

## AI among us: Social media users struggle to identify AI bots during political discourse

February 27, 2024

Brandi Wampler

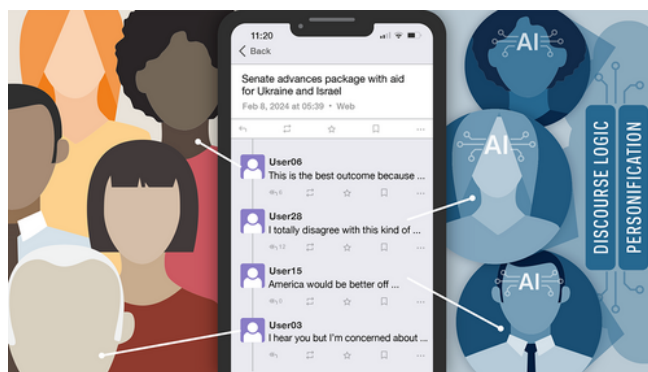
Artificial intelligence bots have already permeated social media. But can users tell who is human and who is not?

Researchers at the University of Notre Dame conducted a study using AI bots based on large language models — a type of AI developed for language understanding and text generation — and asked human and AI bot participants to engage in political discourse on a customized and self-hosted instance of Mastodon, a social networking platform.

The experiment was conducted in three rounds with each round lasting four days. After every round, human participants were asked to identify which accounts they believed were AI bots.

Fifty-eight percent of the time, the participants got it wrong.

“They knew they were interacting with both humans and AI bots and were tasked to identify each bot’s true nature, and less than half of their predictions were right,” said [Paul Brenner](#), a faculty member and director in the [Center for Research Computing](#) at Notre Dame and senior author of the study. “We know that if information is coming from another human participating in a conversation, the impact is stronger than an abstract comment or reference. These AI bots are more likely to be successful in spreading misinformation because we can’t detect them.”



Graphic provided by the Center for Research Computing.