nós ousamos prometer uma *Didática Magna*, isto é, um método universal de ensinar tudo a todos. E de ensinar com tal certeza, que seja impossível não conseguir bons resultados. E de ensinar *rapidamente*, ou seja, sem nenhum enfado e sem nenhum aborrecimento para os alunos e para os professores, mas antes com sumo prazer para uns e para outros. E de ensinar *solidamente*, não superficialmente e apenas com palavras, mas encaminhando os alunos para uma verdadeira instrução (Comenius, 2001 [1638], p.3-4).

O problema, obviamente, não está no ideal de democratização da educação, mas na operacionalização desse objetivo. Comenius propunha uma organização lógica do conhecimento, de tal forma que qualquer assunto pudesse ser aprendido por qualquer pessoa por meio de um método de ensino que ele acreditava ser eficaz e eficiente, em que um/a único/a professor/a leciona para centenas de estudantes:

De tal maneira que um só professor seja suficiente para instruir, ao mesmo tempo, centenas de alunos, com um esforço dez vezes menor que aquele que atualmente costuma dispende-se para ensinar cada um dos alunos. [...] O professor, com os mesmos exercícios, pode, ao mesmo tempo e de uma só vez, ministrar o ensino a uma multidão de alunos sem qualquer incômodo (Comenius, 2001 [1638], p. 46, 96).

A Didática Magna representou uma nova abordagem pedagógica ao propor um modelo de educação que, hoje, poderíamos rotular como massificada. Essa proposta também implicava, ainda que de forma não nomeada na época, o que atualmente reconhecemos como precarização do trabalho docente (Sampaio; Marin, 2004; Lapa; Pretto, 2010).

A ideia de uma educação massificada ganhou força com os processos de automação generalizada decorrentes da Revolução Industrial, em que os métodos de produção artesanal foram gradualmente substituídos por processos mecanizados, movidos a energia a vapor e, posteriormente, elétrica. Desde o século XIX, diversos esforços foram realizados no sentido de criar uma máquina capaz de automatizar o ensino por meio da instrução programada. O objetivo era desenvolver um sistema que apresentasse o conteúdo em pequenas unidades testáveis, oferecesse feedback imediato após a realização de um teste e possibilitasse à/ao estudante aprender no próprio ritmo. Esses esforços culminaram no desenvolvimento das máquinas mecânicas de ensinar, que viabilizaram, pela primeira vez, a automação da instrução em escala.

Paralelamente às máquinas instrucionais, a massificação do ensino encontrou outra aliada poderosa nas chamadas máquinas sensórias (Santaella, 1997), como a fotografia, o rádio, o cinema, o telefone e a televisão. Esses dispositivos permitiram a reprodução de diferentes formas de signos e ampliaram o alcance e as formas de mediação do conhecimento.

Foi nesse contexto que emergiu a tele-educação, modelo que expandiu as possibilidades da educação massiva para além do que o livro impresso já permitia.

A Revolução Industrial, os meios de comunicação de massa, os avanços nas técnicas de autoinstrução e, mais recentemente, as tecnologias digitais em rede, impulsionaram o que hoje chamamos de versão contemporânea da instrução programada. Trata-se do resultado de uma rede histórica de desejos, tecnologias e desenvolvimentos que a sociedade vem perseguindo há séculos. “Se a automação computadorizada aumentou a produção nas fábricas, então o mesmo pode ocorrer no ensino superior” (Plato, s.d., n.p., tradução nossa).

A pedagogia (ciber)tecnicista deve ser compreendida também a partir do desejo histórico de cientificar a educação, por meio da adoção de abordagens mensuráveis, controláveis e alinhadas ao paradigma experimental-positivista, que ganhou força na primeira metade do século XX. É nesse contexto ideológico e científico que emerge a área do currículo, a qual deu sustentação à concepção tecnicista de educação (Silva, T., 1999). Fundamentada em uma racionalidade liberal-científica, essa perspectiva valorizava métodos, técnicas e tecnologias como pilares da prática educativa, consolidando-se na forma de uma pedagogia tecnicista orientada pelos seguintes princípios:

- a aprendizagem escolar deve se restringir aos conteúdos científicos;
- o papel da escola é instruir e preparar a/o estudante para o mundo do trabalho;
- o foco deve recair sobre métodos e tecnologias que tornem o ensino mais eficiente e eficaz;
- o/a professor/a deve atuar como mero/a executor/a de currículos concebidos por especialistas.

Essa abordagem não ficou restrita às décadas de 1960, 1970 e 1980; ao contrário, continuou a ser apropriada por gestoras/es, professoras/es e alunas/os, reconfigurando-se, a partir da década de 1990, sob a forma da pedagogia neotecnicista, impulsionada pela racionalidade neoliberal (Freitas, 2011a). A pedagogia neotecnicista não é apenas uma reformulação

do tecnicismo original, mas é também um alerta para a retomada e o fortalecimento de sua lógica sob novas roupagens, agora marcadas pelo discurso da eficiência, da mensuração e da responsabilização. Essa nova versão mantém a padronização dos processos educativos e aprofunda os mecanismos de controle, sobretudo por meio da “predominância da avaliação externa de larga escala, implementada a partir de uma teoria em que o professor e a escola eram responsabilizados de fora do sistema educacional pelos resultados de aprendizagem” (Freitas, 2021, n.p.). Trata-se, portanto, de uma pedagogia que combina tecnicismo, neoliberalismo e cultura da avaliação, e que encontra no cibertecnicismo sua expressão atualizada e intensificada pelas tecnologias digitais e pelas promessas de automação do ensino.

Com a popularização das tecnologias digitais em rede a partir dos anos 2000, o tecnicismo ingressou em uma nova fase: a da pedagogia neotecnicista digital, ou, como propomos neste trabalho, pedagogia cibertecnicista. Essa abordagem representa um “novo estágio de controle do magistério e dos estudantes que vem na onda das tecnologias digitais. Ele combina demandas [...], métodos de ensino e processos de avaliação – tudo embarcado em plataformas de aprendizagem – híbridas ou não” (Freitas, 2021, n.p.). Um de seus principais objetivos é formar sujeitos capazes de gerir a si mesmos, operando como administradores de sua própria trajetória de aprendizagem, acumulando competências e habilidades exigidas pelo mercado de trabalho. Trata-se de uma pedagogia cada vez mais presente no cotidiano educacional e que precisa ser criticamente interrogada: que subjetividades ela visa produzir? Qual é seu ideal de sujeito e de sociedade? A quem ela serve na construção desse ideal? De que forma ela impacta os processos formativos contemporâneos?

Não estamos fazendo um ataque ao uso das tecnologias computacionais no contexto educacional. Tampouco se trata de uma crítica à Educação a Distância em si, pois sabemos que outras abordagens didático-pedagógicas são possíveis nessa modalidade, como a Educação *Online*, que temos praticado e sobre a qual temos teorizado (Pimentel; Carvalho, 2020). Nosso objetivo é chamar a atenção para o fato de que as máquinas de ensinar não pertencem ao passado. Ao contrário: elas vêm sendo constantemente atualizadas, acompanhando as inovações tecno-

pedagógicas, passando por sucessivas reformulações numa busca contínua por aperfeiçoar a arte de ensinar sem professoras/es. A busca pela massificação e automação da instrução não perdeu força. A IA generativa é a nova tecnologia que irá turbinar o movimento cibertecnista.

20.2 – Professoras/es serão substituídas/os por IA?

Um questionamento que nos assombra há várias décadas e que foi revigorado com o lançamento do ChatGPT: nós, professoras e professores, seremos substituídas/os pelas tecnologias digitais? O fato é que estão emergindo modelos de escola e universidade sem professoras/es, substituindo docentes por IA generativa:

- *“O currículo escolar do Arizona será ensinado por IA, sem professores”* (Feathers, 2024). Em uma escola dos EUA, as aulas serão conduzidas integralmente por IA, com base em programas de aprendizagem personalizada de empresas como Khan Academy e IXL. Nesse modelo de escola, da Unbound Academy, não há professoras/es, mas sim mentoras/es e treinadoras/es de aprendizagem (*coaches*).

Nossos algoritmos de IA criam trajetórias de aprendizagem personalizadas, adaptadas às necessidades, habilidades e progresso individuais de cada estudante. A tecnologia avalia continuamente o desempenho, o engajamento e o estado emocional do aluno para otimizar, em tempo real, o conteúdo, o nível de dificuldade e o ritmo das lições. Essa abordagem garante que os estudantes trabalhem sempre em sua zona de desenvolvimento ideal, recebendo material desafiador na medida certa para manter a motivação e permitindo que aprendam até 2,3 vezes mais rápido do que em escolas tradicionais. Por meio de análises detalhadas e monitoramento contínuo, a IA identifica rapidamente áreas em que o estudante pode precisar de suporte adicional ou onde pode ser desafiado ainda mais para aprimorar sua experiência de aprendizagem. Se um aluno demonstrar dificuldades ou necessitar de material mais avançado, a IA adapta as lições instantaneamente, garantindo suporte e desafios em tempo real. (Unbound Academic Institute, 2024, p. 23, tradução nossa)

- *“Inteligência Artificial: escola no Reino Unido inicia turma sem professores”* (Nabuco, 2024). Uma escola particular em Londres iniciou uma turma sem professores, “com todas as disciplinas principais sendo ensinadas inteiramente por plataformas de aprendizagem adaptativa orientadas por IA” (David Game College, 2024, s.p., tradução nossa). Professoras/es ficam em segundo plano, atuando como treinadoras/es de aprendizagem (*learning coaches*), dando suporte quando necessário e ensinando temas polêmicos ou que a IA apresenta muitas limitações, como educação sexual e aulas de arte.
- *“Professores gerados por inteligência artificial dão aulas em universidade de Hong Kong”* (Exame, 2024). A Universidade de Ciência e Tecnologia de Hong Kong (HKUST) está testando professoras/es virtuais baseadas/os em IA. No curso Redes Sociais para Criativos, cerca de 30 estudantes assistem a aulas realizadas por avatares digitais, incluindo uma versão virtual de Albert Einstein. Esses avatares são personalizáveis e podem ser exibidos em telas ou em realidade virtual. O pesquisador Pan Hui, responsável pelo projeto, acredita que, no futuro, professoras/es geradas/os por IA serão mais confiáveis do que as/os professoras/es humanas/os, mas defende a coexistência entre docentes humanas/os e virtuais, que podem proporcionar mais diversidade e imersão. No entanto, pesquisas mostram que a maioria das/os estudantes ainda prefere professoras/es reais, valorizando empatia, experiência e *feedback* humano.

O que nos resta perguntar não é mais se a IA tem potencial para substituir docentes, pois isso já está sendo efetivado. O que precisamos atentar é o por vir: nos próximos anos, a substituição de professoras/es por tecnologias inteligentes ficará restrita a poucas escolas e universidades, ou se tornará uma realidade em larga escala?

A resposta a essa questão depende de múltiplos fatores: de nossas concepções sobre a finalidade da educação, de como compreendemos o papel de estudantes e docentes em tempos de IA generativa, das políticas públicas que incentivam ou restringem o uso de tecnologias digitais nas instituições de ensino, da expansão ou limitação de técnicas didático-

computacionais automatizadas, entre outros aspectos. Queremos, aqui, refletir sobre essa questão, que precisa ser amplamente debatida por todas as pessoas que lecionam em qualquer nível educacional, pois a revolução promovida pelas tecnologias digitais está apenas no começo e o mundo continuará mudando rapidamente.

A substituição da docência humana não ocorrerá de forma abrupta ou repentina: trata-se de um processo histórico em construção há mais de um século, desde as primeiras experiências com as máquinas de ensinar. Sob essa perspectiva, a IA generativa representa apenas mais um passo rumo à automação de tudo. Antes dela, *chatbots* já eram utilizados em algumas plataformas educacionais. Hoje, a maioria das/os estudantes universitárias/os ao redor do mundo já utiliza essa tecnologia, e sua popularização tende a se expandir também entre alunas/os do ensino médio e fundamental.

A substituição de professoras/es universitárias/os vem sendo efetivada em nosso país há vários anos, antes mesmo do lançamento do ChatGPT. Para compreender esse processo, é preciso considerar que 80% das/os graduandas/os brasileiras/os estão matriculadas/os em universidades privadas (Brasil, 2025) e que, dessas instituições, alguns grandes grupos educacionais com fins lucrativos e sem compromisso real com a qualidade do ensino vêm reduzindo o número de docentes, priorizando a maximização dos lucros. Estamos sofrendo uma consequência da mercantilização e financeirização da educação superior, que tem levado à precarização do trabalho docente, à redução da carga horária de professoras/es, à substituição de aulas presenciais por ensino a distância e ao aumento do controle administrativo sobre o ensino.

Na EaD, é comum a contratação de um/a tutor/a para supervisionar centenas ou milhares de estudantes, substituindo o trabalho que antes era realizado por diversas/os docentes. No Brasil, entre 2015 e 2024, o ensino superior registrou um aumento de 27% no número de estudantes matriculadas/os em cursos de graduação, enquanto, no mesmo período, o número de professoras/es na educação superior foi reduzido em 4% (Brasil, 2025). Como consequência, temos colegas que foram demitidas/os nesse período. Portanto, ao menos no ensino superior, a substituição de docentes

já estava em curso antes do *boom* da IA generativa, que tem potencial para acelerar ainda mais esse processo.

Profissionais da educação tornam-se substituíveis quando a educação é concebida como sinônimo de instrução; quando o papel do/a professor/a se reduz a apresentar conteúdos e corrigir provas; quando estudantes são vistas/os como tábulas rasas. Para quem já acreditava na substitutibilidade de docentes, essa convicção se fortaleceu com a popularização da IA generativa.

Como reação, algumas/uns docentes ignoram as tecnologias digitais, limitando atividades e avaliações à sala de aula sem computadores ou internet. Essa ação ignora o contexto em que vivemos: crianças, jovens, adultas/os e idosos/os aprendem cotidianamente em ecossistemas digitais permeados por buscadores, plataformas educacionais, repositórios acadêmicos e, mais recentemente, sistemas baseados em IA generativa. Nessa realidade hiperconectada, excluir as tecnologias digitais do espaço escolar aprofunda as desigualdades: quem dispõe de capital cultural e acesso a dispositivos computacionais continuará explorando essas tecnologias fora da sala de aula, enquanto as/os demais estudantes continuarão excluídas/os ciberculturalmente.

Acreditamos, assim como António Nóvoa (2018), que “nada substitui um bom professor”, nem as tecnologias digitais, nem a inteligência artificial generativa. No entanto, reconhecemos que a lógica capitalista deseja essa substituição. Precisamos lutar contra a visão reducionista de educação que elimina as relações humanas dos processos de ensino-aprendizagem, tanto de professoras/es quanto de colegas de turma. Essa não é uma luta apenas de professoras/es; trata-se de uma responsabilidade coletiva, que exige mobilização da sociedade, das instituições de ensino e, sobretudo, da esfera política.

A necessidade de lutarmos contra uma automação completa dos processos de ensino-aprendizagem não deve nos impedir de usar as tecnologias digitais, especialmente a IA generativa, nos cotidianos educacionais com o objetivo de promover uma formação condizente com o tempo presente. Precisamos compreender as potencialidades, riscos e limitações da IA generativa, seus usos, implicações e as teorizações a seu respeito. Isso pode envolver transformar crenças, concepções e práticas, o que requer investimentos na formação docente.

Retomando o questionamento que intitula este capítulo: nós, professoras/es, seremos substituídas/os pela IA generativa? Nossa resposta não é um categórico sim ou não, pois depende de diversos fatores: nossos valores e crenças didático-pedagógicas, perspectivas teóricas, conhecimentos sobre as potencialidades e limitações das tecnologias generativas, políticas públicas, interesses capitalistas etc.

Não se trata de uma coisa OU outra, mas sim de uma coisa E outra. A computação vem sendo utilizada para promover situações didáticas inovadoras, compreendendo a IA generativa como uma tecnologia de apoio às práticas pedagógicas; E, também, vem sendo utilizada para acelerar a implantação da arte de ensinar sem professoras/es, compreendendo a IA generativa como uma tecnologia capaz de reduzir os “gastos” com docentes, recursos didáticos e infraestrutura física. Trata-se de uma questão que envolve tecnofilia E tecnofobia, adoção E resistência, indução E desencorajamento, imposição E proibição, entre outras (o)posições.

Que educação desejamos?

Queremos encerrar este livro discutindo três diferentes configurações educacionais em relação ao uso das tecnologias digitais: (1) a educação presencial sem acesso a tecnologias digitais disponíveis para docentes e estudantes; (2) a educação totalmente digital, sem que estudantes tenham acesso a qualquer professor/a humano/a; e (3) a educação híbrida, que combina momentos presenciais e *online*, integrando a mediação de professoras/es e de IA generativa. Alguns questionamentos precisam ser feitos: A quem interessa cada uma dessas configurações educacionais? Que ideal de estudante se pretende formar e que processos subjetivos são produzidos? Quais abordagens didático-pedagógico-curriculares estão sendo adotadas? Quais intencionalidades econômicas as orientam? Quais as potencialidades e desafios de cada configuração?

Educação sem tecnologias digitais



Fonte: dos autores, ilustração de Mônica Lopes

Vivemos na era da IA generativa, que tem transformado diversos aspectos da sociedade e, na educação, tem alavancado novas práticas de aprendizagem e ensino. No entanto, não podemos perder de vista que ainda há escolas e universidades sem amplo acesso a tecnologias digitais em rede de boa qualidade para docentes e estudantes.

Essa configuração educacional acarreta implicações para as práticas pedagógicas e os processos formativos. A ausência de acesso à internet, a conteúdos online e a tecnologias generativas restringe a exploração das múltiplas fontes de informação e das possibilidades do digital em rede para aprender e cocriar durante as aulas. Uma educação sem tecnologias digitais compromete a vivência de fenômenos ciberculturais na formação, dificultando o desenvolvimento do letramento digital necessário para o uso de IA generativa e outras tecnologias do nosso tempo.

Nesse contexto, as políticas de inclusão digital induzida pelo Estado tornam-se fundamentais. Diferentemente da “inclusão espontânea”, que ocorre gradualmente e de forma voluntária, a inclusão induzida é um processo orientado e, muitas vezes, compulsório de inserção das pessoas na sociedade em rede. Esse processo exige um trabalho educativo planejado e a implementação de políticas públicas que “visam dar oportunidades a uma grande parcela da população excluída do uso e dos benefícios da sociedade da informação. É o que conhecemos por projetos de inclusão digital” (Lemos, 2011, p. 19).

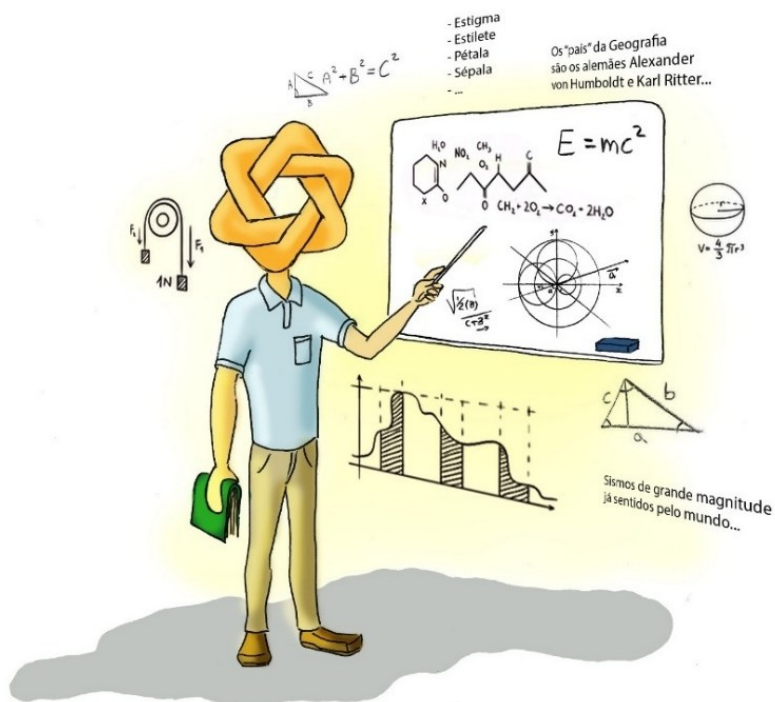
É evidente que o uso de tecnologias digitais na educação, por si só, não garante uma sociedade mais justa e menos desigual. No entanto, a exclusão cibercultural perpetua o abismo das desigualdades no acesso aos melhores recursos de aprendizagem-ensino de nosso tempo. A quem interessa uma educação desplugada, *infopobre* e alheia aos fenômenos da cultura digital?

Enquanto o filho do rico acessa todos os recursos disponíveis em rede, a partir de casa, com equipamentos de áudio e vídeo de última geração, os filhos dos pobres acessam a partir da escola ou dos centros públicos, onde as proibições são muito mais destacadas que as possibilidades de comunicação, acesso às informações e produção oferecidas pelas redes (Bonilla; Pretto, 2011, p.41).

Desde 2011, a ONU reconhece o acesso à internet como um direito humano. Contudo, não basta dispor de um bom computador com conexão de banda larga; isso é condição para a entrada no ciberespaço. É igualmente necessário promover o letramento digital, imprescindível ao exercício da cidadania contemporânea.

No extremo oposto está a educação totalmente digital, mediada exclusivamente por tecnologias digitais, sem que estudantes tenham acesso a professoras/es humanas/os. Essa é a configuração típica de cursos online massivos e de alguns cursos a distância em nosso país até 2025. Quem defende esse modelo de educação prioriza a redução de custos e a maximização de lucros em detrimento da qualidade da formação.

Educação sem professoras/es



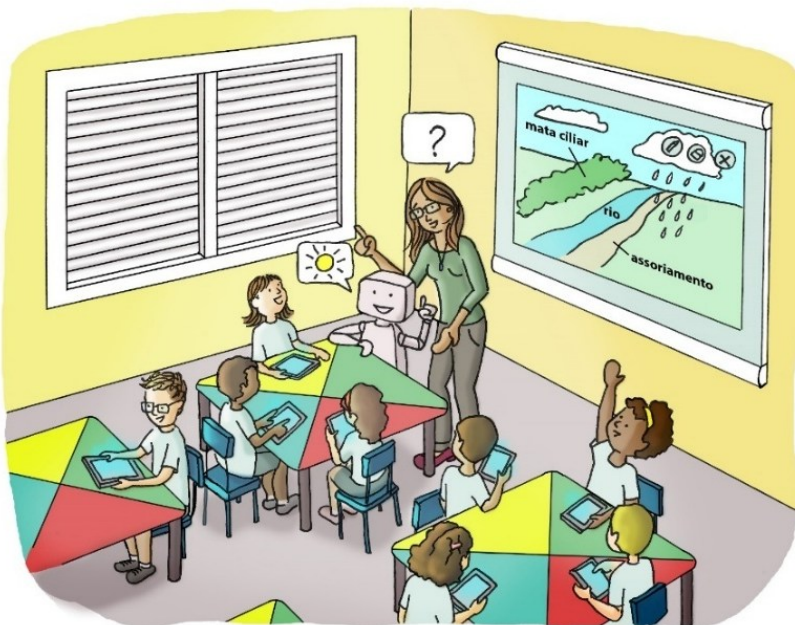
Fonte: dos autores, ilustração de Mônica Lopes

Uma educação sem professoras/es impõe desafios à aprendizagem das/os estudantes, especialmente pela ausência de interação social, empatia e envolvimento afetivo, fatores que podem levar à desmotivação e à evasão. Nesse modelo educacional, a ênfase recai sobre a assimilação de conteúdos, negligenciando as subjetividades das/os estudantes, suas singularidades, conhecimentos, interesses e experiências de vida. Ao substituir professoras/es e colegas de turma por uma educação totalmente digital, perde-se a oportunidade de aprender a negociar sentidos coletivamente e de se constituir pela diferença com a/o outra/o. A presença de um/a professor/a humano/a e o convívio com uma turma ampliam as possibilidades de interatividade, conversação, coautoria, colaboração, experiências compartilhadas, relações sociais, vínculos afetivos, convivência com as diferenças e a diversidade de pontos de vista — elementos essenciais para o desenvolvimento do pensamento crítico, da criatividade e da tecedura de conhecimentos.

Além disso, a ausência de professoras/es no processo educacional pode facilitar o controle ideológico e a manipulação das informações trabalhadas com as/os estudantes, atendendo aos interesses de determinados grupos. A mediação humana na tecedura do conhecimento pode alertar para esse problema e apoiar as/os estudantes na atribuição de sentidos às informações que acessam na vida cotidiana.

Reconhecemos que a educação em nosso tempo precisa assumir uma configuração híbrida, que valorize tanto os momentos presenciais quanto os *online*, assim como a atuação conjunta de professoras/es e de tecnologias generativas como mediadoras/es do conhecimento. O objetivo é criar experiências de aprendizagem ainda mais dinâmicas, contextualizadas e engajantes, em sintonia com as exigências da sociedade contemporânea.

Educação híbrida



Fonte: dos autores, ilustração de Mônica Lopes

A educação híbrida também apresenta alguns desafios, como os custos envolvidos na criação e manutenção de uma infraestrutura física e tecnológica adequada, além da necessidade de promover a formação inicial e continuada de professoras/es para o uso pedagógico das tecnologias digitais em rede. Uma formação que inclua a vivência de estratégias didático-curriculares que mobilizem as tecnologias digitais, incluindo as generativas, para promover atividades autorais e colaborativas, ampliar os espaços de conversação e interatividade, potencializar os letramentos digitais e ciberculturais, entre outras ações. Apostamos em uma educação híbrida que reconheça docentes e estudantes como sujeitas/os situadas/os culturalmente, críticas/os, criativas/os, autoras/es em suas práticas e inventivas/os na criação de situações didático-curriculares, com e sem o uso de tecnologias digitais.

As três configurações de educação — sem tecnologias digitais, sem professoras/es humanas/os e híbrida — têm nos ajudado a refletir sobre as tensões, os riscos e as potencialidades das tecnologias generativas no processo formativo. Com as reflexões apresentadas neste livro, esperamos ter contribuído para ampliar a compreensão sobre as implicações e teorizações do uso pedagógico da IA generativa. Que possamos, juntas/os, construir uma educação mais inovadora, inclusiva e humana, tendo as tecnologias generativas como aliadas nos processos de aprender e ensinar. O futuro da educação está sendo escrito agora, e todas/os nós somos coautoras/es dessa transformação.

“

Uma técnica não é nem boa, nem má (isto depende dos contextos, dos usos e dos pontos de vista), tampouco neutra (já que é condicionante ou restritiva, já que de um lado abre e de outro fecha o espectro de possibilidades). Não se trata de avaliar seus “impactos”, mas de situar as irreversibilidades às quais um de seus usos nos levaria, de formular os projetos que explorariam as virtualidades que ela transporta e de decidir o que fazer dela. Contudo, acreditar em uma disponibilidade total das técnicas e de seu potencial para indivíduos ou coletivos supostamente livres, esclarecidos e racionais seria nutrir-se de ilusões. Muitas vezes, enquanto discutimos sobre os possíveis usos de uma dada tecnologia, algumas formas de usar já se impuseram. Antes de nossa conscientização, a dinâmica coletiva escavou seus atratores. Quando finalmente prestamos atenção, é demasiado tarde... Enquanto ainda questionamos, outras tecnologias emergem na fronteira nebulosa onde são inventadas as ideias, as coisas e as práticas. Elas ainda estão invisíveis, talvez prestes a desaparecer, talvez fadadas ao sucesso. Nestas zonas de indeterminação onde o futuro é decidido, grupos de criadores marginais, apaixonados, empreendedores audaciosos tentam, com todas as suas forças, direcionar o devir.

”

(Lévy, 1999, p. 26-27)

Referências

*Todos os links das referências foram acessados e checados em 23 out. 2025.

AI BOOM. In: Wikipedia: the free encyclopedia. Disponível em:

https://en.wikipedia.org/wiki/AI_boom *

ALBERGOTTI, Reed. The secret history of Elon Musk, Sam Altman, and OpenAI.

Semafor, 24 mar. 2023. Disponível em:

<https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai> *

ALFANO, Bruno. Governo de São Paulo adota inteligência artificial para corrigir dever de casa dos alunos. O Globo, Rio de Janeiro, 19 maio 2025. Disponível em:

<https://oglobo.globo.com/brasil/educacao/noticia/2025/05/19/governo-de-sao-paulo-adota-inteligencia-artificial-para-corrigir-dever-de-casa-dos-alunos.ghtml> *

ALLEN, Liz; O'CONNELL, Alison; KIERMER, Veronique. How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. **Learned Publishing**, Nova Jersey, v. 32, n. 1, p. 71-74, 2019. Disponível em:

<https://onlinelibrary.wiley.com/doi/epdf/10.1002/leap.1210> *

ALMEIDA, Fernando José de Almeida; SILVA, Maria da Graça Moreira da Silva.

Reflexões sobre tecnologias, educação e currículo: conceitos e trajetórias. In: VALENTE, J. A.; FREIRE, F. M. P.; ARANTES, F. L. (org.). **Tecnologia e educação: passado, presente e o que está por vir**. Campinas, SP: Nied: Unicamp, 2018. p.122-148.

ALMEIDA, Gilberto Martins. Quando a IA e a Propriedade Intelectual se cruzam, como ficam a Ética e o Direito? **SBC Horizontes**, 1 fev. 2023. Disponível em:

<https://horizontes.sbc.org.br/index.php/2023/02/quando-a-ia-e-a-propriedade-intelectual-se-cruzam-como-ficam-a-etica-e-o-direito> *

ALTMAN, Sam. We will have to monetize it somehow at some point; the compute costs are eye-watering. Twitter, 5 dez. 2022. Disponível em:

<https://twitter.com/sama/status/1599669571795185665> *

ALVES, Lynn (org.). **Inteligência artificial e educação**: refletindo sobre os desafios contemporâneos. Salvador: EDUFBA; Feira de Santana: UEFS Editora, 2023.

ALVES, Nilda. Decifrando o pergaminho: o cotidiano das escolas nas lógicas das redes cotidianas. In: OLIVEIRA, Inês B.; ALVES, Nilda (org.). **Pesquisa**

- nos/dos/com os cotidianos das escolas:** sobre redes de saberes. Petrópolis, RJ: DP et Alii, 2001, p. 13-38.
- ALVES, Nilda. Cultura e cotidiano escolar. **Revista Brasileira de Educação**, v. 23, p. 62-74, 2003.
- AMARAL, Mirian Maia; VERGARA, S. C. O “aluno-cliente” nas instituições de ensino superior: uma metáfora a ser banida do discurso educacional? **Boletim Técnico do Senac**, v. 37, n. 1, p. 4-15, 2011. Disponível em: <https://www.bts.senac.br/bts/article/view/197/180> *
- AMARAL, Mirian Maia. **Autorias docente e discente:** pilares de sustentabilidade na produção textual e imagética em redes educativas presenciais e online. Tese (Doutorado em Educação). Rio de Janeiro: Universidade Estácio de Sá – UNESA, 2014.
- AMARAL, Mirian Maia; VELOSO, Maristela Midlej Silva de Araujo; ROSSINI, Tatiana Stofella Sodré. A autoria coletiva no contexto da educação em tempos de cibercultura. In: SANTOS, E. O.; PIMENTEL, M.; SAMPAIO, F. F. (org.). **Informática na educação:** autoria, linguagens, multiletramentos e inclusão. Porto Alegre: Sociedade Brasileira de Computação, 2021. Disponível em: <https://ieducacao.ceie-br.org/autoriacoletiva> *
- APEOESP. Inteligência Artificial é instrumento auxiliar, jamais poderá substituir o professor! APEOESP – Sindicato dos Professores do Ensino Oficial do Estado de São Paulo, 18 abr. 2024. Disponível em: <https://www.apeoesp.org.br/publicacoes/professor/inteligencia-artificial-e-instrumento-auxiliar-jamais-podera-substituir-o-professor> *
- AUE, Jeniffer. The Relationship between Art and AI. Medium, [s.l.], 15 maio 2018. Disponível em: <https://medium.com/design-ibm/the-role-of-art-in-ai-31033ad7c54e> *
- BACKES, Luciana. As manifestações da autoria na formação do educador em espaços digitais virtual. **Educação, Ciência e Cultura**, v. 17, n. 2, p. 71-85, 2012.
- BAHDANAU, Dzmitry; CHO, KyungHyun; BEGIO, Yoshua. Neural Machine Translation by Jointly Learning to Align and Translate. **International Conference on Learning Representations**, Vancouver, 2015.
- BARANAUSKAS, Maria Cecília Calani; ROCHA, Heloísa Vieira; MARTINS, Maria Cecília; D'ABREU, João Vilhete. Uma taxonomia para ambientes de aprendizado baseados no computador. In: VALENTE, J. A. (org.). **O computador na sociedade do conhecimento**. Brasília: MEC, 1999. p. 45-70.
- BARBIER, René. **A pesquisa-ação**. Trad. Lucie Didio. Brasília: Liber Livro. 2007.
- BARTHES, Roland. **A Morte do Autor**. In: O rumor da língua. São Paulo: Brasiliense, 2004 [1967].
- BBC. A história da informação - Documentário (2012). YouTube, 27 ago. 2023. Disponível em: <https://www.youtube.com/watch?v=ppNCQ5cC5uA> *

- BEIGUELMAN, Giselle. Máquinas companheiras. **Morel**, n. 7, p. 76-86, 2023.
- BELLONI, M. L. **Educação a distância**. 6. ed. Campinas, SP: Autores Associados, 2012..
- BLOOM, Harold. **The anxiety of influence**: a theory of poetry. New York: Oxford University Press, 1973.
- BLOOM, Benjamin S. The 2 sigma problem: the search for methods of group instruction as effective as one-to-one tutoring. **Educational Researcher**, v. 13, n. 6, 1984, p. 4-16.
- BLOOMBERG Originals. Inside OpenAI, the Architect of ChatGPT | The Circuit. YouTube, 15 junho de 2023. Disponível em: <https://www.youtube.com/watch?v=p9Q5a1Vn-Hk> *
- BODEN, Margaret A. **Inteligência Artificial**: uma brevíssima introdução. São Paulo: Editora Unesp, 2020.
- BONILLA, Maria Helena Silveira; PRETTO, Nelson de Luca (orgs.). **Inclusão digital**: polêmica contemporânea [online]. Salvador: EDUFBA, 2011, 188p. Disponível em: <http://books.scielo.org> *
- BOZKURT, Aras et al. The Manifesto for Teaching and Learning in a Time of Generative AI: A Critical Collective Stance to Better Navigate the Future. **Open Praxis**, v.16, n.4, p.487-513, 2024. Disponível em: <https://openpraxis.org/articles/10.55982/openpraxis.16.4.777> *
- BRAND, Amy et al. Beyond authorship: attribution, contribution, collaboration, and credit. **Learned Publishing**, Nova Jersey, v. 28, n. 2, p. 151-155, 2015. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1087/20150211> *
- BRASIL. Decreto nº 75.699, de 06 de maio de 1975. Promulga a Convenção de Berna para a Proteção das Obras Literárias e Artísticas. **Diário Oficial da União**: seção 1, Brasília, DF, p. 5137, 07 maio 1975. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto/1970-1979/d75699.htm *
- BRASIL. Presidência da República. Lei nº 9.610, de 19 de fevereiro de 1998. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Brasília: Presidência da República, 1998. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l9610.htm *
- BRASIL. Lei Nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. [Marco Civil da Internet]. Brasília, DF: **Diário Oficial da União**, 2014.
- BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília: MEC, 2018.
- BRASIL. Lei Nº 14.358, de 1º de junho de 2022. Dispõe sobre o valor do salário-mínimo a vigorar a partir de 1º de janeiro de 2022. **Diário Oficial da União**, 2022a.

BRASIL. Ministério da Educação. **Computação**: complemento à BNCC. Brasília: MEC, 2022c.

BRASIL. Projeto de Lei nº 2.338, de 2023. Dispõe sobre o uso de Inteligência Artificial. Câmara dos Deputados, 3 maio 2023b. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233> *

BRASIL. **Censo da educação superior 2024: notas estatísticas**. Brasília-DF: Inep/MEC, 2025. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_escolar_da_educacao_superior_2024.pdf *

BROWN, Tom B. et al. Language Models are Few-Shot Learners. **Advances in Neural Information Processing Systems**, v. 33, p. 1877-1901, 2020. Disponível em: <https://arxiv.org/pdf/2005.14165.pdf> *

BRUNS, Axel. **Blogs, Wikipedia, Second life, and Beyond**: from production to produsage. New York: Peter Lang, 2008.

BRYSBART, Marc. How many words do we read per minute? A review and meta-analysis of reading rate. **Journal of Memory and Language**, v. 109, 2019.

BUSH, Vannevar et al. As we may think. **The Atlantic Monthly**, v. 176, n. 1, p. 101-108, 1945.

BUZATO, Marcelo El Khouri. Multiletramentos e informática na escola. In: SANTOS, E. O.; PIMENTEL, M.; SAMPAIO, F. F. (Org.). **Informática na Educação**: autoria, linguagens, multiletramentos e inclusão. Porto Alegre: Sociedade Brasileira de Computação, 2021. Disponível em: <https://ieducacao.ceie-br.org/multiletramentos> *

CALVÃO, Leandro Dantas; PIMENTEL, Mariano; FUKS, Hugo. **Do email ao Facebook**: uma perspectiva evolucionista sobre os meios de conversação da internet. Rio de Janeiro: UNIRIO, 2014.

CARVALHO, Felipe; PIMENTEL, Mariano. Atividades autorais online: aprendendo com criatividade. **SBC Horizontes**, 12 nov. 2020. Disponível em: <http://horizontes.sbc.org.br/index.php/2020/11/atividadesautorais> *

CARVALHO, Felipe. **#Pedagogiasciberculturais**: como aprendemos-ensinamos a nos tornar o que somos? Tese de Doutorado (em Educação). Rio de Janeiro: UERJ, 2021. Disponível em: <https://www.bdt.uerj.br:8443/handle/1/17428> *

CARVALHO, Felipe; PIMENTEL, Mariano. Desculpa a interrupção, professor, eu nem sei se eu poderia te interromper: quais os sentidos da conversação em aula? **Revista Periferia**, v. 14, n. 3, 2022, p. 127-148.

CARVALHO, Felipe; PIMENTEL, Mariano. Estudar e aprender com o ChatGPT. **Revista Educação e Cultura Contemporânea**, v. 20, 2023.

CARVALHO, Felipe; PIMENTEL, Mariano. Diferentes concepções de uso do ChatGPT na educação. In: PORTO, C.; SANTOS, E.; BOTTENTUIT JR., J. B. (org.) **ChatGPT e**

- outras inteligências artificiais:** práticas educativas na Cibercultura. São Luís: EDUFMA, 2024, p. 37-49.
- CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e terra, 2005.
- CERTEAU, Michel de. **A invenção do cotidiano:** artes de fazer. Petrópolis, Rio de Janeiro: Editora Vozes, 15 edição, 2008.
- CETIC.br. **TIC Educação 2022:** pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras [livro eletrônico]. São Paulo: CGI.br, 2023. Disponível em:
https://cetic.br/media/docs/publicacoes/2/20231122132216/tic_educacao_2022_livro_completo.pdf
- CETIC.br. **TIC Educação 2024:** coletiva de imprensa. São Paulo: CGI.br, 16 set. 2025. Disponível em:
https://cetic.br/media/analises/tic_educacao_2024_principais_resultados.pdf *
- CHEGG.ORG. Global Student Survey 2023. Disponível online:
<https://www.chegg.org/global-student-survey-2023> *
- CHOMSKY, Noam. Chomsky: ChatGPT contra o pensamento crítico. **Outras Palavras**, 2 maio 2023. Disponível em:
<https://outrapalavras.net/tecnologiaemdisputa/chomsky-o-chatgpt-contra-o-pensamento-critico> *
- CIBERNÉTICA. In: Wikipedia, a enciclopédia livre. Flórida: Wikimedia Foundation, s.d. Disponível em: <https://pt.wikipedia.org/wiki/Cibernética> *
- CLARCKE. Arthur C. **Profiles of the Future:** An Inquiry into the Limits of the Possible. New York: Harper & Row, 1962.
- COELHO, Leonardo; SCHURIG, Sofia. Livros escritos por IA começam a pipocar na Amazon brasileira. Núcleo, 22 ago. 2023. Disponível em:
<https://nucleo.jor.br/reportagem/2023-08-22-livros-ia-amazon> *
- COLTON, Simon; WIGGINS, Geraint A. Computational creativity: The final frontier? In: ECAI 2012. IOS Press, 2012. p. 21-26. Disponível em:
<https://ebooks.iospress.nl/doi/10.3233/978-1-61499-098-7-21> *
- COMENIUS, I. A. **Didactica Magna** (1621-1657). Lisboa: Fundação Calouste Gulbenkian, 2001.
- CORREIA, Ana Paula. É o ChatGPT uma nova tendência no Ensino Superior? **Revista Docência e Cibercultura**, seção Notícias, abr. 2023. Disponível em:
<https://www.e-publicacoes.uerj.br/index.php/re-doc/announcement/view/1622> *
- COUTO JUNIOR, Dilton Ribeiro; AMARO, Ivan; TEIXEIRA, Marcelle Medeiros; RUANI, Ruann Moutinho. Do face a face às dinâmicas comunicacionais em/ na rede: a conversa online como procedimento metodológico da pesquisa em educação. **Educação em Foco**, v. 25, n. 1, p. 82-98, 2020. Disponível em:
<https://periodicos.ufjf.br/index.php/edufoco/article/view/29955> *

- DAILY MAIL. Cheeky monkey! Macaque borrows photographer's camera to take hilarious self-portraits. MailOnline, Londres, 4 jul. 2011. Disponível em: <https://www.dailymail.co.uk/news/article-2011051/Black-macaque-takes-self-portrait-Monkey-borrows-photographers-camera.html> *
- DALL-E 2. OpenAI (em inglês). Disponível em: <https://openai.com/dall-e-2> *
- DAVID GAME COLLEGE. GCSE AI Adaptive Learning Programme. Disponível em: <https://www.davidgamecollege.com/courses/courses-overview/item/102/gcse-ai-adaptive-learning-programme> *
- DEEP BLUE. In: Wikipédia, a enciclopédia livre. Flórida: Wikimedia Foundation. Disponível em: https://pt.wikipedia.org/wiki/Deep_Blue *
- DEEPSEEK-AI et al. DeepSeek-V3 Technical Report. 18 fev. 2025. Disponível em: <https://arxiv.org/pdf/2412.19437> *
- DELEUZE, Gilles; GUATTARI, F. **Mil platôs**. Rio de Janeiro: Editora 34, 1997 [originalmente publicado em 1980]. v.4
- DELEUZE, Gilles. **Diferença e repetição**. Editora Paz e Terra, 2018 [originalmente publicada em 1968].
- DELORS, Jacques et al. **Educação: um tesouro a descobrir**; relatório para a UNESCO da Comissão Internacional sobre Educação para o Século XXI. São Paulo: Cortez, 1998.
- DERRIDA, Jacques. **Torres de babel**. Belo Horizonte: Editora UFMG, 2002.
- DERRIDA, Jacques. Nada de hospitalidade, passo da hospitalidade. In: Anne Dufourmantelle convida Jacques Derrida a falar da hospitalidade / Jacques derrida [Entrevistado]. São Paulo: Escuta, 2003.p. 67-135.
- DIGITAL Education Council Global AI Student Survey 2024. **AI or not AI: what students want**. 2024. Disponível em: <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024> *
- Digital Education Council Global AI Faculty Survey 2025. **AI Meets Academia: What Faculty Think**. 2024. Disponível em: <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-faculty-survey> *
- DORIA, Pedro. Os concorrentes do GPT chegaram. O Globo, 01 out. 2024. Disponível em: <https://oglobo.globo.com/opiniao/pedro-doria/coluna/2024/10/os-concorrentes-do-gpt-chegaram.ghtml> *
- DRAXLER, Fiona et al. The AI ghostwriter effect: When users do not perceive ownership of AI-generated text but self-declare as authors. **ACM Transactions on Computer-Human Interaction**, v. 31, n. 2, p. 1-40, 2024. Disponível online: <https://dl.acm.org/doi/full/10.1145/3637875> *
- ENDICOTT, Sean. GPT-4 was just announced, and Microsoft confirmed that it powers the new Bing. Windows Central, 14 mar. 2023.

- ESCOLAS de NY proíbem aplicativo de inteligência artificial que escreve redação para alunos. Bom Dia Brasil, TV Globo. 25 jan. 2023. Disponível em: <https://globoplay.globo.com/v/11309490> *
- EXAME. Professores gerados por inteligência artificial dão aulas em universidade de Hong Kong. Exame, 5 fev. 2024. Disponível em: <https://exame.com/tecnologia/professores-gerados-por-inteligencia-artificial-dao-aulas-em-universidade-de-hong-kong> *
- FEATHERS, Todd. Arizona School's Curriculum Will Be Taught by AI, No Teachers. Gizmodo [blog], 19 dez. 2024. Disponível em: <https://gizmodo.com/arizona-schools-curriculum-will-be-taught-by-ai-no-teachers-2000540905> *
- FIGUEIREDO, Ana Luiza. ChatGPT ganha acesso à internet e responde perguntas atuais. Olhar Digital, 16 maio 2023.
- FIGURE. Figure Status Update - OpenAI Speech-to-Speech Reasoning. YouTube, Figure. Disponível em: <https://www.youtube.com/watch?v=Sq1QZB5baNw> *
- FIRTH, J. R. **Papers in Linguistics 1934-1951**. Londres: Oxford University Press, 1957
- FORBES. Brasil está entre os 4 países que mais usam o ChatGPT; veja ranking. Forbes Tech, 8 mar. 2024. Disponível em: <https://forbes.com.br/forbes-tech/2024/03/brasil-esta-entre-os-4-paises-que-mais-usam-o-chatgpt-veja-ranking> *
- FOUCAULT, Michel. O que é um autor? In: **Ditos e escritos**: estética, literatura e pintura, música e cinema. Rio de Janeiro: Forense Universitária, 2001, v. 3, p. 264-298.
- FOUCAULT, Michel. **Ditos e escritos V**: Ética, sexualidade, política. Rio de Janeiro: Forense Editora, 2006
- FREIRE, Paulo. **Pedagogia do oprimido**. Rio de Janeiro: Paz e Terra, 1970.
- FREIRE, P.; FAUNDEZ, A. **Por uma pedagogia da pergunta**. Rio de Janeiro: Paz e Terra, 1985.
- FREIRE, Paulo. **A Educação na Cidade**. 5 ed. São Paulo: Cortez, 2001 [originalmente publicado em 1991].
- FREIRE, Paulo. **Pedagogia da autonomia**: saberes necessários à prática educativa. 52 ed. Rio de Janeiro: Paz e Terra, 2015 [originalmente publicado em 1996].
- Freitas, Aline. Governo de SP avalia utilizar inteligência artificial para “aprimorar” conteúdo digital nas escolas estaduais. G1, São Paulo, 17 abr. 2024. Disponível em: <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> *
- FREITAS, L. C. Conseguiremos escapar ao neotecnicismo? In: SOARES, M. B.; KRAMER, S.; LUDKE, M. (orgs.). Escola Básica (Anais da 6ª CBE). Campinas: Papirus. 1992.

- FREITAS, L. C. Responsabilização, meritocracia e privatização: conseguiremos escapar ao neotecnicismo? III Seminário de Educação Brasileira, fev. 2011a, p. 1-35.
- FREITAS, L. C. Neotecnicismo e formação do educador. In.: ALVES, Nilda (org.). **Formação de Professores: pensar e fazer**. São Paulo: Cortez, 11. ed., 2011b, p. 95-108.
- FREITAS, L. C. **A reforma empresarial da educação**: nova direita, velhas ideias. São Paulo: Expressão Popular, 2018.
- FREITAS, L. C. Neotecnicismo digital. Blog Avaliação Educacional, 2021. Disponível em: <https://avaliacaoeducacional.com/2021/07/11/neotecnicismo-digital> *
- G1. Quem é o engenheiro demitido pelo Google por dizer que programa de inteligência artificial tem consciência. G1, 24 jul. 2022. Disponível em: <https://g1.globo.com/tecnologia/noticia/2022/07/24/quem-e-o-engenheiro-demitido-pelo-google-por-dizer-que-programa-de-inteligencia-artificial-tem-consciencia.ghtml> *
- GALLO, Sílvio. Eu, o outro e tantos outros: educação, alteridade e filosofia da diferença. In: Anais do II Congresso Internacional Cotidiano: Diálogos sobre Diálogos. Universidade Federal Fluminense, Rio de Janeiro. 2008. p. 1-16.
- GARLAND, Alex. *Ex_Machina*: instinto artificial. Reino Unido: DNA Films, 2014.
- GARTNER Hype Cycle. Disponível em: <https://www.gartner.com/en/chat/gartner-hype-cycle> *
- GEMINI. Crie especialistas personalizados com os Gems. Gemini, 2024. Disponível em: <https://gemini.google/overview/gems> *
- GENERATIVE pre-trained transformer. In: Wikipedia, the free encyclopedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: https://en.wikipedia.org/wiki/Generative_pre-trained_transformer *
- GOMES, Alex Sandro; PIMENTEL, Edson Pinheiro. ambientes virtuais de aprendizagem para uma educação mediada por tecnologias digitais. In: PIMENTEL, M.; SAMPAIO, F. F.; SANTOS, E. (org.). **Informática na educação**: ambientes de aprendizagem, objetos de aprendizagem e empreendedorismo. Porto Alegre: Sociedade Brasileira de Computação, 2021.
- GOOD, Irving John. Speculations concerning the first ultraintelligent machine. **Advances in Computers**, v. 6, p. 31-88, 1966. Disponível em: [https://doi.org/10.1016/S0065-2458\(08\)60418-0](https://doi.org/10.1016/S0065-2458(08)60418-0) *
- GPT-2. In: Wikipedia, the free encyclopedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: <https://en.wikipedia.org/wiki/GPT-2> *
- GPT-3. In: Wikipedia, the free encyclopedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: <https://en.wikipedia.org/wiki/GPT-3> *
- GPT-4. In: Wikipedia, the free encyclopedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: <https://en.wikipedia.org/wiki/GPT-4> *

- GRANT, Nico. ChatGPT and Other Chat Bots Are a 'Code Red' for Google Search. The New York Times, 18 jul. 2023. Disponível em: <https://www.nytimes.com/2022/12/21/technology/ai-chatgpt-google-search.html> *
- GROS, Théotime. Can Artificial Intelligence Create Art? ResearchGate, Berlin, June 2019. Disponível em: <https://doi.org/10.13140/RG.2.2.10238.41287> *
- HALL, Stuart. **Cultura e representação**. Rio de Janeiro: PUC-Rio: Apicuri, 2016.
- HARARI, Yuval Noah. **Nexus**: uma breve história das redes de informação, da Idade da Pedra à inteligência artificial. São Paulo: Companhia das Letras, 2024.
- HENDRYCKS, Dan et al. Measuring Mathematical Problem Solving With the MATH Dataset. arXiv, 8 nov. 2021. Disponível em: <https://arxiv.org/pdf/2103.03874v2> *
- HOFFMAN, Liz; ALBERGOTTI, Reed. Microsoft eyes \$10 billion bet on ChatGPT. Semafor, 10 jan. 2023. Disponível em: <https://www.semafor.com/article/01/09/2023/microsoft-eyes-10-billion-bet-on-chatgpt> *
- HU, Krystal. ChatGPT sets record for fastest-growing user base – analyst note. Reuters, 2 fev. 2023. Disponível em: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01> *
- HUBERT, K. F.; AWA, K. N.; ZABELINA, D. L. The current state of artificial intelligence generative language models is more creative than humans on divergent thinking tasks. **Scientific Reports**, v. 14, p. 3440, 10 fev. 2024. Disponível em: <https://www.nature.com/articles/s41598-024-53303-w> *
- IBGE. Síntese de Indicadores Sociais. 2022. Disponível em: <https://www.ibge.gov.br/estatisticas/multidominio/condicoes-de-vida-desigualdade-e-pobreza/9221-sintese-de-indicadores-sociais.html> *
- IDS (Instituto de Defesa do Software). Pela primeira vez, tribunal chinês decide que imagens geradas por IA podem ser protegidas por direitos autorais. 2023. Disponível em: <https://ids.org.br/noticia/pela-primeira-vez-tribunal-chines-decide-que-imagens-geradas-por-ia-podem-ser-protetidas-por-direitos-autorais> *
- INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Notas estatísticas**: Censo Escolar 2023. Brasília: Inep, 2024. Disponível em: https://download.inep.gov.br/publicacoes/institucionais/estatisticas_e_indicadores/notas_estatisticas_censo_escolar_2023.pdf *
- INTELIGÊNCIA artificial pode ter sentimentos? Veja o que a ciência diz sobre robô do ChatGPT que fez declaração de amor para humano. Fantástico, 12 mar. 2023. Disponível em: <https://globoplay.globo.com/v/11441732> *
- INTERIM Measures for the Management of Generative AI Services. In: Wikipedia, the free encyclopedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em:

https://en.wikipedia.org/wiki/Interim_Measures_for_the_Management_of_Generative_AI_Services *

CBL. Regulamento - 66º Prêmio Jabuti (2024). Câmara Brasileira do Livro, 2024. Disponível em: <https://pt.scribd.com/document/742748735/REGULAMENTO-66-JABUTI-2024-ad> *

JAQUES, Patrícia Augustin; NUNES, Maria Augusta S. N. Computação Afetiva aplicada à Educação. In: SAMPAIO, F. F.; PIMENTEL, M.; SANTOS, E. O. (org.). **Informática na educação**: games, inteligência artificial, realidade virtual/aumentada e computação ubíqua. Porto Alegre: Sociedade Brasileira de Computação, 2021. Disponível em: <https://ceie.sbc.org.br/livrodidatico/index.php/computacaoafetiva> *

JONES, Cameron R.; BERGEN, Benjamin K. People cannot distinguish GPT-4 from a human in a Turing test. arXiv, 9 maio 2024. Disponível em: <https://arxiv.org/abs/2405.08007> *

KARNAL, Leandro. Leandro Karnal desafia o ChatGPT. Prazer, Karnal, 9 mar. 2023. Disponível em: <https://youtu.be/OONpZy0sTqo> *

KASPAROV, Garry. The chess master and the computer. The New York Review of Books, v. 57, n. 2, p. 16-19, 2010. Disponível em: <https://www.nybooks.com/articles/2010/02/11/the-chess-master-and-the-computer> *

KASTRUP, Virgínia. Políticas cognitivas na formação do professor e o problema do devir-mestre. **Educação & Sociedade**, v. 26, p. 1273-1288, 2005. Disponível em: <https://www.scielo.br/j/es/a/bG374G5nJQ6jtVgCbb7Vsvb> *

KEIRSEY, Kent. Invoke Secures Copyright in Landmark Ruling for AI-Assisted Artwork. Invoke Blog, 10 fev. 2025. Disponível em: <https://www.invoke.com/post/invoke-receives-copyright-in-landmark-ruling-for-ai-assisted-artwork> *

KHAN, Salman. How AI Could Save (Not Destroy) Education - TED. YouTube, 1 maio 2023. Disponível em: <https://www.youtube.com/watch?v=hJP5GqnTrNo> *

KIELA, Douwe; THRUSH, Tristan; ETHAYARAJH, Kawin; SINGH, Amanpreet. Plotting Progress in AI. Contextual AI Blog, 31 jul. 2023. Disponível em: <https://contextual.ai/blog/plotting-progress> *

KIM, Tae. The Age of AI Is Still in the Early Innings. These 3 Stocks Will Keep Winning. Barron's, 15 nov. 2024. Disponível em: <https://www.barrons.com/articles/ai-technology-stocks-d289e665> *

KLEIN, Naomi. AI machines aren't 'hallucinating'. But their makers are. The Guardian, 8 maio 2023.

KOSMYNA, Nataliya et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv, [s.l.], 2025. Disponível em: <https://arxiv.org/pdf/2506.08872v1> *

- KURZWEIL, Ray. **A singularidade está próxima**: quando os humanos transcendem a biologia. São Paulo: Itaú Cultural: Iluminuras, 2018.
- KURZWEIL, Ray. **A Singularidade está mais próxima**: a fusão do ser humano com o poder da inteligência artificial. São Paulo: Goya, 2024. Edição do Kindle.
- LAFRANCE, Yvon. O conhecimento: ciência e opinião. In: FRONTEROTTA, F.; BRISSON, L. **Platão**: leituras. São Paulo: Edições Loyola, 2011. p.149-166.
- LAPA, Andrea; PRETTO, Nelson De Luca. Educação a distância e precarização do trabalho docente. **Em Aberto**. Brasília, v. 23, n. 84, p.79-97, nov. 2010. Disponível em: <https://rbep.inep.gov.br/ojs3/index.php/emaberto/article/view/2469> *
- LARGE Language Model. In: Wikipedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: https://pt.wikipedia.org/w/index.php?title=Modelo_de_linguagem_grande&olvid=66235795 *
- LEMOS, André. Cibercultura como território recombinante. In: MARTINS, C. D. et al. (org.) **Territórios recombinantes**: arte e tecnologia — debates e laboratórios. São Paulo: Instituto Sérgio Motta, 2007. p.35-48.
- LEMOS, André. Prefácio. In: BONILLA, Maria Helena; PRETTO, Nelson (org.). **Inclusão digital**: polêmica contemporânea. Salvador: EDUFBA, 2011, p. 15-21.
- LÉVY, Pierre. **As tecnologias da inteligência**: o futuro do pensamento na era da informática [1990]. Rio de Janeiro: Editora 34, 1993.
- LÉVY, Pierre. **Cibercultura**. Rio de Janeiro: Ed.34, 1999.
- LIGHTMAN, Hunter et al. Let's Verify Step by Step. OpenAI, Research, 31 mar. 2023. Disponível em: <https://arxiv.org/abs/2305.20050> *
- LINGUAGEM INCLUSIVA. Ministério da Gestão e da Inovação em Serviços Públicos, 11 jul. 2023. Disponível em: <https://www.gov.br/gestao/pt-br/assuntos/inovacao-governamental/gestao-de-carreiras/lins/linguagem-inclusiva> *
- MACEDO, Roberto Sidnei. **Currículo, diversidade e equidade**: luzes para uma educação intercítica. Salvador: EdUFBA, 2007.
- MACEDO, Roberto Sidnei. **Compreender/mediar a formação**: o fundante da educação. Brasília: Liber Livro, 2010.
- MACEDO, Roberto Sidnei. **Atos de currículo e autonomia pedagógica**: o socioconstrucionismo curricular em perspectiva. Petrópolis, RJ: Vozes, 2013.
- MACEDO, Roberto Sidnei. Formação de professores e políticas de sentido experienciais e acontecimentais. In: Formação de professores, educação online e democratização do acesso às redes [live], Congresso Virtual UFBA, 2020. Disponível em: <https://www.youtube.com/watch?v=UD0KrPkHbiY&t=295s> *

- MADDALENA, Tania L. A hiperescrita de si: memória, experiência e invenção digital na formação de professores. **Revista Teias**, v. 21, n. 60, p. 203-217, 2020.
- MARIN, Alda Junqueira; SILVA, Aída Maria Monteiro; SOUZA, Maria Inês Marcondes de. **Situações didáticas**. São Paulo: Junqueira & Marin Editores, 2003.
- MARKOWITZ, Dale. Explicação sobre transformadores: entenda o modelo por trás de GPT, BERT, e T5. YouTube, 2022. Disponível em: <https://www.youtube.com/watch?v=SZorAJ4l-sA> *
- MARQUES, Fabrício. O plágio encoberto em textos do ChatGPT. Pesquisa FAPESP, 9 mar. 2023. Disponível em: <https://revistapesquisa.fapesp.br/o-plagio-encoberto-em-textos-do-chatgpt> *
- MARSHALL, Frank. The Man vs. The Machine. EUA: ESPN, 2014. Disponível em: https://youtube.com/watch?v=A_q2fNc-7-0&si=EnSikalECMiOmarE *
- MCCARTHY, J., MINSKY, M., ROCHESTER, N., SHANNON, C.E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. 1955. Disponível online: <http://raysolomonoff.com/dartmouth/boxa/dart564props.pdf> *
- MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, n. 4, p. 115-133, 1943.
- MCLUHAN, Marshall. **The Gutenberg Galaxy**: The Making of Typographic Man [A galáxia de Gutenberg: a formação do homem tipográfico]. Toronto: University of Toronto Press, 1962.
- MCLUHAN, Marshall. **Understanding Media**: The Extension of Man [Os meios de comunicação como extensões do homem]. New York: McGraw Hill, 1964.
- MICRODADOS do Enem 2022. INEP, 2022. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem> *
- MICRODADOS do Enem 2023. INEP, 2023. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem> *
- MICRODADOS do Enem 2024. INEP, 2024. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem> *
- MIDJOURNEY. Terms of Service. 7 mar. 2024. Disponível em: <https://docs.midjourney.com/docs/terms-of-service> *
- MIT Technology Review. As ferramentas de detecção de texto da Inteligência Artificial são realmente fáceis de enganar. 4 ago. 2023. Disponível em: <https://mittechreview.com.br/as-ferramentas-de-deteccao-de-texto-da-inteligencia-artificial-sao-realmente-faceis-de-enganar> *
- NABUCO, Laura. Inteligência Artificial: escola no Reino Unido inicia turma sem professores. Livre [blog]., 8 set. 2024. Disponível em:

<https://olivire.com.br/inteligencia-artificial-escola-no-reino-unido-inicia-turma-sem-professores> *

NELSON, Theodor Holm. Complex information processing: a file structure for the complex, the changing and the indeterminate. In: ACM '65: Proceedings of the 1965 20th national conference, 1965, p. 84-100. Disponível em: <https://dl.acm.org/doi/10.1145/800197.806036> *

NOLASCO-SILVA, Leonardo. A professora artífice ou Sobre Dramaturgias 'docentesdiscentes'. **Arcos Design**, v. 15, n. 1, p. 70-86, 2022. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/arcosdesign/article/view/65411> *

NOTTOMBROWN. Languages_by_word_count.csv. GitHub, OpenAI, GPT-3, 2020. Disponível em: https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv *

NÓVOA, António. Nada substitui um bom professor [vídeo, excerto do seminário Mais Sucesso, realizado por Projecto Fénix 2011, Montagem: Ágora Gaia 2015]. Facebook, 5 abr. 2018. Disponível em: <https://www.facebook.com/watch/?v=1722057514521527> *

NUNES, Desnes et al. Evaluating GPT-3.5 and GPT-4 Models on Brazilian University Admission Exams. arXiv preprint arXiv:2303.17003, 2023. Disponível em: <https://arxiv.org/abs/2303.17003> *

OECD. Results from TALIS 2024 - Country notes: Brazil. Paris: OECD Publishing, 2025. Disponível em: https://www.oecd.org/en/publications/results-from-talis-2024-country-notes_e127f9e2-en/brazil_1e93d3b5-en.html *

OLHAR DIGITAL. Educação: governo de São Paulo anuncia uso de IA na formulação de aulas. YouTube, 17 abr. 2024. Disponível online: https://youtu.be/is-xwBfx5WI?si=ZcWY-0Ndq_XiPsDv *

OLIVEIRA, W.; BITTENCOURT, I. I.; DERMEVAL, D.; ISOTANI, S. Gamificação e Informática na Educação. In: SAMPAIO, F. F.; PIMENTEL, M.; SANTOS, E. O. (org.). **Informática na Educação**: games, inteligência artificial, realidade virtual/aumentada e computação ubíqua. Porto Alegre: Sociedade Brasileira de Computação, 2021. Disponível em: <https://ieducacao.ceie-br.org/gamificacao> *

OPENAI. OpenAI API. 11 jun. 2020a. Disponível em: <https://openai.com/index/openai-api> *

OPENAI. GPT-3 Model Card. 18 set. 2020b. Disponível em: <https://github.com/openai/gpt-3/blob/master/model-card.md> *

OPENAI. New GPT-3 capabilities: Edit & insert. 15 mar. 2022a. <https://openai.com/index/gpt-3-edit-insert> *

OPENAI. Introducing ChatGPT. 30 nov. 2022b. <https://openai.com/index/chatgpt> *

OPENAI. Introducing ChatGPT Plus. 1 fev. 2023a. Disponível em: <https://openai.com/index/chatgpt-plus> *

OPENAI. GPT-4 Technical Report. 1 mar. 2023b. Disponível em: <https://arxiv.org/pdf/2303.08774v5> *

OPENAI. GPT-4. 14 mar. 2023c. Disponível em: <https://openai.com/index/gpt-4-research> *

OPENAI. GPT-4 System Card. 23 mar. 2023d. Disponível em: <https://cdn.openai.com/papers/gpt-4-system-card.pdf> *

OPENAI. Improving mathematical reasoning with process supervision. Research, 31 mar. 2023e. Disponível em: <https://openai.com/research/improving-mathematical-reasoning-with-process-supervision> *

OPENAI. ChatGPT & Teste de Turing. ChatGPT-4, 2 jul. 2023f. Disponível online: <https://chat.openai.com/share/2e02d3d2-f460-47c0-8800-ec096f6874e9> *

OPENAI. Dall-e 3. 20 set. 2023g. Disponível em: <https://openai.com/index/dall-e-3> *

OPENAI. Presidente do Brasil. ChatGPT-3.5, 6 out. 2023h. Disponível online: <https://chat.openai.com/share/d26cdc70-524a-47f3-8dab-9464cadabc0a> *

OPENAI. Modelo de Linguagem GPT-4. ChatGPT-4, 7 out. 2023i. Disponível em: <https://chat.openai.com/share/d4229781-7114-4e49-8016-45f2e79ae277> *

OPENAI. Introducing GPTs. 6 nov. 2023j. Disponível em: <https://openai.com/index/introducing-gpts> *

OPENAI. Termos de Uso. 14 nov. 2023l. Disponível em: <https://openai.com/pt-BR/policies/terms-of-use> *

OPENAI. Introducing the GPT Store. OpenAI, 10 jan 2024a. Disponível em: <https://openai.com/blog/introducing-the-gpt-store> *

OPENAI. Tutor socrático. OpenAI, 5 fev 2024b. Disponível em: <https://chatgpt.com/share/7759fd29-1e3a-4eb8-990b-81e2853f3516> *

OPENAI. Hello GPT-4o. OpenAI, 13 maio 2024c. Disponível em: <https://openai.com/index/hello-gpt-4o> *

OPENAI. Introducing ChatGPT Edu. OpenAI, 30 maio 2024d. Disponível em: <https://openai.com/index/introducing-chatgpt-edu> *

OPENAI. Introducing OpenAI o1-preview. OpenAI, 12 set. 2024e. Disponível em: <https://openai.com/index/introducing-openai-o1-preview> *

OPENAI. Learning to Reason with LLMs. OpenAI, 12 set. 2024f. Disponível em: <https://openai.com/index/learning-to-reason-with-llms> *

OPENAI. Pesquisadores Mariano e Felipe. 11 out. 2024g. Disponível em: <https://chatgpt.com/share/67092d3b-d904-800a-8c70-2bf1fd770cfb> *

OPENAI. Cibercultura. ChatGPT-o1. 17 dez. 2024h. Disponível em: <https://chatgpt.com/share/67615a68-d570-800a-955e-1801465646c4> *

OPENAI. Afirmações. ChatGPT-4o. 17 dez. 2024i. Disponível em: <https://chatgpt.com/share/6761b463-41d0-800a-8d23-937b78ca6392> *

- OPENAI. Paralelepípedo. ChatGPT-4o. 16 jan. 2025a. Disponível em: <https://chatgpt.com/share/6788feed-d938-800a-b9f5-8e7c8adedfda> *
- OPENAI. Imagestesia. ChatGPT-4o. 29 jan. 2025b. Disponível em: <https://chatgpt.com/share/679a3f84-8c7c-800a-9c47-6f5320143a5d> *
- OPENAI. OpenAI o3-mini. OpenAI, 31 jan. 2025c. Disponível em: <https://openai.com/index/openai-o3-mini> *
- OPENAI. Equação do segundo grau. ChatGPT-o1. 2 fev. 2025d. Disponível em: <https://chatgpt.com/share/67a01cf5-0200-800a-bacc-f244babf22cf> *
- OPENAI. Introducing GPT-4.5. OpenAI, 27 fev. 2025e. Disponível em: <https://openai.com/index/introducing-gpt-4-5> *
- OPENAI. Educação e IA. OpenAI, 15 mar. 2025f. Disponível em: <https://chatgpt.com/share/67d57da9-2f70-800a-b8ce-3c6a40166e56> *
- OPENAI. Apresentando OpenAI o3 e o4-mini. 16 abr. 2025g. Disponível em: <https://openai.com/index/introducing-o3-and-o4-mini> *
- OPENAI. Enem 2024 – GPT-o1. 18 abr. 2025h. Disponível em: <https://chatgpt.com/share/6802ab96-cc88-800a-b299-fe1b5e2b63a1> *
- OPENAI. Enem 2024 – GPT-4.5. 25 abr. 2025i. Disponível em: <https://chatgpt.com/share/6802a91b-aaf0-800a-8210-78d6828cb3df> *
- OPENAI. Resumo do livro IA generativa e Educação – GPT-o3. 29 maio 2025j. Disponível em: <https://chatgpt.com/share/6838c46e-e0bc-800a-8543-0957a25e1848> *
- OPENAI. Enem 2024 – GPT-o3. 3 jun. 2025l. Disponível em: <https://chatgpt.com/c/6862df38-79b8-800a-aa13-c5dac1b666a4> *
- OPENAI. Imaginando futuros possíveis da educação com IA. 6 jul. 2025m. Disponível em: <https://chatgpt.com/share/686ae0cb-ef80-800a-8e32-e670394b85a9> *
- OPENAI. Introducing GPT-5. 7 ago. 2025n. Disponível em: <https://openai.com/pt-BR/index/introducing-gpt-5> *
- OPENAI. Enem 2024 – GPT-5 thinking. 5 set. 2025o. Disponível em: <https://chatgpt.com/share/68bb3e87-6d9c-800a-be32-4d6eb4a240ed> *
- PAUSA nos experimentos gigantes de IA: uma carta aberta. Future of life, 22 mar. 2023. Disponível em: <https://futureoflife.org/pt-br/carta-aberta/pausar-experimentos-gigantes-de-ia> *
- PERRIGO, Billy. Exclusive: OpenAI used kenyan workers on less than \$2 per hour to make ChatGPT less toxic. Time, 18 jan. 2023. Disponível online: <https://time.com/6247678/openai-chatgpt-kenya-workers> *
- PESQUISA TIC Educação 2022: Pesquisa sobre o uso das Tecnologias de Informação e Comunicação nas escolas brasileiras. CETIC.br, nov. 2022. Disponível em:

<https://www.cetic.br/pt/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-nas-escolas-brasileiras-tic-educacao-2022> *

PIMENTEL, Mariano. Linguagem Logo: teoria e prática. Canal Didática Online [YouTube]. Disponível em: <https://www.youtube.com/watch?v=qQXmMkJz8AM> *

PIMENTEL, Mariano; AZEVEDO, Viviane; CARVALHO, Felipe. ChatGPT substituirá professoras e professores? **SBC Horizontes**, 10 mar. 2023a. Disponível em: <http://horizontes.sbc.org.br/index.php/2023/03/chatgpt-substituira-professoras-e-professores> *

PIMENTEL, Mariano; AZEVEDO, Viviane; CARVALHO, Felipe. ChatGPT é realmente inteligente? **SBC Horizontes**, 17 mar. 2023b. Disponível em: <http://horizontes.sbc.org.br/index.php/2023/03/chatgpt-e-realmente-inteligente> *

PIMENTEL, Mariano; AZEVEDO, Viviane; CARVALHO, Felipe. ChatGPT: a era da autoria híbrida humana/o-IA. **SBC Horizontes**, 21 mar. 2023c. Disponível em: <http://horizontes.sbc.org.br/index.php/2023/03/chatgpt-a-era-da-autoria-hibrida> *

PIMENTEL, Mariano; BERINO, Aristóteles de Paula. GPT Paulo Freire: conversando com uma simulação do patrono da educação brasileira. **Revista Artes de Educar**, 2025.

PIMENTEL, Mariano; CARVALHO, Felipe. Princípios da Educação Online: para sua aula não ficar massiva nem maçante! **SBC Horizontes**, maio 2020. Disponível em: <http://horizontes.sbc.org.br/index.php/2020/05/23/principios-educacao-online> *

PIMENTEL, Mariano; CARVALHO, Felipe. Instrução (re)programada, máquinas (digitais em rede) de ensinar e a pedagogia (ciber)tecnicista. **SBC Horizontes**, 26 jul. 2021. Disponível em: <http://horizontes.sbc.org.br/index.php/2021/07/maquinas-de-ensinar> *

PIMENTEL, Mariano; CARVALHO, Felipe. Cibertecnicismo. **Revista de Educação Pública**, v. 31, n. jan-dez, p. 1–22, 2022. Disponível em: <https://periodicoscientificos.ufmt.br/ojs/index.php/educacaopublica/article/view/13919> *

PIMENTEL, Mariano; CARVALHO, Felipe. ChatGPT: concepções epistêmico-didático-pedagógicas dos usos na educação. **SBC Horizontes**, 6 jun. 2023a. Disponível em: <http://horizontes.sbc.org.br/index.php/2023/06/chatgpt-concepcoes> *

PIMENTEL, Mariano; CARVALHO, Felipe. Implicações e teorizações dos usos das IA generativas na educação. **Revista Brasileira de Cultura (CULT)**. São Paulo - SP, p. 20 - 22, 01 set. 2023b. Disponível em: <https://revistacult.uol.com.br/home/ia-generativas-educacao> *

PIMENTEL, Mariano; CARVALHO, Felipe. Autoria em tempos de inteligência artificial generativa: implicações para (re)pensar a educação. In: PORTO, C.; VASCONCELOS, A. D.; LINHARES, R. (OrgS.). **Educiber**: educação e inteligência

- artificial: travessias. Aracaju: Editora Universidade Tiradentes, 2024, v. 6, p. 193-216. Disponível em: <https://doi.org/10.17564/2022.88303-32-0> *
- PIMENTEL, Mariano; CARVALHO, Felipe; CHATGPT-4, OpenAI. ChatGPT: potencialidades e riscos para a Educação. **SBC Horizontes**, 8 maio 2023. Disponível em: <http://horizontes.sbc.org.br/index.php/2023/05/chatgpt-potencialidades-e-riscos-para-a-educacao> *
- PIMENTEL, Mariano; CARVALHO, Felipe; SILVEIRA, Victor Junger. IA Generativa pode ser coautora? **Triade**: Comunicação, Cultura e Mídia, Sorocaba, SP, v. 12, n. 25, p. 1-23, 2024. Disponível em: <https://periodicos.uniso.br/triade/article/view/5569> *
- PIRES, Ramon; ALMEIDA, Thales Sales; ABONIZIO, Hugo; NOGUEIRA, Rodrigo. Evaluating GPT-4's vision capabilities on brazilian university admission exams. arXiv:2311.14169v1 [cs.CL] 23 nov 2023. Disponível online: <https://arxiv.org/pdf/2311.14169> *
- PLATÃO. **A República**. Introdução, tradução e notas de Maria Helena da Rocha Pereira, 9ª ed. Lisboa: Fundação Calouste Gulbenkian, 1996.
- PLATO (computer system). **Wikipédia**, a enciclopédia livre. Disponível em: [https://en.wikipedia.org/wiki/PLATO_\(computer_system\)](https://en.wikipedia.org/wiki/PLATO_(computer_system)) *
- PORTO, Cristiane; SANTOS, Edméa; BOTTENTUIT JR., João Batista. (Orgs.) **ChatGPT e outras inteligências artificiais**: práticas educativas na Ciberultura. São Luís: EDUFMA, 2024. Disponível online: <https://www.edufma.ufma.br/index.php/produto/chatgpt-e-outras-inteligencias-artificiais-praticas-educativas-na-cibercultura> *
- POSSA, Julia. NY derruba proibição ao ChatGPT e passa a incentivar IA nas escolas. *_Tecnologia, giz_br*, UOL, 20 maio 2023. Disponível em: <https://gizmodo.uol.com.br/ny-derruba-proibicao-ao-chatgpt-e-passa-a-incentivar-ia-nas-escolas> *
- PRADELLA, Lyncon. Aumento de 5% e IA regulamentada: veja como foi acordo de roteiristas para encerrar greve. CNN Brasil, Florianópolis, 27 set. 2023. Disponível em: <https://www.cnnbrasil.com.br/entretenimento/aumento-de-5-e-ia-regulamentada-veja-como-foi-acordo-de-roteiristas-para-encerrar-greve> *
- QWEN Team. Qwen2.5-Max: Exploring the Intelligence of Large-scale MoE Model. 28 jan. 2025. Disponível em: <https://qwenlm.github.io/blog/qwen2.5-max> *
- RADFORD, Alec et al. Language models are unsupervised multitask learners. OpenAI, 14 fev. 2019. Disponível em: https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf *
- RADFORD, Alec; NARASIMHAN, Karthik; SALIMANS, Tim; SUTSKEVER, Ilya. Improving language understanding by generative pre-training. 2018. Disponível em: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf *

- RAULINO, Fabiana; RAMOS, Reinaldo. 30 formas de utilizar o ChatGPT em Aula. Disponível em: <https://view.genial.ly/63d9bbc140e03c0017ca1603> *
- REIN, David; HOU, Betty Li; STICKLAND, Asa Cooper et al. GPQA: A Graduate-Level Google-Proof Q&A Benchmark. arXiv, 20 nov. 2023. Disponível em: <https://arxiv.org/pdf/2311.12022> *
- RIBEIRO, Tiago; SKLIAR, Carlos. Escolas, pandemia e conversação: notas sobre uma educação inútil. **Série-Estudos**: Periódico do Programa de Pós-Graduação em Educação da UCDB, p. 13-30, 2020. Disponível em: <https://serie-estudos.ucdb.br/serie-estudos/article/view/1484> *
- ROBÔ. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, s. d. Disponível em: <https://pt.wikipedia.org/wiki/Rob%C3%B4> *
- ROOSE, Kevin. The Brilliance and Weirdness of ChatGPT. The New York Times, 5 dez. 2022. Disponível em: <https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html> *
- ROSENBLATT, Frank. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, v. 65, n. 6, p. 386, 1958.
- ROTH, Emma. Microsoft spent hundreds of millions of dollars on a ChatGPT supercomputer. The Verge, 13 mar. 2023. Disponível em: <https://www.theverge.com/2023/3/13/23637675/microsoft-chatgpt-bing-millions-dollars-supercomputer-openai> *
- RUSSEL, Stuart J.; NORVIG, Peter. **Inteligência Artificial**: uma abordagem moderna. 4ª ed. [edição Kindle]. Rio de Janeiro: GEN, 2022.
- SACKS, Harvey; SCHEGLOFF, Emanuel A.; JEFFERSON, Gail. Sistemática elementar para a organização da tomada de turnos para a conversa. **Veredas-Revista de Estudos Linguísticos**, v. 7, n. 1 e 2, 2003.
- SAMPAIO, Maria M. F.; MARIN, Alda J. Precarização do trabalho docente e seus efeitos sobre as práticas curriculares. **Educação & Sociedade**, Campinas, v. 25, n. 89, p. 1203-1225, set./dez. 2004. Disponível em: <https://www.scielo.br/j/es/a/t7pjz85czHRW3GcKpB9dmNb/?format=pdf&lang=pt> *
- SANTAELLA, Lucia. **Navegar no ciberespaço**: o perfil cognitivo do leitor imersivo. São Paulo: Paulus, 2004.
- SANTAELLA, Lucia. **Comunicação ubíqua**: repercussões na cultura e na educação. São Paulo: Paulus, 2013.
- SANTAELLA, Lucia. **Neo-humano**: a sétima revolução cognitiva do Sapiens. São Paulo: Paulus, 2022.
- SANTAELLA, Lúcia. **A inteligência artificial é inteligente?** São Paulo: Edições 70, 2023a.
- SANTAELLA, Lúcia. Por que é imprescindível um manual ético para a Inteligência Artificial Generativa? **TECCOGS – Revista Digital de Tecnologias Cognitivas**,

- São Paulo, n. 28, p. 7-24, 2023b. Disponível em: <https://doi.org/10.23925/1984-3585.2023i28p7-24> *
- SANTAELLA, Lucia. Balanço crítico do ChatGPT: contribuições da semiótica para a IA [curso remoto]. Espaço Cult, 2023c. Disponível em: <https://www.cultloja.com.br/produto/semiotica-inteligencia-artificial> *
- SANTAELLA, Lucia. IA generativa e o perfil semiótico-cognitivo do leitor iterativo. Sociotramas [blog], 2 jan. 2024a. Disponível em: <https://sociotramas.wordpress.com/2024/01/02/ia-generativa-e-o-perfil-semiotico-cognitivo-do-leitor-iterativo> *
- SANTAELLA, Lucia. A criatividade nos prismas da Inteligência Artificial Generativa. **Triade: Comunicação, Cultura e Mídia**, Sorocaba, SP, v. 12, n. 25, p. 1-21, 2024b. DOI: 10.22484/2318-5694.2024v12i25588. Disponível em: <https://periodicos.uniso.br/triade/article/view/5588> *
- SANTAELLA, Lucia; KAUFMAN, Dora. A inteligência artificial generativa como quarta ferida narcísica do humano. **MATRIZES**, São Paulo, v. 18, n. 1, p. 37-53, jan./abr. 2024. DOI: 10.11606/issn.1982-8160.v18i1p37-53.
- SANTOS, Edméa Oliveira dos. **Educação online: cibercultura e pesquisa-formação na prática docente**. Tese de Doutorado. Bahia, UFBA/FACED, 2005.
- SANTOS, Edméa. **Pesquisa-formação na cibercultura**. Teresina: EDUFPI, 2019. Disponível online: http://www.edmeasantos.pro.br/assets/livros/Livro%20PESQUISA-FORMA%C3%87%C3%83O%20NA%20CIBERCULTURA_E-BOOK.pdf
- SANTOS, Edméa. As pessoas estão falando tanta besteira sobre o ChatGPT, IA. Facebook, 17 mar. 2023. Disponível em: <https://www.facebook.com/edmea.santos/posts/6503379073028292> *
- SANTOS, Edméa; CHAGAS, Alexandre; BOTTENTUIT JR., João Batista. (Orgs.) **ChatGPT e educação na cibercultura: fundamentos e primeiras aproximações com Inteligência Artificial**. São Luís: EDUFMA, 2024. Disponível online: <https://www.edufma.ufma.br/index.php/produto/chatgpt-e-educacao-na-cibercultura-fundamentos-e-primeiras-aproximacoes-com-inteligencia-artificial> *
- SANTOS, Rosemary dos; CARVALHO, Felipe da Silva Ponte de; MADDALENA, Tânia Lúcia. Conversas ubíquas via Whatsapp: ambiências formativas multirreferenciais. In: PORTO, C.; OLIVEIRA, K. E.; CHAGAS, A. (Org.). **Whatsapp e educação: entre mensagens, imagens e sons**. Ilhéus, BH: EDUFBA, 2017, v. 1, p. 193-214.
- SARAIVA, Karla. A Babel eletrônica: hospitalidade e tradução no ciberespaço. In: SKLIAR, Carlos (org.). **Derrida & a educação**. Belo Horizonte: Autêntica, 2005. p. 43-58.
- SAVIANI, Dermeval. **Escola e Democracia**. 1. ed. Campinas: Cortez; Autores Associados, 1986.

- SAVIANI, Dermeval. **Escola e democracia**: polêmicas do nosso tempo. 32. ed. Campinas: Autores Associados, 1999.
- SEARLE, John R. Minds, Brains and Programs. **Behavioral and Brain Sciences**, n. 3, v.3, p. 417–457, 1980.
- SELFIES da macaca. In: Wikipedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: https://pt.wikipedia.org/wiki/Selfies_da_macaca *
- SILVA, Amós Coelho da; MONTAGNER, Aírto Ceolin. **Dicionário latino-português**: etimologia, gramática, derivações, exemplos. 2 ed. Petrópolis: Vozes, 2012.
- SILVA, João Augusto Ramos; BERNARDO JÚNIOR, Ronaldo; OLIVEIRA, Fátima Bayma. Abandono e conclusão de alunos inscritos em cursos MOOC. In: Anais do 20º Congresso Internacional ABED de Educação a Distância. 2014.
- SILVA, Marco. **Sala de aula interativa**. Rio de Janeiro: Quartet, 2000.
- SILVA, Marco. Interatividade na educação híbrida. In: PIMENTEL, M.; SANTOS, E.; SAMPAIO, F. F. (org.). **Informática na educação**: interatividade, metodologias e redes. Porto Alegre: Sociedade Brasileira de Computação, 2021.
- SILVA, Tomaz T. da et al. **A produção social da identidade e da diferença. Identidade e diferença**: a perspectiva dos estudos culturais. Petrópolis (RJ): Vozes, 2000, p. 73-102.
- SILVA, Tomaz Tadeu da. **Documentos de Identidade**: uma introdução às teorias do currículo. Belo Horizonte: Autêntica, 1999.
- SIMON, Hebert A. **Models of Man**: Social and Rational. John Wiley, 1957.
- SKINNER, Burrhus Frederic. Teaching machine and programmed learning, 1956. Disponível em: <https://www.youtube.com/watch?v=CFYruzWeFwQ> *
- SKINNER, Burrhus Frederic. Teaching Machines. **Science**, New Series, v.128, n.3330, out. 1958, p. 969-977. Disponível em: <https://www.science.org/doi/10.1126/science.128.3330.969> *
- SKINNER, Burrhus Frederic. **Tecnologia do ensino**. São Paulo: Edusp, 1972.
- SONTAG, Susan. **Sobre fotografia**. São Paulo: Companhia das Letras, 2004.
- SOUZA, Renata. Racismo algorítmico! Publicação no Instagram, 2023. Disponível em: https://www.instagram.com/p/Cy3jaTnuYWe/?img_index=1 *
- SPINAK, Ernesto. Inteligência Artificial e a comunicação da pesquisa. Scielo em Perspectiva. 30 ago. 2023. Disponível online: <<https://blog.scielo.org/blog/2023/08/30/inteligencia-artificial-e-a-comunicacao-da-pesquisa> *
- STOKEL-WALKER, Chris. ChatGPT listed as author on research papers: many scientists disapprove. Nature, News, 18 jan. 2023. Disponível em: <https://www.nature.com/articles/d41586-023-00107-z> *

- TARDIF, Maurice. **Saberes docentes e formação profissional**. Petrópolis: Vozes, 2002.
- TAY (chatbot). In: Wikipedia. Flórida: Wikimedia Foundation, [s. d.]. Disponível em: [https://en.wikipedia.org/wiki/Tay_\(chatbot\)](https://en.wikipedia.org/wiki/Tay_(chatbot)) *
- TEST scores of AI systems on various capabilities relative to human performance. Our World in Data, 2024. Disponível em: <https://ourworldindata.org/grapher/test-scores-ai-capabilities-relative-human-performance> *
- THE WHITE HOUSE (EUA). Fact sheet: president Biden issues executive order on safe, secure, and trustworthy artificial intelligence. Washington, D.C., 30 out. 2023. Disponível em: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence> *
- THÉÂTRE D'opéra Spatial. In: Wikimedia Commons. Disponível em: https://en.wikipedia.org/wiki/Théâtre_D'opéra_Spatial *
- TURING, Alan. Computing Machinery and Intelligence. **Mind**, LIX, v. 236, p. 433-460, 1950.
- TYLDUM, Morten. O jogo da imitação. EUA: Black Bear Pictures, 2014. 115 min.
- ULAM, Stanislaw. John von Neumann 1903-1957. **Bulletin of the American Mathematical Society**, n. 64, p. 1-49, 1958. Disponível em: <https://www.ams.org/journals/bull/1958-64-03/S0002-9904-1958-10189-5> *
- UNBOUND ACADEMIC INSTITUTE. Educational Plan, 12 fev. 2024. Disponível online: https://drive.google.com/file/d/1uEzeZCe-GSjBGN1_xxpN-3zEq1TAkylY/view *
- UNESCO. **ChatGPT e Inteligência Artificial na educação superior**: guia de início rápido. 2023. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000385146_por *
- UNESCO. **Marco referencial de competências em IA para professores**. 2025a. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000394280> *
- UNESCO. **Marco referencial de competências em IA para estudantes**. 2025b. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000394281> *
- UNIÃO EUROPEIA. Diretiva (UE) 2019/790 do Parlamento Europeu e do Conselho, de 17 de abril de 2019, relativa aos direitos de autor e direitos conexos no mercado único digital [...]. **Jornal Oficial da União Europeia**, Bruxelas, 2019. Disponível em: <http://data.europa.eu/eli/dir/2019/790/oj> *
- UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024 que cria regras harmonizadas em matéria de inteligência artificial [...]. **Jornal Oficial da União Europeia**, Bruxelas, 2024. Disponível em: <http://data.europa.eu/eli/reg/2024/1689/oj> *
- VALENTE, José Armando. Análise dos diferentes tipos de softwares usados na Educação. In: VALENTE, J. A. (org.). **O computador na sociedade do**

- conhecimento** [Também reformatado e republicado pelo MEC]. Campinas: Unicamp/NIED, 1999.
- VALENTE, José. Diferentes usos do computador na educação. In: VALENTE, J. A. (org.). **Computadores e conhecimento**: repensando a educação. Campinas, SP: Unicamp: Nied, 1993. p. 1-28.
- VARANASI, Lakshmi. ChatGPT creator OpenAI is in talks to sell shares in a tender offer that would double the startup's valuation to \$29 billion. Insider, 5 jan. 2023. Disponível em: <https://www.businessinsider.com/chatgpt-creator-openai-talks-for-tender-offer-at-29-billion-2023-1> *
- VASWANI, Ashish et al. Attention is all you need. Computation and Language, 2017. Disponível em: <https://doi.org/10.48550/arXiv.1706.03762> *
- VEEN, Fjodor van. The Neural Network Zoo. The Asimov Institute, 14 set. 2016. Disponível em: <https://www.asimovinstitute.org/neural-network-zoo> *
- VERMA, Pranshu; TAN, Shelly. A bottle of water per email: the hidden environmental costs of using AI chatbots. The Washington Post, 18 set. 2024. Disponível em: <https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers> *
- VICARI, R. M. Inteligência Artificial aplicada à Educação. In: PIMENTEL, Ma-riano; SAMPAIO, F. F.; SANTOS, E. O. (org.). **Informática na Educação**: games, inteligência artificial, realidade virtual/aumentada e computação ubíqua. Porto Alegre: Sociedade Brasileira de Computação, 2021. Disponível em: <https://ieducacao.ceie-br.org/inteligenciaar-tifical> *
- VICARI, Rosa; BRACKMANN, Christian; MIZUSAKI, Lucas; LOPES, Daniel; BARONE, Dante; CASTRO, Henrique. **Referencial Curricular**: Inteligência Artificial no Ensino Médio. 2022. Disponível em: <https://www.inf.ufrgs.br/ciars/index.php/aceso-ao-curso> *
- VILLALOBOS, Pablo; HO, Anson; SEVILLA, Jaime; BESIROGLU, Tamay; HEIM, Lennart; HOBBAHN, Marius. Will we run out of data? Limits of LLM scaling based on human-generated data. arXiv, 4 jun 2024. Disponível em: <https://arxiv.org/pdf/2211.04325> *
- VINCHON, Florent et al. Artificial intelligence & creativity: a manifesto for collaboration. **The Journal of Creative Behavior**, n. 57, v. 4, p. 472–484, 2023. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/jocb.597> *
- VINGE, Vernor. The coming technological singularity: how to survive in the post-human era. Proceeding of symposium Vision-21: Interdisciplinary Science and Engineering in the Era of Cyberspace. Ohio, NASA, 1993, p. 11-22.
- WALSH, Matt. ChatGPT Statistics (2023) — The Key Facts and Figures. Style Factory, 12 abr. 2023. Disponível em: <https://www.stylefactoryproductions.com/blog/chatgpt-statistics> *
- WEIZENBAUM, Joseph. **Computer power and human reason**: from judgment to calculation. New York: W. H. Freeman, 1976.

- WEIZENBAUM, Joseph. ELIZA—a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 9, n. 1, p. 36-45, 1966.
- WIENER, Norbert. **Cybernetics**; or, Control and communication in the animal and the machine. Cambridge, MA: The MIT Press, 2019 [1948].
- YUE, Xiang et al. MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI. 2024. Disponível online: <https://mmmu-benchmark.github.io/#leaderboard> *
- ZARIPOV, Rudolf. Об алгоритмическом описании процесса сочинения музыки [Sobre a descrição algorítmica do processo de composição musical]. Academia de Ciências da USSR. 1960.

*Todos os links das referências foram acessados e checados em 23 out. 2025.

Contribuidores

Mariano Pimentel – conceituação, escrita, imagens e revisão do texto.

Felipe Carvalho – conceituação, escrita e revisão do texto.

Lucia Castro – revisão final do texto.

Monica Lopes – capa e ilustrações.

GPT 3.5, 4o, o1, o3, 4.5 e 5 – conceituação, escrita, imagens e revisão do texto.

Professoras/es que realizaram cursos e oficinas conosco sobre IA generativa e educação – discussão dos conceitos e revisão do texto.

Sobre os autores



Mariano Pimentel é doutor em Informática e atua como professor na UNIRIO (Universidade Federal do Estado do Rio de Janeiro). Recebeu prêmio Jabuti em 2012 pelo livro *Sistemas Colaborativos*. Realiza pesquisas em sistemas de informação, educação, cibercultura e colaboração.

Currículo Lattes: <http://lattes.cnpq.br/1920411639358905>

G.Acadêmico: https://scholar.google.com/citations?user=_UDWLSwAAAAJ

Felipe Carvalho é doutor em Educação e atua como professor no Programa de Pós-Graduação em Educação da Universidade Estácio de Sá (PPGE/UNESA) e no Programa de Pós-Graduação em Ensino de Ciências e Saúde da Afya UNIGRANRIO (PPGECS). É coordenador do Grupo de Pesquisa em Educação e Cibercultura (GPEC/CNPq). Realiza pesquisas em informática na educação, formação de professores, didática e cibercultura.



Currículo Lattes: <http://lattes.cnpq.br/8539464540238508>

G.Acadêmico: <https://scholar.google.com/citations?user=ENw6bhYAAAAJ>