Safeguarding Newsletter – October 2024

Spotlight on Deepfakes

What are Deepfakes?

A deepfake is an image, video, sound, voice, or GIF which has been manipulated by a computer to superimpose someone's face, body or voice onto something else.
This could be done with or without the person's consent.

Deepfakes have become popular due to the accessibility of apps. This means anyone can make Deepfakes without sophisticated technology.

Facts about Deepfakes

- In 2019, AI company Sensity warned that **96%** of deepfakes they surveyed online were non-consensual 'pornographic' materials.
- From 2019 to 2021, the number of deepfakes online grew from roughly 14,000 to 145,000.

The harmful side of Deepfakes

Cyberbullying – they can be used to deliberately mock, taunt, or inflict public embarrassment on victims.

Extortion and Exploitation - they can be used to create incriminating, embarrassing, or suggestive material. They can be very realistic and convincing. These images can then be used to extort money or additional 'real' images.

Homophobic Abuse – An image is created of a person in 'gay pornography' and be used to 'out' the person.

Image-based Sexual Abuse- Images of children can be used to generate sexualized Deepfakes. Some software can remove clothing from pictures

How to spot a Deepfake

Deepfakes can vary in their quality. Some will be quite obviously fake. For others, it can be tricky to spot whether they are real or not.

- *Glitches* – There are typically signs if you look closely at the video itself. Is there rippling, pixelation or blurring around key facial features, like the neck, eyes, or mouth? This may become more obvious when a person moves, blinks, or turns their head or body.

- *Audio* – There may be an indication that lip movements do not match what you are hearing. Look closely for natural mouth movements.

What you can do about Deepfakes

- Talk about deepfakes to your child/ children and the importance of image consent. Ensure they know why they should ask someone before using an image of them to create a deepfake or manipulated picture.
- Check privacy settings on social media.
- Speak to your child/children about their safety and privacy settings online. You should also check that they limit public access to their social media images.
- Report the images:
    - CEOP Safety Centre
    - Report Remove | Childline
    - Take It Down (ncmec.org)
    - Deepfakes: what you can do | Police.uk (www.police.uk)

Further information

Information from: A Beginners Guide to Deepfakes - Safer Schools (oursaferschools.co.uk)

Support and Advice for Parents and Carers on Synthetic Media & Deepfakes (swgfl.org.uk)

Navigating the Digital Realm: A Parent's Guide to Discussing AI Deepfakes with Kids - Natterhub

What is a deepfake? | Internet Matters