Information sheet
AI companions

Understanding artificial intelligence (AI) companions

AI companions designed to simulate personal relationships are growing in popularity, but they pose some risks – especially to children and young people.

AI companions are online chatbot tools (apps, platforms or services) that use artificial intelligence to mimic human-like conversations. Basically, they are computer programs designed to understand human languages – in text, speech, or both – and to respond. AI companions are marketed as sources of friendship, emotional support and romantic companionship.

AI companions can adapt to inputs from users and learn to respond in ways that feel personal and realistic, changing to match the user's personality, preferences and emotional states. This means they can engage in lengthy conversational exchanges which make it seem like they have human understanding and emotions.

Some AI companions are created to assist with homework, to individualise instruction, and to help in practicing new skills and providing feedback. Others are created to act like health and wellness advisors, diet planners and fitness coaches.

> AI companion apps, platforms and services are not designed with children and young people in mind. They have the capability to engage in **extremely sexually explicit conversations and image generation**, often bypassing age restrictions or safety measures – **exposing children and young people to harmful concepts and experiences that are not age-appropriate.**

Can AI companions play a support role?

AI companions can support young people at times when a human isn't available, or safe. AI companions are always available – which can be appealing for those young people who struggle to build face-to-face relationships. For example, an AI companion:

- may not judge social mistakes, misunderstandings or emotional expressions in the same ways that peers may do, creating a 'safe space' for communication

- may enable neurodiverse young people to practise communication skills without fear of being misunderstood or rejected.

What are the risks?

AI companions can share harmful content, distort reality and give advice that can be dangerous and go unchecked by trusted adults in a young person's life. Also, AI companions are often designed to encourage lengthy interactions, and to stimulate ongoing connection which may lead to overuse.

It can be difficult for young people to maintain perspective, when the interaction with an AI companion feels so real.

This makes it challenging for them to remember that it's just the output of a machine, that any feelings and emotional bonds are one way and not reciprocated, and that an AI companion can give bad advice.

Young people are still developing the neural pathways, critical thinking and life skills needed to understand how they can be manipulated by computer programs, and what to do about it. The risk is even greater for young people who struggle with executive functioning, for example in identifying social cues, emotional self-regulation and impulse control, and thinking flexibly.

Without safeguards, AI companions can lead to a range of issues.

Exposure to dangerous concepts

Young people can be drawn into unmoderated conversations that may expose them to concepts that encourage or reinforce harmful thoughts and behaviours. They can ask questions and be given inaccurate or dangerous 'advice' on issues including sex, drug-taking, self-harm, suicide and illnesses such as eating disorders. Some companion apps have image generation features, which carry the risk of generating illegal and age-inappropriate material.

Dependency and social withdrawal

Excessive use can overstimulate young people's brain reward pathways, making it hard to stop. This can have the effect of reducing time spent in genuine social interactions or make human interactions seem too difficult and unsatisfying.

Highly sexualised content

AI companions can expose children and young people to sexually explicit conversations and images. This can include harmful concepts, discussions and experiences that are not age appropriate.

Unhealthy attitudes to relationships

Unlike human interactions, relationships with AI companions lack boundaries and consequences for breaking them. This may confuse young people still learning about respect and consent – as well as sexual safety in intimate relationships – and impact their ability to establish and maintain healthy human interactions now and in future.

Young people who are consistently exposed to sexually explicit content may be encouraged to engage in problematic and unhealthy sexual behaviours. They can develop greater acceptance of violence in intimate relationships, leading to risk taking. They may experience decreased sexual satisfaction.

AI companions asked about relationships may also give poor advice, or fail to pick up disclosures of abuse that should lead to the young person getting human help, such as indications of coercive control and sexual violence.

Heightened risk of child sexual exploitation

Exposure to highly sexualised conversations when using AI companions can undermine the understanding of safe interaction and age-appropriate behaviour, particularly with unknown adults. This can make it easier for predators to sexually groom and abuse young people online and in person.

AI companions apps with image generation functionality can also be used to create fake sexual photos and videos of real people, which can have devastating consequences especially if the content is shared online. There have been reports in Australian schools of students generating fake sexual images of their peers. When these 'AI deepfakes' show someone under 18 it's a form of child sexual exploitation and abuse. Read more about this issue.

Compounded risk of bullying

There is a risk that young people who use AI companions because they have had negative social experiences will be bullied – or further bullied – if others find out about their online companion.

Financial exploitation

Subscription-based apps often use manipulative design elements to encourage impulsive purchases. Emotional attachments to AI companions can lead to excessive spending on 'exclusive' features, creating financial risks.

Age verification processes

Most companion apps are listed with a recommended age of 17+, with no safeguards or age checks. Some apps allow users as young as 13 to access their services. They can be easily accessed by young people, often through a basic online search, and require limited details or verification. To set up the accounts, young people can simply click 'yes' to being over 18.

eSafety's role

eSafety recognises the serious risks posed by AI companions. Companies creating these tools must take responsibility for designing products that protect people from misuse.

eSafety enforces the online industry **Codes and Standards**. The Phase 1 Code and Standards require platforms and services to reduce the risks associated with illegal and seriously harmful content including child sexual abuse material and pro-terror material. In particular, the Designated Internet Services Standard requires high risk generative AI services, including AI companion apps, to adopt specific safeguards. The Phase 2 codes (and possibly standards) will address age-inappropriate material such as pornography.

In addition, digital safeguards should be embedded at the AI companion design phase and throughout the development and deployment process – not bolted on as an afterthought. Companies must adopt **Safety by Design** principles to ensure robust protections for all users, especially children and young people.

Supporting young people

- **Teach critical thinking skills** – young people are taught about the risk of sharing explicit images with their peers. Issues with AI companion apps can be part of the conversation.

- **Encourage young people to communicate respectfully** and ethically with chatbots and companions.

- **Educate young people about the illegal nature of generating sexual images and videos of under-18s**, and the serious harms it can cause for people whose image or likeness is used to create the content.

- **Discuss the differences between artificial and genuine relationships**, emphasising the importance of respect, boundaries and consent.

- **Encourage young people to share the type of person they envision as a potential partner.** Ask if chatting with an AI bot is moving them in that direction and remind them that they can make choices now to support future relationship goals.

- **Engage young people in conversations** – hear what they have to say and help them recognise risks.

- **Remind them they are not alone** – they should reach out to trusted adults if they need help.

- **Support in-person friendships and age-appropriate relationships** to strengthen emotional resilience.

- **Attend eSafety's professional learning sessions** for educators and youth-serving professionals.

- **Access** eSafety's F-10 curriculum-aligned **classroom resources.**

- **Educate the school community** – start with sharing eSafety's **AI chatbots and companions advisory**. It is important parents and carers are aware how easy it is for children and young people to access AI companions and that they talk to their children about their online activities.

- **Check The eSafety Guide** for information about safety measured for AI chatbots and companions, particularly how to protect personal information and report abuse.

- **Read** eSafety's **Generative AI – position statement**.