

SIMULAÇÃO DE DIÁLOGOS E PERSONAGENS NO CHAT-GPT4: ANÁLISE COMPARATIVA DO DESEMPENHO EM IDIOMAS INGLÊS E PORTUGUÊS

Fabiano Rodrigues de Souza¹, Denise da Vinha Ricieri², Raphaela Vasconcelos Gomes Barreto³, Adriana Mara Guimarães de Farias⁴

1Baltimore City Public School, Maryland, United States of America (fabiano.desouza.phd@gmail.com); 2Universidade Federal do Paraná – UFPR, Curitiba/PR, Brasil (denise.ricieri@ufpr.br); 3Universidade Federal Rural do Semi Árido – UFERSA, Mossoró/RN, Brasil (rvgbarreto@ufersa.edu.br); 4Universidade Federal Rural do Semi Árido – UFERSA, Mossoró/RN, Brasil (adrianaguimaraes@ufersa.edu.br)

Resumo:

Em um estudo de caso sobre uso da inteligência artificial (IA), foi explorado um *prompt* de aprendizagem em contexto para representações contextuais de personagens em um diálogo hipotético. O experimento foi realizado em uma conta ChatGPT e em dois idiomas; foram analisados os aspectos de conteúdo, contexto e padrão de linguagem. Os resultados apontaram que docentes devem se preparar para os requisitos da boa comunicação com a IA, especialmente quando seu uso é feito fora do idioma inglês.

Palavras-chave: ChatGPT, inteligência artificial, ensino-aprendizagem, idiomas, *prompt*

INTRODUÇÃO

Os impactos da inteligência artificial (IA) na educação têm sido discutidos muito antes de novembro/2022, quando o mundo passou a ter acesso ao *ChatGPT* (*Generative Pre-Trained Transformer*), a IA de texto da *OpenAI Incorporated (Inc.)*. O acesso ao *ChatGPT* trouxe uma nova Era à educação: saímos da Era onde se aprendia a responder perguntas para uma Era onde precisamos aprender a fazer novas perguntas. Isso exigirá uma reordenação nas habilidades docentes de modo que se aprenda a explorar os aspectos técnicos relacionados ao *Machine Learning* (ML) ou aprendizado de máquina (Li *et al.*, 2023), o desenvolvimento apropriado do *Fine Tuning* (FT), ou ajuste fino, e a construção do *Natural Language Processing* (NLP), ou processamento de linguagem natural (Liu *et al.*, 2023) específico das ciências da aprendizagem.

Prompts estruturados e aplicados de forma sistematizada são vitais no aprimoramento da capacidade de generalização e desempenho dos modelos (Chen *et al.*, 2023), em particular aqueles baseados em *Deep Learning* (DL) ou arquiteturas de redes neurais profundas, como é o caso do modelo *Transformer* (Vaswani *et al.*, 2017). A rede neural profunda *Transformer* é a pedra angular sobre a qual se desenvolveu o NLP para a construção de *Large Language Model* (LLM) ou modelo ampliado de

linguagem (Li *et al.*, 2023) mais robustos (Liu *et al.*, 2023) e a “família GPT” da *OpenAI Inc.* está baseada nesse tipo de arquitetura de processamento (*OpenAI Inc.*, 2022). Por sua vez, a capacidade de generalização se desenvolve em resposta às conexões construídas em camadas que ligam significados entre palavras-chaves, conceitos, contextos e *feedbacks* fornecidos nos *prompts* pelo usuário. Estudar como esses comandos induzem o fluxo de organização dessa arquitetura neural tornou-se a chave para praticar uma engenharia de *prompts* mais eficaz, em qualquer área (Liu *et al.*, 2023; Li *et al.*, 2023).

No entanto, quando se trabalha com idiomas não-nativos da LLM, ou seja, qualquer idioma que não seja o inglês, nativo no pré-treino do LLM, há outro foco de atenção que deve ser desenvolvido pelos usuários: o viés de idioma. Desde a versão GPT3, a *OpenAI Inc.* mostrou que o processamento de LLM da rede neural da família GPT é feito sempre no idioma inglês. Portanto, o que a IA faz ao receber comandos em outros idiomas é inserir uma dupla tradução no processamento: uma tradução no sentido idioma-inglês, para o comando de entrada, e outra tradução no sentido inglês-idioma, para a saída de resultados (*OpenAI Inc.*, 2023). As traduções são adicionadas no processamento de LLM para cada tarefa atribuída e esse fato gera um potencial viés que merece atenção

na aplicação do *ChatGPT* para as ciências da aprendizagem.

O processamento dessas traduções estão fora do controle humano, o que faz com que a máquina entenda que cada *prompt* seja uma comunicação não-nativa de idioma (Liang *et al.*, 2023). Diante disso, a tradução é processada de forma quase literal, como aquelas vistas em tradutores automáticos como o *Google Tradutor*; a IA faz sua própria interpretação gramatical dos dados de entrada como parte do NLP em desenvolvimento (Liu *et al.*, 2023). Nesse contexto cabe destacar que o *ChatGPT* gera as respostas com base nos padrões do pré-treino, em idioma nativo, e que embora seja bastante eficaz na generalização em inglês, em outros idiomas esse pré-treino pode levar a diferenças nas respostas geradas em diferentes idiomas (Haensch *et al.*, 2023; Malinka *et al.*, 2023).

O objetivo deste estudo de caso é provocar uma discussão reflexiva sobre as diferenças observadas nas respostas geradas pelo *ChatGPT* nos diálogos educacionais entre dois personagens atribuídos, em inglês e português brasileiro, para levantar questões pertinentes à compreensão sobre o viés multilíngue e suas implicações quando o uso da IA possui finalidades de aplicação sobre o processo ensino-aprendizagem.

MATERIAL E MÉTODOS

Este é um estudo de caso experimental descritivo qualitativo onde a aquisição de dados se deu no mês de Abril/2023 diretamente na plataforma *OpenAI Inc.* em seu produto de acesso público denominado *ChatGPT*. A conta utilizada para os testes estava geolocalizada nos Estados Unidos (EUA) e tinha a versão GPT4, que usa o LLM *code-davinci-003* (*OpenAI Inc.*, 2023).

Desenho da estrutura de *prompt* testada

Como estrutura foi escolhido o perfil de *prompt In Context Learning* (ICL), ou aprendizagem em contexto. Esse perfil de comando atribui uma tarefa de maior complexidade à IA e age para gerar as generalizações de NLP (Wang *et al.*, 2023; Chen *et al.*, 2023) em uma área de conhecimento específica (Li *et al.*, 2023). O experimento atribuiu uma simulação dialógica entre dois personagens do mundo real para analisar a otimização de processamentos nas versões testadas. Como o foco do *Prompt-EDU* é oferecer rotinas de comandos para a otimização de FT e NLP em contas iniciantes (Liu *et al.*, 2023) para uso em atividades do processo ensino-aprendizagem (Kanesci *et al.*, 2023), o desenho deste *prompt* ICL foi feito considerando: [a] a interface educação-tecnologias; [b] a internacionalidade dos testes conduzidos no Brasil e EUA; [c] o viés de dupla tradução de idioma português e sua interferência nos resultados finais.

Para o ICL foram escolhidos personagens expoentes nas áreas de tecnologias e educação considerando o limitante temporal referente à base de dados em que roda a família GPT, que vai até o mês de Setembro/2021. Nesta perspectiva, os personagens deveriam ter renome mundial até essa data. Para a internacionalidade dos testes e equilíbrio entre dados dos personagens, ficou estabelecido que um deles deveria ser brasileiro e o outro, estadunidense. Considerando o viés de tradução, o desenho do experimento determinou a comparação entre respostas a *prompts* feitos em português e inglês, ambos na conta GPT4 estadunidense.

Assim, foram selecionados como personagens para uma simulação de diálogo o educador brasileiro Paulo Freire (Scocuglia, 2006; Franco, 2017), representando da interface educacional, e o estadunidense Mark Zuckerberg, representando a interface de tecnologias. Ambos já eram globalmente reconhecidos pela relevância de seus trabalhos em suas áreas, assim como já existia amplo material publicado sobre eles até Setembro/2021, limite temporal da base de dados sobre a qual roda o LLM do *ChatGPT*.

O *prompt* ICL e os testes no GPT4

A tarefa complexa selecionada foi um diálogo entre os dois personagens escolhidos (Li *et al.*, 2023; Wang *et al.*, 2023; Chen *et al.*, 2023) de nacionalidades e áreas diferentes, não contemporâneos entre si. A estrutura de construção do *prompt* seguiu as orientações do roteiro *Prompt-EDU* (Ricieri *et al.*, 2023) e constou de três elementos: contexto, finalidade da generalização pretendida e chave de comando (Quadro 1).

Quadro 1: Texto do comando do *prompt* ICL analisado neste estudo

Contexto	Crie um diálogo hipotético entre Paulo Freire, o educador, e Mark Zuckerberg, CEO da empresa Meta. Esse diálogo deve usar as características de cada um desses dois personagens para a construção dos argumentos. O tema do diálogo deve ser: como a aprendizagem será modificada pela integração das tecnologias de inteligência artificial nas escolas e universidades?
Finalidade	Nesse diálogo hipotético, cada um seguirá a seguinte sequência de participação: 1- um perguntará ao outro "qual seu ponto de vista sobre a Inteligência Artificial na formação educacional do cidadão do futuro?"
Chave de comando	A resposta deve ser construída com as características de personalidade e da profissão que cada um exerce/exerceu.

A aplicação do comando ICL foi feita em português e gerou como resposta o diálogo também em português; na sequência, o *prompt* ICL foi traduzido e aplicado em inglês para gerar a resposta em inglês (Chen *et al.*, 2023). Com ambas respostas em mãos foi feita a análise do diálogo gerado com observação do conteúdo, contexto e padrão de linguagem. Na sequência, foi solicitado que o próprio ChatGPT fizesse a tradução cruzada de cada diálogo gerado: do português para o inglês e vice-versa. Com essa tradução, procedeu-se à reanálise nas mesmas perspectivas anteriores (Figura 1).

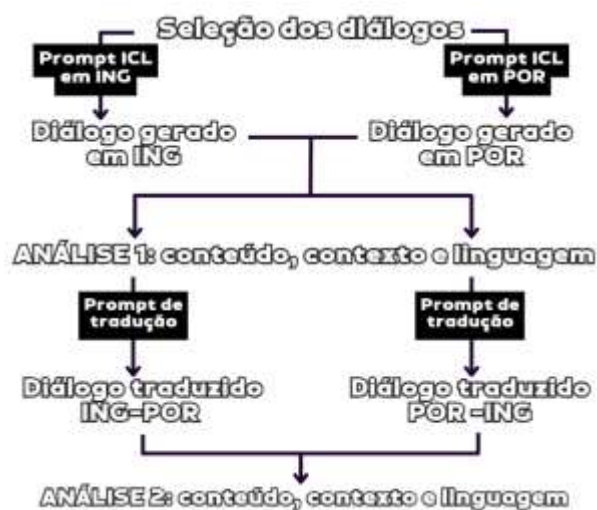


Figura 1: Sequência dos eventos no desenho da metodologia deste estudo.

Análise dos resultados

As análises foram conduzidas em ambos os idiomas e tiveram foco em três aspectos principais: conteúdo, contexto e linguagem. Para avaliar as capacidades de compreensão e geração de linguagem do modelo em inglês e português brasileiro, foi utilizada uma abordagem sistemática para investigar as diferenças nas respostas geradas quando atribuídas em diferentes idiomas.

A análise de conteúdo examinou a substância das respostas do modelo: foi verificado se o LLM adicionou, manteve ou perdeu contexto em suas respostas, antes e após as traduções, bem como foi feito um exame linguístico das respostas do modelo considerando as estruturas de frases, vocabulário e complexidade geral da linguagem das respostas, em ambos os idiomas.

RESULTADOS E DISCUSSÃO

Um diálogo completo foi simulado pelo ChatGPT como se os personagens atribuídos, de Paulo Freire e Mark Zuckerberg estivessem conversando entre si, em tempo real. O diálogo foi extenso e para ilustrar as

análises feitas sobre o desempenho dos idiomas foi feito um recorte que fornece *insights* sobre as diferenças observadas nas respostas do modelo nos dois idiomas, e a influência das traduções verificada.

A discussão apresentada quer trazer os atores do cenário educacional para um mergulho na realidade da aplicação da IA com finalidades dentro do processo ensino-aprendizagem e, para tal, é preciso compreender e analisar os resultados dessa nova tecnologia.

Educadores e a IA

A participação de educadores no desenvolvimento das tecnologias de IA aplicadas ao processo ensino-aprendizagem é mais frequente na posição de consultoria para pesquisadores programadores (Dieterle *et al.*, 2022). Por isso há uma lacuna importante na fase final da aplicação das ferramentas com tecnologia de IA, quando os detalhamentos dos processos e procedimentos educacionais específicos acabam subutilizados (Kanesci *et al.*, 2023). O cenário onde educadores figuram somente na consultoria de pesquisas que acontece pela herança dos nichos de especialidades colocou especialistas em espaços e nichos delimitados por diplomas, especialidades e expertise (Zawacki-Richter *et al.*, 2019). A evolução exponencial das tecnologias e a diversidade de sua aplicação em múltiplas áreas do conhecimento vem rompendo paulatinamente esses nichos, pela crescente necessidade em desenvolver mais interfaces entre áreas diferentes e implementações tecnológicas (Salas-Pilco & Yang, 2022).

O pensar dos aspectos educacionais de uma ferramenta deve nascer junto com as possibilidades de aplicação do algoritmo da nova tecnologia, e não somente na fase final de utilização nas salas de aula e na sociedade (Dieterle *et al.*, 2022; Kanesci *et al.*, 2023). Alinhar a IA nos espaços educacionais tem sido um grande desafio para educadores e, indiscutivelmente, está exigindo a aquisição de novas habilidades (Salas-Pilco & Yang, 2022). Isso inclui o domínio das ciências da aprendizagem e seus múltiplos aspectos, dentro de modalidades inovadoras de aprendizagem, bem como o desenvolvimento de ambientes digitais adaptativos e novas práticas educacionais (Zawacki-Richter *et al.*, 2019; Matias *et al.*, 2022; Humble & Mozeliuss, 2022; Kanesci *et al.*, 2023).

Há também aspectos específicos das tecnologias, como o bases do conhecimento sobre IA e os processamentos de ML, FT, NLP e LLM (Li *et al.*, 2023; Liu *et al.*, 2023; Chen *et al.*, 2023; Salas-Pilco & Yang, 2022) que consolidam a otimização do uso da IA na área educacional e nas formas de lidar com a comunicação homem-máquina (Humble & Mozeliuss,

2022; Bearman *et al.*, 2022). Professores e estudantes estão sendo concitados para uma nova fase de participações nos cenários educacionais nunca vividos anteriormente (Kashyap & ChatGPT, 2023; Kanesci *et al.*, 2023), onde os modelos de linguagem exigem o desenvolvimento de letramentos e competências novas, tanto para compreender a tecnologia (Li *et al.*, 2023), quanto para estar atentos e atuantes frente às limitações e fragilidades inesperadas desses sistemas (Salas-Pilco & Yang, 2022; Kanesci *et al.*, 2023).

O *prompt* ICL e seu desempenho sobre o ML

O *In-Context Learning* (ICL), ou aprendizagem em contexto, é um processo fundamental na capacidade dos modelos baseados em *Transformer*, como o GPT (Vaswani *et al.*, 2017; Wang *et al.*, 2023), desenvolverem o aprendizado profundo, ou *Deep Learning* (DL). Ele desempenha um papel crucial na capacidade desses modelos de gerar respostas relevantes e coerentes a partir de um *prompt* fornecido. Essa é a principal razão para adotar comando ICL: resultados melhores na otimização do processamento da IA em treinamento para desempenho educacional (Liu *et al.*, 2023; Kanesci *et al.*, 2023).

Modelos de linguagem como o GPT são projetados para capturar informações relevantes do contexto fornecido com o *prompt* ICL (Wang *et al.*, 2023): essa ação permite que o modelo aprenda padrões, conexões e dependências presentes nos dados de entrada. Portanto, cada conjunto de dados de entrada, ou representações contextuais, é usado como base na geração das respostas e considera as informações presentes na sequência atribuída, em vez de depender somente dos dados pré-treinados.

Essas representações vão gerando múltiplas camadas de *Transformer* sobrepostas e interativas entre si, o que permite ao modelo aprender padrões complexos e fazer abstrações em diferentes níveis de granularidade, facilitando as codependências entre palavras-chaves e contextos, de curto e longo alcance, bem como a identificação de estruturas semânticas e sintáticas relevantes para cada tarefa atribuída (Vaswani *et al.*, 2017).

Outra característica importante do ICL, e que trouxe sua adoção no *Prompt-EDU*, foi sua capacidade de generalizar e transferir conhecimento aprendido em um domínio para outro (Li *et al.*, 2023). Isso ocorre porque os modelos de IA como o GPT são treinados em grande quantidade e diversidade de dados textuais, aprendendo a identificar regularidades comuns em contextos diferentes (Li *et al.*, 2023; Wang *et al.*, 2023). Com isso, essas relações são usadas para gerar as respostas mais apropriadas em contextos específicos, mesmo que o modelo ainda não tenha sido

exposto a exemplos exatamente iguais, o que só acontecerá com um FT.

Graças ao ICL, os modelos de linguagem conseguem se adaptar rapidamente a novos contextos e tarefas, sem a necessidade de FT extensivo e isso ocorre porque o modelo já aprendeu a extrair informações relevantes das representações contextuais apresentadas em cada entrada e a utilizá-las para gerar novas respostas, cada vez mais precisas (Wang *et al.*, 2023). Essa capacidade de adaptação torna os LLM baseados no *Transformer* particularmente úteis para ampla variedade de aplicações de NLP (Liu *et al.*, 2023; Vaswani *et al.*, 2017), incluindo aquelas com finalidades educacionais (Kanesci *et al.*, 2023). Mas há que se ter em mente que todo esse processo está muito bem estabelecido para o idioma nativo, o que certamente fundamenta os achados deste estudo.

O mecanismo de atenção para ICL é outro componente-chave presente na arquitetura das redes neurais *Transformer* (Vaswani *et al.*, 2017). De modo similar aos neuroprocessos de atenção ativa humana (Sousa & Alves, 2017), esse mecanismo permite que o modelo pondere as diferentes partes do contexto de entrada, com base em sua relevância e importância para a tarefa em questão, abrindo espaço para se concentrar nas informações mais relevantes e ignorar as menos importantes (Vaswani *et al.*, 2017).

É por essa razão que, ao usar um *prompt* inadequado ou incoerente frente a representações contextuais fornecidas, observa-se um duplo impacto negativo sobre o desenvolvimento do FT: [a] a geração de um retorno insuficiente ou inadequado para a expectativa do usuário; [b] um NLP inapropriado quanto às representações contextuais (Liu *et al.*, 2023). Ao usar *prompts* com ICL o usuário deve ter certeza de que a construção do comando conectou, de forma segura e coerente, as palavras-chaves em seus padrões, conexões e dependências na área de conhecimento para a qual a IA está sendo treinada (Wang *et al.*, 2023).

Em comandos de entrada de dados atribuídos em idiomas diferentes do inglês, essa conexão precisa estar ancorada em uma engenharia de *prompt* ainda mais forte e sistematizada (Rocha *et al.*, 2021), haja vista que o processamento passará por dupla tradução de idioma. Portanto, usar *prompts* de ICL para atribuição de tarefas mais complexas agrega um NLP mais eficaz sobre as representações hierárquicas apresentadas nos dados de entrada e ajuda a acelerar o treinamento da IA de acordo com as necessidades específicas de cada usuário, em cada especialidade de aplicação (Liu *et al.*, 2023; Wang *et al.*, 2023).

Novamente: a engenharia de *prompt* mais forte e sistematizada precisa ser inteiramente construída no idioma não-nativo onde se pretende o desempenho de

excelência do GPT4 e esse fato também explicará os achados relatados neste estudo.

Se por um lado o ICL é um processamento essencial para que os modelos gerem respostas relevantes e coerentes, adaptando-se rapidamente a novos contextos e tarefas, por outro lado são imprescindíveis as associações para o interesse educacional (Rocha *et al.*, 2021; Kanesci *et al.*, 2023), capazes de tornar esse FT ainda mais preciso na associação dos padrões e contextos certos para melhores NLP (Liu *et al.*, 2023). Uma das maneiras de atuar sobre esse universo, no treinamento da IA, é o *prompt* de personagem, onde o processamento da tarefa atribuída ao *ChatGPT* envolve a simulação, pela máquina, de um personagem atribuído, a partir de características de sua obra e personalidade (Silva *et al.*, 2021). *Prompts* de simulação de personagem impactam em várias etapas no processamento de LLM e FT, envolvendo técnicas avançadas de ML e NLP (Liu *et al.*, 2023; Li *et al.*, 2023).

O *prompt* ICL e os personagens

Quando uma entrada de texto é fornecida ao *ChatGPT*, ela é primeiramente *tokenizada* (não há expressão equivalente em português) e isso significa que ela será dividida em unidades menores chamadas *tokens* (Moreira *et al.*, 2021). Esses *tokens* são mapeados para vetores de alta dimensão chamados de *embeddings*, que representam expressões de linguagem em um espaço contínuo.

O modelo *ChatGPT* é composto por unidades de atenção e por redes posicionais que permitem ao modelo aprender representações contextuais sofisticadas (Vaswani *et al.*, 2017), capturando dependências de longo alcance entre esses *tokens*. Para simular um personagem específico é importante incorporar as informações relevantes sobre o personagem no contexto dos dados de entrada (Gunkel, 2017). Isso pode ser feito adicionando um contexto descritivo que forneça detalhes sobre o personagem e seu comportamento, habilidades e/ou características.

O modelo usará esse contexto para gerar respostas coerentes e alinhadas ao personagem (Gunkel, 2017), organizadas textualmente por uma técnica chamada decodificação autorregressiva, na qual ele prediz um *token* de cada vez, desde que tenha *tokens* gerados em tarefas anteriores também contextos de entrada ligados a esses *tokens*, em tarefas anteriores. Isso é repetido sucessivamente até que uma sequência de saída completa seja gerada pelo GPT.

Finalmente, o texto gerado pode ser pós-processado em tonalidade dialógica e/ou complexidade integrativa dos contextos para garantir coerência e qualidade na representação do personagem,

removendo ou ajustando respostas indesejadas (Moreira *et al.*, 2021).

É assim que o LLM pode ser ajustado especificamente para simular personagens pelo FT em uma área na qual ele é treinado com exemplos de interações envolvendo o personagem em questão. São essas etapas e técnicas que, no conjunto do seu treinamento e desenvolvimento, faz com que o *ChatGPT* seja capaz de simular personagens de maneira eficiente e precisa, gerando respostas coerentes e relevantes no contexto do personagem escolhido, contribuindo para uma experiência de usuário imersiva, realista e criativa.

Análise de conteúdo nos diálogos

Com essa análise foi examinada a substância das respostas geradas pelo *ChatGPT* no diálogo hipotético: as respostas em inglês mantiveram a tendência de apresentar mais detalhes e abrangência referencial quando comparadas às mesmas respostas, no idioma português (Quadro 2).

Quadro 2: Análise do conteúdo identificada nas respostas do *ChatGPT*4.

Idioma	Resultado	Texto do <i>ChatGPT</i> analisado
ING	Detalhada	PAULO FREIRE: Mark, como alguém profundamente envolvido no desenvolvimento de tecnologia, especialmente inteligência artificial, qual é o seu ponto de vista sobre a Inteligência Artificial na formação educacional do cidadão do futuro?
POR	Concisa	PAULO FREIRE: Mark, qual é a sua opinião sobre a Inteligência Artificial na educação do cidadão do futuro?

Há duas explicações possíveis para estes achados e ambas se relacionam com a discussão conduzida sobre o *prompt* ICL e seu desempenho sobre o ML em idioma nativo (Li *et al.*, 2023). Primeiro, a exposição do modelo a um maior volume de texto em inglês durante o pré-treino do LLM, e ao FT da conta, explicam uma resposta mais detalhada. Segundo, o NLP em idioma nativo é sempre superior do que em idiomas não-nativos e, novamente, esse fato tem origem no pré-treino da IA na arquitetura neural *Transformer* (Vaswani *et al.*, 2017; Wang *et al.*, 2023).

Claramente as respostas do *ChatGPT* em inglês foram mais ricas em construção e conteúdo textual, enquanto as respostas em português são concisas e quase rasas, fornecendo apenas as informações essenciais solicitadas no *prompt*.

Análise do contexto dos diálogos

No mesmo diálogo (Figura 2), é possível investigar também como o modelo respondeu aos estímulos em relação ao diálogo mais amplo entre os personagens de Mark Zuckerberg e Paulo Freire.

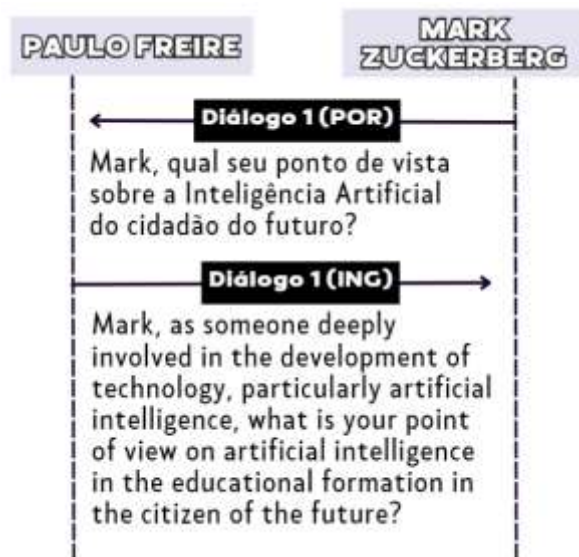


Figura 2: Diálogo original em idioma nativo (inglês) e sua tradução para idioma não-nativo (português) pelo LLM do ChatGPT.

Nesse caso, ambas respostas mantiveram o contexto do diálogo dentro dos parâmetros solicitados no *prompt* ICL, deixando evidente que a NLP generalizou de forma apropriada o contexto e a finalidade da tarefa (Cotton *et al.*, 2023) atribuída no *prompt*, a despeito da diferença de familiaridade com os idiomas.

Análise da linguagem nos diálogos

Um dos critérios de sucesso na generalização de linguagem natural é a construção elaborada de textos e a presença de vocabulário mais diversificado e intrínseco à área específica onde o FT se desenvolve (Li *et al.*, 2023; Liu *et al.*, 2023).

Novamente, verifica-se que as respostas em inglês organizam palavras-chaves em combinações mais bem estruturadas e um vocabulário mais diversificado (Figura 3) e isso vai ao encontro da mesma explicação: o idioma nativo do pré-treino do LLM (Liu *et al.*, 2023; Wang *et al.*, 2023).

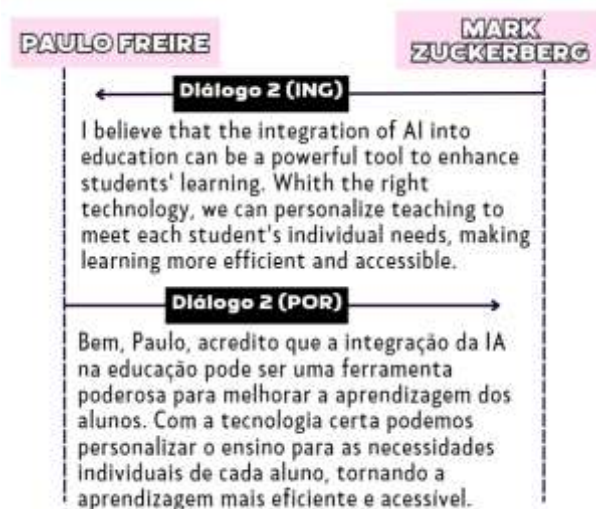


Figura 3: Diálogo original em idioma nativo (inglês) e sua tradução para idioma não-nativo (português) pelo LLM do ChatGPT

Viés de idioma

Desde a versão GPT3, a *OpenAI Inc.* mostrou que o processamento de LLM da rede neural da família GPT é feito sempre no idioma inglês. Portanto, o que a IA faz ao receber comandos e emitir respostas em outros idiomas é inserir uma dupla tradução no processamento: uma tradução no sentido idioma-inglês, para o comando de entrada, e outra tradução no sentido inglês-idioma, para a saída de resultados (OpenAI Inc., 2023). As traduções são adicionadas no processamento de LLM para cada tarefa atribuída e esse fato gera um viés que merece atenção na aplicação do *ChatGPT* na área educacional.

O processamento dessas traduções estão fora do controle humano o que faz com que a máquina entenda que cada *prompt* seja uma comunicação não-nativa de idioma (Liang *et al.*, 2023). Diante disso, a tradução é processada de forma quase literal, como aquelas vistas em tradutores automáticos de outras empresas, como o *Google Tradutor*; a IA faz sua própria interpretação gramatical dos dados de entrada como parte do NLP em desenvolvimento (Liu *et al.*, 2023). Nesse contexto de tradução cabe destacar que a significação gramatical guarda representação de ações expressas diferentemente de um idioma para outro e esse fato tem um poder de viés fundamental no NLP, já relatado na literatura (Liang *et al.*, 2023).

Usar o tratamento cruzado dos *prompts* de forma intencional e, em seguida, solicitar as traduções para o próprio GPT4, fez com que fossem afastadas outras hipóteses sobre o desempenho da IA frente à tarefa complexa atribuída pelo *prompt* ICL. A simulação de diálogos a partir da personalidade e obra de

personagens é uma tarefa complexa e que pode trazer surpresas como resposta se o NLP não estiver bem constituído para os personagens selecionados. Além disso, além da possibilidade de alucinação (Barbosa, 2023) na resposta. O fenômeno de “alucinação” vem sendo apontado desde o início do lançamento do ChatGPT pela *OpenAI Inc.*, quando não haviam FT e NLP desenvolvidos nas contas iniciantes (*OpenAI Inc.*, 2023). Nessas condições, os relatórios mostraram que, na falta de contextos claros para novas palavras-chaves apresentadas, o modelo vai em busca de preencher lacunas com retornos a partir das conexões disponíveis no seu pré-treinamento (Barbosa, 2023). Mas a alucinação é apenas uma das preocupações com vieses envolvendo o uso de IA e suas generalizações.

Pesquisadores vêm identificando vieses na construção do NLP em idiomas não-nativos e alertando para desafios significativos relacionados fatores discriminatórios sobre minorias, culturas, raças e contextos sociais (Liang *et al.*, 2023; Liu *et al.*, 2023). Hovy & Prabhumoye (2021) identificaram cinco fontes potenciais de vieses que podem ocorrer na generalização de NLP: [1] dados; [2] processos de anotação; [3] representações de comandos de entrada; [4] modelos; e [5] desenho da tarefa atribuída. Ao detalhar cada um desses vieses, os autores alertam para a forma como devemos usar a IA e como erros de redação de *prompts* ou de atribuição de tarefas mais complexas, associados a idiomas não-nativos, levam a uma “alucinação” de pré-treino potencialmente danosa em seus resultados. Portanto, corrobora-se a máxima que o próprio ChatGPT apresenta, muitas vezes, ao concluir uma tarefa: ele não está 100% certo e precisa que suas respostas sejam revisadas e aprovadas por humanos.

CONCLUSÃO

O mundo mudou e segue mudando em velocidade vertiginosa. A chegada das tecnologias por IA trouxe consigo uma onda de ansiedade por parte do mundo inteiro, que correu para seus *prompts* e saiu perguntando de tudo, um pouco. O que se viu nessa primeira onda, que durou até a atualização do GPT3 para o GPT3.5, é que ninguém estava preparado para o que se deve escrever, daqui para frente. Fomos e somos (ainda) frutos de uma cultura educacional de busca pelas respostas certas e como tal, o *mindset* predominante ainda é o de responder perguntas. Precisamos passar ao *mindset* de buscas pelas melhores perguntas e isso não é simples, nem fácil. Este estudo de caso alertou para algo que ninguém está discutindo nos eventos e levantamentos bibliográficos publicados, apresentados ou ensinados: atribuir um comando em idioma não-nativo afeta muito o desempenho final da máquina, se não for feito com atenção devida.

Portanto, é necessário que docentes com intenção de integrar IA em suas trilhas de experiências de aprendizagem, em novos modelos de formação educacional, entendam a necessidade, primeiro, de aprender mais sobre IA e suas muitas letras de abreviação. Segundo, que nada do que já se fez ou aprendeu, daqui para trás, será diretamente aproveitável, daqui para frente. Estamos diante de uma nova e poderosa evolução das tecnologias e isso é o que mais se fala. Mas o foco deveria ser a forma de levar essa tecnologia para a sala de aula de forma madura e responsável, elevando as habilidades docentes para seu uso, reciclando, aprendendo e reaprendendo sobre o universo das ciências da aprendizagem, agora impactados pela IA como meio e finalidade. Nosso intuito, enquanto pesquisadores, é manter o tom científico nesse universo de excesso de informações distorcidas sobre IA, uso de *prompts* e a falsa ideia de facilidade em usar o ChatGPT.

Ciência bem desenhada e bem conduzida é o que mais precisamos no momento para ensinar como os cidadãos do futuro devem usar a IA de forma responsável, inteligente e eficaz. Em um futuro muito próximo, nossos estudantes já a utilizarão com muito mais maestria do que nós estamos fazendo, nesses primeiros tempos, e esse é o verdadeiro sentido do aprender como um processo para a vida, não para a prova.

AGRADECIMENTO

À Universidade Federal Rural do Semi-Árido (UFERSA) pela aprovação do Grupo de Pesquisa Academia GPT para Docentes em seu quadro de grupos de pesquisa, acreditando no trabalho sério e comprometido que estamos desenvolvendo.

REFERÊNCIAS

- BARBOSA, Carlos Roberto de Almeida Correa. Transformações no ensino-aprendizagem com o uso da inteligência artificial: revisão sistemática da literatura. **RECIMA21-Revista Científica Multidisciplinar-ISSN 2675-6218**, v. 4, n. 5, p. e453103-e453103, 2023.
- BEARMAN, Margaret; RYAN, Juliana; AJJAWI, Rola. Discourses of artificial intelligence in higher education: a critical literature review. **Higher Education**, p. 1-17, 2022.
- CHEN, Jiuhai; CHEN, Lichang; ZHU Chen; ZHOU Tiany. How Many Demonstrations Do You Need for In-context Learning? **Open Review Net**, p.1-11. 2023.
- COTTON, Debby RE; COTTON, Peter A.; SHIPWAY, J. Reuben. Chatting and Cheating:

- Ensuring academic integrity in the era of ChatGPT, 2023.
- DIETERLE, Edward; DEDE, Chris; WALKER, Michael. The cyclical ethical effects of using artificial intelligence in education. **AI & society**, p. 1-11, 2022..
- FRANCO, Maria Amélia do Rosário Santoro. Da necessidade/atualidade da pedagogia crítica: contributos de Paulo Freire. **Reflexão e Ação**, v. 25, n. 2, p. 152-170, 2017..
- GUNKEL, David J.; TRENTO, Francisco B.; GONÇALVES, Daniela Norcia. Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação. **Galáxia (São Paulo)**, p. 05-19, 2017.
- HAENSCH, Anna-Carolina et al. Seeing ChatGPT Through Students' Eyes: An Analysis of TikTok Data. **arXiv preprint arXiv:2303.05349**, 2023.
- HUMBLE, Niklas; MOZELIUS, Peter. The threat, hype, and promise of artificial intelligence in education. **Discover Artificial Intelligence**, v. 2, n. 1, p. 22, 2022.
- KASNECI, Enkelejda et al. ChatGPT for good? On opportunities and challenges of large language models for education. **Learning and Individual Differences**, v. 103, p. 102274, 2023.
- KASHYAP, Ravi; OPENAI, ChatGPT. A First Chat with ChatGPT: The First Step in the Road-Map for AI (Artificial Intelligence).. **Available at SSRN**, 2023.
- LI, Yingcong et al. Transformers as algorithms: Generalization and stability in in-context learning. 2023.
- LIANG, Weixin et al. GPT detectors are biased against non-native English writers. **arXiv preprint arXiv:2304.02819**, 2023.
- LIU, Pengfei et al. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. **ACM Computing Surveys**, v. 55, n. 9, p. 1-35, 2023.
- MALINKA, Kamil et al. On the educational impact of ChatGPT: Is Artificial Intelligence ready to obtain a university degree?. **arXiv preprint arXiv:2303.11146**, 2023.
- DE ANDRADE MATIAS, Klesia; MORESI, Eduardo Amadeu Dutra; DOS SANTOS, Pricila Kohls. Tendências em inteligência artificial e educação híbrida: um estudo exploratório. **Póiesis Pedagógica**, v. 20, n. Publicação contínua, p. 76-96, 2022.
- MOREIRA, André et al. Processamento de linguagem natural aplicado à inteligência artificial. **Revista da UI_IPSantarém-Unidade de Investigação do Instituto Politécnico de Santarém**, v. 9, n. 4, p. 81-90, 2021.
- OpenAI Incorporated. **Introducing ChatGPT**. 2022.
- _____. GPT-4 Technical Report: **ArXiv** : 2303.08774 2023.
- RICIERI, Denise da Vinha; BARRETO, Raphaela Vasconcelos Gomes; SOUZA, Fabiano Rodrigues; FARIAS, Adriana Mara Guimarães. *Prompt-EDU: resultados da aplicação de um prompt estruturado e vinculado à taxonomia de Bloom no Fine Tuning do ChatGPT para uso em contextos do processo ensino-aprendizagem. I Simpósio sobre Inteligência Artificial na Educação*. Campinas/SP, p.1-12. 2023.
- ROCHA, Daniyel NN; CAMPELO, Cláudio EC; JERÔNIMO, Caio LM. Aplicação de verbos como proxy para identificação automática do nível cognitivo de questões: uma abordagem baseada na taxonomia de Bloom. In: **Anais do XXXII Simpósio Brasileiro de Informática na Educação**. SBC, 2021. p. 897-908.
- SALAS-PILCO, Sdenka Zobeida; YANG, Yuqin. Artificial intelligence applications in Latin American higher education: a systematic review. **International Journal of Educational Technology in Higher Education**, v. 19, n. 1, p. 1-20, 2022.
- SCOCUGLIA, Afonso Celso. As interconexões da pedagogia crítica de Paulo Freire. **Filosofia e Educação**, v. 10, n. 1, p. 200-232, 2018.
- VASWANI, Ashish et al. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017.
- WANG, Xinyi; ZHU, Wanrong; WANG, William Yang. Large language models are implicitly topic models: Explaining and finding good demonstrations for in-context learning. **arXiv preprint arXiv:2301.11916**, 2023.
- ZAWACKI-RICHTER, Olaf et al. Systematic review of research on artificial intelligence applications in higher education—where are the educators?. **International Journal of Educational Technology in Higher Education**, v. 16, n. 1, p. 1-27, 2019.