KNOW2PROTECT®

TIPS2PROTECT AGAINST AI AND DEEPFAKES

KNOW THE THREATS

Artificial Intelligence (AI) technology is rapidly evolving — and with it, there are new threats to online safety. For example, online predators use generative AI (GAI) to create morphed or fake child sexual abuse material (CSAM) and then sometimes engage in blackmail schemes targeting kids. This technology is also being used by kids and teens to target their peers.

GAI CSAM IS STILL CSAM

Let's be clear: **GAI child sexual abuse imagery is still CSAM.** It is illegal, emotionally and psychologically harmful and deeply traumatic. Understanding this evolving threat is the first step in protecting kids and teens.

Creating a deepfake used to require technical expertise and dedicated tools; today, anyone with an internet connection can use AI to turn a normal photo into a nude image[1].

NEW GAI TECHNOLOGIES AND ASSOCIATED RISKS

Understanding the latest technologies and their associated risks is crucial to prevent, protect and empower victims and survivors.

AI technology allows a user to create new images, videos, audio and text based on user requests or prompts[2]. This could include any visual depiction, such as computer-generated images, that appear to depict a minor engaging in sexually explicit conduct.

Deepfake technology uses AI to create realistic images or videos of people doing or saying things they never did. The term "deepfake" combines "deep learning" — a type of AI used to generate the content — and "fake."

Nudifying Services allow an individual to upload an innocent clothed image of a person to create a virtual depiction of that person unclothed.

Reports show these services are promoted as tools on nudifying websites and forums — sometimes even disguised as "pedophile guides" — that coach predators on how to exploit children using AI. Online predators use nudifying services to upload a fully clothed image of a child or teen and use AI to generate a fake, sexualized version of that image. These fake images are then used for blackmail or other forms of exploitation.

[1] commonsensemedia.org/kids-action/articles/deepfakes-can-be-a-crime-teaching-ai-literacy-can-prevent-it
[2] missingkids.org/theissues/generative-ai

TIPS2PROTECT AGAINST AI AND DEEPFAKES

THE RISE IS ALARMING:

In 2024, the National Center for Missing and Exploited Children (NCMEC) received **67,000 reports involving GAI**, a 1,325% increase from just 4,700 in 2023. Predators use GAI to create fake but realistic CSAM and coerce children into compliance or further exploitation[3].

According to THORN, **1 in 10 minors report that they know of cases where their friends and classmates have created "deepfake nudes"** of other kids using generative AI tools[4].

In a recent study, THORN surveyed 1,200 young people aged 13 to 20 regarding the awareness of deepfake nudes, lived experiences and involvement in creating such content. The study found that **1 in 6 respondents indicated they personally knew someone** who had encountered at least one of the following experiences:

- Had a deepfake nude made of them by someone else
- Discovered a deepfake nude of them was being (re)shared
- Created a deepfake nude of themselves
- Created a deepfake nude of a minor
- Reshared a deepfake nude of a minor[5]

ACCORDING TO NCMEC, THESE ARE WAYS GAI IS USED TO EXPLOIT KIDS AND TEENS:

Text to Chat: Entering text to get a chat model to pretend it is a child and engage in a sexually explicit chat.

Text to Text: Entering text to generate guides, tutorials or suggestions on how to groom and harm kids and teens both on and offline.

Text to Image: Entering text prompts to generate CSAM imagery or to alter previously uploaded files to make them sexually explicit.

Image to Image:

- Uploading known CSAM imagery and attempting to generate entirely new CSAM based on these known images.
- Uploading known CSAM and altering or adding new abusive elements to existing images.
- Uploading innocent images of a child and generating sexually explicit or exploitative images of the child for personal gratification or sharing and to use to perpetrate financial sextortion against a child[6].

Tip2Protect: Encourage kids and teens to think critically about the content they encounter online and teach them to question the authenticity of images and videos, understanding that AI can fabricate realistic looking — but fake — content.

[3] missingkids.org/gethelpnow/cybertipline/cybertiplinedata
[4] thorn.org/press-releases/report-1-in-10-minors-say-peers-have-used-ai-to-generate-nudes-of-other-kids
[5] thorn.org/blog/deepfake-nudes-are-a-harmful-reality-for-youth-new-research-from-thorn
[6] missingkids.org/content/dam/missingkids/pdfs/final-written-testimony-john-shehan-house-oversight-subcommittee-hearing.pdf

TIPS2PROTECT AGAINST AI AND DEEPFAKES

LEGAL AND ETHICAL IMPLICATIONS OF GAI CSAM

Any sexually explicit content of a minor is considered illegal. Creating, sharing or even possessing this type of content on a device is illegal and can have life-long consequences. Both adults and minors can be held accountable when they commit crimes online. In addition to possible federal criminal violations, many states have specific criminal statutes regarding sextortion and online harassment of a minor that may have a range of legal ramification such as a verbal warning from a juvenile or family court judge, court ordered community service or counseling, to more serious consequences including incarceration. Check state and local government resources in your area for more information.

Many young people may not realize that "just messing around" with AI tools to generate fake nudes of classmates or others can cross legal lines, with life-altering consequences[7].

HOW2AVOID GAI CSAM

- Ensure your and your child's profiles are set to private and that all images are for friends-only and not publicly accessible.
- Use parental control settings. View **Connect2Protect: Parental Controls** for more information.
- Maintain open lines of communication with children about their online activities.
- Encourage them to come to you with any concerns or questions they may have about AI, deepfakes and online safety.

HOW2RESPOND

Take immediate action if you know or suspect a child or teen in your care is a victim of online CSEA.

- Remain calm.
- Consider putting the device in airplane mode and ensure Bluetooth and Wi-Fi are turned off so evidence cannot be remotely erased, deleted or altered by the offender.
- Save all communication — do not delete.
- Contact law enforcement and ensure they can access the device.
- Do not send images from your child's device to any other device. It is illegal to disseminate CSAM.
- Report the incident to law enforcement.

[7] commonsensemedia.org/kids-action/articles/deepfakes-can-be-a-crime-teaching-ai-literacy-can-prevent-it

TIPS2PROTECT AGAINST AI AND DEEPFAKES

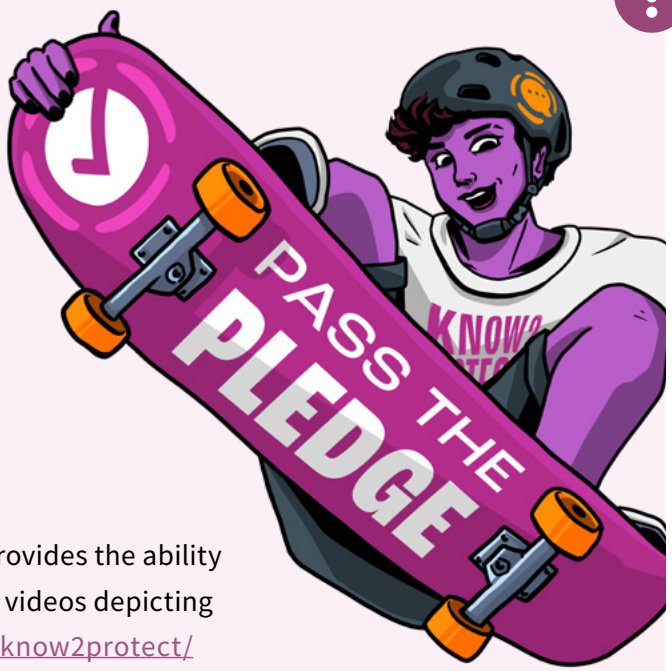
HOW2REPORT

You can submit a report in one of the following ways:

- Call the Know2Protect Tipline at **833-591-KNOW (5669)**. All information received via the Tipline will be reviewed by appropriate personnel and referred to Homeland Security Investigations field offices for potential investigation.
- Contact your local, state or tribal law enforcement officials directly. In an emergency, call **911**.
- Submit a **CyberTipline** report with the National Center for Missing and Exploited Children at report.cybertip.org.

Visit dhs.gov/know2protect/how-to-report to learn more.

Resources



- **Survivor Support Page:** Know2Protect shares resources and tips on how to help victims of online CSEA. Visit dhs.gov/know2protect/survivor-support to learn more.
- **Take It Down:** The National Center for Missing and Exploited Children offers a free service to help remove online nude, partially nude or sexually explicit photos and videos taken before age 18. Visit takeitdown.ncmec.org to learn more.
- **Generative AI Informational Bulletin:** AI provides the ability to produce exponentially more pictures and videos depicting child sexual abuse. Learn more at dhs.gov/know2protect/publication/generative-ai-csea.
- **THORN:** Check out "Navigating Deepfake Nudes: A Guide to Talking to Your Child About Digital Safety" at info.thorn.org/navigating-deepfake-nudes#get-the-guide and "Deepfake Nudes & Young People" at thorn.org/research/library/deepfake-nudes-and-young-people.