AI tools are already supercharging online harassment and hate campaigns. Can they be stopped?

Here are several problems resulting from A.I. tools that experts discovered on 4chan — and what regulators and technology companies are doing about them.

Artificial images and A.I. pornography

A.I. tools like Dall-E and Midjourney generate novel images from simple text descriptions. But a new wave of A.I. image generators are made for the purpose of creating fake pornography, including removing clothes from existing images.

Follow the latest news and policy debates on sustainable agriculture, biomedicine, and other 'disruptive' innovations. Subscribe to our newsletter.
SIGN UP

Illinois expanded its law governing revenge pornography to allow targets of nonconsensual pornography made by A.I. systems to sue creators or distributors. California, Virginia and New York have also passed laws banning the distribution or creation of A.I.-generated pornography without consent.

Cloning voices

Late last year, ElevenLabs, an A.I. company, released a tool that could create a convincing digital replica of someone's voice saying anything typed into the program.

Almost as soon as the tool went live, users on 4chan circulated clips of a fake Emma Watson, the British actor, reading Adolf Hitler's manifesto, "Mein Kampf."

…

ElevenLabs rushed to impose limits, including requiring users to pay before they could gain access to voice-cloning tools. But the changes did not seem to slow the spread of A.I.-created voices, experts said.

This is an excerpt. Read the full article here