

What Parents & Educators Need to Know about

DEEPIFAKES

A "deepfake" generates photos, videos and audio via AI models trained on the subject to mimic their look and sound, making it appear they said words or committed acts that never occurred. While initially unsophisticated and easy to identify, deepfakes are increasingly more convincing as the technology behind them continues to improve at a solid pace.

WHAT ARE THE RISKS?

FAKE NEWS

With the technology getting easier to use and public figures having lots of photos and clips to train AI from, deepfakes are often used to spread fake news, propaganda and scams. For example, in 2023, an audio deepfake of Labour leader Sir Kier Starmer designed to make him sound like he was berating an aid was pushed by propagandists on X.

SCAMS

With the right training material, deepfakes can be made from anyone's voice and appearance, and that means scammers can take advantage. In 2019, criminals faked the voice of an energy company's CEO to steal €220,000 from the company, and private citizens have also been targeted. In 2023, an Arizona mother was the subject of a fake kidnapping phone call, using deepfaked audio of her daughter pleading for her life.

EXTORTION

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 could become a target without even doing anything. A deepfake video showing a person in an uncompromising position could make the victim pay up, even if the scenes depicted are entirely fictitious.

EXPLICIT MATERIAL

A 2019 report claimed that 96% of deepfake videos were pornographic in nature. In other words, users were realistically superimposing the faces of other people onto the bodies of pornographic actors. This is creepy in its own right but can also lead to the kind of extortion highlighted above if the video looks convincing enough. Indeed, the FBI warned about this phenomenon in 2023.

Advice for Parents & Educators

KEEP THOSE PROFILES PRIVATE

To make a convincing deepfake, you need a supply of images, audio, or video of the subject. This 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 away.

KNOW THE SIGNS

While deepfake technology has come on in leaps and bounds in recent years, there are still telltale signs to look and listen out for. The process of creating deepfakes can leave blurry edges and flickering textures, especially around hair and teeth. If the mouth doesn't seem to be moving in relation to the words spoken, that's another telltale sign. For audio, listen out for mispronounced words and a slightly unusual, robotic rhythm of speech.

USE RESEARCH AND COMMON SENSE

More sophisticated deepfakes won't have obvious signs but can still be spotted with critical thinking and investigation. If you see a friend or family member in what you think is a deepfake, you could contact them directly and clear it up with them. For public figures, look them up to discover if what you saw was legitimate. Also employ critical thinking – think about why this clip could have been made.

INFORM AND EMPOWER CHILDREN

As well as emphasising the need for privacy and not trusting 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 easier to use, children can start using the technology to make deepfakes of schoolmates, teachers, or other adults. Make sure they're aware of the harm this technology can cause and the ethics of manipulating someone's image in this way.

Meet Our Expert

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

