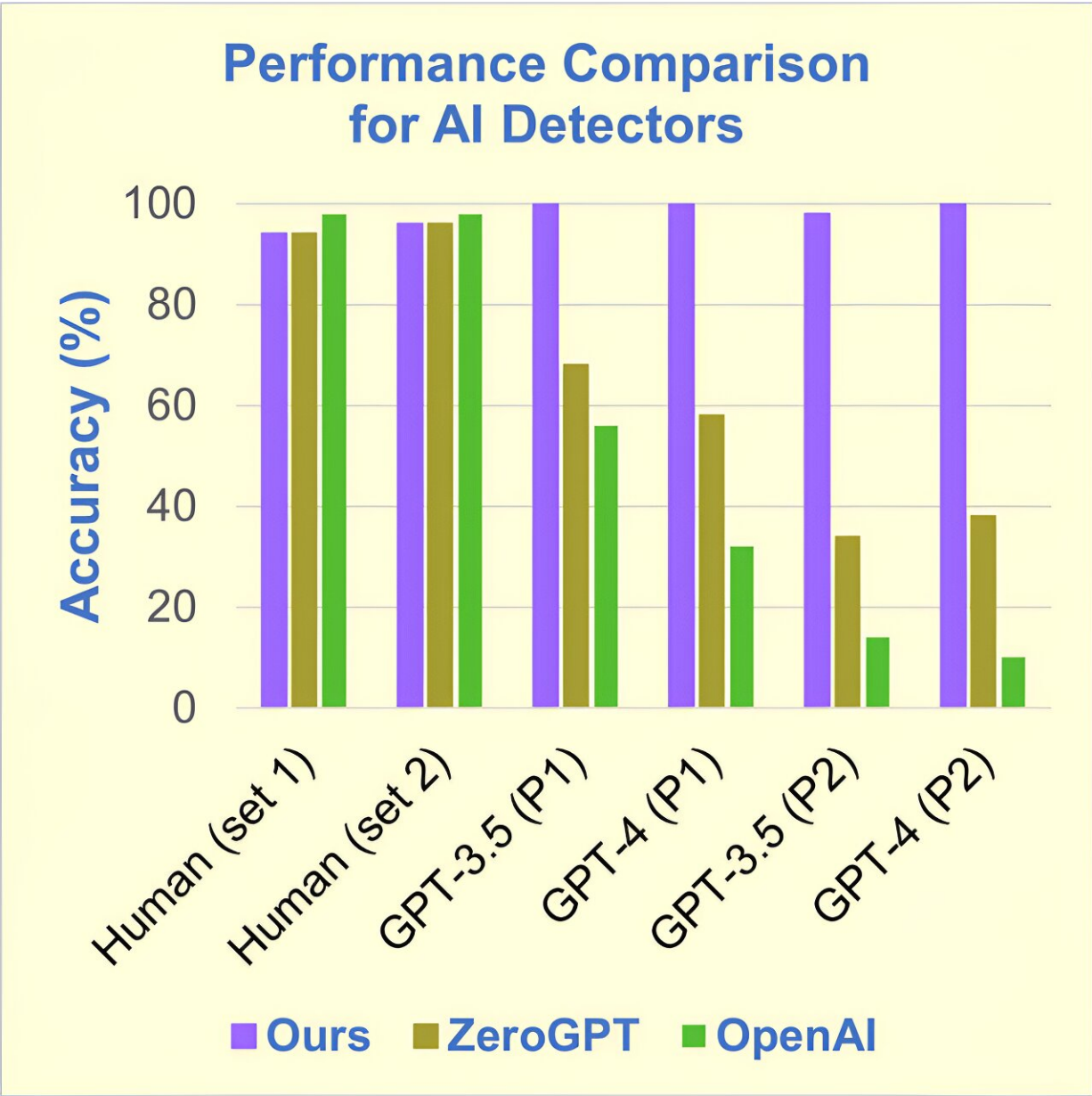


Tool detects AI-generated text in science journals

November 7 2023, by Peter Grad



Graphical abstract. Credit: *Cell Reports Physical Science* (2023). DOI: 10.1016/j.xcrp.2023.101672

In an era of heightened concern in academia regarding AI-generated essays, there is reassuring news from the University of Kansas.

Researchers have developed an AI text [detector](#) for scientific essays that can distinguish between human-written and computer-generated content almost 100% of the time.

In a [study](#) appearing Nov. 6 in *Cell Reports Physical Science*, Professor Heather Desaire noted that while there are several AI detectors for general content currently available, none are particularly good when applied to [scientific papers](#).

"Most of the field of text analysis wants a really general detector that will work on anything," Desaire said. Instead, her team zeroed in on reports written specifically for [scientific journals](#) on the subject of chemistry.

"We were really going after accuracy," she said.

The team's detector was trained on journals published by the American Chemical Society. They collected 100 introductory passages authored by professionals and then programmed ChatGPT to write its own introductions based either on journal abstracts or simply on the titles of reports.

When the ChatGPT detector scanned the three categories of reports, it correctly identified the human-authored passages 100% of the time, as well as reports generated from prompts including only report titles. The

results were almost as good with reports relying on introductory passages, with correct identification 98% of the time.

A competing classifier such as ZeroGPT, which boasts 98% accuracy for detecting general AI-written essays, however, did poorly when it came to the chemistry-related reports. It achieved only an average 37% accuracy rate on the same reports based on titles, and only a few percentage points better on reports based on introductory text.

A second competitor, OpenAI, fared even worse, failing to correctly identify authorship of essays an average of 80% of the time.

"Academic publishers are grappling with the rapid and pervasive adoption of new artificial intelligence text generators," Desaire said. "This new detector will allow the [scientific community](#) to assess the infiltration of ChatGPT into chemistry journals, identify the consequences of its use, and rapidly introduce mitigation strategies when problems arise."

Scientific journals are rewriting their rules regarding article submissions, with most banning AI-generated reports and requiring disclosure of any other AI processes used in composing a report.

Desaire listed several concerns about the risks of AI-generated content infiltrating scientific journals: "Their overuse may result in a flood of marginally valuable manuscripts. They could cause highly cited papers to be over-represented and emerging works, which are not yet well known, to be ignored."

"Most concerning," she added, "is these tools' tendency toward 'hallucination,' making up facts that are not true."

As an illustration, Desaire included a personal anecdote about the results

of a ChatGPT-written biographical sketch about her. It said she "graduated from the University of Minnesota, is a member of the Royal Society of Chemistry, and won the Biemann Medal." Impressive accomplishments, but all false.

"While this example is amusing," Desaire said, "infiltrating the [scientific literature](#) with untruths is far from funny."

But she remains optimistic. Some say resisting the emergence of AI-generated content is inevitable, she said, and they argue "developing tools like this one is participating in an [arms race](#) [against AI] that humans will not win."

She said editors must take the lead in detecting AI contamination.

"Journals should take reasonable steps to ensure that their policies about AI writing are followed, and we think it is fully feasible to stay ahead of the problem of detecting AI," she said.

More information: Heather Desaire et al, Accurately detecting AI text when ChatGPT is told to write like a chemist, *Cell Reports Physical Science* (2023). [DOI: 10.1016/j.xcrp.2023.101672](https://doi.org/10.1016/j.xcrp.2023.101672)

© 2023 Science X Network

Citation: Tool detects AI-generated text in science journals (2023, November 7) retrieved 27 April 2024 from

<https://techxplore.com/news/2023-11-tool-ai-generated-text-science-journals.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.