

What Parents & Educators Need to Know about DEEPFAKES

A deepfake is a piece of artificial intelligence (AI) trickery designed to make you believe that a person has said or done something they haven't. AI models are able to use audio, video and photos of a subject to produce misleading mimics. Initially unsophisticated and easy to spot, deepfakes are improving every day. And with the spread of user-friendly AI, it's easier than ever for children to make their own, or be the unsuspecting target of them.

WHAT ARE THE RISKS?

A BULLYING RISK

AI has got easier to use and more widespread, making it a "gift" to bullies. Children can generate fake audio of peers saying damaging things, or pictures of them doing something embarrassing. More disturbingly, Metro recently highlighted cases where fake sexualised images of teens are spread through social media. Mainstream apps don't allow this usage, but with AI becoming widespread, easy-to-use sites without guardrails keep emerging for darker uses.

FAKE FAMILY SCAMS

With a few seconds of audio — from an Instagram reel or TikTok — scammers can clone a person's voice. They can then use this to scam relatives with phone calls asking for money or assistance. This is a widespread tactic for scammers, with the FBI reporting 22,000 such cases in 2025, defrauding victims out of \$893 million that year alone.

FAKED CELEBRITY ENDORSEMENTS

With the technology getting easier to use and public figures having lots of clips for training, celebrities are a frequent target. According to McAfee, the most deep faked celebrities are Taylor Swift, Scarlett Johansson, Jenna Ortega, Sydney Sweeney and Sabrina Carpenter, and their AI dopelgangers are usually pushing scams. Even YouTuber MrBeast has been deepfaked by scammers on TikTok. It can be used for propaganda too, making politicians appear to do or say things to improve or harm their reputation.

EXTORTION RISK

You might not be the one being tricked by a deepfake — you could also be the star of it. Blackmail videos are as old as film itself, but with deepfake technology you don't have to have done anything wrong to be a target. A deepfake video showing a person in an uncompromising position could make the victim pay up, even when the content never happened.

Advice for Parents & Educators

KEEP PROFILES PRIVATE

To make a convincing deepfake, you need a supply of images, audio or video of the subject. That means that most non-celebrity victims are targeted via social media where there's a ready-made repository of media to train a deepfake from. Most social media companies have privacy settings to prevent unauthorised access. Enable these to keep strangers and bullies away.

USE RESEARCH AND COMMON SENSE

Even perfect deep fakes rely on trust, so think critically. If it's someone you know claiming to need help, then ask them to prove they're who they say they are with an inside joke. For public figures, try Googling the person along with a bit of the contentious speech in quotation marks. Also question why you're seeing the video. If it's a paid promotion, then ask yourself who's paying and why. If someone is sharing the clip, do they dislike the target?

KNOW WHAT TO LOOK FOR

While deepfakes are getting harder to spot, there are still telltale signs to look out for. Deepfake creation can leave blurry edges and flickering textures, especially around hair and teeth. It can also look too sleek, with models often trained on professional video. Make sure the mouth shapes match the spoken words, too. As for audio, listen out for mispronounced phrases and a slightly unusual, robotic rhythm of speech.

TALK TO YOUR CHILDREN

As well as emphasising the need for privacy and to not trust everything they see online, it's important to ensure children realise why deepfakes are problematic in the first place. With deepfake technology getting more accessible and easy to use, children can start using the technology to make deepfakes of schoolmates, teachers or other adults. Ensure they're aware of the serious trouble the technology can cause, and the ethical problems of manipulating someone's image without consent.

Meet Our Expert

Alan Martin is an experienced technology journalist who has written for the likes of The Guardian, The Telegraph, The i, Wired, TechRadar, Tom's Guide, The Evening Standard and The New Statesman.



The National College