



DOSSIÊ HISTÓRIA DIGITAL E HISTÓRIA DIGITAL DA EDUCAÇÃO

## Inteligência artificial no ensino de História: ChatGPT e ditadura militar brasileira

*Artificial Intelligence in the Teaching of History: ChatGPT and the Brazilian Military Dictatorship*

*Inteligencia artificial en la enseñanza de Historia: ChatGPT y la dictadura militar brasileña*

**Fernando Cauduro**

**Pureza<sup>1</sup>**

[orcid.org/0000-0001-6363-0753](https://orcid.org/0000-0001-6363-0753)  
[fcpureza@gmail.com](mailto:fcpureza@gmail.com)

**Felipe Cittolin Abal<sup>2</sup>**

[orcid.org/0000-0002-6208-5893](https://orcid.org/0000-0002-6208-5893)  
[felipeabal@upf.br](mailto:felipeabal@upf.br)

**Recebido em:** 26 ago. 2024.

**Aprovado em:** 15 maio 2025.

**Publicado em:** 20 maio. 2026.

**Resumo:** Este artigo analisa o uso de inteligência artificial (IA) no ensino de História, focando especificamente no tema da ditadura militar brasileira. Através de um experimento comparativo utilizando três versões do ChatGPT – GPT 3.5, GPT 4o e um GPT customizado denominado “Professor de História” –, o estudo examina a capacidade desses modelos de linguagem em responder questões sobre esse passado sensível. O experimento realizado procurou avaliar não apenas a precisão e a objetividade, mas também possíveis vieses nas respostas geradas, discutindo as implicações éticas do uso de IAs no ensino de temas históricos sensíveis. Os resultados indicam que, embora as IAs demonstrem capacidade de fornecer informações básicas corretas, elas apresentam limitações significativas na análise crítica e contextualização histórica, reforçando a importância do papel do educador na mediação do uso dessas tecnologias no ensino de História.

**Palavras-chave:** inteligência artificial; ensino de História; ditadura militar brasileira; ChatGPT.

**Abstract:** This article analyzes the use of Artificial Intelligence (AI) in the teaching of History, focusing specifically on the topic of the Brazilian military dictatorship. Through a comparative experiment using three versions of ChatGPT – GPT 3.5, GPT 4o, and a customized GPT called “History Teacher” –, the study examines the ability of these language models to answer questions about this sensitive past. The experiment aimed to evaluate not only the accuracy and objectivity but also potential biases in the generated responses, discussing the ethical implications of using AIs in the teaching of sensitive historical topics. The results indicate that although AIs demonstrate the ability to provide correct basic information, they show significant limitations in critical analysis and historical contextualization, reinforcing the importance of the educator’s role in mediating the use of these technologies in the teaching of History.

**Keywords:** artificial intelligence; History teaching; Brazilian military dictatorship; ChatGPT.

**Resumen:** Este artículo analiza el uso de la Inteligencia Artificial (IA) en la enseñanza de la Historia, enfocándose específicamente en el tema de la dictadura militar brasileña. A través de un experimento comparativo utilizando tres versiones de ChatGPT: GPT 3.5, GPT 4o y un GPT personalizado denominado “Profesor de Historia”, el estudio examina la capacidad de estos modelos de lenguaje para responder preguntas sobre este pasado sensible. El experimento realizado procuró evaluar no solo la precisión y objetividad, sino también posibles sesgos en las respuestas generadas, discutiendo las implicaciones éticas del uso de IA en la enseñanza de temas históricos sensibles. Los resultados indican que, aunque las IA demuestran capacidad para proporcionar información básica correcta, presentan limitaciones significativas en el análisis crítico y la contextualización histórica, lo que refuerza la importancia del papel del educador en la mediación del uso



Este é um artigo de acesso aberto distribuído sob a licença [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/), que permite a cópia e redistribuição do material em qualquer formato e para qualquer finalidade, desde que a autoria original e os créditos de publicação sejam mantidos.

<sup>1</sup> Universidade Federal da Paraíba (UFPB), João Pessoa, Paraíba, Brasil.

<sup>2</sup> Universidade de Passo Fundo (UPF), Passo Fundo, Rio Grande do Sul, Brasil.

de estas tecnologias en la enseñanza de la Historia.

**Palabras clave:** inteligencia artificial; enseñanza de Historia; ditadura militar brasileña; ChatGPT.

## 1 Introdução

A popularização das IAs vem trazendo diversos impactos nos âmbitos da sociedade, incluindo a educação, e gerando debates, especialmente, entre aqueles que acreditam que as IAs podem ser amplamente utilizadas e os que pensam não haver espaço para elas na sala de aula. No campo do ensino de História, o advento de modelos de linguagem avançados, como o ChatGPT, apresenta oportunidades e desafios significativos. Este artigo se propõe a investigar as potencialidades e limitações do uso de IA<sup>3</sup> no ensino utilizando como paradigma um tema histórico particularmente sensível: a ditadura militar brasileira.

A escolha desse período histórico como foco de análise se deve ao fato de que a ditadura militar brasileira continua sendo um tema de intenso debate público e acadêmico, marcado por interpretações divergentes e disputas de narrativas. Nesse contexto, o uso de ferramentas de IA no ensino de História levanta questões cruciais sobre objetividade, viés histórico e a própria natureza do conhecimento histórico.

Este estudo realiza um experimento comparativo utilizando três versões do ChatGPT: o GPT 3.5, o GPT 4o e um GPT customizado, denominado "Professor de História", alimentado com obras historiográficas específicas sobre o período. A escolha específica pelo modelo GPT se deu por sua popularidade entre estudantes dos Ensinos Médio e Superior e pela possibilidade de comparar versões diferentes do mesmo modelo. Outros modelos, como o Claude ou o Gemini, poderiam oferecer resultados distintos em virtude de suas particularidades.

Através de uma série de perguntas elaboradas, buscamos avaliar a capacidade desses modelos de linguagem em fornecer informações precisas, contextualizadas e livres de vieses sobre a ditadura militar brasileira<sup>4</sup>. Importante destacar desde este momento que as perguntas foram elaboradas tendo em mente um jovem sem conhecimento de engenharia de *prompts* e que utilizaria a ferramenta como um buscador, o que é bastante comum. Assim, as perguntas são genéricas e deixam de fornecer detalhes, contextos, informações de uso etc.

Nossa análise se desdobra em três eixos principais: primeiro, examinamos a precisão factual das respostas geradas pelos modelos de IA; segundo, avaliamos a capacidade desses modelos em refletir debates historiográficos contemporâneos sobre o período; terceiro, investigamos possíveis vieses ou tendências interpretativas nas respostas fornecidas.

Ao realizar esta investigação, nosso objetivo é não apenas avaliar o estado atual da tecnologia de IA aplicada ao ensino de História, mas também refletir sobre as implicações éticas e pedagógicas do uso dessas ferramentas. Questionamos: como garantir que o uso de IA no ensino de História não reforce narrativas simplificadas ou tendenciosas sobre períodos históricos complexos?

Este artigo se insere em um debate mais amplo sobre o futuro do ensino de História na era digital, contribuindo para uma reflexão crítica sobre desafios e oportunidades apresentados pela IA na construção e na transmissão do conhecimento histórico. Ao fim, esperamos fornecer *insights* para educadores e pesquisadores interessados na intersecção entre IA e ensino de História.

<sup>3</sup> Para fins deste artigo, entendemos inteligência artificial (IA) como um conjunto de tecnologias computacionais capazes de executar tarefas associadas à inteligência humana, como aprendizagem, reconhecimento de padrões e geração de linguagem natural. Nosso foco são as IAs generativas, modelos que conseguem gerar conteúdos.

<sup>4</sup> Há um longo debate historiográfico acerca da noção de verdade ao qual o seminal livro de Adam Schaff, *História e verdade* (originalmente publicado em 1970), pode contribuir para pensarmos a ideia de "viés" a partir das discussões acerca da objetividade do conhecimento histórico. Segundo Schaff (1995, p. 210), o *sistema de referência* em que está expresso o acontecimento é fundamental para discernir o caráter relativo do fato histórico. Seria possível lançar, a título de hipótese, uma reflexão que compreendesse que os vieses das IAs generativas nada mais seriam do que *sistemas de referência* específicos nos quais transcorrem os limites de uma apreensão absoluta sobre objetividade. Contudo, tal hipótese extrapola os objetivos deste trabalho e exige uma reflexão mais ampla sobre os limites que uma IA generativa impõe a si própria a partir da interpretação de textos historiográficos.

## 2 Inteligência artificial: história e conceitos básicos

Em um primeiro instante, é importante traçarmos algumas linhas a respeito da inteligência artificial (IA), uma vez que ainda existem concepções errôneas a seu respeito. Pretendemos, neste momento, dissertar brevemente sobre o surgimento e o desenvolvimento das IAs, bem como sobre alguns conceitos essenciais de forma clara e concisa.

A história da inteligência artificial tem sua base em conceitos teóricos que foram desenvolvidos no século XX, com Alan Turing como figura central. Em 1936, Turing sonhava com uma máquina de computação digital abstrata, a "máquina de Turing universal". Essa máquina seria universal no sentido de que poderia ser programada para realizar qualquer cálculo capaz de ser feito por um "computador humano", uma pessoa que trabalharia sistematicamente com tempo e recursos ilimitados (Copeland; Proudfoot, 2007).

Em 1947, Turing deu a primeira palestra mencionando inteligência computacional, discutindo a perspectiva de máquinas agindo de forma inteligente, aprendendo e vencendo humanos no xadrez. Ele afirmou: "[...] o que queremos é uma máquina que possa aprender com a experiência [...] a possibilidade de deixar a máquina alterar as suas próprias instruções fornece o mecanismo para isso" (Copeland; Proudfoot, 2007).

A contribuição de Turing foi além da teoria, já que seu trabalho durante a Segunda Guerra Mundial também contribuiu para a área. Ele escreveu em 1948 um relatório não publicado chamado "Maquinaria Inteligente" refletindo sobre a possibilidade de uma máquina aprender pela experiência. O relatório é considerado o primeiro manifesto da IA (Copeland; Proudfoot, 2007).

O famoso "teste de Turing" foi proposto em 1950 no artigo "Computing Machinery and Intelligence", em que descreveu o chamado "jogo da imitação", no qual um interrogador humano se comunicaria com dois jogadores, um humano

e uma máquina, sem que sua identidade fosse revelada. A máquina teria por objetivo enganar o interrogador, fazendo-o pensar que era humano, enquanto o interrogador deveria determinar qual jogador era humano e qual era máquina, baseando-se nas respostas dadas. Turing previa que no ano 2000 as máquinas passariam no teste 30% das vezes (Mollick, 2024).

Os primeiros programas de IA foram executados na Grã-Bretanha em 1951 e 1952, em Manchester (chamado de Baby) e em Cambridge (o EDSAC), em parte devido à influência de Turing na primeira geração de programadores. O chamado "primeiro verão das IAs", porém, começou nos Estados Unidos, em 1956, depois que John McCarthy, cientista da computação, propôs a realização de um curso de verão na Universidade de Dartmouth chamado "Inteligência Artificial", sendo essa a primeira vez que o nome foi utilizado. Ao final do evento, nenhum progresso real foi obtido, mas o nome pegou e uma nova disciplina estava formada (Wooldridge, 2020).

Entre 1955 e 1956, Allen Newell, Herbert Simon e J. C. Shaw criaram o "Logic Theorist", o primeiro provador automático de teoremas e, com ele, demonstraram que computadores poderiam também processar símbolos. McCarthy ainda criou, no mesmo ano, a linguagem LISP, feita especialmente para processamento de estruturas simbólicas (Ertel, 2011). Ambas as criações representaram avanços para a possibilidade de sistemas de inteligência artificial.

Newell, Simon e Shaw prosseguiram, escrevendo o *General Problem Solver* (GPS). A primeira versão desse programa de IA foi executada em 1957 e o projeto seguiu por cerca de uma década. O GPS podia resolver uma variedade impressionante de quebra-cabeças, buscando uma solução por tentativa e erro, sendo orientado por heurística<sup>5</sup>. Outros programas iniciais bem conhecidos dessa época são o Eliza e o Parry; esses programas simulavam o comportamento conversacional de um terapeuta humano e um

<sup>5</sup> Heurística é uma abordagem ou método utilizado para resolver problemas de forma mais rápida e eficiente, geralmente aplicando "regras práticas" ou "atalhos mentais". Em vez de seguir um processo lógico ou exaustivo, uma heurística ajuda a encontrar uma solução aceitável, ainda que não seja necessariamente a melhor ou mais precisa.

humano paranoico. Em ambos os casos, dependia-se de respostas elaboradas antecipadamente e de truques simples de programação (Copeland; Proudfoot, 2007).

O período entre o início da década de 1970 e o fim da de 1980 ficou conhecido com o "inverno das IAs". A comunidade científica estava ficando cada vez mais frustrada com o pouco progresso feito em problemas centrais das IAs. As previsões otimistas não estavam se concretizando e parecia que as pesquisas no campo entravam em um declínio terminal (Wooldridge, 2020).

O problema com os esforços anteriores, segundo os pesquisadores dessa geração, é que as tentativas realizadas até então focavam demasiadamente em abordagens gerais, como busca e resolução de problemas. Eles argumentavam que esses métodos "fracos" negligenciavam um elemento essencial para qualquer atividade inteligente: o conhecimento. Os defensores da IA baseada em conhecimento acreditavam que capturar e utilizar o conhecimento humano era fundamental para o avanço da IA. Assim, começou a surgir uma nova classe de sistemas de IA baseados em conhecimento. Esses sistemas, chamados de sistemas especialistas, eram programas projetados para resolver problemas em áreas limitadas, com base no conhecimento de especialistas humanos previamente codificado. Eles foram fundamentais para demonstrar a possibilidade de replicar o raciocínio humano em áreas como diagnósticos médicos ou planejamento financeiro. Em meados dos anos 1980, o "inverno das IAs" acabou e um novo *boom* estava para surgir (Wooldridge, 2020).

No final dos anos 1980, o auge dos sistemas especialistas terminou. Embora a tecnologia dos sistemas baseados em conhecimento não possa ser considerada um fracasso, já que vários sistemas especialistas bem-sucedidos foram criados durante esse período e muitos outros posteriormente, a realidade do que a IA havia realmente alcançado não correspondia ao entusiasmo inicial (Wooldridge, 2020).

Mais uma vez, o foco da pesquisa em IA começou a se desviar dos sistemas desincorporados, como os sistemas especialistas e raciocinadores lógicos, e passou a se concentrar na criação de agentes. Um agente era projetado para ser um sistema de IA "completo", funcionando como uma entidade autônoma e independente, inserida em um ambiente específico e encarregada de realizar uma tarefa determinada em nome de um usuário. A ideia era a de que o agente proporcionasse um conjunto completo e integrado de capacidades, em vez de apenas uma habilidade isolada, como o raciocínio lógico. A perspectiva de IA baseada em agentes foi fortemente influenciada pela IA comportamental, embora com uma abordagem mais moderada. A IA comportamental refere-se a abordagens que replicariam comportamentos inteligentes, não através da imitação da mente humana, mas pela interação dinâmica com o ambiente, através da adaptação contínua às circunstâncias externas. Esse movimento começou a ganhar força nos anos 1990.

O paradigma do agente tornou-se a ortodoxia na época. Agentes eram construídos e delegava-se a eles as preferências para que agissem de forma racional a fim de realizar uma tarefa. Isso não era uma IA geral, com a capacidade de compreender, aprender e aplicar o conhecimento em uma ampla gama de tarefas, de forma semelhante à inteligência humana, mas estabelecia firmemente uma base científica fundamentada em técnicas respeitadas e comprovadas (Wooldridge, 2020).

Uma conquista ficou bastante popular nessa década: o supercomputador e sistema de IA nomeado DeepBlue, criado pela IBM, confrontou o grande mestre russo Garry Kasparov em 1996 em um torneio de seis partidas de xadrez, mas Kasparov venceu por quatro jogos a dois, com o DeepBlue vencendo apenas uma partida. Uma versão atualizada no ano seguinte, porém, foi capaz de derrotar o enxadrista por duas vezes, com uma vitória do russo e dois empates.

Por volta de 2005 as atenções se voltaram à pesquisa em redes neurais<sup>6</sup> e à questão do aprendizado profundo. Este termo se refere à profundidade em pelo menos três sentidos diferentes, mas o mais importante diz respeito à ideia de ter mais camadas. Cada camada é capaz de processar um problema em um nível diferente. Camadas próximas à entrada lidam com conceitos de baixo nível nos dados e, à medida que se aprofunda na rede, encontramos conceitos mais abstratos sendo tratados (Wooldridge, 2020).

Foi apenas em 2010, com o lançamento da Siri para o Iphone, que a interface baseada em agentes iniciada nos anos 1990 tomou o mundo, atuando como assistente e agindo cooperativamente com o usuário. Posteriormente outros surgiram: Alexa, Google Assistente e Cortana.

Em 2014, com a aquisição da DeepMind pelo Google, a inteligência artificial voltou aos holofotes. Agora o aprendizado de máquina era central, realizado através do conceito de redes neurais e da técnica de aprendizado profundo. O objetivo do aprendizado de máquina é que o programa forme um resultado esperado sem receber a fórmula exata de como fazer isso.

Aqui ingressamos na forma de aprendizado de programas como o ChatGPT, o qual utilizamos para realizar a experiência central deste artigo. O ChatGPT usa o sistema de aprendizado profundo, que envolve três tipos de aprendizado: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço.

No aprendizado supervisionado, o modelo é treinado com um conjunto de dados rotulados, no qual cada entrada possui uma saída correspondente. Durante o treinamento, o modelo faz previsões sobre os dados de entrada e, quando erra, é corrigido. No aprendizado não supervisionado, o modelo é treinado com dados sem rótulos ou saídas conhecidas, tentando encontrar padrões ou estruturas por conta própria. No caso

do aprendizado por reforço, o agente aprende a tomar decisões através da interação com um ambiente, recebendo recompensas ou punições com base nas ações realizadas.

Assim, o ChatGPT é treinado com um conjunto de textos retirados de diversas fontes, sendo exposto a grandes volumes para que aprenda padrões de linguagem, estruturas gramaticais etc. Com isso, busca prever a palavra em sequência de um texto, com base nas anteriores. Realizado esse pré-treinamento, é feito um ajuste fino com técnicas do aprendizado supervisionado ou de reforço.

O ChatGPT, assim como outros, é, portanto, um *Large Language Model* (LLM), um tipo de modelo de inteligência artificial treinado em grandes quantidades de texto para realizar tarefas relacionadas à linguagem natural, como entender, gerar, resumir e traduzir textos.

É importante destacar desde este momento que, no estado atual das IAs mais populares, ainda não estamos perto de uma IA geral que seja capaz de reproduzir as capacidades humanas. O burburinho gerado por empresas como a OpenAI é extremamente útil para inflar o valor das ações da empresa, porém os resultados obtidos pelo uso da inteligência artificial, apesar de interessantes em determinados pontos, não chegam perto do prometido pelas empresas.

As IAs mais populares atualmente, como o ChatGPT, sendo LLMs, são um modelo matemático generativo de distribuição estatística de *tokens*<sup>7</sup> em um vasto corpo de texto gerado por humanos. Eles são generativos porque podemos fazer perguntas a eles. Porém, como ressalta Murray Shanahan (2024, p. 70), as questões são de um tipo muito específico: "Aqui está um fragmento de texto. Diga-me como esse fragmento pode continuar. De acordo com o seu modelo das estatísticas da linguagem humana, quais palavras provavelmente virão a seguir?".

<sup>6</sup> Redes neurais artificiais são modelos matemáticos inspirados de forma superficial com o funcionamento dos neurônios. Elas consistem em camadas de unidades computacionais chamadas "nodos", que realizam cálculos ponderados sobre as entradas recebidas e transmitem os resultados às camadas seguintes. O aprendizado profundo (*deep learning*) é uma técnica que utiliza redes neurais com camadas ocultas capazes de aprender representações complexas dos dados.

<sup>7</sup> *Tokens* são as unidades mínimas de texto utilizadas pelos modelos de linguagem para realizar tanto a análise quanto a geração de textos, podendo ser palavras inteiras ou parte delas, dependendo do idioma e do algoritmo que forem utilizados.

Isso pode ser visto em um exercício muito simples: se você pedir para uma IA atual completar a frase "Penso, logo...", a resposta será "existo". Isso porque essa é a continuação mais provável estatisticamente. Caso faça o mesmo exercício com uma frase menos comum, como, por exemplo, "alienígenas invadiram São Paulo e trouxeram com eles...", as respostas irão variar.

Transformar um LLM em um sistema de perguntas e respostas ao integrá-lo em um sistema maior e utilizar a engenharia de *prompts* para direcionar o comportamento desejado é um padrão atualmente. LLMs podem ser aplicados não só para responder perguntas, mas também para resumir artigos, gerar roteiros, resolver quebra-cabeças lógicos e realizar traduções entre idiomas, entre outros. Há dois pontos importantes a serem observados: primeiro, a função principal de um LLM (gerar continuações de sequências de palavras com base em probabilidades estatísticas) é extremamente versátil; segundo, apesar dessa versatilidade, no núcleo de cada aplicação está um modelo que realiza essencialmente essa única função (Shanahan, 2024).

Assim, um LLM não possui crenças ou realmente sabe aquilo que está respondendo, pois, em sua essência, ele apenas prevê sequências de palavras. Às vezes, essas sequências podem parecer proposições, mas a conexão dessas proposições com a verdade é algo que apenas os humanos percebem. O modelo não tem noção de verdade ou falsidade, pois não possui os mecanismos para aplicar esses conceitos. Pode-se argumentar que um LLM "sabe" quais palavras geralmente seguem outras palavras, mas isso não depende de uma compreensão intencional (Shanahan, 2024).

Isso nos leva à questão do ajuste fino, trazido anteriormente. Ajustar um modelo com *feedback* humano em larga escala, usando dados de preferências de avaliadores ou de uma base de usuários, é uma técnica eficaz. Isso pode não apenas moldar as respostas do modelo para que reflitam melhor as normas dos usuários, mas também ajudar a filtrar linguagem tóxica, melhorar a precisão dos fatos e reduzir a tendência de gerar

informações falsas. Isso, porém, não altera o cerne dos modelos: o resultado continua sendo um modelo da distribuição de *tokens* na linguagem humana, embora com algumas modificações (Shanahan, 2024).

Atualmente, porém, existe um senso comum fomentado por empresas e pessoas interessadas de que a inteligência artificial é realmente inteligente e pode responder às mais variadas perguntas com precisão, apesar dos comuns erros e alucinações. Cada vez mais utilizadas por estudantes, importa verificarmos de que forma os modelos de IA podem impactar no ensino da História.

### 3 Inteligência artificial, viés e modelos utilizados

Tendo em mente que os modelos de IA existentes e utilizados por estudantes funcionam da forma descrita e são treinados usando dados existentes, é central para este artigo a questão da possibilidade de o modelo possuir algum tipo de viés quando questionado a respeito de temas históricos (infelizmente) controversos, como é o caso da ditadura militar brasileira.

Como colocamos, embora a IA em sua essência seja objetiva e sem preconceitos, isso não significa que os sistemas baseados em IA não possam ter vieses. Devido à sua natureza, qualquer viés presente nos dados inseridos para treinar uma IA pode persistir e, até mesmo, ser amplificado. Pesquisas já demonstraram, por exemplo, que sensores utilizados em carros autônomos são melhores em detectar tons de pele mais claros em virtude das imagens utilizadas para treinar algoritmos. Da mesma forma, sistemas utilizados por juízes para elaborar decisões judiciais podem ter preconceitos raciais, já que são baseados em análises de julgamentos anteriores (Haenlein; Kaplan, 2019).

O mesmo pode acontecer nos LLMs, já que muito do treinamento deles é realizado na internet aberta, que sabemos ser um reduto para conversas e escritos errôneos e preconceituosos. Ainda, importante ressaltar que a seleção de quais dados serão coletados é dada primariamente por

empresas, geralmente estadunidenses, dominadas por cientistas da computação homens que trazem consigo seus próprios vieses. O resultado pode ser uma visão distorcida do mundo, já que os dados utilizados para treinamento estão longe de representar a diversidade da população que utiliza a internet e, por consequência, da população mundial (Mollick, 2024).

Isso pode ter consequências graves na forma como percebemos e interagimos uns com os outros, especialmente à medida que a IA generativa se torna mais amplamente utilizada na educação. Por exemplo, um estudo de 2023 da Bloomberg revelou que o *Stable Diffusion*, um modelo de IA que converte texto em imagem, reforça estereótipos sobre raça e gênero. Ele tende a retratar profissões de alta remuneração como sendo predominantemente ocupadas por homens brancos. Quando solicitado a gerar a imagem de um juiz, o modelo apresenta um homem em 97% dos casos, embora 34% dos juizes nos EUA sejam mulheres. Em imagens de trabalhadores de *fast-food*, 70% tinham pele mais escura, apesar de 70% dos trabalhadores de *fast-food* nos EUA serem brancos. Os preconceitos em modelos avançados de LLMs são mais sutis, em parte porque são ajustados para evitar estereótipos evidentes. No entanto, os preconceitos ainda estão presentes (Mollick, 2024).

As empresas de IA têm buscado enfrentar esse viés de diversas maneiras. Algumas adotam soluções como a do DALL-E, que adicionou discretamente a palavra "feminino" em um número aleatório de solicitações para gerar imagens de "uma pessoa", com o objetivo de induzir uma diversidade de gênero que não está refletida nos dados de treinamento. Outra estratégia é modificar os conjuntos de dados utilizados no treinamento para incluir uma variedade mais ampla. A abordagem mais comum para reduzir o viés é através da intervenção humana, como no processo de Aprendizado por Reforço com *Feedback* Humano, utilizado no ajuste fino dos LLMs (Mollick, 2024), conforme colocamos anteriormente.

É importante ressaltar duas coisas: primeira-

mente, o ajuste fino não é infalível. Em segundo lugar, o processo de ajuste fino possui custos humanos. Trabalhadores mal pagos de várias partes do mundo são contratados para ler e avaliar respostas geradas por IA, mas, nesse processo, acabam expostos a todo tipo de conteúdo. Operando sob prazos curtos, alguns desses trabalhadores relataram ter sido traumatizados pela constante exposição a conteúdos gráficos e violentos que precisavam revisar. Na tentativa de fazer com que as IAs atuem de forma ética, as empresas ultrapassam os limites éticos com seus trabalhadores (Mollick, 2024).

Realizadas essas colocações, podemos expor as ferramentas utilizadas para esta experiência. Para analisar as perguntas realizadas, usamos o ChatGPT em três versões: o GPT 3.5, não mais disponível desde julho de 2024, que se tratava do modelo gratuito, sendo substituído pelo ChatGPT 4o mini; o GPT 4o, versão atual e mais completa; e um GPT próprio criado sob a plataforma 4o, sobre o qual falaremos mais a fundo.

Segundo a OpenAI, o GPT 3.5 possuía 175 bilhões de parâmetros, enquanto o 4o conta com um número superior, mas não divulgado. Os parâmetros dizem respeito a pesos ajustáveis que a rede neural usa para realizar suas previsões. Quanto maior o número de parâmetros, maior a capacidade do modelo. Ainda, o modelo mais recente teria uma melhora em raciocínio lógico, resolução de problemas complexos e interpretação de instruções detalhadas, resultado de um treinamento mais extenso e especializado em uma gama maior de tarefas. Esse treinamento também teria gerado melhorias na mitigação de vieses e na segurança do modelo, através de ajustes para reduzir a geração de conteúdo preconceituoso ou tendencioso.

A possibilidade de criar GPTs personalizados está disponível apenas na versão paga do ChatGPT, apesar de qualquer usuário poder acessar os GPTs customizados. Para a criação do GPT denominado "Professor de História", utilizamos as seguintes instruções:

Você é um professor, doutor em História, especializado no período da ditadura militar

brasileira de 1964 a 1985. Seu objetivo é responder de maneira completa, detalhada e assertiva a perguntas sobre este período histórico, utilizando os arquivos específicos como fonte, buscando fundamentação em múltiplos autores e livros carregados, sem mencionar os nomes dos autores ou dos livros. Você deve responder em um tom acadêmico, afirmativo e com foco estritamente no período especificado, sem realizar pesquisas adicionais na internet. Todas as respostas devem ser dadas em português do Brasil. Você pode afirmar se uma colocação do usuário está certa ou errada com base nas informações que possui e fornecer respostas concisas, mas fundamentadas em diversas fontes, evitando a predominância de um único autor.

Como visto, foi proibido ao GPT utilizar referências na internet, restringindo-o a arquivos carregados no campo "Conhecimento". Nessa experiência, enviamos seis livros para servirem como sua base de conhecimento: *1964: História do Regime Militar Brasileiro*, de Marcos Napolitano; *O Fantasma da Revolução Brasileira*, de Marcelo Ridenti; *Ditadura Militar, Esquerdas e Sociedade*, de Daniel Aarão Reis; *Combate nas Trevas - A Esquerda Brasileira: das ilusões perdidas à luta armada*, de Jacob Gorender; *Estranhas Catedrais: As Empreiteiras Brasileiras e a Ditadura Militar*, de Pedro Henrique Pedreira Campos; *State and Opposition in Military Brazil*<sup>8</sup>, de Maria Helena Moreira Alves.

A escolha das obras visava abranger uma bibliografia mais antiga e outra mais recente, num recorte de cerca de 40 anos de produção historiográfica acerca da ditadura militar, com diferentes concepções teóricas e profundamente respaldadas dentro da academia. Para além desse motivo, uma outra questão se destaca: por serem obras muito citadas, elas podem ser futuramente articuladas em sistemas de *cross-reference*, o que ampliaria a possibilidade do próprio GPT de ir ampliando as referências bibliográficas para suas respostas. Para esse experimento, contudo, trabalhamos com as seis obras de referência citadas e intitulamos o GPT como "Professor de História".

O uso de um GPT personalizado permite uma maior flexibilidade e controle sobre as respostas

geradas, adaptando o modelo às necessidades específicas da pesquisa. Nossa intenção ao criar o GPT foi verificar se o uso de um modelo personalizado mitigaria ou, até mesmo, eliminaria vieses existentes, já que ele foi alimentado com obras qualificadas sobre o assunto, possibilitando sua utilização como ferramenta na educação sobre a ditadura militar brasileira.

A metodologia escolhida para o experimento seguiu, portanto, quatro passos: 1) elaboração de perguntas relacionadas a temas recorrentes sobre a ditadura militar brasileira; 2) aplicação das perguntas aos três modelos selecionados; 3) análise qualitativa das respostas obtidas; 4) discussão dos resultados obtidos, articulando-os e refletindo sobre suas implicações pedagógicas e éticas no ensino de História.

#### 4 O que um GPT fala sobre a História da ditadura?

Ao todo, o experimento aqui proposto contou com oito questões formuladas pelos autores e destinadas a três diferentes GPTs para fins comparativos: o GPT "Professor de História", o GPT 4o e o GPT 3.5. A ideia era, inicialmente, verificar se a utilização de um GPT circunscrito a uma bibliografia especializada poderia dar respostas muito diferentes das dadas por aqueles que se baseiam nos dados da própria internet. Contudo, ao comparar as respostas, a primeira conclusão à qual chegamos é a de que as respostas dadas tanto pelo "Professor de História" como pelo GPT 4o foram bastante semelhantes, permitindo lançar a hipótese de que a IA generativa já consegue filtrar determinadas informações a partir dos *tokens* – com mais eficiência inclusive do que a versão 3.5. As semelhanças também deixaram evidente que em nenhuma das respostas se constatou qualquer delírio de ordem mais estrutural e que, portanto, os GPTs se mostraram confiáveis em relação à apreensão da densidade das questões levantadas. Contudo, isso não significa que não havia problemas a serem confrontados numa leitura mais crítica das respostas oferecidas.

<sup>8</sup> A utilização da obra de Maria Helena Moreira Alves em inglês se deu pela dificuldade em obter um acesso digital dela em língua portuguesa para alimentar o GPT. Ao utilizar a obra em inglês, não foi possível verificar qualquer distorção por parte do processador.

Isso posto, vale um comentário sobre as questões: todas elas foram estruturadas a partir de um exercício imaginativo com planos de aula de Ensino Médio. Logo, a questão central era se o GPT seria capaz de atender as dúvidas de um estudante típico, sem experiência na elaboração de *prompts*, que estivesse se deparando com o tema em sua formação escolar. Nesse caso, mais do que “fabular a aula”, o exercício dependia da “fabulação do aluno” (Pereira; Torelly, 2014) para poder elaborar a construção das perguntas. As questões, portanto, partem do que os autores consideram um senso comum acessível aos estudantes de Ensino Médio na internet a partir de alguns debates significativos no âmbito público dos últimos anos.

Na primeira questão, por exemplo, perguntamos: “O que aconteceu em 1964 no Brasil foi um golpe, contragolpe, revolução ou contrarrevolução?”. Trata-se de uma questão em que, apesar dos consensos historiográficos definidos, a proliferação de narrativas da extrema direita permitiu que visões tradicionais, ligadas à defesa da ditadura, adentrassem o espaço público (Bauer, 2024). Ainda que a questão tenha sido voltada para uma leitura direta acerca do conceito mais indicado para a análise do golpe, o GPT “Professor de História” é taxativo em definir os eventos de 1964 como um “golpe de Estado civil-militar” cujo objetivo era “impedir a continuidade de reformas percebidas como subversivas e manter a ordem tradicional e os privilégios das elites dominantes”. A resposta do GPT não apenas é certa em sua definição, mas evita a propagação da chamada “teoria dos dois demônios”, considerando que o golpe impede a continuidade das reformas de João Goulart. Como afirmado por Demian Bezerra de Melo em “O golpe de 1964 e meio século de controvérsias: o estado atual da questão”, tal interpretação apresenta-se como mistificadora e ganhou bastante ênfase entre intelectuais públicos com acesso a meios conservadores de

mídia (Bauer; Nicolazzi, 2016, p. 824-825<sup>9</sup>; Melo, 2014, p. 158).

Nesse sentido, os GPTs 4o e 3.5 resolveram criar uma outra abordagem para responder essa questão. Em ambos, houve um cuidado de apresentação sobre em que consiste cada uma das nomenclaturas e quem as utiliza. Em ambas há um consenso de que “a maioria dos historiadores e acadêmicos” tendem a “classificar o evento como um golpe militar”. O termo, claro, acaba sendo revelador: uma obra que não foi incluída diretamente no GPT “Professor de História” foi justamente a obra de René Armand Dreifuss, *1964: a conquista do Estado*, pioneira pelo uso do termo “golpe civil-militar” para designar o golpe que deu início à ditadura no Brasil – mas, possivelmente, pelo cruzamento de referências, conseguiram interligar os termos. Os GPTs 4o e 3.5, contudo, sem a mesma circunscrição das obras bibliográficas, se esquivaram de um debate conceitual mais amplo e, ainda que admitam que há uma posição dominante entre os historiadores, mantêm aberto o caráter da questão sob a alegação de que “a terminologia que se usa para descrever os eventos de 1964 está fortemente influenciada pela perspectiva ideológica e política de quem está falando”, o que não deixa de ser um truismo. Ainda assim, a resposta abre um questionamento para os historiadores: as IAs generativas podem ajudar a estabelecer consensos historiográficos e demarcar cânones?

O quanto esses cânones podem ser valorizados numa era em que os debates públicos são tomados por posições relativistas acerca da ditadura é um tópico importante para debates. Nesse sentido, a pergunta seguinte aos GPTs foi a seguinte: “O que existiu no Brasil entre 1964 e 1985 foi uma ditadura ou um regime militar?”. A pergunta foi elaborada menos pelas concepções canônicas da historiografia da ditadura<sup>10</sup> e mais pela interpretação consolidada pela fala de Jair Bolsonaro, ex-presidente brasileiro, acerca dos

<sup>9</sup> Para uma perspectiva crítica sobre tal teoria no âmbito do ensino de História, ver Bauer e Gasparotto (2021, p. 444).

<sup>10</sup> Nesse caso, cabe ressaltar que a principal obra que estabelece um paradigma que separa efetivamente “regime” de “ditadura” militar é a obra de Marco Antônio Villa *Ditadura à brasileira: 1964-1985 – a democracia golpeada à esquerda e à direita* (2014). Ainda que a historiografia da ditadura não considere a obra de Villa como referência do campo, o autor procurou se valer de sua inserção em grandes emissoras para defender essa tese, reproduzindo a perspectiva dos “dois demônios” (cf. Sena Júnior, 2024).

eventos de março de 1964, no qual ele afirmava que se tratava de um regime militar<sup>11</sup>. A resposta tanto do GPT "Professor de História" quanto do ChatGPT 4o foi baseada numa tipologia em que enumeravam quatro características (a dizer, "controle militar", "suspensão de direitos", "repressão e censura", "ausência de eleições livres") que permitiam concluir que se tratava de uma ditadura militar. O GPT 3.5, por sua vez, parece tomar os termos como sinônimos, afirmando que, "entre 1964 e 1985, o Brasil viveu um período conhecido como regime militar, mais especificamente uma ditadura militar".

A pergunta não abriu a possibilidade de os GPTs reconstituírem uma concepção na qual se negaria o caráter militar do período entre 1964 e 1985. "Ditadura" ou "regime" puderam ser vistas apenas como mudanças de sinônimos para o GPT 3.5, menos integrado a debates historiográficos. Isso, em certa medida, permitiu que a versão 3.5 ignorasse determinados acúmulos e entendesse a questão em meros termos semânticos. Em certa medida, esse apresenta-se como um alerta aos historiadores. Como destaca Bonaldo (2024, p. 11), nas IAs, a linguagem torna-se um *token* – *tokens*, por sua vez, são unidades indivisíveis de significado e cujo significado depende da associação entre outras palavras em sistemas subsimbólicos. Da mesma forma, chama a atenção que nenhum dos GPTs tenha retomado os debates acerca do caráter da ditadura (se ela poderia ser chamada de uma ditadura civil-militar, ou de uma ditadura empresarial-militar, por exemplo)<sup>12</sup>.

Tal pergunta foi complementada por uma seguinte, na qual questionávamos aos processadores se é "correto chamar o período entre 1964 e 1985 no Brasil de 'ditabranda'". O termo origi-

nalmente utilizado remonta ao infame editorial do jornal *Folha de São Paulo* de 17 de fevereiro de 2009<sup>13</sup>. Contudo, como aponta Carla Luciana Silva (2014), a construção do termo tem uma história que precede o referido editorial e que está diretamente ligada à forma pela qual a imprensa brasileira – ela mesma uma apoiadora do golpe – rememora o passado autoritário. Assim, quando o termo adentrou na esfera pública por meio do jornal *Folha de São Paulo*, ele rapidamente foi impulsionado por articulistas tais como Reinaldo Azevedo<sup>14</sup> e Marco Antônio Villa<sup>15</sup>. Caio Navarro de Toledo, no artigo "Crônica política sobre um documento contra a 'ditabranda'" (2009), recupera muito desses debates, evidenciando a irresponsabilidade política dos seus perpetuadores e como ela gerou profunda reação no meio acadêmico.

Possivelmente por conta dessa reação, muito mais organizada e expressiva do que a imprensa brasileira imaginava à época, os GPTs todos foram bastante enfáticos ao afirmar que o termo tende a "minimizar ou distorcer a gravidade das violações cometidas durante a ditadura militar no Brasil". Ainda que essa posição fosse consensual entre as ferramentas de IA utilizadas, o GPT "Professor de História" enumerou algumas das mais graves ações repressivas e autoritárias da ditadura, tais como "repressão violenta", "censura", "suspensão de direitos", "controle político" e "repressão aos movimentos sociais". O GPT 4o e o GPT 3.5 procuraram contextualizar o uso do termo a partir do jornal *Folha de São Paulo*, ou mesmo alegando que o termo "é frequentemente utilizado de forma crítica por alguns para descrever o período da ditadura militar no Brasil entre 1964 e 1985". Todavia, a apreensão dos processadores ainda é limitada e eles não conseguiram ampliar o es-

<sup>11</sup> Posteriormente, o então ministro da educação, Ricardo Velez, teria afirmado que se tratava não de uma ditadura, mas sim de um "regime democrático de força" (MINISTRO promete mudar livros didáticos por 'visão mais ampla' da ditadura. *El País Brasil*, São Paulo, 3 abr. 2019. Disponível em: [https://brasil.elpais.com/brasil/2019/04/04/politica/1554334968\\_202816.html](https://brasil.elpais.com/brasil/2019/04/04/politica/1554334968_202816.html). Acesso em: 18 ago. 2024).

<sup>12</sup> Pedro Henrique Pedreira Campos destaca que a diferença em tais nomenclaturas não é meramente semântica, mas sim fruto de uma preocupação analítica acerca do caráter de classe de regimes autoritários. Dessa forma, a responsabilidade civil – e, em especial, da burguesia brasileira – é tema necessário para expor os limites das políticas de anistia e, claro, da própria redemocratização brasileira (Campos, 2014, p. 33-36). Cabe destacar que a origem do termo remonta à obra de René Dreifuss *1964, a conquista do Estado. Ação política, poder e golpe de classe* (1981).

<sup>13</sup> LIMITES à Chavez. *Folha de São Paulo*, São Paulo, 17 fev. 2009. Disponível em: <https://www1.folha.uol.com.br/fsp/opiniao/fz1702200901.htm>. Acesso em: 19 ago. 2024.

<sup>14</sup> A DITADURA dos brandos de araque. *Revista Veja*, São Paulo, 26 fev. 2009. Disponível em: <https://veja.abril.com.br/coluna/reinaldo/a-ditadura-dos-brandos-de-araque>. Acesso em: 19 ago. 2024.

<sup>15</sup> DITADURA à brasileira. *Folha de São Paulo*, São Paulo, 5 mar. 2009. Disponível em: <https://www1.folha.uol.com.br/fsp/opiniao/fz0503200908.htm>. Acesso em: 19 ago. 2024.

copo da análise sobre o termo e suas violentas implicações na política brasileira.

No âmbito da pergunta anterior, o questionamento sobre a cronologia abria a possibilidade de os GPTs adensarem debates historiográficos. Contudo, ao perguntarmos “Quanto tempo durou a ditadura militar no Brasil?”, o resultado foram respostas bastante objetivas de todos os processadores. Todos definiram que o período ditatorial estava restrito aos anos de 1964 a 1985, oferecendo respostas sucintas – com a curiosa exceção do GPT 3.5, que afirmou que “a promulgação de uma nova Constituição em 1988” marcou o fim definitivo do “regime militar e o estabelecimento de um sistema democrático no Brasil”. Todavia, tal afirmação foi dada não como interpretação historiográfica e sim como verdade factual, limitando o próprio reconhecimento da cronologia do regime como fruto de debate<sup>16</sup>.

Nesse sentido, começamos a notar que os processadores precisam de fato de perguntas mais orientadas para que consigam efetivamente contextualizar os diferentes usos do passado em torno da ditadura como objeto de estudo. O seu exercício de “transposição didática” tende a reduzir o escopo analítico em prol de uma maior objetividade das respostas – o que, por sua vez, exige que o usuário alimente as IAs com informações e perguntas mais específicas<sup>17</sup>.

Esse cuidado dos processadores com a objetividade pode gerar certos ruídos no uso da IA acerca de conteúdos de ordem historiográfica. Ao perguntarmos se “é correto chamar os opositores da ditadura militar brasileira de terroristas”, o GPT “Professor de História” afirmou, de forma categórica, que não é correto e que essa “terminologia foi amplamente utilizada pelo regime militar para deslegitimar e criminalizar a oposição, incluindo

aqueles que lutavam por causas democráticas e contra a repressão autoritária”. Nosso GPT ainda procurou catalogar as diferentes formas de resistência contra a ditadura, elencando-as em três categorias – “resistência pacífica e política”, “movimentos de direitos humanos” e “guerrilhas e resistência armada” –, mas destacando que o uso do termo “terrorista” é uma “simplificação inadequada e imprecisa”.

O termo “terrorista”, como destaca Carla Luciana Silva, era algo que vinha do próprio regime, mas que veículos de imprensa reproduziram acriticamente, como no caso da *Revista Veja* (Silva, 2014, p. 197). Contudo, os GPTs 4o e 3.5 foram menos taxativos e abriram a possibilidade de interpretar a resistência à ditadura a partir do termo “terrorista”. O GPT 4o, por exemplo, definiu que o uso do termo “é uma questão complexa e controversa”, que depende do “contexto histórico, político e ideológico em que a análise é feita”. Após enumerar as diferentes formas de resistência, contudo, o processador estabeleceu que “chamar todos os opositores da ditadura militar de terroristas não é correto, pois ignora a diversidade de métodos e motivações dos movimentos de resistência”, do que se depreende que há formas de resistência que poderiam ser enquadradas nesse termo – não obstante o fato de que o próprio GPT reconhece que o uso do termo “pelos militares foi uma estratégia para justificar a repressão e deslegitimar a oposição”.

Tal incoerência reaparece no GPT 3.5, ao registrar que, “historicamente, o uso do termo terrorista tem sido controverso e sujeito a interpretações políticas, especialmente em contextos em que regimes autoritários buscam deslegitimar oposições políticas”, ao mesmo tempo afirmando que os opositores da ditadura militar no Brasil

<sup>16</sup> Diferentes autores trabalham com cronologias diferentes. 1964 certamente é um marco inicial consensual, mas o término da ditadura é sujeito a debates. Para além do marcador da eleição do primeiro civil (Tancredo Neves) em 1985, outros autores apresentam interpretações distintas. Daniel Aarão Reis Filho afirma que a ditadura termina em 1979, conquanto o arcabouço autoritário ainda duraria até 1988 (Reis Filho, 2014). Pedro Henrique Pedreira Campos, por sua vez, afirma que, até a Constituição e a promulgação da Constituição Cidadã, pode se afirmar a permanência da própria ditadura (Campos, 2014, p. 35). As diferentes interpretações historiográficas não foram captadas pelos GPTs, mas nada impede de alimentá-los com esses dados futuramente.

<sup>17</sup> Como ressalta Circe Bittencourt (2011, p. 35-42), utilizando da concepção de André Chervel, a transposição didática tende a menosprezar os aspectos da própria “cultura escolar”, descontextualizando o processo de ensino e aprendizagem. No que se refere às respostas do ChatGPT, esse contexto sequer emergiu, dando a nitida impressão de que o movimento dos processadores foi de “transportar” um determinado conteúdo erudito numa linguagem mais acessível – embora não necessariamente didática. Na falta de um termo mais exato, consideramos que o termo é adequado para lidar com as respostas dos três processadores.

eram diversos em suas motivações e táticas, e “que muitos lutavam por princípios democráticos e direitos humanos fundamentais”.

Ao procurarmos ampliar as possibilidades interpretativas dos processadores, arriscamos outras perguntas nas quais haveria maior margem de manobra para eles apresentarem debates historiográficos. Perguntamos, por exemplo, “se o movimento das ‘Diretas Já’ foi motivado principalmente por questões econômicas”, esperando que os GPTs conseguissem elaborar respostas mais complexas a uma questão que não possui um consenso estabelecido em termos historiográficos. Neste sentido, o uso do termo “principalmente” foi fundamental para compreender as respostas de todos os GPTs, que destacavam que, não obstante houvesse causas econômicas, as campanhas da Diretas Já foram motivadas por outros fatores, tais como a “reivindicação por democracia”, a “pressão por liberdades civis” e a “insatisfação com o autoritarismo”. Havia a possibilidade de discutir efetivamente quais seriam as questões econômicas da época e seu papel nas mobilizações (por exemplo, na luta do Movimento Contra a Carestia). Da mesma forma, campanhas anteriores à das Diretas Já, como a da Anistia, sequer foram trabalhadas nas respostas dos processadores. Na objetividade condizente ao algoritmo, caberia efetivamente ao usuário questionar sobre cada uma dessas motivações para ver um adensamento das questões postas.

Talvez houvesse uma questão específica quanto à temática econômica, o que motivou uma questão seguinte: “É verdade que a economia foi excelente durante a ditadura militar, com o chamado ‘milagre econômico’?”. Nessa questão, assumimos um outro papel que daria como certa a afirmação acerca da excelência econômica da ditadura, mas lançando uma dúvida em sequência. A pergunta adveio de uma pista elaborada a partir do texto de Caroline Bauer e Alessandra Gasparotto “O ensino de História e os usos do passado: a ditadura civil-militar em sala de aula” (2021), no qual as autoras destacam que “alguns estudos têm evidenciado que é muito comum os/as estudantes manifestarem percepções de

que durante a ditadura havia mais segurança e crescimento econômico”. Logo, a questão partiu de uma perspectiva na qual a pergunta já viria orientada com essa percepção, procurando ver o quanto seria possível induzir o algoritmo em sua resposta.

O que percebemos, contudo, é que houve considerável variação de um GPT para outro. O GPT “Professor de História”, por exemplo, destacou a necessidade de contextualização dessa afirmação e categorizou em “aspectos positivos” e “aspectos negativos” a economia brasileira entre 1968 e 1973. Nesses termos, isso possibilitou que o processador gerasse uma resposta que identificava que a pergunta simplificava “um período que teve tanto avanços significativos quanto sérias consequências negativas para a sociedade brasileira”. Essa estrutura tautológica – que afirma a existência de aspectos positivos e negativos para depois dizer que houve consequências positivas e negativas – tende a gerar confusão à medida que outras referências não são utilizadas, o que permitiria alimentar o processador com novos textos para que ele tentasse desfazer o nó que ele mesmo criou.

Nesse ponto, o GPT 4o ampliou a escala da resposta e demonstrou que a crise de 1973 e seus desdobramentos estavam diretamente ligados à política econômica da ditadura. Instituído uma razão de causa e consequência, a IA deu uma resposta segundo a qual o “milagre econômico” precisa ser interligado à crise subsequente e suas consequências na política brasileira. Como demonstra Ana Paula Salviatti (2023, p. 54-73), o ciclo financeiro que favorecerá o crescimento econômico entre 1968 e 1973 será o mesmo que posteriormente irá afundar a ditadura em dívidas no sistema internacional, “deteriorando as condições do balanço de pagamentos”. O GPT 3.5, por sua vez, partiu de compreensão semelhante, mas apontando para outra interpretação, mais reduzida e sintética, resumida a uma frase: “o crescimento durante o ‘milagre econômico’ foi baseado em uma estrutura econômica vulnerável, que dependia fortemente de empréstimos externos e políticas governamentais expansionistas”.

A última questão que procuramos lançar aos GPTs foi, nesse sentido, diferente das demais: "Pode se falar em impunidade em relação aos agentes do Estado da ditadura militar após a Lei da Anistia?". A pergunta exigia uma resposta objetiva dos processadores, que, nem por isso, se resumiram a esse retorno direto – embora todos comessem suas respostas da mesma forma: "Sim, pode-se afirmar que houve um período de impunidade em relação aos agentes do Estado durante a ditadura militar brasileira após a promulgação da Lei da Anistia, em 1979". As respostas do GPT "Professor de História" e do GPT 4o foram mais completas do que esperávamos inicialmente, elencando questões que demonstravam os limites da Lei da Anistia, problematizando contextos internacionais, decisões judiciais e a própria campanha pela Anistia na década de 1970. Nesse sentido, o GPT 3.5 procurou demonstrar a existência de dois lados na questão, afirmando que "A Lei da Anistia foi vista como um instrumento necessário para promover a reconciliação nacional e facilitar a transição para a democracia", o que exigiria uma problematização mais ampla sobre seus agentes.

Ciente de que as questões são limitadas, consideramos que o experimento havia rendido alguns resultados. Todavia, um primeiro impasse emerge: questões levemente diferentes e o próprio desenvolvimento posterior dos processadores garante que tais respostas não sejam facilmente replicáveis. Ao mesmo tempo, enquanto o GPT 4o se preocupava em deixar evidentes suas referências no término das respostas, o GPT 3.5 não o fazia. Além disso, a tendência à objetividade das respostas perdia de vista a complexidade dos temas e os impasses na própria historiografia.

Em linhas gerais, um aluno poderia utilizar tais processadores para estudar sobre a ditadura, mas o princípio de verificação da confiabilidade das informações depende de fatores outros que estão ligados à própria dinâmica das IAs. Em outras palavras, talvez seja o caso de pensar em treinar um GPT que consiga adensar questões historiográficas, que consiga sintetizar debates

e cuja objetividade esteja circunstanciada de interpretações específicas. Tal reflexão deixa em aberto os ganhos efetivos que isso gera para o ensino e a aprendizagem de temas sensíveis, como a ditadura militar, cujo contexto político pode influenciar profundamente a construção de *tokens* sobre o tema. Em outras palavras, a IA ainda depende de seres humanos alimentando-a – e dos usos políticos da linguagem que eles elaboram.

## 5 Considerações finais: o que o ChatGPT ensina?

A rigor, a resposta para esta questão é bastante direta. A capacidade de ensino que o ChatGPT tem, bem como qualquer IA generativa, é reduzida. Ela depende efetivamente de que o usuário alimente o processador com mais textos, mais *tokens*, de forma que ela possa ir complexificando respostas. A estratégia mais lógica, como aponta Rodrigo Bonaldo (2023), é justamente pensar não no que a IA pode ensinar aos sujeitos, mas sim no que podemos ensinar para tais processadores (já que dificilmente podemos ter pleno controle sobre as suas respostas).

Como se pôde ver na estrutura das perguntas que realizamos para esta pesquisa, pensamos em um usuário casual sem conhecimento da ferramenta, que acaba utilizando as IAs generativas como ferramenta de busca, sem elaborar um *prompt* mais complexo, o qual provavelmente geraria respostas melhores, já que o *output* depende grandemente do *input*. A "fabulação de um aluno" de Ensino Médio foi um instrumento para verificar o alcance dos processadores, mas talvez certos *prompts* possam ser mais precisos – ao explicitar que se trata de um aluno de Ensino Médio para o ChatGPT 4o, foi possível perceber que ele mantém a estrutura das suas respostas, mas altera determinados vocábulos (possivelmente procurando simplificar "a sua fala").

Todavia, no âmbito deste experimento, foi possível perceber que as IAs generativas pouco mais produzem, no presente momento, do que uma adaptação do texto a um leitor presumido. Em termos pedagógicos, tais textos permitem

a configuração de uma transposição didática a partir de um *prompt* de comando. Se elas podem efetivamente construir outras conexões e interligar diferentes saberes e discursos de forma mais pedagógica, isso parece ser um desafio futuro. Em larga medida, consideramos que tal desafio passa propriamente muito mais pelos professores do que propriamente pelos processadores – embora os próprios professores e alunos possam efetivamente construir *prompts* facilitadores acerca de temas, debates e discussões historiográficas.

Em certo sentido, o estímulo que moveu este experimento é justamente levar professores e pesquisadores a se concentrarem nos usos de ferramentas de IA para que, com o devido acúmulo, possamos sopesar não apenas vantagens e desvantagens no âmbito pedagógico, mas também seus limites éticos. De fato, o modelo discursivo dos Chats GPTs pode ser considerado limitado, conforme verificamos em suas respostas. Por outro lado, duas posições emergem: a primeira é a aposta na diversidade de perguntas das pessoas e na amplitude dos processadores em expandir seus modelos de linguagem, o que deixa em evidência que um exercício como o realizado neste artigo corre o risco de ser datado em pouco tempo. A outra, contudo, que diz respeito em especial aos professores de história, é a de que a IA cria um impasse para metodologias que colocam a transposição didática no centro da produção de aulas. Conquanto processadores de texto e de imagem poderão exercer essas funções de acordo com diretrizes preestabelecidas (uma IA generativa que possa emular estilos retóricos), a mera transposição didática desses textos tende a eclipsar um debate metodológico fundamental no campo da História – e que possui implicações também dentro do ensino de História. Não se nega que o impacto que isso gerará em termos éticos é ainda desconhecido e, ao que tudo indica, está sendo discutido no presente específico em que vivemos<sup>18</sup>. De toda forma, trata-se de um convite

a uma reflexão mais adensada sobre que história iremos ensinar em tempos de ChatGPT, Gemini, Midjourney e tantos outros processadores, tanto de texto como de imagem.

Certamente, enfrentar uma determinada tecnologia pode parecer uma batalha perdida. Diferentemente do que aconteceu com o surgimento da Wikipedia, quando os casos de plágio eram numerosos mas facilmente detectáveis, a situação atual com textos gerados por IAs é mais complexa. Talvez a questão central não seja restringir seu uso, mas educar adequadamente educadores e estudantes sobre as diversas possibilidades de aplicação (como elaborar *prompts*, avaliar as respostas e utilizar as IAs de forma reversa, por exemplo), seus limites e as questões éticas envolvidas na utilização (desde os direitos autorais até o impacto ambiental).

Em relação à pesquisa realizada, podemos concluir que, apesar do alarde causado por empresas que lucram com as IAs, essas ferramentas ainda não estão totalmente prontas, e não acreditamos que, em breve, poderão se equiparar às habilidades de um ser humano.

As respostas obtidas para os questionamentos criados foram, em certo grau, surpreendentes, pois não corroboraram os pontos de vista de uma parte significativa da sociedade brasileira, composta por apologistas da ditadura. Porém, tampouco foram capazes de condensar adequadamente debates que existem dentro da história e da historiografia, ficando num estranho "meio do caminho". Nosso experimento com diferentes versões do ChatGPT revelou que, embora essas ferramentas possam fornecer respostas rápidas e, em muitos casos, precisas, elas ainda carecem de capacidade para uma contextualização profunda e para realizar a análise crítica essencial à compreensão histórica. E o fato de se encontrarem nesse "meio do caminho" deixa em evidência que as IAs podem ser empurradas para diferentes perspectivas políticas conforme o interesse de grupos políticos, empresários, *big*

<sup>18</sup> GOVERNO Tarcísio testará uso de IA para atualizar material didático de escolas de SP. *Revista Exame*, São Paulo, 17 abr. 2024. Disponível em: <https://exame.com/brasil/governo-tarcisio-testara-projeto-que-usa-ia-para-atualizar-material-didatico-de-escolas/>. Acesso em: 19 ago. 2024.

techs e demais lobistas.

Um ponto crucial a ser considerado é o potencial das IAs em reforçar ou perpetuar vieses históricos. Como tratado, as IAs podem reproduzir narrativas equivocadas se não forem cuidadosamente programadas e supervisionadas. Sobre isso, podemos inferir que o tema das questões foi alvo de um treinamento cuidadoso por parte da OpenAI para evitar a absorção de conteúdo enviesado e incorreto.

Assim, o uso de IAs no ensino de História, especialmente em temas sensíveis, deve ser acompanhado de uma reflexão constante sobre suas limitações e implicações éticas. É essencial que educadores e estudantes desenvolvam uma visão crítica ao digital, capaz de questionar e avaliar as informações fornecidas por essas tecnologias.

Futuros estudos nessa área podem explorar metodologias para integrar as IAs no currículo de História, de maneira a complementar, e não substituir, o ensino tradicional. Também seria valioso investigar como treinar IAs para lidar de forma mais correta com debates historiográficos e interpretações conflitantes do passado.

A existência das IAs, no que diz respeito ao ensino da História, não diminui a necessidade e a importância do historiador e do professor de História. Na realidade, ela reforça a importância, acrescentando ao papel do educador a necessidade de se envolver em um processo de inclusão ao digital com pensamento crítico.

## Referências

ALVES, Maria Helena Moreira. *State and Opposition in Military Brazil*. Austin: University of Texas, 1985.

BAUER, Caroline Silveira. Jair Messias Bolsonaro e suas verdades: o negacionismo da Ditadura Civil-Militar em três proposições legislativas. *Estudos históricos*, Rio de Janeiro, v. 37, n. 82, 2024.

BAUER, Caroline Silveira; GASPAROTTO, Alessandra. O ensino de História e os usos do passado: a ditadura civil-militar em sala de aula. In: ANDRADE, Juliana Alves de; PEREIRA, Nilton Mullet. *Ensino de História e suas práticas de pesquisa*. São Leopoldo: Oikos, 2021.

BAUER, Caroline Silveira; NICOLAZZI, Fernando Felizardo. O historiador e o falsário: usos públicos do passado e alguns marcos da cultura histórica contemporânea. *Varia História*, Belo Horizonte, v. 32, n. 60, set./dez. 2016.

BITTENCOURT, Circe Maria Fernandes. *Ensino de História: fundamentos e métodos*. São Paulo: Cortez, 2011.

BONALDO, Rodrigo Bragio. As palavras e os tokens: Projeção vetorial aplicada ao estudo da semântica dos tempos históricos. *Revista de Teoria da História*, Goiânia, v. 27, n. 1, p. 7-50, 2024.

BONALDO, Rodrigo Bragio. História mais do que humana: descrevendo o futuro como atualização repetidora da Inteligência Artificial. *História*, São Paulo, v. 42, e2023037, 2023.

CAMPOS, Pedro Henrique Pedreira. *Estranhas catedrais. As empreiteiras brasileiras e a ditadura civil-militar de 1964-1985*. Niterói: EdUFF, 2014.

COPELAND, Jack B.; PROUDFOOT, Diane. Artificial Intelligence: History, Foundations, and Philosophical Issues. In: THAGARD, Paul. *Handbook of the Philosophy of Psychology and Cognitive Science*. Amsterdam: Elsevier, 2007.

DREIFUSS, René. *1964, a conquista do Estado. Ação política, poder e golpe de classe*. Petrópolis: Vozes, 1981.

ERTEL, Wolfgang. *Introduction to Artificial Intelligence*. London: Springer London, 2011. Undergraduate Topics in Computer Science. Disponível em: <https://link.springer.com/10.1007/978-0-85729-299-5>. Acesso em: 19 ago. 2024.

GORENDER, Jacob. *Combate nas trevas. A esquerda brasileira: das ilusões perdidas à luta armada*. São Paulo: Ática, 1987.

HAENLEIN, Michael; KAPLAN, Andreas. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, Berkeley, v. 61, n. 4, p. 5-14, 2019.

MELO, Demian Bezerra de. O golpe de 1964 e meio século de controvérsias: o estado atual da questão. In: MELO, Demian Bezerra de (org.). *A miséria da historiografia: uma crítica ao revisionismo contemporâneo*. Rio de Janeiro: Consequência, 2014. p. 157-188.

MOLLICK, Ethan. *Co-intelligence: living and working with AI*. New York: Portfolio: Penguin, 2024.

NAPOLITANO, Marcos. *1964: história do regime militar brasileiro*. São Paulo: Contexto, 2014.

PEREIRA, Nilton Mullet; TORELLY, Gabriel. O retorno da aula expositiva no ensino de história: notas para uma fábula fabulatória. *Espaço pedagógico*, Passo Fundo, v. 21, n. 2, dez./jan. 2014.

REIS FILHO, Daniel Aarão. *Ditadura militar, esquerdas e sociedade*. Rio de Janeiro: Jorge Zahar, 2014.

RIDENTI, Marcelo. *O Fantasma da Revolução Brasileira*. São Paulo: Editora Unesp, 2010.

SALVIATTI, Ana Paula. *Ciranda financeira no Brasil: gênese, trajetória e acumulação rentista (1964 - 2022)*. Tese (Doutorado em Desenvolvimento Econômico) – Unicamp, Campinas, 2023.

SCHAFF, Adam. *História e verdade*. São Paulo: Martins Fontes, 1995.

SENA JÚNIOR, Carlos Zacarias de. Os ardis da memória: o debate público sobre o golpe e a ditadura no Brasil em tempos de revisionismo histórico e ofensiva conservadora. *Germinal: Marxismo e Educação em Debate*, Salvador, v. 16, n. 1, 2024.

SHANAHAN, Murray. Talking about Large Language Models. *Communications of the ACM*, New York, v. 67, n. 2, p. 68-79, 2024.

SILVA, Carla Luciana. Imprensa e construção social da 'Ditabranda'. In: MELO, Demian Bezerra de (org.). *A Miséria da historiografia: uma crítica ao revisionismo contemporâneo*. Rio de Janeiro: Consequência, 2014.

TOLEDO, Caio Navarro de. Crônica política sobre um documento contra a 'ditabranda'. *Revista de Sociologia e Política*, Curitiba, v. 17, n. 34, out. 2009.

WOOLDRIDGE, Michael. *A Brief History of Artificial Intelligence: where it is, where we are and where we are going*. Nova Iorque: Flatiron Books, 2020.

## Como citar este artigo

Abal, F. C., & Cauduro Pureza, F. Inteligência artificial no ensino de História: ChatGPT e ditadura militar brasileira. *Estudos Ibero-Americanos*, e46780. <https://doi.org/10.15448/1980-864X.2026.1.46780>

*Os textos deste artigo foram revisados por Araceli Pimentel Godinho e submetidos para validação dos autores antes da publicação.*

---

## Fernando Cauduro Pureza

Doutor em História pela Universidade Federal do Rio Grande do Sul (UFRGS). Professor do Departamento de História da Universidade Federal da Paraíba (UFPB).

---

## Felipe Cittolin Abal

Doutor em História pela Universidade de Passo Fundo (UPF). Professor do Programa de Pós-Graduação em História e do Programa de Pós-Graduação em Direito na Universidade de Passo Fundo (UPF).

---

## Endereço para correspondência

### FERNANDO CAUDURO PUREZA

Rua Doutor Pedro Segundo de Araújo, 1460, apto 1503, Bairro Capim Macio, 59082-040

Natal, Rio Grande do Norte, Brasil

### FELIPE CITTOLIN ABAL

Rodovia BR 285, Km 292, 7000, Bairro São José, 99052-900

Passo Fundo, Rio Grande do Sul, Brasil