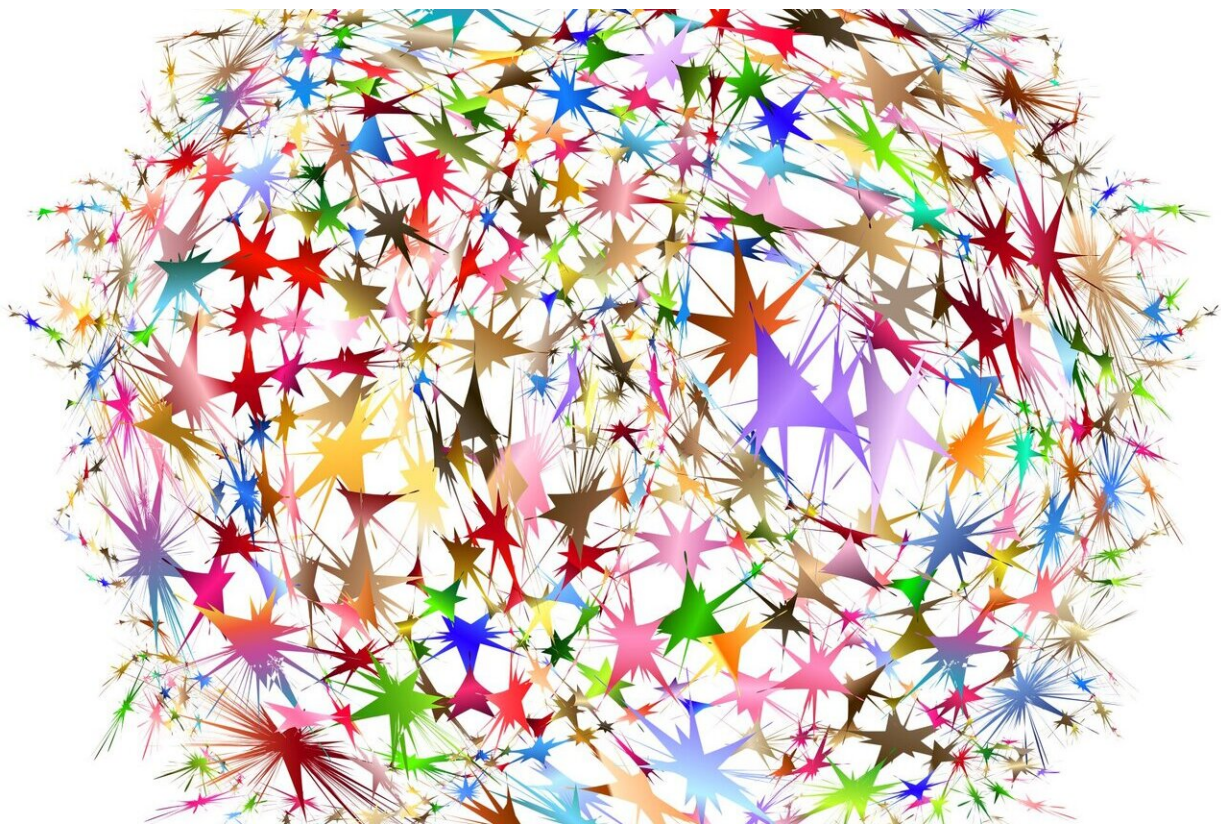


Medical researchers show AI can easily generate large volumes of health-related disinformation

November 13 2023



Credit: CC0 Public Domain

Government and industry guardrails are urgently needed for Generative AI to protect the health and well-being of our communities, say Flinders

University medical researchers who put the technology to the test and saw how it failed.

Rapidly evolving Generative AI, the cutting-edge domain prized for its capacity to create text, images, and video, was used in the study to test how [false information](#) about health and medical issues might be created and spread—and even the researchers were shocked by the results.

In the study, the team attempted to create disinformation about vaping and vaccines using Generative AI tools for text, image, and video creation.

In just over an hour, they produced over 100 misleading blogs, 20 deceptive images, and a convincing deep-fake video purporting health disinformation. Alarmingly, this video could be adapted into over 40 languages, amplifying its potential harm.

Bradley Menz, first author, registered pharmacist, and Flinders University researcher, says he has serious concerns about the findings, drawing upon prior examples of disinformation pandemics that have led to fear, confusion, and harm.

"The implications of our findings are clear: society currently stands at the cusp of an AI revolution, yet in its implementation governments must enforce regulations to minimize the risk of malicious use of these tools to mislead the community," says Mr. Menz.

"Our study demonstrates how easy it is to use currently accessible AI tools to generate large volumes of coercive and targeted misleading content on critical health topics, complete with hundreds of fabricated clinician and patient testimonials and fake, yet convincing, attention-grabbing titles."

"We propose that key pillars of pharmacovigilance—including transparency, surveillance, and regulation—serve as valuable examples for managing these risks and safeguarding [public health](#) amidst the rapidly advancing AI technologies," he says.

The research investigated OpenAI's GPT Playground for its capacity to facilitate the generation of large volumes of health-related disinformation. Beyond large-language models, the team also explored publicly available generative AI platforms, like DALL-E 2 and HeyGen, for facilitating the production of image and [video content](#).

Within OpenAI's GPT Playground, the researchers generated 102 distinct blog articles, containing more than 17,000 words of disinformation related to vaccines and vaping, in just 65 minutes. Further, within 5 minutes, using AI avatar technology and [natural language](#) processing the team generated a concerning deepfake video featuring a health professional promoting disinformation about vaccines. The video could easily be manipulated into over 40 different languages.

The investigations, beyond illustrating concerning scenarios, underscore an urgent need for robust AI vigilance. It also highlights important roles health care professionals can play in proactively minimizing and monitoring risks related to misleading health information generated by artificial intelligence.

Dr. Ashley Hopkins from the College of Medicine and Public Health and senior author says that there is a clear need for AI developers to collaborate with [health care professionals](#) to ensure that AI vigilance structures focus on public safety and well-being.

"We have proven that when the guardrails of AI tools are insufficient, the ability to rapidly generate diverse and large amounts of convincing [disinformation](#) is profound. Now, there is an urgent need for transparent

processes to monitor, report, and patch issues in AI tools," says Dr. Hopkins.

The paper—"Health Disinformation Use Case Highlighting the Urgent Need for Artificial Intelligence Vigilance"—will be published in *JAMA Internal Medicine*.

More information: Bradley D. Menz et al, Health Disinformation Use Case Highlighting the Urgent Need for Artificial Intelligence Vigilance, *JAMA Internal Medicine* (2023). [DOI: 10.1001/jamainternmed.2023.5947](https://doi.org/10.1001/jamainternmed.2023.5947)

Peter J. Hotez, Health Disinformation—Gaining Strength, Becoming Infinite, *JAMA Internal Medicine* (2023). [DOI: 10.1001/jamainternmed.2023.5946](https://doi.org/10.1001/jamainternmed.2023.5946)

Provided by Flinders University

Citation: Medical researchers show AI can easily generate large volumes of health-related disinformation (2023, November 13) retrieved 29 April 2024 from <https://medicalxpress.com/news/2023-11-medical-ai-easily-generate-large.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--