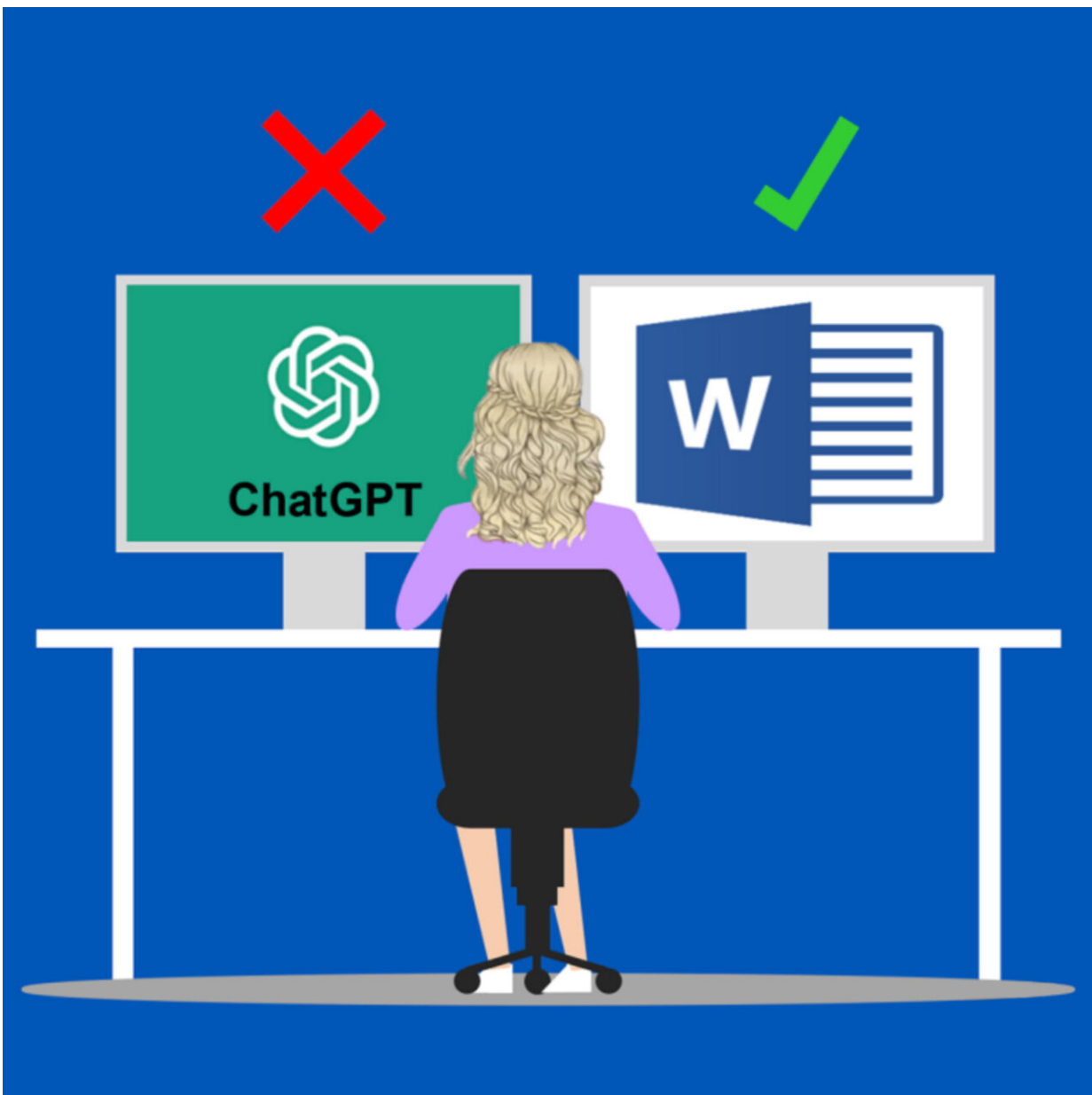


AI-generated academic science writing can be identified with over 99% accuracy

June 7 2023



ChatGPT vs Human. Credit: Heather Desaire and Romana Jarosova, University of Kansas

The debut of artificial intelligence chatbot ChatGPT has set the world abuzz with its ability to churn out human-like text and conversations. Still, many telltale signs can help us distinguish AI chatbots from humans, according to a study published on June 7 in the journal *Cell Reports Physical Science*. Based on the signs, the researchers developed a tool to identify AI-generated academic science writing with over 99% accuracy.

"We tried hard to create an accessible method so that with little guidance, even [high school students](#) could build an AI detector for different types of writing," says first author Heather Desaire, a professor at the University of Kansas. "There is a need to address AI writing, and people don't need a computer science degree to contribute to this field."

"Right now, there are some pretty glaring problems with AI writing," says Desaire. "One of the biggest problems is that it assembles text from many sources and there isn't any kind of accuracy check—it's kind of like the game Two Truths and a Lie."

Although many AI text detectors are available online and perform fairly well, they weren't built specifically for academic writing. To fill the gap, the team aimed to build a tool with better performance precisely for this purpose. They focused on a type of article called perspectives, which provide an overview of specific research topics written by scientists. The team selected 64 perspectives and created 128 ChatGPT-generated articles on the same research topics to train the [model](#). When they compared the articles, they found an indicator of AI writing—predictability.

Contrary to AI, humans have more complex paragraph structures, varying in the number of sentences and total words per paragraph, as well as fluctuating sentence length. Preferences in punctuation marks and vocabulary are also a giveaway. For example, scientists gravitate towards words like "however," "but" and "although," while ChatGPT often uses "others" and "researchers" in writing. The team tallied 20 characteristics for the model to look out for.

When tested, the model aced a 100% accuracy rate at weeding out AI-generated full perspective articles from those written by humans. For identifying individual paragraphs within the article, the model had an accuracy rate of 92%. The research team's model also outperformed an available AI text detector on the market by a wide margin on similar tests.

Next, the team plans to determine the scope of the model's applicability. They want to test it on more extensive datasets and across different types of academic science writing. As AI chatbots advance and become more sophisticated, the researchers also want to know if their model will stand.

"The first thing people want to know when they hear about the research is 'Can I use this to tell if my students actually wrote their paper?'" said Desaire. While the model is highly skilled at distinguishing between AI and scientists, Desaire says it was not designed to catch AI-generated student essays for educators. However, she notes that people can easily replicate their methods to build models for their own purposes.

More information: Heather Desaire, Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools., *Cell Reports Physical Science* (2023). [DOI: 10.1016/j.xcrp.2023.101426](https://doi.org/10.1016/j.xcrp.2023.101426). [www.cell.com/cell-reports-phys ... 2666-3864\(23\)00200-X](https://www.cell.com/cell-reports-phys/2666-3864(23)00200-X)

Provided by Cell Press

Citation: AI-generated academic science writing can be identified with over 99% accuracy (2023, June 7) retrieved 19 April 2024 from <https://techxplore.com/news/2023-06-ai-generated-academic-science-accuracy.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.