

SOCIAL SCIENTISTS BATTLE BOTS TO GLEAN INSIGHTS ONLINE

Automated production of social-media posts can confound research studies.

By Heidi Ledford

Social-media bots that pump out computer-generated content have been accused of swaying elections and damaging public health by spreading misinformation. Now, some social scientists have a fresh accusation: bots meddle with research studies that mine popular sites such as Twitter, Reddit and Instagram for information on human health and behaviour.

Data from these sites can help scientists to understand how natural disasters affect mental health, why young people have flocked to e-cigarettes in the United States and how people join together in complex social networks. But such work relies on distinguishing the real voices from the automated ones.

“Bots are designed to behave online like people,” says Jon-Patrick Allem, a social scientist at the University of Southern California in Los Angeles. “If a researcher is interested in describing public attitudes, you have to be sure that the data you’re collecting on social media is actually from people.”

Computer scientist Sune Lehmann designed his first bots in 2013, as a social-network experiment for a class that he was teaching at the

Technical University of Denmark in Kongens Lyngby. Back then, he says, Twitter bots were simple, obscure and mainly meant to increase the number of followers for specific accounts. Lehmann wanted to show his students how such bots could manipulate social systems, so together they designed simple bots that impersonated fans of the singer Justin Bieber.

The ‘Bieber Bots’ quickly attracted thousands of followers. But social-media bots have continued to evolve, becoming more complex and harder to detect. They surged into the spotlight after the 2016 US presidential election – amid accusations that bots had been deployed on social media in an attempt to sway the vote. “All of a sudden, it became something of interest to people,” Allem says.

Since then, Allem has shown that tweets generated by bots are twice as likely as their real counterparts to attest that e-cigarettes help people to give up smoking¹. Bots are also more likely to tout the unproven health benefits of cannabis². These studies rely on algorithms that estimate the likelihood that a Twitter account is automated. But despite bot-detecting tools with names like BotSlayer, Allem says that many social-science and public-health researchers still fail to filter out

probable automated content from their data.

That omission can pollute a data set, says Amelia Jamison, who studies health disparities at the University of Maryland in College Park and has mined social media for posts that oppose vaccination. “You might be artificially giving the bots a voice by treating them as if they are really part of the discussion, when they are actually just amplifying something that may not be voiced by the community,” she says.

One problem that the field must grapple with is how to define a bot, says Katrin Weller, an information scientist at the Leibniz Institute for the Social Sciences in Cologne, Germany. Not all bots are dispensing misinformation: some provide data from weather stations, or general news updates. Some researchers define Twitter bots as those accounts that send out more than a certain number of messages each day – a loose definition that could rope in prolific human tweeters.

Other definitions are more complex, but bot detectors are locked in an arms race with bot developers. First-generation social-media bots were relatively simple programs that retweeted others’ posts at regular intervals. Now, advances in machine learning have enabled the creation of more sophisticated bots that post original content. Some post at random intervals and mimic human patterns, such as not tweeting when a person would probably be asleep. Some developers will mix in human-generated content with automated content to better camouflage their bots.

“Once you know more about the bots and how to detect them, then this knowledge is also available for the bot creators,” says Oliver Grübner, who studies quantitative health geography at the University of Zurich in Switzerland. “It’s a really tricky field.”

Like Lehman, some social scientists are creating their own bots to conduct social experiments. Kevin Munger, a political scientist at Pennsylvania State University in University Park, and his colleagues built bots that chided Twitter users who used racist language. One set of bots had profile pictures of white men; the other had profile pictures of black men. Munger found that Twitter users were more likely to tone down their racist rhetoric after being called out by bots with a white male profile picture³.

After his Bieber Bot success, Lehmann designed more sophisticated bots to study how behaviours spread from one group to another. But bots now have such a bad reputation that he is leaning towards abandoning the approach, for fear of a public backlash. “I kind of thought: ‘I’ll find another quiet corner and do my research without courting controversy,’” he says.

1. Allem, J.-P. et al. *JMIR Public Health Surveill.* **3**, e98 (2017).

2. Allem, J.-P. et al. *Am. J. Public Health* <https://doi.org/10.2105/AJPH.2019.305461> (2019).

3. Munger, K. *Political Behav.* **39**, 629–649 (2017).



Separating real online voices from automated ones can be a problem for researchers.