

A inteligência artificial no combate à desinformação: limites e potencialidades do projeto Fátima do Aos Fatos¹

Gabriel Bhering ²

Raí de Castro ³

Thuany Menezes ⁴

Jean Almeida ⁵

Introdução

Atualmente, a desinformação não falseia apenas a forma da notícia, mas é capaz de manipular, por exemplo, voz e rosto a partir da Inteligência Artificial (IA). Ou seja, o trabalho do jornalista, que foi dobrado para conseguir checar as “fake news”, agora precisa ser ainda mais intensificado diante dos conteúdos falsos advindos da IA. Diferente do falseamento de notícias, as desinformações produzidas com IA geram muito mais dúvida pela qualidade que apresentam, como a foto do Papa Francisco vestindo um casaco branco de Filipo Sorcinelli, que chegou a virar uma matéria na Vogue republicada no site da Globo com o seguinte título: “Papa Francisco viraliza ao usar jaqueta branca, midi e cheia de detalhes: ‘Muito icônico’”. Em seguida, o grupo Globo desmentiu a desinformação, mas até esse momento também foi vítima das armadilhas da IA. Apesar dos perigos dessa nova tecnologia para o exercício do jornalismo profissional, já existem casos de checagens sendo

¹ Trabalho apresentado no GT 2 - Estratégias de comunicação em ambientes digitais do IV Encontro Virtual da ABCiber – Associação Brasileira de Pesquisadores em Ciberultura. Perspectivas Interdisciplinares e Reconfigurações na Ciberultura: Dados, Algoritmos e Inteligência Artificial. Realização da UNIFAE, nos dias 20 e 21 de junho de 2024.

² Bacharel em Jornalismo, mestrando em Comunicação pelo PPGCOM-UFJF, bhering.gabriel@estudante.ufjf.br

³ Bacharel em Jornalismo, mestrando em Comunicação pelo PPGCOM-UFJF, raidecastro9@gmail.com.

⁴ Estudante do Curso de Jornalismo da Universidade Federal de Juiz de Fora, integrante do NJA - Núcleo de Jornalismo e Audiovisual

⁵ Estudante do Curso de Jornalismo da Universidade Federal de Juiz de Fora, Bolsista PIBIC (CNPq), integrante do NJA - Núcleo de Jornalismo e Audiovisual, almeidapaulo.jean@estudante.ufjf.br

realizadas com apoio da IA, como o projeto Fátima da agência de checagem “Aos Fatos”, que será refletido neste artigo a partir da revisão bibliográfica e a pesquisa documental feita, a fim de entender quais são as potencialidades e os limites dessa estratégia de checagem feita com o apoio da IA.

Fundamentação Teórica

Em primeiro lugar, é importante esclarecer que o fenômeno da desinformação está inserido, na perspectiva de (Derakhshan & Wardle, 2017), na “Desordem Informacional”. Este conceito abrange, então, a desinformação que está ligada com os conteúdos produzidos com a intenção de desinformar, como acontece com frequência no período eleitoral; a informação incorreta que diz respeito aos casos em que a transmissão mesmo sendo errada, não foi arquitetada com a intenção de prejudicar terceiros; e também a informação maliciosa que, mesmo sendo verdadeira, não agrega em nada no debate público, como dados bancários de terceiros.

Com o progresso das IAs surgiram novas ferramentas que podem ajudar tanto a combater a desinformação quanto complexificar o fenômeno. Uma das inovações mais preocupantes nesse sentido são os deep fakes, conteúdos de mídia gerados por IA que podem recriar áudio e vídeo, gerando representações falsas e convincentes de eventos e declarações. De acordo com Santaella e Salgado (2021), os deep fakes têm a capacidade de distorcer, intencionalmente ou não, a percepção pública sobre indivíduos, associando seus nomes e identidades a ideias ou atos que não correspondem à realidade. Essa tecnologia pode ser usada para difamar, enganar ou manipular audiências, criando falsidades difíceis de detectar sem ferramentas avançadas de verificação.

Nota-se como um exemplo positivo do uso da IA para combater a desinformação, o projeto Fátima desenvolvido pelo portal Aos Fatos. Esse projeto utiliza algoritmos avançados de análise de dados e processamento de linguagem natural para monitorar,

identificar e desmentir informações falsas disseminadas em diversas plataformas digitais. O projeto Fátima destaca-se em um contexto onde a verificação de fatos é essencial para preservar a informação. A IA permite analisar rapidamente diferentes tipos de dados e responder com eficiência, algo que seria inalcançável manualmente.

A Fátima tem um grande potencial para combater a desinformação. Ela pode processar rapidamente diversos tipos de dados, identificar padrões suspeitos e oferecer verificações em tempo real, no entanto, ocorre às vezes a incapacidade da IA de responder alguma notícia suspeita por falta de recursos no banco de dados, quando é o caso, a própria inteligência artificial encaminha para um grupo de jornalistas do portal para que os mesmos verifiquem a informação. Essas ferramentas são essenciais para responder rapidamente a surtos de desinformação e diminuir seus impactos antes que se espalhem amplamente.

Reflexões sobre Fátima

O robô Fátima — cujo nome é adaptado de “factma”, abreviação para fact machine — é um projeto lançado em 2019 pelo site jornalístico independente de verificação Aos Fatos. Inicialmente disponível apenas no site do veículo e no Twitter, o projeto abrange, hoje, outras plataformas digitais, como o WhatsApp, o Facebook e o Telegram. A ferramenta, que está na sua quarta fase — em uma versão beta, utilizada para testes —, tem como função, segundo descrição do próprio Aos Fatos, ensinar o usuário a “procurar informações confiáveis”⁶.

No formato *chatbot* — uma espécie de assistente virtual que usa inteligência artificial e programação para se comunicar, por texto, com o usuário —, disponível nos aplicativos de mensagem, o robô se apresenta como uma ferramenta útil para “separar

⁶ Disponível em

<https://www.youtube.com/watch?v=yziW9TWgMXU&list=PLn92GYEvD3xeE9Vra3u0lt0KcNFd5VrrM&index=2>

notícia de opinião, procurar dados confiáveis e saber procurar por indícios de credibilidade de uma notícia, de um meme ou mesmo de um vídeo”⁷. No Twitter, a promessa é outra: o robô trabalha para “rastrear usuários que compartilham links falsos”⁸.

A versão beta da Fátima foi apresentada em novembro de 2023 sob a nomenclatura FátimaGTP e incorpora a tecnologia de LLM (sigla em inglês para Grande Modelo de Linguagem) para interpretar as perguntas dos usuários e dar respostas mais relevantes em linguagem natural. A versão do robô utiliza um banco de dados com reportagens e checagens já publicadas pelo Aos Fatos conectado ao GPT-4 — modelo de inteligência artificial da empresa OpenAI —, e as respostas são formuladas usando apenas o conteúdo produzido pelo Aos Fatos, com recorte de conteúdos publicados até março de 2023⁹.

Segundo Fávero (apud Ajour, 2023), a estratégia de aliar a Fátima ao projeto GTP minimiza a ocorrência de resoluções erradas e permite atualizar as informações diariamente. Para o diretor de inovação do Aos Fatos, o aproveitamento do conteúdo produzido pelo veículo é um destaque, já que, em seu modelo convencional, o ChatGTP utiliza coletadas apenas até 2021.

Sobre a potencialidade da Fátima de impactar o campo jornalístico, Fávero pontua que a aliança da robô de checagem com o chatbot americano torna a ferramenta brasileira ainda mais útil para levar informação verificada à sociedade e que ela pode se tornar um bom exemplo de como aplicar IA generativa para ajudar a aumentar o alcance do jornalismo. Nessa esteira, Nalon (apud Terra, 2023) aponta que os recorrentes problemas envolvendo desinformação e ódio em aplicativos reforçam a demanda por jornalismo

⁷ Disponível em

<https://www.youtube.com/watch?v=yziW9TWgMXU&list=PLn92GYEvD3xeE9Vra3u0lt0KcNFd5VrrM&index=2>

⁸ Disponível em

<https://www.youtube.com/watch?v=yziW9TWgMXU&list=PLn92GYEvD3xeE9Vra3u0lt0KcNFd5VrrM&index=2>

⁹ O dado foi fornecido pela própria Fátima, em resposta a uma pergunta feita à robô sobre o recorte da base de dados utilizada por ela. Acesso em 28 mai 2022.

confiável e inovador nesses dispositivos.

A versão beta da FátimaGTP, no entanto, pode apresentar falhas. Segundo o Aos Fatos, “nesta nova fase, o objetivo é coletar impressões dos leitores sobre a precisão e segurança do chatbot [e], por se tratar de uma versão de testes, é possível que o chatbot produza respostas imprecisas ou que fique momentaneamente indisponível” (Aos Fatos, 2023).

Não obstante a fase de testes, a robô checadora insere o Brasil numa espécie de esforço internacional para o combate à desinformação. A FátimaGTP tem o apoio da fundação Heinrich Böll, e a ferramenta vem sendo desenvolvida desde que o Aos Fatos foi selecionado para participar do programa JournalismAI, da LSE (*London School of Economics and Social Sciences*). O projeto da LSE, inclusive, conta com jornalistas brasileiros, que a redação do Aos Fatos colabora com os sites Newtral, da Espanha, e Núcleo Jornalismo, do Brasil — pertencentes ao projeto, que pesquisa soluções de inteligência artificial para *chatbots* jornalísticos.

Em 2024, o projeto brasileiro ganhou o concurso Digital Media Américas 2024, organizado pela Associação Mundial de Editores de Notícias (WAN-IFRA), que premia os melhores trabalhos jornalísticos do continente americano. A premiação recente se somou a duas conquistas anteriores do projeto Fátima: em 2023, a estratégia para as eleições de 2022 do chatbot Fátima no Telegram venceu a categoria engajamento de audiência no Digital Media Américas, e em 2019, a primeira versão do projeto saiu vencedora do Prêmio Cláudio Weber de Jornalismo de Dados na categoria inovação. Em suas diferentes fases, a Fátima já recebeu financiamentos de cinco organizações, dentre elas a Meta.

Considerações Finais

Durante este estudo, percebem-se dificuldades para encontrar projetos que utilizam Inteligência Artificial (IA) no combate à desinformação, evidenciando uma escassez significativa nesse campo. Essa carência sublinha a necessidade urgente de maior investimento e desenvolvimento de tecnologias voltadas para essa finalidade, além de incentivar novos estudos e pesquisas na área.

Este novo campo da IA aplicada ao jornalismo necessita de uma compreensão mais aprofundada e de mais investigações para determinar como a IA pode ser integrada de maneira eficaz nas práticas jornalísticas.

Além disso, o estudo destaca o potencial da IA para o jornalismo, especialmente com base nas experiências já realizadas com o projeto Fátima, desenvolvido pela agência de checagem "Aos Fatos". Este projeto demonstra como a IA pode ser utilizada para monitorar, identificar e desmentir informações falsas, comprovando-se uma ferramenta essencial no combate à desinformação.

Para o futuro, é sugerido expandir os estudos sobre a aplicação da IA no jornalismo, com foco no desenvolvimento de novas tecnologias e metodologias que possam ser incorporadas nas práticas jornalísticas cotidianas. Também é fundamental incentivar a colaboração entre pesquisadores, jornalistas e desenvolvedores de tecnologia para criar soluções mais robustas e eficazes no combate à desinformação.

Palavras-chave

Desinformação; checagem ; Inteligência Artificial ; jornalismo

Referências

AOS FATOS. **Ajude a testar a FátimaGPT, nova versão da robô checadora do Aos Fatos.** 2023.



IV Encontro Virtual da ABCiber

Associação Brasileira de Pesquisadores em Ciberultura

Perspectivas Interdisciplinares e Reconfigurações na Ciberultura:

Dados, Algoritmos e Inteligência Artificial

Online — 20 e 21/06/2024

Disponível em:
<<https://www.aosfatos.org/noticias/ajude-a-testar-a-fatimagpt-nova-versao-da-robo-checadora-do-aos-fatos/>>. Acesso em: 28 mai 2024.

DERAKHSHAN, H.; WARDLE, C. **Information disorder: definitions**. In: Understanding and addressing the desinformation system, 1., 2017, Filadélfia. Annals... Filadélfia: University of Pennsylvania, 2017. p. 5-12. Disponível em: <https://bit.ly/2GbeyJ2>. Acesso em: 24 set. 2019.

TERRA. **Aos Fatos ganha prêmio internacional por combate à desinformação no Telegram. 2023.** Disponível em:
<https://www.terra.com.br/noticias/checamos/aos-fatos-ganha-premio-internacional-por-combate-a-desinformacao-no-telegram_62aa42b991ea682cef32c402f06ee44cvcyzfdnt.html>. Acesso em: 28 mai 2024.

SANTAELLA, Lúcia; SALGADO, Marcelo de Mattos. **Deepfake e as consequências sociais da mecanização da desconfiança**. TECCOGS - Revista Digital de Tecnologias Cognitivas, n. 23, p. 90-103, jan./jun. 2021. Disponível em:
<https://revistas.pucsp.br/index.php/teccogs/article/view/55981/37929> . Acesso em: 28 mai. 2024.