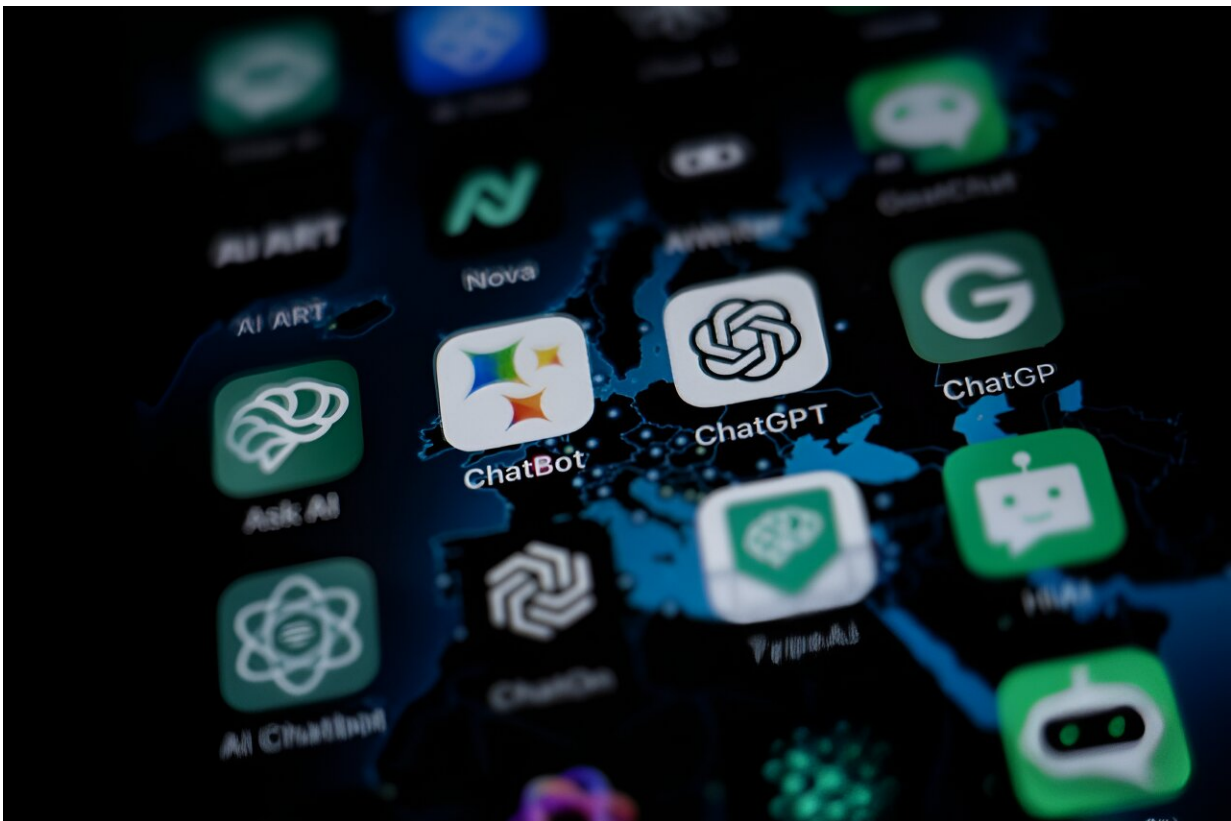


Fearing 'pillaging', news outlets block an OpenAI bot

August 30 2023, by Carole GUIRADO



A consortium of news outlets issued an open letter earlier in August saying AI firms must ask for permission before using copyrighted text and images to generate content.

A growing number of media outlets are blocking a webpage-scanning tool used by ChatGPT creator OpenAI to improve its artificial

intelligence models.

The New York Times, CNN, Australian broadcaster ABC and news agencies Reuters and Bloomberg have taken steps to thwart GPTBot, a web crawler launched on August 8.

They were followed by French news organizations including France 24, RFI, Mediapart, Radio France and TF1.

"There's one thing that won't stand: it's the unauthorized pillaging of content," Radio France president Sibyle Veil said at a news conference on Monday.

Nearly 10 percent of the top 1,000 websites in the world blocked access to GPTBot just two weeks after it was launched, according to plagiarism tracker Originality.ai.

They include Amazon.com, Wikihow.com, Quora.com and Shutterstock. Originality.ai said it expects the list to grow by five percent per week.

On its website, OpenAI says that "allowing GPTBot to access your site can help AI models become more accurate and improve their general capabilities and safety".

But the California startup also provides directions on how to block the bot.

"There is no reason for them to come and learn about our content without compensation," Laurent Frisch, director of digital and innovation strategy at Radio France, told AFP.

Fair remuneration

AI tools like chatbot ChatGPT and image generators DALL-E 2, Stable Diffusion and Midjourney exploded in popularity last year with their ability to generate a wealth of content from just brief text prompts.

However, the firms behind the tools, including OpenAI and Stability AI, already face lawsuits from artists, authors and others claiming their work has been ripped off.

"Enough with being plundered by these companies that turn profits on the back of our production," added Vincent Fleury, director of digital space at France Medias Monde, the parent company of France 24 and RFI.

French media executives also voiced concern about their content being associated with fake information.

They said talks are needed with OpenAI and other generative AI groups.

"Media must be remunerated fairly. Our wish is to obtain licensing and payment agreements," said Bertrand Gie, director of the news division at newspaper Le Figaro and president of the Group of Online Services Publishers.

'Maintain public trust'

US news agency Associated Press reached an agreement with OpenAI in July authorizing the startup to tap its archives dating back to 1985 in exchange for access to its technology and its AI expertise.

OpenAI has also committed \$5 million to back the expansion of the American Journalism Project, an organization that supports local media.

It also offered the non-profit up to \$5 million in credits to help

organizations assess and deploy AI technologies.

A consortium of news outlets, including AFP, the Associated Press and Gannett/U.S. Today, issued an open letter earlier in August saying AI firms must ask for permission before using copyrighted text and images to generate content.

The organizations said that, while they support the responsible deployment of generative AI technology, "a [legal framework](#) must be developed to protect the content that powers AI applications as well as maintain public trust in the media that promotes facts and fuels our democracies."

© 2023 AFP

Citation: Fearing 'pillaging', news outlets block an OpenAI bot (2023, August 30) retrieved 5 May 2024 from <https://techxplore.com/news/2023-08-pillaging-news-outlets-block-openai.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.