# **DESINFORMAÇÃO E GERAÇÃO AUTOMÁTICA DE CONTEÚDO: A IA COMO FATOR DE RISCO NA DIVULGAÇÃO CIENTÍFICA**

Cristiane Menezes Ferreira, Doutoranda, UFF, email: cristianemf@id.uff.br

Francisco Gilson Rebouças Pôrto Junior, Doutor, UFT, email: gilsonportouft@gmail.com

# **PALAVRAS-CHAVE:** Inteligência Artificial. Desinformação. Divulgação Científica.

# **INTRODUÇÃO**

A acelerada integração de sistemas de Inteligência Artificial (IA) em diversas áreas da ciência e da tecnologia, vem impulsionando avanços expressivos nos últimos anos. Entretanto, o uso de ferramentas de IA também gera preocupações, especialmente no que se refere à falta de transparência de seus algoritmos (Almeida, 2024). Esse contexto favorece a proliferação de desinformação no campo da ciência, comprometendo nossa capacidade de distinguir fatos em meio a um grande volume de conteúdos enganosos. Modelos avançados de linguagem (LLMs) são capazes de elaborar textos que simulam a escrita científica, ao mesmo tempo em que *deepfakes* criam imagens e vídeos falsos que parecem reais. Combinações deletérias como essa, encontram potencial de distribuição ampliado em redes sociais, abalando a confiança nas instituições e nas publicações acadêmicas, podendo confundir tanto especialistas, quanto o público leigo (De Souza; Santaella, 2021). Diante disso, este relato de pesquisa se propõe a investigar como a IA pode favorecer a promoção de inverdades na divulgação científica e o quanto o seu uso irresponsável ou malicioso tem criado um cenário sem precedentes, com consequências desastrosas para a reputação da ciência e para a sociedade como um todo.

# **METODOLOGIA**

Para o desenvolvimento da pesquisa foi adotada uma metodologia de revisão bibliográfica com base em documentos publicados entre os anos de 2017 e 2025, em bases indexadas como SciELO, Google Acadêmico e *Web of Science*. Foram selecionados artigos que investigam a interseção entre IA e desinformação, com foco na divulgação científica (Telles, 2025). O trabalho foi centrado na busca de evidências da existência de mecanismos pelos quais os sistemas automatizados contribuem para a veiculação de publicações enganosas.

# **RESULTADOS E DISCUSSÃO**

A análise dos dados obtidos evidenciou que, embora a IA seja um recurso potencializador do progresso científico, os riscos por ela oferecidos, entre eles a geração e disseminação de informações falsas, impactam profundamente a ciência. Os riscos são diversos e incluem a capacidade que LLMs têm em produzir textos que podem conter interpretações incorretas de informações ou menções a documentos inexistentes; e a elaboração de *deepfakes* científicos, mediante a criação ou manipulação de imagens e vídeos realistas, que distorcem fatos e produzem material fraudulento (Cai *et al*., 2025). Os danos são expandidos nas redes sociais, onde conteúdo científico falso, gerado com o auxílio de IA, é facilmente compartilhado e curtido em níveis comparáveis aos de artigos verdadeiros, especialmente quando impulsionado por contas automatizadas. Além disso, a criação de "câmaras de eco" e de "bolhas de filtro", alavancada por algoritmos gerados por IA, fortalece discursos de ódio e narrativas negacionistas, inviabilizando fontes confiáveis e levando à disseminação desenfreada de narrativas falsas (Menezes *et al*., 2023). Assim, a desinformação gerada ou amplificada por IA, alimenta um ceticismo generalizado, já que a exposição recorrente a informações contraditórias ou falsas, fomenta a desconfiança excessiva do público (Natal, 2024). Ademais, ainda sobrecarrega os profissionais comprometidos com a popularização da ciência que, além de comunicá-la, também precisam se ocupar em desmentir mitos e a educar sobre como identificar os formatos sofisticados de geração de conteúdos imprecisos e/ou enganosos – o que torna o trabalho muito mais complexo. Apesar dos grandes riscos, a IA também pode ser uma aliada contra a desinformação, atuando na detecção de *deepfakes* e na checagem de fatos (Bhering, 2024) No entanto, esses recursos constantemente disputam espaço com os geradores de conteúdos falsos, além de possuírem certas limitações que podem comprometer sua eficiência, como vieses algorítmicos e vulnerabilidade à influência de fontes de dados incompletas ou desatualizadas (Almeida, 2024). Diante desse cenário, torna-se fundamental exigir transparência das plataformas quanto ao funcionamento dos algoritmos, capacitar o público para reconhecer e questionar informações produzidas por IA e fomentar tecnologias eficazes contra a manipulação de informações.

**CONSIDERAÇÕES FINAIS**

A capacidade da IA de gerar e propagar desinformação configura um panorama de risco considerável para a divulgação científica, que pode comprometer cada vez mais a credibilidade da ciência. Para enfrentar esse desafio, é essencial a criação de estratégias que combinem a implementação de diretrizes de utilização responsável de IA, o fortalecimento da educação midiática e da alfabetização científica, juntamente com um investimento contínuo no desenvolvimento de tecnologias de combate à desinformação. Em um mundo em que a distinção entre fatos e conteúdos falsos se torna cada vez mais complexa, a divulgação científica deve ser um meio eficaz de garantir que o conhecimento científico se torne cada vez mais acessível e confiável.

**REFERÊNCIAS BIBLIOGRÁFICAS**

ALMEIDA, Virgílio; NAS, Elen. Desafios da IA responsável na pesquisa científica. **Revista USP**, n. 141, p. 17-28, 2024. Disponível em: <https://revistas.usp.br/revusp/article/view/225204>. Acesso em 12 jul. 2025.

BHERING, Gabriel et al. A inteligência artificial no combate à desinformação: limites e potencialidades do projeto Fátima do Aos Fatos. In: **ENCONTRO VIRTUAL DA ABCIBER 2024**. 2024. Disponível em: <https://abciber.org.br/simposios/index.php/virtualabciber/virtual2024/paper/view/2486>. Acesso em 15 jul. 2025.

CAI, Zhixi et al. AV-Deepfake1M: A large-scale LLM-driven audio-visual deepfake dataset. In: **Proceedings of the 32nd ACM International Conference on Multimedia**. 2024. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3664647.3680795> . Acesso em 10 jul. 2025.

MENEZES, Vera Lúcia Moraes Araujo et al.. Câmaras de eco e filtros-bolha no contexto da gestão algorítimica da atenção. 2023. Disponível em: <https://cajapio.ufma.br/index.php/littera/article/view/22587>. Acesso em 15 jul. 2025.

DE SOUZA, Carlos Eduardo; SANTAELLA, Lucia. Deepfakes na perspectiva da semiótica. **TECCOGS: Revista Digital de Tecnologias Cognitivas**, n. 23, 2021. Disponível em: <https://revistas.pucsp.br/teccogs/article/view/55976>. Acesso em 11 jul. 2025.

NATAL, Bruno. Ceticismo digital na era da desinformação por IA. **MIT Technology Review - Brasil**, 19 abr. 2024. Disponível em: <https://mittechreview.com.br/ceticismo-digital-na-era-da-desinformacao-por-ia/>. Acesso em 10 Jul. 2025.

TELLES, Marina Klein; MONTARDO, Sandra Portella. Inteligência artificial e desinformação:: uma análise dos trabalhos apresentados em congressos da SBPJOR de 2015 a 2022. **Conexão-Comunicação e Cultura**, v. 22, n. 1, 2025. Disponível em: <https://sou.ucs.br/etc/revistas/index.php/conexao/article/view/11419>. Acesso em 10 jul. 2025.