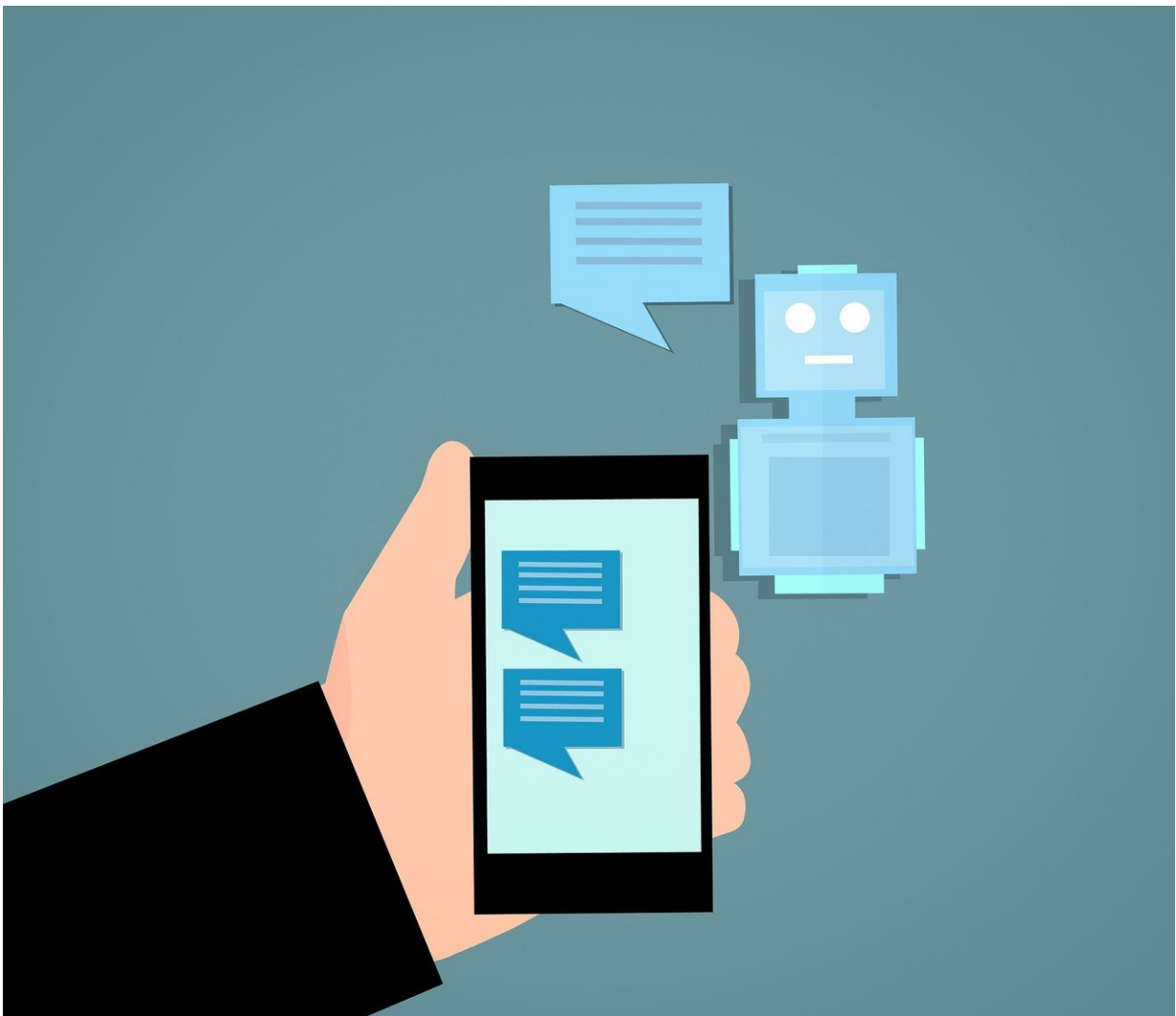


Chatbots tell people what they want to hear, researchers find

May 13 2024



Credit: CC0 Public Domain

Chatbots share limited information, reinforce ideologies, and, as a result, can lead to more polarized thinking when it comes to controversial issues, according to new Johns Hopkins University–led research.

The study challenges perceptions that chatbots are impartial and provides insight into how using conversational search systems could widen the public divide on hot-button issues and leave people vulnerable to manipulation.

"Because people are reading a summary paragraph generated by AI, they think they're getting unbiased, fact-based answers," said lead author Ziang Xiao, an assistant professor of computer science at Johns Hopkins who studies human-AI interactions. "Even if a chatbot isn't designed to be biased, its answers reflect the biases or leanings of the person asking the questions. So really, people are getting the answers they want to hear."

Xiao and his team share their findings at the [Association of Computing Machinery's CHI conference on Human Factors in Computing Systems](#) at 5 p.m. ET on Monday, May 13.

To see how chatbots influence online searches, the team compared how people interacted with different search systems and how they felt about controversial issues before and after using them.

The researchers asked 272 participants to write out their thoughts about topics including health care, [student loans](#), or sanctuary cities, and then look up more information online about that topic using either a chatbot or a traditional search engine built for the study.

After considering the [search results](#), participants wrote a second essay and answered questions about the topic. Researchers also had participants read two opposing articles and questioned them about how

much they trusted the information and if they found the viewpoints to be extreme.

Because chatbots offered a narrower range of information than traditional web searches and provided answers that reflected the participants' preexisting attitudes, the participants who used them became more invested in their original ideas and had stronger reactions to information that challenged their views, the researchers found.

"People tend to seek information that aligns with their viewpoints, a behavior that often traps them in an echo chamber of like-minded opinions," Xiao said. "We found that this echo chamber effect is stronger with the chatbots than traditional web searches."

The echo chamber stems, in part, from the way participants interacted with chatbots, Xiao said. Rather than typing in keywords, as people do for traditional search engines, chatbot users tended to type in full questions, such as, "What are the benefits of universal health care?" or "What are the costs of universal [health care](#)?" A chatbot would answer with a summary that included only benefits or costs.

"With chatbots, people tend to be more expressive and formulate questions in a more conversational way. It's a function of how we speak," Xiao said. "But our language can be used against us."

AI developers can train chatbots to extract clues from questions and identify people's biases, Xiao said. Once a chatbot knows what a person likes or doesn't like, it can tailor its responses to match.

In fact, when the researchers created a chatbot with a hidden agenda, designed to agree with people, the echo chamber effect was even stronger.

To try to counteract the echo chamber effect, researchers trained a chatbot to provide answers that disagreed with participants. People's opinions didn't change, Xiao said. The researchers also programmed a [chatbot](#) to link to source information to encourage people to fact-check, but only a few participants did.

"Given AI-based systems are becoming easier to build, there are going to be opportunities for malicious actors to leverage AIs to make a more polarized society," Xiao said. "Creating agents that always present opinions from the other side is the most obvious intervention, but we found they don't work."

Authors include Johns Hopkins graduate student Nikhil Sharma and Microsoft principal researcher Q. Vera Liao.

Provided by Johns Hopkins University

Citation: Chatbots tell people what they want to hear, researchers find (2024, May 13) retrieved 26 June 2024 from <https://techxplore.com/news/2024-05-chatbots-people.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.