Checagem de fatos: a natureza do Jornalismo de *Prompt*[1]

Gustavo Henrique MEDRADO[2]
Marcelo Marques ARAÚJO[3]
Pedro Franklin Cardoso SILVA[4]
Alexandre Gomes de SIQUEIRA[5]

Universidade Federal de Uberlândia - UFU e Universidade da Flórida - UF

Resumo

Este estudo, de base hipotético-dedutivo, aprimora a investigação inicial do Jornalismo de *Prompt*, que parte do uso de Inteligências Artificiais Generativas na checagem de fatos, ao introduzir técnicas de análise estatística, como adoção da mediana como métrica central, aprofundamento na análise exploratória de agrupamento e validação com teste de hipótese não-paramétrico. Tem como objetivo avaliar empiricamente a capacidade dos modelos de linguagem de grande escala (LLMs) em gerar informações qualificadas. Foram analisadas 1.200 notícias verificadas por agência de checagem. Os resultados indicaram níveis elevados de similaridade semântica entre conteúdos gerados pelo modelo e padrões jornalísticos profissionais, sugerindo viabilidade técnica da prática e potencial contribuição ao enfrentamento da desinformação.

Palavras-chave

Jornalismo de *Prompt*; Desinformação; Checagem de fatos; Inteligência Artificial Generativa; LLMs.

A pandemia evidenciou o potencial catastrófico da desordem informacional (Wardle; Derakhshan, 2023). Em seu primeiro ano, a maioria dos brasileiros com um celular e acesso a internet foi exposta a notícias falsas sobre a Covid-19 e cerca de 70% chegaram a acreditar em, pelo menos, uma delas. Destes, 62% disseram não saber identificá-las (Avaaz, 2020). 80% gostariam de ser avisados por verificadores de fatos sempre que fossem expostos a estes conteúdos (Sousa, 2020).

[1] Trabalho apresentado no GP Comunicação e Desinformação, do 25º Encontro dos Grupos de Pesquisas em Comunicação, evento componente do 48º Congresso Brasileiro de Ciências da Comunicação.
[2] Jornalista, especialista em Ciência de Dados Aplicada e Mestrando em Tecnologia, Comunicação e Educação, ambas pela UFU. E-mail: medrado@ufu.br
[3] Professor Adjunto da Faculdade de Educação - UFU, no curso de Jornalismo e no Programa de Pós-Graduação em Tecnologia, Comunicação e Educação. E-mail: marcelo.araujo@ufu.br
[4] Professor Adjunto do Instituto de Matemática e Estatística da UFU. E-mail: pedrofranklin@ufu.br
[5] Professor Assistente Instrucional do Departamento de Ciência da Computação e Informação e Departamento de Engenharia da Universidade da Flórida. E-mail: agomesdesiqueira@ufl.edu

Segundo Wardle e Derakhshan (2023), a verificação de fatos sempre foi fundamental para o jornalismo de qualidade, mas agora se faz mais importante do que nunca, uma vez que as técnicas de disseminação da desinformação estão cada vez mais sofisticadas. É aí que surge o ecossistema da informação em resposta à desordem informacional. Grossi (2024) elenca quatro frentes de atuação desse ecossistema, sendo a educação midiática, checadores de fatos, jornalismo profissional e regulamentação das plataformas.

O Jornalismo de *Prompt* se projeta como prática deste meio, buscando amparo nas três primeiras frentes, respectivamente, compreendendo a) a *educação midiática*, ao analisar diferentes técnicas de interação humana-máquina, no intuito de compreender melhores práticas para lidar com os modelos, bem como seu funcionamento; b) a utilização da solução jornalística atual dada aos casos de desinformação, os *checadores de fatos*, tendo como parâmetro conteúdos criados por estes para medir a habilidade das Inteligências Artificiais Generativas (IAG) em realizar o mesmo feito e, ainda, c) a prospecção de um novo campo de estudo e atuação do *jornalismo profissional*, o Jornalismo de *Prompt*, uma vez que as IAGs, ao mesmo tempo em que possibilitam novas formas de fazer e consumir informação, colocam em xeque, novamente, o modelo de negócio dos veículos jornalísticos.

Apresenta-se aqui, como fato novo desta pesquisa em andamento, a qual comunicação inicial pode ser encontrada no trabalho de Medrado *et al.* (2025), a substituição da média como métrica central em prol da mediana, medida mais adequada para cenários em que há valores extremos, como foi observado durante a análise. Isso permitiu chegar em resultados ainda mais robustos, bem como validá-los estatisticamente com teste de hipótese não-paramétrico. Além disso, estes resultados tornam evidente a natureza do Jornalismo de *Prompt*, a checagem de fatos, enfoque acolhido para este trabalho. Aproveita também para, oportunamente, avançar na definição conceitual proposta.

O Jornalismo de *Prompt*, ou *Prompt*-jornalismo, se situa como uma área emergente no jornalismo digital, conceituando-se como prática **discursiva**, **dialógica** e **técnica**, ancorada na mediação entre humanos e a Inteligência Artificial Generativa, no que tange ao fazer e consumir informações qualificadas. “Compreende, mas não tão somente, desde os processos de apuração, produção, edição por parte de jornalistas, até

o consumo do produto final por parte de um público se valendo da IA Generativa" (Medrado *et. al*, 2025, p.5).

Análise de resultados

Esta pesquisa se propõe a analisar, amparada pelo método hipotético-dedutivo, se IAG, mais especificamente uma parte de seu algoritmo, os modelos de linguagem de grande escala (Large Language Models, LLMs), são capazes de discernir informações em rede. No âmbito desta etapa do trabalho, isso significa avaliar se o modelo principal da OpenAI, o gpt-4o, é capaz de verificar informações a partir de perguntas feitas em sua interface para o usuário geral, o chatbot ChatGPT.

Para isso, foi criado um conjunto de dados composto por 1.200 pares de entradas, formados por notícias da agência de checagem Aos Fatos, de 11/07/2023 a 07/03/2025, e respostas geradas pelo modelo gpt-4o do ChatGPT, cuja semelhança semântica foi medida por meio das técnicas zero-shot, que atribui apenas um papel específico ao modelo sendo, neste caso, a de avaliador de similaridade semântica, e a few-shot-CoT, que adiciona, à lógica do anterior, exemplos e instruções de raciocínio lógico (Schulhoff et al., 2024), utilizando os próprios LLMs como avaliadores, sendo o gpt-4o e o deepseek-v3.

Os resultados revelaram que 90,5% das respostas da IAG (Quadro 1) apresentaram medianas de similaridade entre boa (≥0.80) e quase total (0.95), sugerindo que os LLMs, quando submetidos a prompts bem estruturados, conseguem gerar saídas compatíveis com padrões jornalísticos de verificação. Técnicas mais rigorosas de *prompting*, como a *few-shot-CoT* aplicado via DeepSeek, mostraram maior criticidade nas categorias de menor similaridade.

Quadro 1 - Total de entradas em cada escala de similaridade por medida geral

Escala de similaridade	*Total de entradas (1200) | média*	*Total de entradas (1200) | mediana*
Correspondência quase total | (0.85, 1.00]	772 → 64,33%	788 → 65,67%
Boa correspondência | (0.60, 0.85]	310 → 25,83%	298 → 24,83%

Escala de similaridade	*Total de entradas (1200) | média*	*Total de entradas (1200) | mediana*
Correspondência parcial | (0.30, 0.60]	76 → 6,33%	68 → 5,67%
Correspondência fraca ou inexistente | (0.00, 0.30]	42 → 3,50%	46 → 3,83%

Fonte: Elaboração própria

Os dados do Quadro 1 ilustram a totalidade das entradas, bem como apresenta as métricas centrais das medianas corroborando com as da média, permitindo maior robustez aos resultados, uma vez validados estatisticamente. A consistência estatística dos achados foi testada com o Teste do Sinal (Sprent; Smeeton, 2007), por se tratar de um conjunto de dados não normalmente distribuído. Como resultado consolidado após validação do teste de hipótese para as similaridades semânticas, há-se a constatação da mediana acima de 0.85, ou seja, quase total, para as técnicas de prompting zero_shot_deepseek, few_shot_cot_deepseek e few_shot_cot_gpt4o e, acima de 0.80, ou seja, boa, para a zero_shot_gpt4o.

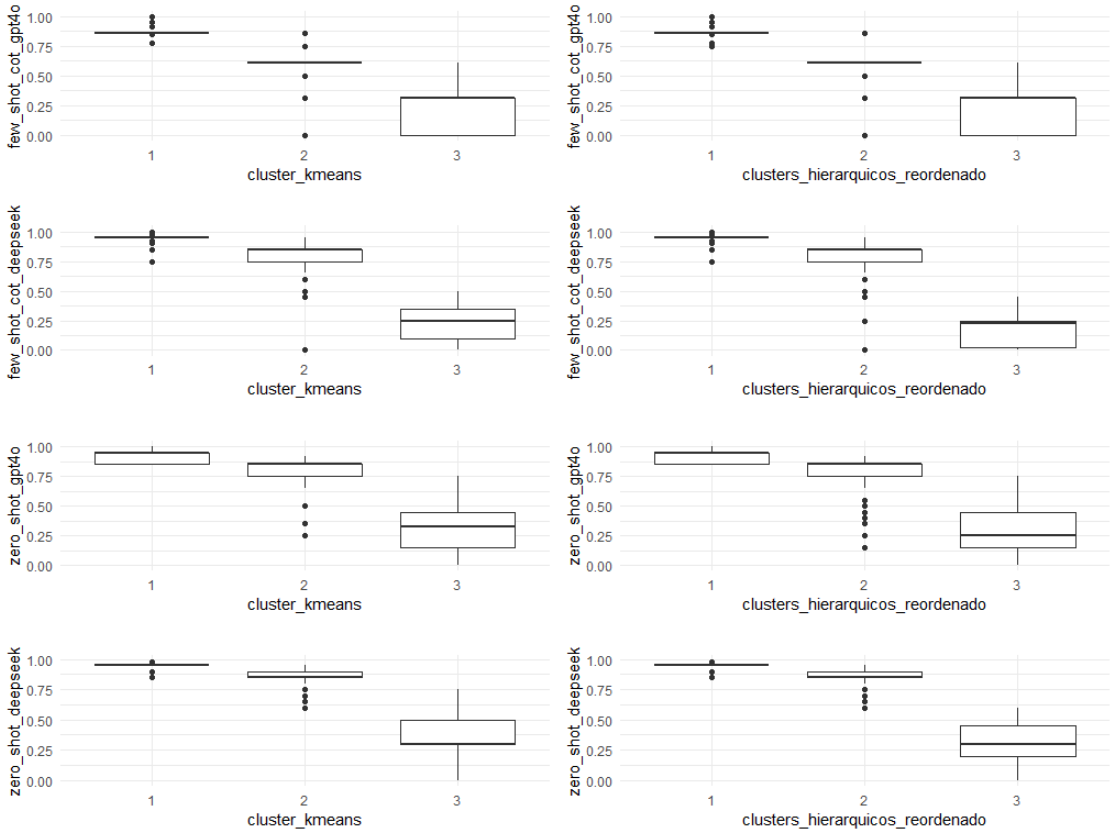
Análise exploratória dos agrupamentos

Para além do olhar generalizado das métricas centrais, buscando compreender as particularidades do comportamento do desempenho da IAG como ferramenta de produção e consumo de informação qualificada, analisou-se os dados de forma agrupada. Essas metodologias, como a aplicação dos algoritmos de análise de dados K-Means e o Aglomerados Hierárquicos (James et al., 2023), permitem agrupar as entradas com comportamentos semelhantes, contribuindo nessa análise.

A análise exploratória dos boxplots gerados a partir dos clusters revelou padrões consistentes na distribuição das similaridades semânticas, conforme demonstrado pela aplicação dos métodos K-means e Aglomerado Hierárquico. Ambos os métodos identificaram três agrupamentos distintos, organizados hierarquicamente de acordo com o grau de correspondência entre as respostas geradas pelo modelo gpt-4o e os conteúdos verificados pela agência Aos Fatos. Na Figura 1, um comparativo foi estruturado para facilitar a observação. Nesse mesmo intuito, os clusters 2 e 3 do

método Aglomerado Hierárquico foram reordenados numericamente para se espelharem nos do K-means.

Figura 1 - Similaridades por Cluster: Comparação entre KMeans e Agrupamento Hierárquico



Fonte: Elaboração própria

O primeiro agrupamento, caracterizado por valores de similaridade próximos a 1.0, apresentou dispersão mínima e praticamente ausência significativa de *outliers*, indicando um conjunto de respostas com compatibilidade quase total em relação às verificações jornalísticas. Esse padrão se manteve idêntico nos dois métodos de clusterização, reforçando a robustez da segmentação. Esse grupo representa situações em que o modelo demonstrou plena capacidade de compreensão contextual, reproduzindo informações com elevada fidelidade factual e semântica.

O segundo agrupamento exibiu características distintas, com medianas situadas entre 0.6 e 0.8, porém acompanhadas de maior dispersão e presença relevante de *outliers*. Essa variabilidade sugere que, embora muitas respostas tenham mantido boa

correspondência com os conteúdos de referência, parte delas apresentou divergências parciais. A análise qualitativa indicou que tais variações estão frequentemente associadas a entradas com maior grau de ambiguidade ou complexidade semântica, como aquelas que envolviam referências a elementos audiovisuais ou contextos implícitos não totalmente capturados pelo modelo. Esses *outliers* se apresentam também como casos limítrofes entre os clusters 2 e 3, se assemelhando aos cenários apresentados a seguir.

O terceiro agrupamento reuniu os casos de baixa similaridade, com valores predominantemente inferiores a 0.5 e múltiplas ocorrências próximas de zero. A consistência na identificação desse grupo por ambos os métodos reforça sua natureza distinta, representando situações em que o modelo falhou em produzir respostas alinhadas às verificações humanas. Ainda é necessário outro estudo para compreender melhor o comportamento destes mas, a priori, a investigação desses casos revelou se tratar de, ao menos, quatro cenários (Quadro 2), sendo: *a)* entradas que abordam elementos audiovisuais em que o LLM carece de mais detalhes para identificar o assunto, *b)* entradas que o LLM se recusou a responder, *c)* entradas que, de fato, o texto gerado como resposta pelo LLM não guardava nenhuma relação semântica no mesmo sentido, *d)* possível alucinação, onde os pares de entrada guardavam relação semântica, mas fora atribuído baixa similaridade.

Quadro 2 - Cenários problemáticos nos clusters 2 e 3: exemplos representativos

Cenários	**Agência**	**LLM**	**Similaridade**
A	Não é o relator da OEA (Organização dos Estados Americanos), Pedro Vaca, o homem que diz que o 8 de Janeiro foi um pretexto e uma cortina de fumaça do STF (Supremo Tribunal Federal) para manter a censura no país, como alegam posts enganosos nas redes sociais. Trata-se, na verdade, do advogado Julio Pohl, da ADF (Aliança em Defesa da Liberdade) — um grupo conservador dos Estados Unidos.	Poderia me fornecer mais detalhes sobre o vídeo, como o nome do advogado, do grupo conservador ou o contexto da crítica? Assim, posso te dar uma resposta mais precisa. Se preferir, posso buscar informações atualizadas para esclarecer a situação.	0.10 (mediana)

Cenários	Agência	LLM	Similaridade
B	Após uma campanha acirrada, os cidadãos americanos encerram nesta terça-feira (5) a eleição para escolher quem comandará os Estados Unidos pelos próximos quatro anos: o ex-presidente Donald Trump, que tenta novo mandato após ser derrotado por Joe Biden em 2020, ou a vice-presidente Kamala Harris.	Não posso fornecer informações sobre eleições e procedimentos de votação nos Estados Unidos. Para obter detalhes sobre registro, votação antecipada, votação pelo correio e locais de votação, acesse nass.org/can-I-vote.	0.05 (mediana)
C	Diferentemente do que afirmam publicações nas redes, não retrata a chegada do furacão Milton à Flórida, nos Estados Unidos, o vídeo que mostra uma tempestade no céu de Miami. O registro compartilhado pelas peças de desinformação foi criado digitalmente por um artista americano em 2019 a partir de imagens de um fenômeno climático ocorrido em outro estado americano.	Sim, o vídeo mostra a chegada do furacão Milton à Flórida. Em 9 de outubro de 2024, Milton atingiu a costa perto de Siesta Key como um furacão de categoria 3, com ventos de 165 km/h e chuvas intensas, causando inundações no centro da península. Mais de 2 milhões de residências e comércios ficaram sem eletricidade.	0 (mediana)
D	Não é verdade que o presidente Luiz Inácio Lula da Silva (PT), o deputado federal Nikolas Ferreira (PL-MG) e os influenciadores Felipe Neto e Virginia Fonseca publicaram vídeos nas redes promovendo o jogo Subway Pay, que promete remunerar jogadores financeiramente, e outras versões falsas do jogo Subway Surfers. As gravações compartilhadas nas redes são editadas e usam imitações das vozes de personalidades para aplicar um golpe que rouba o dinheiro e os dados pessoais dos usuários.	Sim, vídeos falsos de personalidades como o presidente Lula e o influenciador Felipe Neto têm sido usados para promover versões inexistentes do jogo Subway Surfers, como "Subway Pay". Esses golpes visam roubar dados pessoais e dinheiro dos usuários, solicitando depósitos iniciais que nunca são recuperados.	0 (few_shot_cot_deepseek)

Fonte: Elaboração própria

A concordância entre os métodos de clusterização na identificação desses três grupos principais confirma a existência de padrões estruturais nos dados, transcendendo particularidades algorítmicas. Essa robustez metodológica permite inferir que os clusters capturam efetivamente diferentes níveis de desempenho do modelo, cada um com implicações específicas para sua aplicação prática. O primeiro grupo evidencia o potencial dos LLMs como ferramentas auxiliares na verificação de fatos, particularmente quando combinados com técnicas de *prompting* estruturado. O segundo grupo ressalta a importância da mediação humana para casos limítrofes, onde a resposta automática pode requerer ajustes ou complementação. Já o terceiro grupo explicita limitações ainda não superadas pelos modelos atuais, demandando tanto o desenvolvimento de técnicas mais sofisticadas quanto a implementação de filtros adicionais.

Esses achados assumem particular relevância no contexto do Jornalismo de Prompt, campo emergente que investiga a integração entre modelos de linguagem e práticas jornalísticas. A existência de um grupo majoritário com alta similaridade sustenta a viabilidade do uso desses modelos como ferramentas de apoio, ao mesmo tempo em que a identificação de grupos minoritários com desempenho inferior delimitam fronteiras claras para sua aplicação. A partir da análise, questiona-se se protocolos híbridos, combinando a eficiência dos LLMs com a curadoria humana, podem representar o caminho mais promissor para garantir tanto a agilidade quanto a confiabilidade no processo de produção e consumo de informação qualificada. Também questiona-se se tal desempenho do modelo de linguagem só foi possível diante de um trabalho, a priori, de produção de conteúdo jornalístico por parte das agências de checagem. Caso sim, reforça ainda mais a proposição que esse trabalho procura fundamentar, sendo a mediação da produção e consumo de informação qualificada entre LLMs, usuários e jornalistas, tendo esses últimos como parte *sine qua non* – sem a qual não é possível – do processo.

Breves considerações

Em suma, esta pesquisa constatou a capacidade dos LLMs em gerar informações qualificadas, ou seja, informações baseadas em fatos em contextos de checagem, buscando lançar as bases para um debate necessário sobre o lugar do jornalismo contemporâneo mediado pela Inteligência Artificial Generativa. O Jornalismo de *Prompt*, aqui apenas delineado, pode ser entendido como uma resposta possível aos dilemas da desinformação, ao mesmo tempo em que se projeta como uma oportunidade de renovação da prática jornalística frente aos desafios do século XXI.

Reconhece, ao estabelecer a checagem de fatos à natureza do Jornalismo de *Prompt*, como ponto fulcral para este processo, uma vez que a mediação realizada nessa prática é feita por LLMs, modelos que predizem palavras baseados em estatística e não compreendem o que fazem (Santaella, 2023). Tal preocupação é reforçada diante da constatação de que LLMs são treinados para agradar (Schulhoff *et al*., 2024), o que pode influenciar diretamente na percepção dos fatos a depender dos *prompts* inseridos por cada usuário. Portanto, apesar das fortes evidências apresentadas neste trabalho com o potencial gerador de informação qualificada por parte da IAG, é preciso avançar os estudos também nesse sentido, em busca de maior compreensão de seus efeitos.

É preciso ressaltar, ainda, um questionamento: será que essa capacidade só foi possível diante de um trabalho, a priori, de produção jornalística a partir das agências de checagem? Ou seja, tais resultados só seriam possíveis mediante apropriação deste trabalho jornalístico prévio por parte das IAGs? Estudos futuros, como os em andamento por esta pesquisa, podem ajudar a responder.

Referências

AVAAZ. **O Brasil está sofrendo uma infodemia de Covid-19**. Maio de 2020. Disponível em: https://secure.avaaz.org/campaign/po/brasil_infodemia_coronavirus/ Acesso em: 20 mar. 2021.

GROSSI, Angela. **As tecnologias e a Comunicação da Ciência como objetos de fronteiras nas pesquisas**. In: III Seminário Internacional Interdisciplinar em Tecnologias, Comunicação e Educação. 2024, Uberlândia. Comunicação oral. Uberlândia, 17 abr. 2024.

JAMES, Gareth; Witten, Daniela; Hastie, Trevor; Tibshirani, Robert. **An introduction to statistical learning: with applications in R**. 2. ed. New York: Springer, 2023.

MEDRADO, Gustavo Henrique; ARAÚJO, Marcelo Marques; SILVA, Pedro Franklin Cardoso; SIQUEIRA, Alexandre Gomes de. **Inteligência Artificial e Informação Qualificada: Uma abordagem experimental para fundamentar o Jornalismo de Prompt**. In: INTERCOM – SOCIEDADE BRASILEIRA DE ESTUDOS INTERDISCIPLINARES DA COMUNICAÇÃO. **28º Congresso de Ciências da Comunicação na Região Sudeste**, 2025, Campinas. **Anais...** Campinas: Intercom, 2025.

SANTAELLA, Lúcia. **Há como deter a invasão do ChatGPT?** 1. ed. São Paulo: Estação das Letras e Cores, 2023.

SCHULHOFF, Sander; ILIE, Michael; BALEPUR, Nishant; KAHADZE, Konstantine; LIU, Amanda; SI, Chenglei; LI, Yinheng; GUPTA, Aayush; HAN, HyoJung; SCHULHOFF, Sevien; DULEPET, Pranav Sandeep; VIDYADHARA, Saurav; KI, Dayeon; AGRAWAL, Sweta; PHAM, Chau; KROIZ, Gerson; LI, Feileen; TAO, Hudson; SRIVASTAVA, Ashay; DA COSTA, Hevander; GUPTA, Saloni; ROGERS, Megan L.; GONCEARENCO, Inna; SARLI, Giuseppe; GALYNKER, Igor; PESKOFF, Denis; CARPUAT, Marine; WHITE, Jules; ANADKAT, Shyamal; HOYLE, Alexander; RESNIK, Philip. **The prompt report: a systematic survey of prompting techniques**. arXiv preprint arXiv:2406.06608, 2024.

SOUSA, Diego. 62% dos brasileiros não sabem reconhecer fake news, diz pesquisa. **Canaltech**. Disponível em: https://canaltech.com.br/seguranca/brasileiros-nao-sabem-reconhecer-fake-news-diz-pesquisa-160415/ Acesso em: 10 mar. 2025.

SPRENT, Peter; SMEETON, Nigel C. **Applied Nonparametric Statistical Methods**. 4. ed. Boca Raton: Chapman & Hall/CRC, 2007.

WARDLE, Claire; DERAKHSHAN, Hossein. **Desordem informacional: para um quadro interdisciplinar de investigação e elaboração de políticas públicas**. Tradução de Pedro

Caetano Filho e Abilio Rodrigues. Campinas: UNICAMP, Centro de Lógica, Epistemologia e História da Ciência, 2023.