

Online safety trends: AI risks and ‘sextortion’

What is AI?

Artificial intelligence (AI) is the use of computer systems to solve problems and make decisions. It’s already a part of everyday life – you’ve probably come across it in the form of personalised suggestions on social media, shopping sites or route-planning apps.

Generative AI takes a written prompt and runs it through an algorithm to generate new, ‘natural’-seeming content. Tools include:

- Chatbots such as ChatGPT, Google Gemini and GrammarlyGO, which generate text
- Text-to-image programs like DALL-E and Midjourney, which create images

AI technology is developing rapidly, and these tools will only become more sophisticated over time. For example, they’ll be able to create more convincing images or videos.

How does AI pose a safeguarding risk?

Rather than being its own safeguarding issue, AI can impact other safeguarding issues:

- **Hacking and scams** – text generation tools can write convincing emails and text messages to trick pupils into giving malicious actors access to their accounts
 - **AI-generated child sexual abuse images** – some text-to-image tools could be used to create child sexual exploitation material for sexual gratification or as a means of bullying another pupil
 - **‘Deepfake’ pornography** – super-imposing a person’s face into pornographic videos for sexual gratification or to humiliate the person being put in the images. AI technology is used to alter the person’s facial expressions to make the video look more convincing
 - **‘Catfishing’ and sextortion** – criminals can use AI-generated profile pictures to appear younger than they are to befriend and groom children and young people, and then solicit information and/or images from them (e.g. nude or semi-nude photos)
 - **Fake news and misinformation** – text-to-image tools can be used to create convincing fake photos of world events, which could be used to promote certain beliefs (including hateful ones)
-

What is ‘sextortion’?

Financially motivated sexual extortion (often referred to as ‘sextortion’) is a type of online blackmail. An adult (or group of adults) threatens to release nude or semi-nude images of a child unless they pay money or do something else to benefit them.

Sextortion is often carried out by offenders in an organised crime group overseas, and is motivated by profit.

Sometimes adults pose as children to make contact. They might

- Groom or coerce the child into sending nudes or semi-nudes and financially blackmail them
- Use images that have been stolen from the child, taken through hacking their account
- Use digitally manipulated images, including AI-generated images, of the child

When it comes to safeguarding issues involving the sharing of nudes and/or semi-nudes:

Do:

- ✓ Talk to our designated safeguarding lead (DSL), or deputy, as soon as possible
- ✓ Explain to the child that you need to report it to our DSL (or deputy)
- ✓ Reassure the child that they will receive support and help

Don't:

- ✗ View, share, download or print the imagery – if you view the imagery by accident (e.g. if a child shows it to you before you can ask them not to), report this to our DSL or deputy and seek support
- ✗ Ask the pupil to disclose any information about the incident – this is the responsibility of our DSL
- ✗ Share information about the incident with other members of staff, the pupils it involves, or their parents/carers
- ✗ Promise confidentiality
- ✗ Say or do anything to blame or shame any of the children involved
- ✗ Delete the image or ask the young person to delete it

Signs to look out for

If a child is facing a safeguarding issue online, they might:

- Spend more time online, or more time offline. This might be reported by their parents
- Complain of being tired because they were online all night, or have their phone going off a lot
- Have stronger emotional responses or outburst to being online – for example, the child may get unusually angry, upset or distant after checking their phone or using their computer/tablet
- Be secretive about their use of the internet or a device – they may refuse to hand their phone in if it's part of school policy, or refuse to tell you what they get up to online
- Show more general signs of sexual abuse

If a child tells you that they use their device unsupervised – for example, they play on their iPad when they go to bed – this could be a red flag.

Sources

This factsheet was produced by [The Key Safeguarding](https://thekeysupport.com/safeguarding): thekeysupport.com/safeguarding

- With thanks to education expert Aaron King for his help with this content
- Department for Science, Innovation and Technology and UK Council for Internet Safety, [Sharing nudes and semi-nudes: advice for education settings working with children and young people](#).