



Online safety at home: Generative Artificial Intelligence (AI)

Artificial Intelligence (AI) describes technologies that can perform tasks which usually need human intelligence. This could include learning, reasoning, problem-solving, perception, decision-making and speech recognition. Some of the most well-known are **ChatGPT** and **DALL-E**

Whilst AI can have some great benefits, there are also considerations all parents/carers must have around keeping young people safe. These are outlined below.

Deepfakes

Deepfakes can be used to spread misinformation, often in the form of an AI generated celebrity supporting or endorsing something that could be dangerous or controversial.

Exposure to illegal content

AI can be used to create Child Sexual Abuse Material (CSAM) that can be shared between young people. Some AI tools allow learners to create a sexualised image of another child or young person.

Chatbots

[Research by the University of Cambridge](#) has found that some chatbots have been found to engage in sexually explicit dialogue or provide inappropriate advice

Online Scams

Learners may be tricked by AI-generated messages, chatbots or fake profiles into sharing personal information. One of the fastest growing scams is where criminals use AI technology to replicate a person's voice so they can trick victims into handing over money or personal details by pretending to be someone they know and trust.

Peer Harassment and Abuse

Some AI apps allow users to 'nudify' photos, altering clothed images to appear nude. Some young people have used AI to superimpose a child or young person's face onto explicit adult content. It is especially important for us as adults to consider who can access images of our own children online.

More information?

Check out the [Generative AI guide for Parents and Carers](#) on Hwb by clicking on the link above.

Report any safeguarding concerns linked to AI by speaking to us at school or using our [Online Reporting platform](#)

Steps you can take to help keep your child safer online

Parental controls: Where possible, utilise the parental control features available within AI tools or associated platforms. Activate filters or safe modes that limit the types of content the AI can generate. This can help prevent exposure to inappropriate material and ensure that the AI's outputs align with your child's age and maturity level.

Educate on risk: Help your child understand the potential risks associated with gen AI and how to navigate them safely. Make sure your child understands that AI-generated content is not always accurate and should be verified before use. Encourage them to cross-check information and to think critically about the content they generate or encounter.

Reporting and blocking: Knowing how to report and block inappropriate content or interactions is key to maintaining a safe environment. Familiarise yourself and your child with the reporting mechanisms available within AI tools. Encourage them to report any content that makes them uncomfortable, to either you, another trusted adult or to the platform itself. Some helpful links are the Internet Watch Foundation (IWF) and Childline's [Report Remove tool](#) and [NCMEC's Take It Down tool](#)