

Entrevistando um Robô: notas sobre a aplicação experimental da metodologia EEAF usando a ferramenta ChatGPT de inteligência artificial

Interviewing a Robot: notes on the experimental application of the EEAF methodology using the ChatGPT artificial intelligence tool

Márcio Carneiro dos Santos¹ⁱ

Orcid: 0000-0002-0663-1300

Recebido em: 14/02/2023. Aprovado em: 10/04/2023

Resumo

Apresenta-se a metodologia EEAF - Entrevista Exploratória de Avaliação de Funcionalidades - a ser utilizada por pesquisadores da Comunicação, no estudo de características ou análise de material textual, criado por algoritmos de inteligência artificial. Neste trabalho usamos a ferramenta de conversação ChatGPT e propomos um conjunto de métricas: transparência, estruturação, precisão, variedade e atualização, para avaliação no experimento. No geral, os textos criados via código apresentaram a necessidade de revisão humana, principalmente em materiais que lidavam com fatos e dados objetivos.

Palavras-chave: inteligência artificial; ChatGpt; processamento de linguagem natural.

Abstract

It presents the methodology EEAF - Exploratory Interview for the Evaluation of Functionalities - to be used by Communication researchers, in the study of characteristics or analysis of textual material, created by artificial intelligence algorithms. In this work we use the ChatGPT conversation tool and propose a set of metrics: transparency, structuring, precision, variety and updating, for evaluation in the experiment. In general, texts created via code showed the need for human review, especially in materials that dealt with objective facts and data.

Keywords: artificial intelligence; ChatGpt; natural language processing.

1 Introdução

Uma das vertentes mais promissoras da interface entre comunicação e sistemas computacionais são os modelos de processamento de linguagem natural (*Natural Language Processing - NLP*). Em termos simples, algoritmos de inteligência artificial que são treinados

¹ Universidade Federal do Maranhão. São Luís/Brasil. E-mail: marcio.carneiro@ufma.br



para lidar com texto e, mais especificamente, entender e escrever palavras com estrutura e lógica semelhante à humana; a ponto de oferecer respostas bastante convincentes e de utilização em diversos cenários, incluindo comunicação e jornalismo (AP, 2022) e, no caso do nosso objeto, a solução ChatGPT², que detalharemos a seguir, difíceis de identificar como narrativa automatizada.

Na área da Comunicação, o tópico da inteligência artificial tem recebido atenção crescente neste século, no caso do Brasil através de precursores como Squirra (2016; 2016a); Lima Jr e Vergili (2018) e Dos Santos (2016; 2018; 2019) e, mais recentemente, Lemos (2020). No cenário internacional, com bibliografia mais extensa, poderíamos citar Reiter, Sripada e Robertson (2003); Jones (2013); Mayer-Schönberger, Cukier (2013); Musiani (2013); Napoli (2013); Manovich (2014); Latzer (2016); Guzman e Seth (2019; 2018) e, mais recentemente Veel (2018), entre outros.

O processamento de linguagem natural é uma das grandes subdivisões da inteligência artificial, ao lado do aprendizado de máquina e a visão computacional. Contudo, mais especificamente, a ferramenta usada neste projeto é caracterizada como pertencente à subárea da inteligência artificial generativa ou gerativa, que, em termos simples, tenta simular a própria criatividade humana, a partir da geração de textos e imagens³.

O algoritmo utilizado neste experimento é o *bot* ChatGpt⁴, que tem recebido muito interesse da mídia em geral e também do meio acadêmico, entre outras características por ter conseguido mais de um milhão de usuários inscritos em apenas cinco dias (MOLLMAN, 2022), bem como pela eficiência, em relação a algoritmos de gerações anteriores, para produzir textos, sua função principal.

GPT é um modelo de linguagem pré-treinado (*Pre-trained Language Model, PLM*) desenvolvido pela OpenAI⁵. Ele foi construído usando uma grande quantidade de documentos da internet e é capaz de gerar texto natural e coerente, quando alimentado com uma determinada tarefa ou contexto. GPT pode ser usado em várias aplicações, incluindo chatbots, geradores de conteúdo e sistemas de tradução automática.

Assim, ChatGPT é um chatbot, um algoritmo que simula uma conversa com o usuário, que usa o modelo GPT para gerar respostas para perguntas, solicitações e comentários. Ele

² Bot desenvolvido pela empresa OpenAI que usa inteligência artificial. <https://chat.openai.com/>.

³ No caso de geração de imagens, ver www.midjourney.com.

⁴ Ver detalhes em - <https://openai.com/blog/chatgpt/>.

⁵ Empresa de tecnologia que desenvolve vários produtos usando inteligência artificial. <https://openai.com/>.

pode ser usado em uma variedade de contextos como atendimento ao cliente, treinamento de funcionários e conversas virtuais. O ChatGPT é projetado para ser capaz de entender o contexto da conversa e gerar respostas coerentes, escrevendo em várias línguas, inclusive português.

Neste trabalho relatamos um experimento que levantou dados a partir de uma sequência de conversas com a ferramenta ChatGPT, a fim de estabelecer uma possível abordagem para coleta e análise de dados, quando o objeto de estudo é um algoritmo capaz de simular diálogos e produção de sentido.

A proposta é explorar um tipo de aproximação mais adaptada ao conjunto do ferramental metodológico conhecido pelos pesquisadores da Comunicação, além de bastante comum no tipo de treinamento que recebem em sua formação: as entrevistas em profundidade.

A diferença é que, neste caso, o entrevistado foi, em nosso experimento, não um humano, mas um algoritmo de processamento de linguagem natural, capaz de entender e gerar textos e, por isso, conversar e responder perguntas.

Como resultado do próprio experimento, algumas das suas respostas serão citadas neste texto, marcadas em *itálico* para identificação, entre outras finalidades para demonstrar o grau de precisão com que consegue conversar e lidar com as questões. As limitações e possibilidades da sua performance nessa tarefa também serão apresentadas adiante.

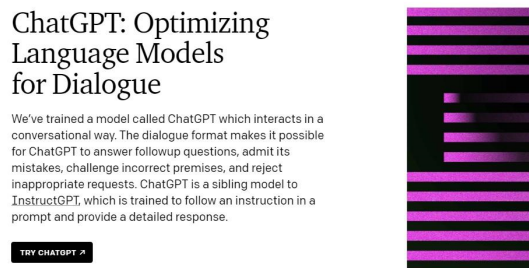
A metodologia EEAF, uma sigla para Entrevista Exploratória de Avaliação de Funcionalidades, é um recorte de um trabalho mais amplo direcionado a desenvolver um protocolo de análise e um conjunto de ferramentas, dedicados a pesquisadores da Comunicação e também das Ciências Sociais e Humanidades, para situações em que precisem lidar com grandes volumes de textos ou automatizar procedimentos de coleta e análise de dados.

O projeto, que tem caráter interdisciplinar e financiamento FAPEMA⁶, consiste num esforço de aproximação entre áreas distintas que, principalmente pelos processos de digitalização e pelo crescimento da adoção e potencialidades de tecnologias emergentes como inteligência artificial, big data, realidade virtual entre outras, faz-se, em nossa visão, cada vez mais necessário.

⁶ Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão.



Figura 1 - Tela inicial da ferramenta ChatGPT da OpenAI.



Fonte: OpenAI

2 Como estudar robôs que conversam - metodologia EEAF

Neste trabalho apresentamos a metodologia EEAF, destinada a coletar dados iniciais para uma abordagem exploratória quando o objeto da investigação é uma solução de inteligência artificial que gera textos e consegue dialogar com um interlocutor humano que pode pedir a ela diversas tarefas. É bom ressaltar que, apesar de estarmos usando aqui o ChatGPT, existem várias outras ferramentas semelhantes que, em tese, podem ser avaliadas da mesma forma.

A inspiração do método foi o experimento conhecido como teste de Turing, concebido pelo matemático e lógico Alan Turing, no final da década de 1940, cuja proposta era determinar se uma máquina seria capaz de exibir inteligência humana. O teste e seu criador, para alguns, são precursores de todo o desenvolvimento posterior do campo científico hoje conhecido como inteligência artificial.

O experimento descrito utilizou a solução ChatGPT da OpenAI que no seu site oficial o descreve como:

Um aperfeiçoamento a partir do GPT-3.5, um modelo de linguagem treinado para produzir texto. O ChatGPT foi otimizado para diálogo usando aprendizado por reforço com feedback humano (RLHF) - um método que usa demonstrações humanas para orientar o modelo para o comportamento desejado (OPENAI, 2022).

Uma das questões levantadas com frequência em estudos que tem como objeto algoritmos de inteligência artificial ou semelhantes é que são caixas opacas, cuja apreensão dos seus processos internos nem sempre é facilitada, até por questões de propriedade industrial.

A proposta metodológica da EEAF, entrevista exploratória para identificação de funcionalidades, é baseada na ideia de um processo de engenharia reversa, através do qual se

busca entender as principais características do sistema em análise, a partir das próprias respostas de texto que o modelo gera.

Como o nome sugere, trata-se de uma entrevista em profundidade direcionada ao modelo em questão para que, através da interface textual, ele possa explicitar ou até revelar informações sobre sua operação que não estão claramente disponíveis na documentação oficial ou pública. Os dados coletados são complementados com as informações oficiais e inclusive, comparados com as mesmas, para efeito de validação.

A proposta da EEAF justifica-se pela dificuldade frequente em se obter informações mais precisas sobre os modos de operação e detalhes do modelo em si, gerada muitas vezes pelo entendimento de que se trata de ativo da empresa, cujo conteúdo deve ser protegido pela legislação de propriedade intelectual ou industrial, o que dificulta o acesso do pesquisador.

Em estudos, como de Amadeu (2019) e Silva (2019), questões bastante sensíveis para as companhias de tecnologia que desenvolvem algoritmos com finalidade semelhante à ChatGPT, tais como racismo algorítmico, controle, vigilância, entre outros, de certa forma explicam (sem justificar, contudo) a eventual dificuldade de acesso aos procedimentos internos para que tal ferramenta funcione.

Para finalidade de estudo, são comuns a coleta e a análise das saídas dos sistemas, como material que o pesquisador utiliza para compreensão do objeto. O diferencial da metodologia EEAF é que ela propõe fazer tanto a coleta, como a análise a partir de parâmetros definidos, o que permite sua utilização por outros pesquisadores e diante de outros modelos de geração de linguagem, desde que, do mesmo tipo e desenvolvimento do ChatGPT.

Seria pouco produtivo usar tais procedimentos em soluções antigas ou limitadas que provavelmente pouco poderiam contribuir para as pesquisas desse campo ainda novo. A metodologia é desenhada para aplicação no estado da arte disponível de ferramentas de geração de texto, baseadas em inteligência artificial que, em tese, vão evoluir e se tornar mais frequentes de agora em diante.

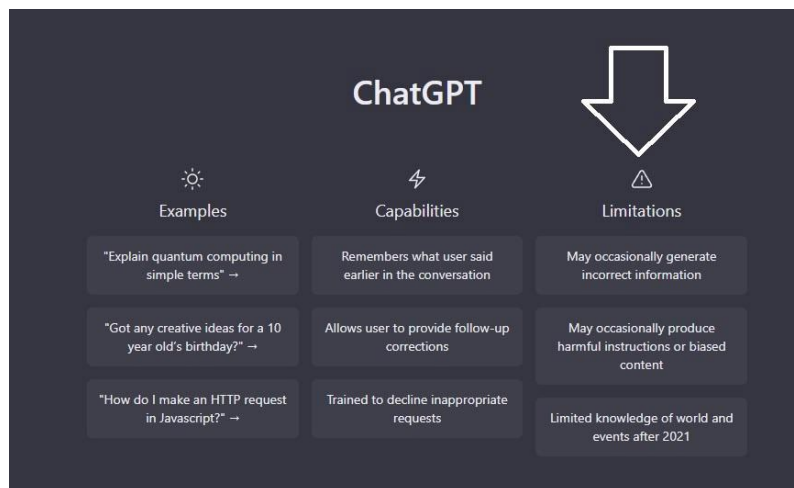
Trata-se, portanto, de uma atualização metodológica, ao nosso modo de ver necessária e útil, considerando a escassez de recursos desse tipo entre pesquisadores da Comunicação e, principalmente, sem obrigá-los a fazer grandes movimentos interdisciplinares, como aprender computação ou a programar, mas sim, uma aplicação diferente de algo que já conhecem e sabem fazer bem, ou seja, entrevistar sujeitos para coletar informações.



2.1 Experimento e resumo de informações coletadas

A EEAF baseia-se na premissa de que modelos de linguagem natural, do tipo GPT são treinados para acertar e ser assertivos nas respostas, o que oferece uma oportunidade para extrair informações ou pelo menos entender de forma mais profunda, a partir das próprias respostas coletadas, suas principais potencialidades e também limitações. Em tese, apesar de ser sempre um fator a ser considerado, pelo que se observou, o modelo GPT não foi treinado para mentir ou dissuadir, o que não significa, como detalharemos a frente, que suas respostas sejam sempre precisas e verdadeiras. Na sua interface de acesso (Fig. 2), inclusive, existe um alerta para a possibilidade de erros e imprecisões nas respostas.

Figura 2 - Tela inicial da interface de ChatGPT



Fonte: OPENAI

Entretanto, a partir do experimento, o contexto dos erros foi identificado principalmente por dois fatores importantes. O primeiro, pelo fato da base de treinamento do algoritmo ser limitada a documentos de 2021, o que o impede simplesmente de saber de coisas que aconteceram depois disso. Por exemplo, solicitado a listar os países campeões do mundo de futebol de 1970 até 2018, ele o fez corretamente. Quando foi perguntado sobre quem ganhou a copa de 2022, ele informou que não poderia responder, inclusive pedindo desculpas.

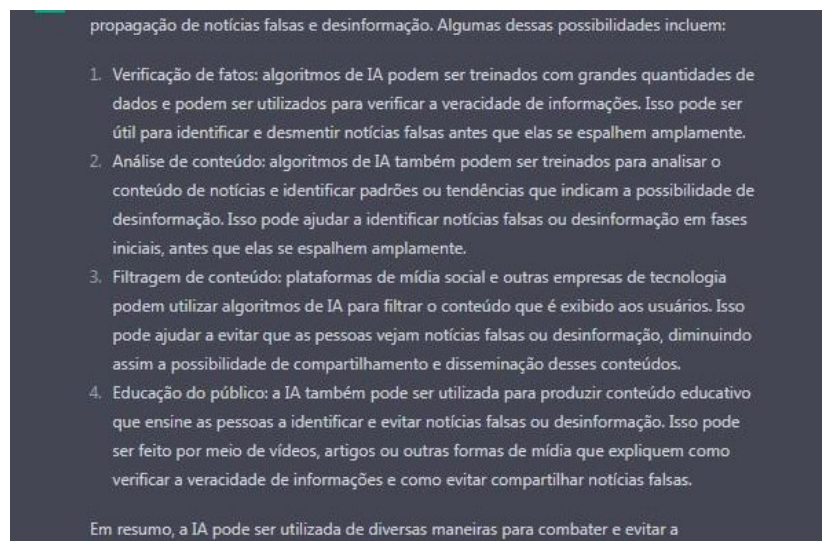
O segundo fator importante para entender o funcionamento do ChatGPT está relacionado ao fato dele não ter acesso à internet e, portanto, não poder fazer consultas ou se atualizar automaticamente. Ele depende sempre de uma atualização da sua base, feita pelos

desenvolvedores, apesar da que utiliza, limitada a itens coletados até 2021, ser bastante extensa e incluir documentos, livros e outros arquivos retirados da internet.

No experimento ficou claro que, principalmente sobre assuntos internos da própria organização onde foi desenvolvido, o algoritmo foi treinado para dar uma resposta educada indicando que não podia falar sobre determinados tópicos ou, mais frequentemente, que tal demanda seria inviável para ele, por não fazer parte da base de textos ou informações a partir da qual foi treinado.

O experimento (Fig. 3) baseou-se na versão 3.5 do modelo GPT original que teve sua última atualização feita em 15 de dezembro de 2022. As entrevistas relatadas neste artigo foram feitas entre os dias 16 e 20 de dezembro, em horários diversos e usando tanto acesso via desktop, quanto via celular.

Figura 3 - Tela da interface da ferramenta ChatGPT da OpenAI, exemplificando funcionalidades, possibilidades de uso e limitações.



Fonte: Do autor

Em vários momentos, a própria plataforma também informou que poderiam ocorrer paralisações nos serviços, devido ao grande volume de solicitações simultâneas de usuários testando o ChatGPT. É fato que a solução tem recebido enorme atenção por diversos setores, muito além dos pesquisadores apenas. Mesmo assim, em todas as vezes que o fluxo da conversa foi interrompido por problemas desse tipo, bastou atualizar a página e reiniciar normalmente.

Em algumas consultas no meio da geração do texto aconteceram também interrupções. Sobre isso, uma das questões foi justamente sobre limites do tamanho de texto que poderia ser

inserido no prompt e das respostas geradas. Sobre esse aspecto, em uma das respostas o ChatGPT informou que, em termos de resposta, escrevia de acordo com a necessidade percebida a partir da pergunta ou demanda.

O fato é que, no geral, as respostas são relativamente curtas, mas podem ser aumentadas, ou simplesmente pedindo por isso, ou através da função de regenerar, ou seja, criar uma outra versão textual daquilo que foi pedido.

A proposta da EEIF tenta endereçar as seguintes categorias:

- Autoidentificação - perguntas sobre o que é, seus propósitos ou objetivos, que tecnologias usa, como opera, em que estágio de desenvolvimento está.
- Usos e limitações - perguntas voltadas a descrever cenários em que as funcionalidades podem ser usadas e limitações inerentes à própria ferramenta sejam identificadas.
- Teste de validação através de solicitações voltadas a confirmar as informações oferecidas na etapa anterior.
- Experimentação adicional - solicitações para avaliar que distorções ou resultados inesperados surgem quando a ferramenta é levada para além das suas limitações originais.

Após a coleta dos dados, através de perguntas feitas diretamente à interface de comunicação do modelo, a avaliação é feita com a quantificação das seguintes métricas ora propostas:

Quadro 1

MÉTRICA	DESCRIÇÃO	PONTOS
Transparência	Quantidade e qualidade sobre as informações que o modelo oferece sobre si mesmo e sua operação em comparação com a documentação pública que a empresa proprietária oferece.	0 A 10
Estruturação	Nível de coesão e coerência identificado a partir dos textos gerados.	0 A 10
Precisão	Nível de acurácia identificado nas informações geradas.	0 A 10
Variedade	Maior ou menor extensão dos tipos de textos ou estruturas narrativas que o modelo consegue executar com eficiência.	0 A 10
Atualização	Capacidade do modelo de se atualizar e ampliar seu escopo de geração de texto.	0 A 10



Fonte: Do autor

3 Funcionalidades

Devido às limitações de tamanho deste artigo seria impossível listar e comentar todas as informações extraídas através da metodologia. Abaixo descreveremos os pontos principais do que foi conseguido.

É importante ressaltar que a metodologia EEAF seria uma espécie de versão do que na Ciência de Dados chamamos de *EDA - exploratory data analysis*, isso é, um movimento inicial de aproximação junto a um conjunto de dados ou objeto que ainda desconhecemos ou sobre o qual sabemos pouco.

Em tese, uma EDA não vai endereçar diretamente as questões ou problemas de uma pesquisa específica. O que ela permite é uma identificação do contexto ou cenário em que se apresentam as características mais gerais do objeto e sobre a qual, análises e operações posteriores poderão ser feitas, mas de forma mais assertiva, a partir dessa primeira avaliação.

3.1 Potenciais cenários de utilização

Uma das questões que destacamos na entrevista foi a de cenários possíveis de utilização. A lista abaixo resume as respostas obtidas. Algumas delas foram efetivamente testadas durante o processo e essa diferença está assinalada com a palavra “Testado” ao final da descrição.

- *Atendimento ao cliente: pode ser usado para responder perguntas e solucionar problemas dos clientes, fornecendo uma opção de atendimento automatizado 24 horas por dia.*
- *Treinamento de funcionários: pode ser usado para fornecer treinamento e orientação aos funcionários, permitindo que eles possam aprender e praticar no seu próprio ritmo.*
- *Conversas virtuais: pode ser usado para criar conversas virtuais com outras pessoas, permitindo que as pessoas se conheçam e socializem sem se encontrar pessoalmente. Testado.*
- *Gerador de conteúdo: pode ser usado para gerar conteúdo de texto, como notícias, artigos ou histórias, de forma automatizada. Testado.*
- *Sistema de tradução automática: pode ser usado para traduzir texto de uma língua para outra de forma automatizada. Testado.*
- *Gerador de relatórios: pode ser usado para gerar relatórios de forma automatizada, a partir de dados e informações fornecidas. Testado.*
- *Gerador de resumos: pode ser usado para gerar resumos de textos ou documentos de forma automatizada. Testado.*
- *Gerador de listas: pode ser usado para gerar listas de itens ou tópicos de forma automatizada. Testado.*
- *Gerador de e-mails: pode ser usado para gerar e-mails de forma automatizada, a partir de informações e contexto fornecidos.*
- *Assistente pessoal: pode ser usado como um assistente pessoal virtual, ajudando as pessoas a gerenciar suas tarefas e compromissos diários.*
- *Gerador de diálogos: pode ser usado para gerar diálogos entre personagens em jogos, filmes ou histórias. Testado.*



- *Sistema de recomendação: pode ser usado como um sistema de recomendação, sugerindo produtos ou serviços baseados nas preferências e interesses do usuário.*
- *Gerador de perguntas: pode ser usado para gerar perguntas para testes ou quizzes de forma automatizada. Testado.*
- *Gerador de títulos: pode ser usado para gerar títulos de forma automatizada, baseado em um contexto ou tema específico. Testado.*
- *Gerador de sinopses: pode ser usado para gerar sinopses de filmes, livros ou outras obras de forma automatizada. Testado.*
- *Gerador de resenhas: pode ser usado para gerar resenhas de produtos ou serviços de forma automatizada. Testado.*
- *Gerador de receitas: pode ser usado para gerar receitas de forma automatizada, baseado em ingredientes e restrições alimentares específicas. Testado.*
- *Gerador de programas de computador: pode ser usado para redigir pequenos códigos em diversas linguagens de programação. Testado.*

Esses são apenas alguns exemplos de possíveis usos para o ChatGPT. Como mencionado anteriormente, o modelo GPT é versátil e pode ser adaptado e personalizado para diferentes aplicações e contextos, o que significa que ele pode ser usado em muitas outras situações além dessas.

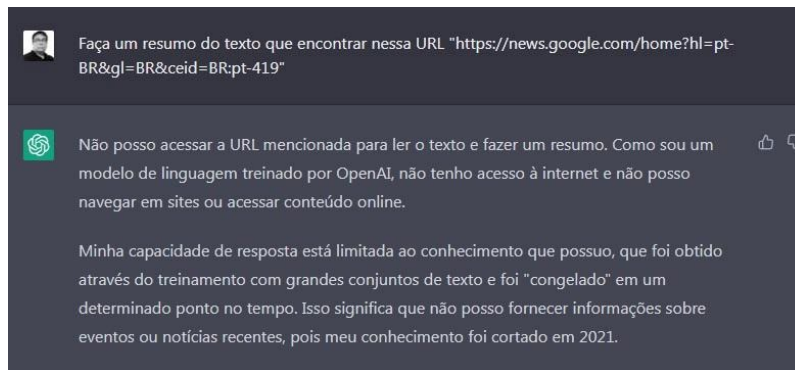
3.2 Formatos de texto

Outro aspecto que consideramos interessante destacar é a possibilidade de a ferramenta simular estilos e referências textuais específicos.

- *Narrativo: um estilo de escrita que conta uma história ou descreve um evento de maneira cronológica. Testado.*
- *Descritivo: um estilo de escrita que descreve e detalha um objeto, lugar, pessoa ou situação. Testado.*
- *Persuasivo: um estilo de escrita que tenta convencer o leitor a tomar uma determinada ação ou a concordar com um ponto de vista específico. Testado.*
- *Expositivo: um estilo de escrita que explica ou esclarece um tópico ou conceito de maneira clara e objetiva. Testado.*
- *Poético: um estilo de escrita que usa ritmo, assonância e rima para criar efeitos literários. Testado.*
- *Diário: um estilo de escrita que relata os pensamentos e acontecimentos de um dia ou período específico.*
- *Jornalístico: um estilo de escrita que apresenta informações de maneira clara e precisa, geralmente com um tom objetivo. Testado.*
- *Acadêmico: um estilo de escrita utilizado em trabalhos científicos ou acadêmicos, que segue uma estrutura padrão e usa linguagem técnica e precisa. Testado.*
- *Ficcional: um estilo de escrita que cria personagens e histórias fictícias. Testado.*

Esses são apenas alguns exemplos dos muitos estilos de escrita diferentes que podem ser utilizados dependendo do contexto e do propósito do texto.

Figura 4 - Tela da interface do diálogo com resposta explicando a limitação de acesso à internet.



Fonte: Do autor

Como demonstrado acima (Fig. 4), apesar do bot conseguir formatar seus textos usando diferentes formatos, ele não consegue aplicar suas funções padrão a solicitações que estejam fora da sua base de dados, como, por exemplo, conteúdo acessível via internet, como o buscador do Google faz corriqueiramente. Essa é uma das suas maiores limitações.

Os testes mostraram, inclusive, que ao ser solicitado, por exemplo, a desenvolver uma notícia com fatos do tempo presente, ele criou números e eventos não existentes, o que indica uma dificuldade do seu uso em algumas funções de redação jornalística.

3.3 Replicação de padrões

Talvez uma das funções mais interessantes do ChatGPT seja a possibilidade de tentar replicar traços narrativos de autores famosos que provavelmente fazem parte da sua base de treinamento.

Solicitado a listar 15 autores que ele poderia tentar utilizar como padrão para a geração do texto, citou *William Shakespeare, Jane Austen, Charles Dickens, Mark Twain, Edgar Allan Poe, Emily Dickinson, Nathaniel Hawthorne, Henry David Thoreau, Herman Melville, Edgar Lee Masters, Emily Brontë, Charlotte Brontë, John Steinbeck, Ernest Hemingway, J.R.R. Tolkien.*

Com a mesma demanda, só que a partir de autores brasileiros, ele respondeu citando os listados abaixo e fazendo considerações sobre seus estilos específicos.

- *. Machado de Assis: Autor considerado o pai da literatura brasileira, conhecido por suas obras satíricas e críticas sociais.*
- *. Graciliano Ramos: Autor que escreveu principalmente sobre a vida no interior do Brasil e abordou temas como a pobreza, a desigualdade social e a opressão.*
- *. Clarice Lispector: Autora que explorou temas profundos e filosóficos em suas obras, como a identidade, o amor e a existência.*

- . João Guimarães Rosa: Autor que utilizou uma linguagem única e poética em suas obras, explorando temas como a identidade, a memória e a relação entre o homem e a natureza.
- . Paulo Coelho: Autor de romances e contos que abordam temas espirituais e filosóficos de uma maneira acessível e inspiradora.

3.4 Informações complementares

Perguntado sobre que informações poderiam gerar textos mais precisos e adequados às solicitações foram listados:

- *Público-alvo*: o público para o qual o texto será destinado, que pode influenciar a linguagem e o nível de complexidade do texto.
- *Propósito*: o objetivo do texto, que pode ser informativo, persuasivo, entreter, etc.
- *Formato*: o formato do texto, que pode ser um artigo, um relatório, uma redação, etc.
- *Comprimento*: o tamanho do texto, que pode ser um curto parágrafo ou um trabalho mais longo.
- *Fonte*: as fontes que devem ser consultadas para escrever o texto, se for o caso.
- *Data de entrega*: o prazo para entrega do texto.
- *Requisitos de formatação*: as normas de formatação que devem ser seguidas, como as regras de estilo APA ou MLA.

3.5 Avaliação das métricas e detalhamento do método de análise

Para a métrica “Transparência” - Nota atribuída: 9.

Uma pontuação de 9 indica que a ferramenta fornece uma alta quantidade e qualidade de informações sobre si mesma e seu funcionamento em comparação com a documentação pública fornecida pela empresa proprietária. Isso sugere que ele é transparente em termos de operação e que os desenvolvedores forneceram informações sobre seu funcionamento interno, o que pode ajudar os usuários a entender como ele gera sua saída. Uma pontuação de 9 também sugere que a ferramenta possui documentação bem escrita e informativa, o que pode ser muito útil para usuários que desejam entender como usá-la de maneira eficaz.

Para a métrica “Estruturação” - Nota atribuída: 9.

Os textos gerados receberam nota 9. Isso indica que os textos apresentam alto nível de coesão e coerência. A coesão refere-se a como as diferentes partes de um texto se encaixam, enquanto a coerência se refere a como o texto como um todo é organizado e compreensível. A pontuação sugere que os textos gerados são bem estruturados e as ideias apresentadas fluem logicamente de uma frase para a outra. Essa métrica também sugere que a ferramenta usada para gerar esses textos foi bem-sucedida em produzir uma saída coerente e coesa.

Para a métrica “Precisão” - Nota atribuída: 7.

É possível supor que houve algum erro ou inconsistência nas informações geradas pelo modelo. Esses erros podem ser devido às limitações dos dados de treinamento ou à complexidade da linguagem usada na entrada. Além disso, é importante observar que a precisão



perfeita nem sempre pode ser alcançada em tarefas de geração de linguagem natural. A pontuação de 7 sugere que o modelo foi bastante preciso, mas pode exigir mais refinamento para melhorar sua precisão na geração de texto.

Para a métrica “Variedade” - Nota atribuída: 9.

Sugere que os textos gerados a partir do modelo são diversos e exibem uma variedade de estruturas narrativas. O modelo é capaz de produzir textos não repetitivos e pode gerar diferentes tipos de conteúdo de forma eficiente. O modelo pode ser treinado em uma ampla variedade de fontes de dados e aprendeu a produzir textos com uma ampla variedade de vocabulário e estilos de linguagem. Isso indica também que o modelo pode ser altamente útil na geração de diversos conteúdos para diferentes aplicações, como marketing, mídia social e criação de conteúdo.

Para a métrica “Atualização” - Nota atribuída: 7.

Sugere que o modelo tem capacidade de se atualizar e expandir seus recursos de geração de texto, mas pode não ser tão eficiente ou eficaz quanto outros modelos com pontuações mais altas nessa métrica. É possível que o modelo tenha algumas limitações em termos de dados de treinamento ou arquitetura, o que pode afetar sua capacidade de atualização e melhoria ao longo do tempo. Além disso, pode exigir mais intervenção manual ou ajuste do usuário para incorporar novos dados e melhorar seu desempenho.

As justificativas acima para as notas, além de indicar como avaliamos o desempenho do modelo no teste, também servem de guia para que outros pesquisadores apliquem a metodologia e entendam como esses parâmetros podem ser considerados na atribuição de valores.

4 Considerações finais

Pelas limitações deste texto, seria impossível discutir ou problematizar as muitas consequências ou impactos desse tipo de solução no mercado de trabalho, na formação futura de profissionais, nos cuidados que professores de agora em diante deverão ter com os textos que recebem para avaliar, entre outros cenários.

Com certeza essa é uma temática emergente que suscitará diversas outras publicações.

Neste texto relatamos o experimento através do qual entrevistamos a ferramenta ChatGPT, no intuito de testarmos nossa proposição de abordagem denominada EEAF.



A iniciativa se justifica pela ausência de métodos ou técnicas consolidadas na área para lidar com objetos do tipo estudado neste caso, ou seja, um modelo de processamento de linguagem natural, treinado a partir de uma extensa base de documentos, capaz de simular razão e discernimento humanos, ao desenvolver respostas textuais para um amplo conjunto de demandas e contextos de utilização, no formato de uma conversa com o usuário.

A abordagem proposta se organiza nas seguintes áreas de avaliação: autoidentificação; usos e limitações; teste de validação e experimentação adicional. Após a coleta dos dados através dos textos das respostas apresentadas, utilizamos as métricas de transparência, estruturação, precisão, variedade e atualização.

Em nossa avaliação, nos quesitos transparência e estruturação o bot recebeu nota 9, apresentando respostas bastante organizadas e com inegável semelhança a uma resposta humana. O conjunto documental da empresa relacionado ao produto também permitiu o entendimento razoável sobre seu funcionamento e modo de operação.

No quesito precisão, a partir dos próprios alertas do sistema sobre eventuais imprecisões e da repetida explicação sobre as limitações cronológicas de sua base de treinamento, foi atribuída nota 7, com a ressalva de que, se o teste tratasse apenas de contexto de geração de textos para uso jornalístico, tal nota seria pior, já que, foi nesse tipo de conteúdo que o ChatGPT demonstrou suas maiores dificuldades, inventando fatos e eventos que não aconteceram.

Nos quesitos variedade e atualização as notas foram 9 e 7 respectivamente, já que o bot apresentou enorme flexibilidade para operar em cenários muito diversos, indo da poesia ao texto publicitário e ainda a partir de uma base relativamente atual que consolidou informações até 2021. A questão da atualização, que teve menor avaliação pela limitação da base de treinamento que foi até 2021, nos parece um item de fácil aprimoramento a ser feito regularmente, o que talvez até já tenha sido feito, dependendo de quando se esteja lendo este artigo.

Considerando a época em que realizamos os testes, ou seja, em dezembro de 2022, o ChatGPT demonstrou ser uma das mais sofisticadas ferramentas da atualidade para executar as tarefas a que se propõe.

Considerando, também, que o bot encontra-se ainda, neste momento, em fase experimental, é provável que em período de tempo curto mais funcionalidades e precisão sejam incorporadas à sua performance.

Entendemos que nossa proposição metodológica se apresenta como abordagem possível, diante de um cenário de rápida evolução tecnológica, no qual as soluções de inteligência

artificial generativa, ou seja, capazes de simular a própria criatividade humana deverão crescer e prosperar.

O experimento demonstrou que apesar de corretos e bem estruturados, a necessidade de uma revisão humana ao final do processo continua sendo essencial, principalmente em materiais que lidem com fatos e dados objetivos, especialmente do tipo factual e jornalístico, em que as principais deficiências e erros nas respostas foram identificadas.

Ficou claro que o ChatGPT funciona bem como um assistente para uma série de atividades, inclusive a acadêmica, principalmente pela capacidade de fazer análises básicas de processamento natural, tais como sumarização, tradução, identificação de tópicos e agrupamentos temáticos. Tais funcionalidades, que serão detalhadas em artigo, com esse recorte específico, são objeto de outro experimento⁷ dentro do projeto atual, que já implementou tais funções num aplicativo pensado para pesquisadores da Comunicação.

A riqueza do aprendizado adquirido no processo demonstrou apenas uma pequena fração do que está por vir e do impacto que a tecnologia da inteligência artificial causará inexoravelmente em vários aspectos da atividade humana.

A conversa com robôs deverá seguir aperfeiçoando-se e tornar-se-á cada vez mais instrutiva e útil, devendo também servir com um constante alerta para as transformações que estão por vir.

Se a primeira onda de impactos da IA, nos mercados de trabalho e na empregabilidade, foi registrada em funções repetitivas e de fácil automatização, nesta segunda onda, caracterizada por modelos mais sofisticados, nos parece que os potenciais atingidos serão as indústrias criativas, da comunicação, do marketing, das artes e, por que não, também da notícia.

Continuaremos conversando com robôs para tentar entender o que está acontecendo. Para fins de esclarecimento adicional, com exceção do que foi marcado em *itálico*, todo o texto deste artigo foi escrito por um humano.

Referências

AMADEU, Sérgio. **Democracia e os códigos invisíveis**: Como os algoritmos estão modulando comportamentos e escolhas políticas. São Paulo: Edições Sesc (Coleção Democracia Digital), 2019.

⁷ Uma ideia inicial desta iniciativa está descrita na parte final deste vídeo, listado no canal de difusão científica do LABCOM/UFMA - <https://youtu.be/wEt30wJ-pO4>.



AP - The Associated Press. **Leveraging AI to advance the power of facts: Artificial Intelligence at the Associate Press**, 2022. Disponível em: <https://www.ap.org/discover/artificial-intelligence> . Acesso em: 10 jan. 2023.

DOS SANTOS, Márcio Carneiro. Internet das Coisas e sistemas inteligentes no jornalismo: o conceito de presença diluído entre as narrativas da complexidade urbana. **Comunicação & Inovação**, v. 17, n. 34, p. 21-39, 2016.

DOS SANTOS, Márcio Carneiro. Inteligência híbrida e análise de sentimentos: integrando curadoria humana e coleta de dados automatizada para avaliar a comunicação de governo. **Conexão-Comunicação e Cultura**, v. 17, n. 33, 2018.

DOS SANTOS, Márcio Carneiro. A datificação de um campo de conhecimento: como algoritmos, números e abordagens quantitativas estão mudando a comunicação. **Organicom**, v. 16, n. 31, p. 145-157, 2019.

GUZMAN, Andrea L; SETH C. Lewis. 2019. Artificial Intelligence and Communication: A Human–Machine Communication Research Agenda. **New Media & Society**. DOI doi:10.1177/1461444819858691.

GUZMAN, Andrea L; SETH C. Lewis. What is human-machine communication, anyway. Human-machine communication: **Rethinking communication, technology, and ourselves**, p. 1-28, 2018.

JONES, Steven. **Against technology: From the Luddites to neo-Luddism**. Routledge, 2013.

LATZER, Michael *et al.* **The economics of algorithmic selection on the Internet**. In: Handbook on the Economics of the Internet. Edward Elgar Publishing, p. 395-425, 2016.

LEMOS, André. Epistemologia da comunicação, neomaterialismo e cultura digital. **Galáxia** (São Paulo), p. 54-66, 2020.

MANOVICH, Lev. **El software toma el mando**. Editorial UOC, 2014.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. **Big data: A revolution that will transform how we live, work, and think**. Houghton Mifflin Harcourt, 2013.

Mollman Steve. ChatGPT has gained 1 million followers in a single week. Here’s why the A.I. chatbot is primed to disrupt search as we know it. **Yahoo Finance**, 2022. Disponível em: <https://finance.yahoo.com/news/chatgpt-gained-1-million-followers-224523258.html> . . Acesso em: 09 dez. 2022.

MUSIANI, Francesca. Governance by algorithms. **Internet Policy Review**, v. 2, n. 3, p. 1-8, 2013.

NAPOLI, Philip. **The algorithm as institution: Toward a theoretical framework for automated media production and consumption**. Fordham University Schools of Business Research Paper, 2013.



REITER, Ehud; SRIPADA, Somayajulu; ROBERTSON, Roma. Acquiring correct knowledge for natural language generation. **Journal of Artificial Intelligence Research**, v. 18, p. 491-516, 2003.

SILVA, Tarcízio. Racismo Algorítmico em Plataformas Digitais: microagressões e discriminação em código. *In*: SIMPÓSIO INTERNACIONAL LAVITS, 6., 2019. **Anais Eletrônicos** [...]. Disponível em: <https://lavits.org/wp-content/uploads/2019/12/Silva-2019-LAVITSS.pdf>. Acesso em: 06 fev. 2023.

SQUIRRA, Sebastião Carlos. A tecnologia e a evolução podem levar a comunicação para a esfera das mentes. **Revista Famecos**, v. 23, n. 1, 2016.

SQUIRRA, Sebastião Carlos. A informação essencial à vida, às máquinas e à comunicação. **Lumina**, v. 10, n. 2, 2016a.

LIMA JR, Walter Teixeira; VERGILI, Rafael. Digital Inclusion and Computational Thinking: New Challenges and Opportunities for Media Professionals. **Information and Technology Literacy: Concepts, Methodologies, Tools, and Applications**, p. 759-773, 2018.

VEEL, Kristin. Make data sing: The automation of storytelling. **Big Data & Society**, v. 5, n. 1, 205395171875668, 2018. DOI doi:10.1177/2053951718756686.

ⁱ Professor da área de Jornalismo em Redes Digitais do Departamento de Comunicação Social da UFMA. Doutor em Tecnologias da Inteligência e Design Digital pela PUC-SP. Coordenador do LABCOM/UFMA.

