

# AI-created child sexual abuse images ‘threaten to overwhelm internet’

Publication Date: 2023-10-25

Author: Dan Milmo

Section: Technology

Tags: Artificial intelligence (AI), Child protection, Internet, Children, Computing, news

Article URL: <https://www.theguardian.com/technology/2023/oct/25/ai-created-child-sexual-abuse-images-threaten-overwhelm-internet>



The “worst nightmares” about artificial intelligence-generated child sexual abuse images are coming true and threaten to overwhelm the internet, a safety watchdog has warned. The Internet Watch Foundation (IWF) said it had found nearly 3,000 AI-made abuse images that broke UK law. The UK-based organisation said existing images of real-life abuse victims were being built into AI models, which then produce new depictions of them. It added that the technology was also being used to create images of celebrities who have been “de-aged” and then depicted as children in sexual abuse scenarios. Other examples of child sexual abuse material (CSAM) included using AI tools to “nudify” pictures of clothed children found online. The IWF had warned in the summer that evidence of AI-made abuse was starting to emerge but said its latest report had shown an acceleration in use of the technology. Susie Hargreaves, the chief executive of the IWF, said the watchdog’s “worst nightmares have come true”. “Earlier this year, we warned AI imagery could soon become indistinguishable from real pictures of children suffering sexual abuse, and that we could start to see this imagery proliferating in much greater numbers. We have now passed that point,” she said. “Chillingly, we are seeing criminals deliberately training their AI on real victims’ images who have already suffered abuse. Children who have been raped in the past are now being incorporated into new scenarios because someone, somewhere, wants to see it.” The IWF said it had also seen evidence of AI-generated images being sold online. Its latest findings were based on a month-long investigation into a child abuse forum on the dark web, a section of the internet that can only be accessed with a specialist browser. It investigated 11,108 images on the forum, with 2,978 of them breaking UK law by depicting child sexual abuse. AI-generated CSAM is illegal under the Protection of Children Act 1978, which criminalises the taking, distribution and possession of an “indecent photograph or pseudo photograph” of a child. The IWF said the vast majority of the illegal material it had found was in breach of the Protection of Children Act, with more than one in five of those images classified as category A, the most serious kind of content, which can depict rape and sexual torture. The Coroners and Justice Act 2009 also criminalises non-photographic prohibited images of a child, such as cartoons or drawings. The IWF fears that a tide of AI-generated CSAM will distract law enforcement agencies from detecting real abuse and helping victims. “If we don’t get a grip on this threat, this material threatens to overwhelm the internet,” said Hargreaves. Dan Sexton, the chief technology officer at the IWF, said the image-generating tool Stable Diffusion – a publicly available AI model that can be adjusted to help produce CSAM – was the only AI product being discussed on the forum. “We have seen discussions around the creation of content using Stable Diffusion, which is openly available software.” Stability AI, the UK company behind Stable Diffusion, has said it “prohibits any misuse for illegal or immoral

purposes across our platforms, and our policies are clear that this includes CSAM". The government has said AI-generated CSAM will be covered by the online safety bill, due to become law imminently, and that social media companies would be required to prevent it from appearing on their platforms.