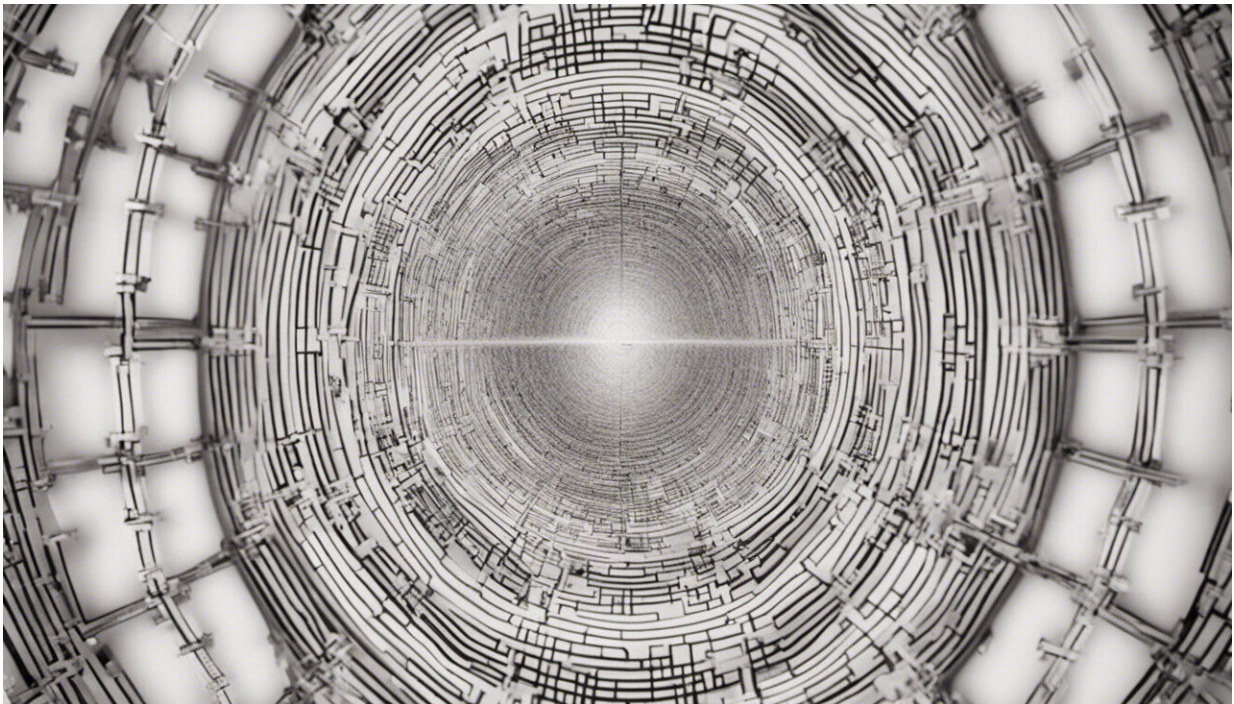


Artificial intelligence may not actually be the solution for stopping the spread of fake news

November 29 2021, by Sze-Fung Lee, Benjamin C. M. Fung



Credit: AI-generated image ([disclaimer](#))

Disinformation has been used in [warfare and military strategy over time](#). But it is undeniably being intensified by the use of smart technologies and social media. This is because these communication technologies provide a relatively low-cost, low-barrier way to disseminate information basically anywhere.

The million-dollar question then is: Can this technologically produced problem of scale and reach also be solved using technology?

Indeed, the continuous development of new technological solutions, such as artificial intelligence (AI), may provide part of the solution.

Technology companies and [social media](#) enterprises are working on [the automatic detection of fake news through natural language processing, machine learning and network analysis](#). The idea is that an algorithm will identify information as "fake news," [and rank it lower to decrease the probability of users encountering it](#).

Repetition and exposure

From a psychological perspective, [repeated exposure to the same piece of information makes it likelier for someone to believe it](#). When AI detects [disinformation](#) and reduces the frequency of its circulation, this can break the cycle of reinforced information consumption patterns.

However, AI detection still remains unreliable. First, current detection is based on the assessment of text (content) and its social network to determine its credibility. Despite determining the origin of the sources and the dissemination pattern of fake news, the fundamental problem lies within how AI verifies the actual nature of the content.

Theoretically speaking, if the amount of training data is sufficient, the AI-backed classification model would be able to interpret whether an article contains fake news or not. Yet the reality is that making such distinctions requires prior political, cultural and social knowledge, or common sense, which [natural language processing](#) algorithms still lack.

In addition, fake news can be highly nuanced when it is deliberately altered to "[appear as real news but containing false or manipulative](#)

[information](#)," as a pre-print study shows.

Human-AI partnerships

Classification analysis is also heavily influenced by the theme—AI often differentiates topics, rather than genuinely the content of the issue to determine its authenticity. For example, articles related to COVID-19 are more likely to be labeled as fake news than other topics.

One solution would be to employ people to work alongside AI to verify the authenticity of information. For instance, in 2018, the Lithuanian defense ministry developed an AI program that ["flags disinformation within two minutes of its publication and sends those reports to human specialists for further analysis."](#)

A similar approach could be taken in Canada by establishing a national special unit or department to combat disinformation, or supporting think tanks, universities and other third parties to research AI solutions for fake news.

Avoiding censorship

Controlling the spread of fake news may, in some instances, be considered censorship and a threat to freedom of speech and expression. Even a human may have a hard time judging whether information is fake or not. And so perhaps the bigger question is: Who and what determine the definition of fake news? How do we ensure that AI filters will not drag us into the false positive trap, and incorrectly label information as fake because of its associated data?

An AI system for identifying fake [news](#) may have sinister applications. Authoritarian governments, for example, may use AI as an excuse to justify the removal of any articles or to prosecute individuals not in

favor of the authorities. And so, any deployment of AI—and any relevant laws or measurements that emerge from its application—will require a transparent system with a third party to monitor it.

Future challenges remain as disinformation—especially when associated with foreign intervention—is an ongoing issue. An algorithm invented today may not be able to detect future [fake news](#).

For example, deep fakes—which are "[highly realistic and difficult-to-detect digital manipulation of audio or video](#)"—are likely to play a bigger role in future [information](#) warfare. And disinformation spread via messaging apps such as WhatsApp and Signal are becoming more difficult to track and intercept because of end-to-end encryption.

A recent study showed that [50 percent of the Canadian respondents received fake news through private messaging apps regularly](#). Regulating this would require striking a balance between privacy, individual security and the clampdown of disinformation.

While it is definitely worth allocating resources to combating disinformation using AI, caution and transparency are necessary given the potential ramifications. New technological solutions, unfortunately, may not be a silver bullet.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Artificial intelligence may not actually be the solution for stopping the spread of fake news (2021, November 29) retrieved 10 April 2024 from <https://techxplore.com/news/2021-11-artificial-intelligence-solution-fake-news.html>

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.