At National Online Safety, we believe in empowering parents, carers and trusted adults with the information to hold an informed conversation about online safety with their children, should they feel it is needed. This guide focuses on one issue of many which we believe trusted adults should be aware of. Please visit www.nationalonlinesafety.com for further guides, hints and tips for adults.

What Parents & Carers Need to Know about DEEPFAKES

A deepfake is an extremely convincing piece of media that is created using artificial intelligence (AI), based on pictures and recordings of the subject. The name comes from the deep learning approach to AI needed to generate them and the fact that they're used to create fake content. Deepfakes can be made as videos, static images and audio – where a person's voice is accurately mimicked to make it seem as though they have said something which, in reality, they have not.

Fake News

The most obvious use to date has been to spread fake news; a politician or well-known figure can be undermined by someone putting damaging words into their mouths. For example, in 2018, a video of Donald Trump advising the people of Belgium on climate change was revealed to be a deepfake, while in 2021 a convincing parody account supposedly belonging to Tom Cruise went viral on TikTok.

Fraudulent Use

Deepfakes are mainly associated with video clips, but audio versions *do* exist and are in fact much simpler to create. These carry obvious criminal potential: to commit financial fraud, for instance. It is possible for someone to be fooled that a trusted person gave specific instructions or authorisation during a phone call, while a fake voice could be generated to bypass phone authentication at some banks.

Potential for Extortion

A sophisticated deepfake video could be used for extortion, even if someone has not done anything to be blackmailed over. A deepfake could believably show a person in a compromising position, and – even if that individual was never actually present – the mere threat of the video being circulated on the internet could be enough to coerce them into paying a ransom.

Pornography

A 2019 study by research company Sensity AI found that 95% of deepfake videos online took the form of pornography. This involves realistically superimposing the faces of other people (usually celebrities or public figures, but members of the public *have* been victims, too) onto the bodies of actors in porn movies. Not only could this be used as a type of extortion, but it is also an appalling invasion of privacy.

Advice for Parents & Carers

Keep Profiles Private

Ensure that you and your family maintain a relatively limited public presence on social media. By enabling privacy restrictions, you can help to prevent scammers from having access to images, video and audio clips from which they could easily copy your voice and facial likeness.

Search for Other Evidence

If the video is supposedly of a prominent public figure, then it is probable that whatever they said on the clip would have also been recorded by someone else (especially if it was a political speech). Google the person's name with a few words of what ever they said in quotation marks. If no matching search results come up – and the video has not been covered by any news outlets – then it is likely you are looking at a deepfake.

Trust Your Instincts

Ask yourself whether the content of a video seems plausible. If the person in the clip is acting out of character (for example, using unexpectedly sensational or divisive language), it could be a sign of a deepfake. Investigate the source: newly created accounts or websites are often suspect. Check if any previous posts from that account display an ideology that would want to either glorify or discredit the speaker.

Look at Details

The process that creates deepfakes sometimes leaves obvious traces. Look out for blurry edges and flickering on faces – especially on textures that are difficult to replicate, such as hair strands. Oddly rendered teeth or a lack of blinking may also be clues. Pay close attention to whether the person's mouth movements directly correspond to the words spoken: many amateur deepfakes fall at this hurdle.

Meet Our Expert

Alan Martin is an experienced technology journalist and the former deputy editor of technology and internet culture website Alphr. Now freelance, he has contributed articles to publications including the *New Statesman*, CNET, the *Evening Standard*, *Wired*, Rock Paper Shotgun, Gizmodo, Pocket Gamer, *Stuff*, *T3*, *PC Pro*, *Macworld*, TechRadar and Trusted Reviews.

National Online Safety®
#WakeUpWednesday

SOURCES: https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them, https://fakeinnews.org/background