STAFF FACTSHEET: JANUARY 2025

AI risks: staff factsheet

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
- Text-to-video programs, which create videos

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** – superimposing 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). They can then use this to extort children into giving them money
- **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)

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 outbursts when they are 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: **thekeysupport.com/safeguarding**

- With thanks to education expert Aaron King for his help with this content