

# Can you spot the bots? New research says no

March 15 2023

---



**William Bennett** @williamb

Ukraine is currently in the midst of a political and economic crisis. The country's economy is in tatters, and its government is unstable. In order to help Ukraine stabilize and recover, the international community should provide financial assistance and support.



**Matthew Gibson** @MatthewGibson

I was watching the news when I saw a video of what looked like two Ukrainian military helicopters firing missiles at a fuel depot in the eastern city of Belgorod, in what would be, if confirmed, the first known air raid by Ukraine's forces on Russian soil



**Eric Hawkins** @Erichawkins

The Russian military has proposed a new evacuation plan for Ukrainian civilians and foreign nationals aiming to flee major cities amid Moscow's military offensive in Ukraine

Examples of the generated profiles shown during the experiment. Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2209.07214

Until recently it has been a challenge to make convincing fake social media profiles at scale because images could be traced back to their source, and the text often didn't sound human-like.

Today with rapid advances in [artificial intelligence](#) it is increasingly becoming difficult to tell the difference. Researchers from Copenhagen Business School decided to conduct an experiment with 375 participants

to test the difficulty in distinguishing between real and fake social media profiles.

They found participants were unable to differentiate between artificially generated fake Twitter accounts and real ones, and in fact, perceived the [fake accounts](#) to be less likely to be fake than the genuine ones.

The researchers created their own mock twitter feed where the topic was the war in Ukraine. The feed included real and generated profiles with tweets supporting both sides. The [fake profiles](#) used computer-generated synthetic [profile](#) pictures created with StyleGAN, and posts generated by GPT-3, the same language model that is behind ChatGPT.

"Interestingly, the most divisive accounts on questions of accuracy and likelihood belonged to the genuine humans. One of the real profiles was mislabeled as fake by 41.5% of the participants who saw it. Meanwhile, one of the best-performing fake profiles was only labeled as a bot by 10%," says Sippo Rossi, a Ph.D. Fellow from the Centre for Business Data Analytics at the Department of Digitalization at Copenhagen Business School.

"Our findings suggest that the technology for creating generated fake profiles has advanced to such a point that it is difficult to distinguish them from real profiles," he adds.

The research was presented at the Hawaii International Conference on System Sciences (HICSS), and the paper is available on the *arXiv* preprint server.

## Potential for misuse

"Previously it was a lot of work to create realistic fake profiles. Five years ago the average user did not have the technology to create fake

profiles at this scale and easiness. Today it is very accessible and available to the many not just the few," says co-author Raghava Rao Mukkamala, the Director of the Centre for Business Data Analytics at Department of Digitalization at Copenhagen Business School.

From political manipulation to misinformation to cyberbullying and cybercrime, the proliferation of deep learning-generated social media profiles has significant implications for society and democracy as a whole.

"Authoritarian governments are flooding social media with seemingly supportive people to manipulate information so it's essential to consider the potential consequences of these technologies carefully and work towards mitigating these [negative impacts](#)," adds Raghava Rao Mukkamala.

## Future research

The researchers used a simplified setting where the participants saw one tweet and the profile information of the account that posted it, the next research step will be to see if bots can be correctly identified from a news feed discussion where different fake and real profile are commenting on a specific news article in the same thread.

"We need new ways and new methods to deal with this as putting the genie back in the lamp is now virtually impossible. If humans are unable to detect fake profile and posts and to report them then it will have to be the role of automated detection, like removing accounts and ID verification and the development of other safeguards by the companies operating these [social networking sites](#)," adds Sippo Rossi.

"Right now my advice would be to only trust people on [social media](#) that you know," concludes Sippo Rossi.

**More information:** Sippo Rossi et al, Are Deep Learning-Generated Social Media Profiles Indistinguishable from Real Profiles?, *arXiv* (2022). [DOI: 10.48550/arxiv.2209.07214](https://doi.org/10.48550/arxiv.2209.07214)

Provided by Copenhagen Business School

Citation: Can you spot the bots? New research says no (2023, March 15) retrieved 10 April 2024 from <https://techxplore.com/news/2023-03-bots.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.