How AI bots spread misinformation online and undermine democratic politics

July 25 2024, by Sophia Melanson Ricciardone

Credit: Unsplash/CC0 Public Domain

Consider a typical morning routine: coffee in hand, you peruse Twitter (now rebranded as X) to catch up on the news. Headlines appear among a flurry of tweets on everything from memes about political leaders to

cultural Marxism, free speech, making America great again and draining the swamp.

Before your day has even begun, a burst of disparate ideas coalesces in your mind in response to the appearance of a single word or catchphrase. It's a scenario repeated daily, where snippets of information mold themselves onto our views and biases, influencing how we interpret online discourse and those who engage in it.

In the heated space of contemporary politics, popularized words and catchphrases wield a lot of influence. Controversial rallying cries like "build the wall" and "Trudeau must go" regularly appear on social media, punctuating debates and discourse with an emotionally palpable fervor.

These phrases are more than mere words; they are ideological shorthand that seek to galvanize people and spark outrage online like never before.

But, in our increasingly digitized world, how do we know whether the accounts we interact with online are other human beings or bots? And given the powerful influence this kind of rhetoric can have, what impact do these bots have on our decision-making and democratic processes?

AI bots

My Ph.D. research focused on the rise of "botaganda"—online content circulated by automated accounts, or bots, for electioneering purposes.

Bots are automated accounts on social media that can be used to post tweets, like and share content or follow users on social media without needing a person to do it manually.

Scholars have highlighted how bots "could be used to covertly exploit weaknesses in [a person's] character and persuade them to take action

against their own best interest."

The advent of artificial intelligence and machine learning has certainly equipped us with several advantages in contemporary life, but it has also made independent political thought much harder to achieve. It is increasing the prevalence of digital misinformation, and demands that we exercise vigilance to ensure we can make informed decisions.

Understanding the social psychology that makes us susceptible to catchphrases like "drain the swamp" is integral to combating the impact of misinformation circulated online. Our social brains are susceptible to these kinds of linguistic triggers in three important ways:

- The fact that humans mimic and synchronize communication styles when interacting;
- We're more likely to remember frequently repeated words and;
- Our brains are more likely to connect unrelated ideas when words or phrases frequently recur together over and over again in the same context.

When we unwittingly engage with bots that repeatedly use certain phrases and terms, this subtly reinforces their association and impact.

What I found

I conducted statistical analysis on how bot-generated content influenced Canadian Twitter users during the SNC Lavalin scandal from March 14 to April 9, 2019. My study found strong correlations between bot-generated and human tweets, suggesting people engaged closely with AI-generated content and stored it in memory for easy retrieval and recall.

My analysis shows that bot-circulated tweets shared a high degree of similarity with human-generated tweets. The similarity in the emotional

salience of bot-generated and human-generated tweets was significantly pronounced.

I first used Spearman's phi coefficient, a statistical tool, to measure how strongly bot tweets related with human tweets. I then applied linear regression to understand this relationship in more detail and to see if changes in bot tweets affected changes in human tweets.

The results show there is a strong correlation between bot and human tweets, and that the content of bot tweets significantly influences linguistic aspects of human-generated tweets. In simpler terms, human tweets replicated bot tweets to a high degree.

On March 14, 2019, for instance, bot tweets shared 75 percent similarity with human tweets, which increased to 92 percent similarity by March 28. The emotional words used in bot tweets were reproduced in human tweets just over 97 percent of the time on March 14. Though the reproduction of emotional words in human tweets decreased over time, the similarity remained significant. This underscores how closely bot- and human-generated content can mirror each other.

Sometimes, bots simply mirror human activity because machine learning and generative AI technology are designed to replicate our behaviors. However, my analysis aims to understand not just how similar bot tweets and human tweets are, but also which one influences the other and in what way.

My examination of prevalent words and phrases like "obstruction of justice," "Trudeau's scandal" and "Liberal coverup" propagated by bots were replicated at high frequency in human-generated tweets, both in unique and retweeted tweets.

Social mimicry

This study's findings support the idea that we are inclined to mimic language structures of tweets posted by bots, which gains traction among users engaged in a resulting ecology of tweets posted both by peers and computer programs, a phenomenon observed in human-computer interactions.

Research underscores our innate tendency to mimic the communication patterns of those around us, particularly those who share our values and beliefs—a phenomenon also evident in our digital interactions. This inclination shapes how we express ourselves online, influencing our language, intonation and even how we make arguments.

Popular words and catchphrases that appear at high frequency within online contexts help our brains organize and remember large amounts of information more easily.

But how do these phrases achieve such resonance?

We know that repeatedly hearing specific words or phrases significantly enhances our ability to recall them more effortlessly, especially when they consistently appear within the same context.

In the landscape of our social brains, single words and phrases don't carry much weight by themselves. What makes them powerful tools for spreading misinformation is their knack for evoking mental connections to familiar ideas, shaping how we understand current political developments.

Consider this: when several disparate concepts are cleverly woven into a memorable catchy phrase, they merge those underlying concepts in our minds, making them feel logical and familiar. For example: "Make America Great Again."

This fusion is especially potent when these ideas strike a chord with our core values and beliefs, making catchphrases highly effective at molding public opinion in profound ways.

As we navigate this era of digital discourse, awareness of blind spots in our social psychology is our best defense. Understanding how cues or triggers affect us can reduce their influence over time. The more aware we are of bots and how they work, the more able we are to protect ourselves from misleading rhetoric, ensuring our democratic processes remain robust and inclusive.

Provided by The Conversation

Citation: How AI bots spread misinformation online and undermine democratic politics (2024, July 25) retrieved 26 July 2024 from https://phys.org/news/2024-07-ai-bots-misinformation-online-undermine.html