



Law enforcement cracking down on creators of AI-generated child sexual abuse images

Nation Oct 25, 2024 10:42 AM EST

WASHINGTON (AP) — A child psychiatrist who altered a first-day-of-school photo he saw on Facebook to make a group of girls appear nude. A U.S. Army soldier accused of creating images depicting children he knew being sexually abused. A software engineer charged with generating hyper-realistic sexually explicit images of children.

Law enforcement agencies across the U.S. are cracking down on a troubling spread of child sexual abuse imagery created through artificial intelligence technology — from manipulated photos of real children to graphic depictions of computer-generated kids. Justice Department officials say they're aggressively going after offenders who exploit AI tools, while states are racing to ensure people generating "deepfakes" and other harmful imagery of kids can be prosecuted under their laws.

READ MORE: New AI rules for national security agencies balance tech's promise with protection against risks

"We've got to signal early and often that it is a crime, that it will be investigated and prosecuted when the evidence supports it," Steven Grocki, who leads the Justice Department's Child Exploitation and Obscenity Section, said in an interview with The Associated Press. "And if you're sitting there thinking otherwise, you fundamentally are wrong. And it's only a matter of time before somebody holds you accountable."

The Justice Department says existing federal laws clearly apply to such content, and recently brought what's believed to be **the first federal case involving purely AI-generated imagery** — meaning the children depicted are not real but virtual. In another case, federal authorities in August arrested a U.S. soldier stationed in Alaska accused of running innocent pictures of real children he knew through an AI chatbot to make the images sexually explicit.

Trying to catch up to technology

The prosecutions come as child advocates are urgently working to curb the misuse of technology to prevent a flood of disturbing images officials fear could make it harder to rescue real victims. Law enforcement officials worry investigators will waste time and resources trying to identify and track down exploited children who don't really exist.

Lawmakers, meanwhile, are passing a flurry of legislation to ensure local prosecutors can bring charges under state laws for AI-generated "deepfakes" and other sexually explicit images of kids. Governors in more than a dozen states have signed laws this year cracking down on digitally created or altered child sexual abuse imagery, according to a review by The National Center for Missing & Exploited Children.

"We're playing catch-up as law enforcement to a technology that, frankly, is moving far faster than we are," said Ventura County, California District Attorney Erik Nasarenko.

Nasarenko pushed legislation signed last month by Gov. Gavin Newsom which makes clear that AI-generated child sexual abuse material is illegal under California law. Nasarenko said his office could not prosecute eight cases involving AI-generated content between last December and mid-September because California's law had required prosecutors to prove the imagery depicted a real child.

AI-generated child sexual abuse images can be used to groom children, law enforcement officials say. And even if they aren't physically abused, kids can be deeply impacted when their image is morphed to appear sexually explicit.

"I felt like a part of me had been taken away. Even though I was not physically violated," said 17-year-old Kaylin Hayman, who starred on the Disney Channel show "Just Roll with It" and helped push the California bill after she became a victim of "deepfake" imagery.

Hayman testified last year at the federal trial of the man who digitally superimposed her face and those of other child actors onto bodies performing sex acts. He was sentenced in May to more than 14 years in prison.

Open-source AI-models that users can download on their computers are known to be favored by offenders, who can further train or modify the tools to churn out explicit depictions of children, experts say. Abusers trade tips in dark web communities about how to manipulate AI tools to create such content, officials say.

A report last year by the Stanford Internet Observatory found that a research dataset that was the source for leading AI image-makers such as Stable Diffusion contained links to sexually explicit images of kids, contributing to the ease with which some tools have been able to produce harmful imagery. The dataset was taken down, and researchers later said they deleted more than 2,000 weblinks to suspected child sexual abuse imagery from it.

Top technology companies, including Google, OpenAI and Stability AI, have agreed to work with anti-child sexual abuse organization Thorn to combat the spread of child sexual abuse images.

But experts say more should have been done at the outset to prevent misuse before the technology became widely available. And steps companies are taking now to make it harder to abuse future versions of AI tools "will do little to prevent" offenders from running older versions of models on their computer "without detection," a Justice Department prosecutor noted in recent court papers.

"Time was not spent on making the products safe, as opposed to efficient, and it's very hard to do after the fact — as we've seen," said David Thiel, the Stanford Internet Observatory's chief technologist.

AI images get more realistic

The National Center for Missing & Exploited Children's CyberTipline last year received about 4,700 reports of content involving AI technology — a small fraction of the more than 36 million total reports of suspected child sexual exploitation. By October of this year, the group was fielding about 450 reports per month of AI-involved content, said Yiota Souras, the group's chief legal officer.

Those numbers may be an undercount, however, as the images are so realistic it's often difficult to tell whether they were AI-generated, experts say.

"Investigators are spending hours just trying to determine if an image actually depicts a real minor or if it's AI-generated," said Rikole Kelly, deputy Ventura County district attorney, who helped write the California bill. "It used to be that there were some really clear indicators ... with the advances in AI technology, that's just not the case anymore."

Justice Department officials say they already have the tools under federal law to go after offenders for such imagery.

The U.S. Supreme Court in 2002 struck down a federal ban on virtual child sexual abuse material. But a federal law signed the following year bans the production of visual depictions, including drawings, of children engaged in sexually explicit conduct that are deemed "obscene." That law, which the Justice Department says has been used in the past to charge cartoon imagery of child sexual abuse, specifically notes there's no requirement "that the minor depicted actually exist."

The Justice Department brought that charge in May against a Wisconsin software engineer accused of using AI tool Stable Diffusion to create photorealistic images of children engaged in sexually explicit conduct, and was caught after he sent some to a 15-year-old boy through a direct message on Instagram, authorities say. The man's lawyer, who is pushing to dismiss the charges on First Amendment grounds, declined further comment on the allegations in an email to the AP.

A spokesperson for Stability AI said that man is accused of using an earlier version of the tool that was released by another company, Runway ML. Stability AI says that it has "invested in proactive features to prevent the misuse of AI for the production of harmful content" since taking over the exclusive development of the models. A spokesperson for Runway ML didn't immediately respond to a request for comment from the AP.

In cases involving "deepfakes," when a real child's photo has been digitally altered to make them sexually explicit, the Justice Department is bringing charges under the federal "child pornography" law. In one case, a North Carolina child psychiatrist who used an AI application to digitally "undress" girls posing on the first day of school in a decades-old photo shared on Facebook was convicted of federal charges last year.

"These laws exist. They will be used. We have the will. We have the resources," Grocki said. "This is not going to be a low priority that we ignore because there's not an actual child involved."

By — Alanna Durkin Richer, Associated Press