

# AI and Child Sexual Exploitation

## OUR POSITION:

**Bravehearts asserts that to combat risks associated with AI technology, education, legislative action, and cooperative efforts are essential**

- **Greater awareness and education.**
- **For Commonwealth, and state and territory laws to include clear definitions of AI-generated CSAM and to ensure legislation is continually reviewed.**
- **For Commonwealth, and state and territory laws to ensure that production of intimate images, including digitally created images, and sharing or threatening to share intimate images are included as crimes.**
- **That Commonwealth, and state and territory governments recognise ‘nudify apps’ as designed to sexually harass, abuse and extort victims, and to ensure these apps are banned.**
- **Technology companies must be held responsible for ensuring safety**
- **To develop a robust research agenda to ensure an informed response to risks of developing AI.**
- **Clear reporting options and support is available for those impacted by online abuse.**

## Background

Artificial intelligence (AI) systems are already having a significant effect on our lives, as they increasingly impact what we perceive, believe, and how we act. With the ongoing advancements in AI technology and the notable surge in investment, we can expect AI to grow even more powerful and impactful in the future.

All forms of technology entail both positive and negative repercussions, yet the consequences of artificial intelligence are remarkably extensive. This technology offers significant opportunities for beneficial developments, but it also presents serious challenges and elevated risks.

The AFP-led Australian Centre to Counter Child Exploitation (ACCCE) has reported a significant increase in the production of AI-generated child abuse materials (Australian Centre to Counter Child Exploitation, 2025). This includes not just adults, but a growing number of young people engaging in the creation of deepfakes for multiple reasons, including the intent to intimidate or embarrass peers. This briefing paper discusses some of the major adverse impacts of AI on children and young people’s safety.

## Positive Applications of AI

We often talk about the negatives of AI, but AI technologies can contribute to the identification and prevention of violence targeting children.

AI has emerged as a valuable asset in safeguarding initiatives, allowing professionals to identify threats, protect victims, and combat child exploitation. Although it cannot substitute for human judgment, AI enhances safeguarding strategies by improving speed, efficiency, and precision.

The rapid growth of digital platforms has created unparalleled opportunities for child abusers to harm and exploit children. AI algorithms can detect and identify suspicious online activities,

such as child sexual exploitation material (CSEM), grooming behaviours, and cyberbullying. By automating the reporting process, AI solutions can expedite the response of tech companies, law enforcement and child protection organisations.

Artificial intelligence plays a pivotal role in the detection of exploitation victims by scrutinizing images, videos, and digital communications. Technologies such as facial recognition and image-matching facilitate the cross-referencing of data with existing databases, aiding in the recovery of missing children and the identification of trafficking victims.

In addition, AI can identify communication patterns and behavioural trends, identifying children who may be at risk of grooming. By accelerating the process of victim identification, these tools assist safeguarding teams in directing resources to areas of greatest need.

### **AI-Generated CSEM and Deepfakes**

The term "deepfakes" is a combination of "deep learning" and "fake." It represents a specific category within the broader scope of generative AI. Typically, deepfakes pertain to media formats such as images, videos, and audio that have been artificially generated to convincingly substitute the appearance of one person with that of another, frequently with the aim of misleading viewers.

AI-generated CSEM refers to the creation of realistic yet entirely artificial explicit content involving children or young people using artificial intelligence algorithms. These technologies exhibit a troubling ability to produce images that can easily be mistaken for real, thereby blurring the lines between real and fake for those engaged in the fight against CSEM and the protection of children.

AI-generated CSEM includes a range of content, including text to chat, text to text, text to image, exploitative image to CSEM (e.g., de-aged images of adults, or modified existing CSEM), innocuous image to CSEM (Bracket Foundation, 2024)

The circulation of AI-generated child sexual abuse videos has raised significant alarm, reflecting the rapid evolution of AI technologies. An increasing number of deepfake videos are being disseminated on dark web platforms, where adult pornographic footage is altered to feature a child's face using sophisticated AI applications (Internet Watch Foundation, 2024).

Arguments are sometimes put forward that AI-generated material is 'victimless', but it is critical to remember that the content of AI-generated child sexual exploitation material is based on material that is sourced from real individuals and/or the AI technology is 'trained' through real images. In the development of AI models, both innocuous images and existing CSEM are utilised within the training datasets. This approach ultimately allows for the replication of many children's images, contributing to the ongoing creation of AI-generated CSEM.

These images can be used to bully, distress and extort children and young people. In addition, these images can be used by predators to groom and trap children into cycles of online sexual abuse. Some victims are being left suffering with depression, anxiety and even suicidal thoughts, with additional risks including safety and further victimisation.

Instances have been documented where offenders have stated that exposure to CSEM has intensified their sexual attraction to minors and emboldened them to commit contact offenses. Additionally, some specialists argue that CSEM has contributed to a broader societal acceptance of the sexualisation of children (Bracket Foundation, 2024). The emergence of AI-generated CSEM further facilitates the "normalisation" of child sexual abuse by catering to the fantasies of perpetrators and integrating such content into mainstream culture.

### **Nudify Apps**

We are also seeing an issue with individuals increasingly targeting children and young people by creating sexual images or nude images through 'nudify apps'; there have been numerous media reports of these apps being used by young people to bully, intimate and threaten peers.

These software applications alter existing images by stripping away clothing and generating hyper-realistic visuals. This technology employs machine learning algorithms to modify the original photographs. The ease of these apps makes it alarmingly easy for anyone to produce and share these images.

Although these images are fake and may even appear artificial, the use of nudify apps can lead to considerable harm. It has the potential to impact on a young person's mental health an individual's reputation and adversely affect their career opportunities. Furthermore, it can result in serious mental and physical health issues, such as social isolation, self-harm, and a diminished trust in others.

### **AI-driven Sextortion**

Offenders are employing AI-manipulated explicit material sourced from innocent images of minors to intimidate them via social media platforms, pressuring them to produce actual sexual images or content, or to provide financial compensation.

According to an online survey of youth in the United States, 40% of those aged 9 to 17 have reported experiencing unsolicited requests for explicit imagery from contacts met exclusively online. The data revealed that boys aged 9 to 12 were more likely to be solicited than girls of the same age, whereas this pattern was reversed among older teenagers. (Thorn, 2022).

### **AI Chatbots and Companions**

AI companions are chatbot applications that use artificial intelligence to replicate personal interactions through conversations (text or spoken) that resemble human dialogue. For many children, AI chatbots and companions are a significant aspect their interactions in the digital realm. They provide opportunities for interactive communication, learning, and various forms of entertainment (eSafety, 2025).

Despite the enjoyment and engagement these virtual friends offer to many children, they also carry certain risks (eSafety, 2025). These risks include AI-driven grooming, exposure to inappropriate content, developing emotional attachments, heightened risk of sexual exploitation, privacy issues, the potential for misinformation, and financial risks.

**AI-driven grooming:** AI-driven bots have reached a level of sophistication that allows them to replicate human conversations with remarkable precision. Malicious individuals exploit these bots to engage with numerous children simultaneously, thereby streamlining the grooming process. The AI is capable of building a connection, with the predator intervening only after a rapport has been formed. This approach complicates the early detection of threats for both children and their parents.

**Inappropriate content:** Certain AI chatbots might produce responses that are unsuitable for younger audiences. Despite the presence of filters, unpredictable exchanges can still take place.

**Emotional dependency:** Children may form significant emotional bonds with AI companions, which could influence their real-world social interactions and overall emotional health.

**Unhealthy perspectives on relationships:** In contrast to human relationships, interactions with AI companions lack established boundaries and the consequences associated with their breach. This can lead to confusion for children and young people who are still learning about the principles of mutual respect and consent, which may hinder their ability to create and uphold healthy relationships, whether they are sexual or non-sexual.

**Amplified danger of sexual abuse:** Regular exposure to sexualized conversations can negatively impact a child's or young person's ability to discern safe interactions and age-appropriate behavior, especially with unfamiliar adults. This can lead to an increased likelihood of being groomed and abused by predators, both in online environments and face-to-face.

**Privacy issues:** AI chatbots frequently gather and retain user information. It is essential to ensure that children do not disclose personal details.

**Misinformation:** The accuracy of responses produced by artificial intelligence is not guaranteed at all times. Consequently, children require assistance to effectively assess and critique the information they receive.

**Financial risks:** Subscription-based applications frequently employ strategies to promote impulsive buying behavior. The emotional bonds formed with AI companions may result in significant expenditures on 'exclusive' features, thereby posing financial risks.

## **Bravehearts Position**

The rapid evolution of AI and machine learning has indisputably resulted in numerous advantages, yet it has not been without its substantial challenges.

At Bravehearts, we assert that education, legislative action, and cooperative efforts are essential. It is imperative to equip young individuals with the knowledge to navigate the ethical dimensions of technology, ensure that legal frameworks adapt to hold offenders accountable, and collaborate with technology firms to establish protective measures against the misuse of AI.

In relation to the growing AI concerns, Bravehearts advocates:

- Greater general community awareness and inclusion of technology-enabled sexual abuse and exploitation in sex and relationship education. Respectful relationship can play an important role in fostering secure and respectful communities both online and offline.

- For Commonwealth, and state and territory laws to include clear definitions of AI-generated CSAM and to ensure legislation is continually reviewed to ensure that emerging risks are included.
- For Commonwealth, and state and territory laws to ensure that production of intimate images, including digitally created images, and sharing or threatening to share intimate images, including digitally created images, are included as crimes.
- That Commonwealth, and state and territory governments recognise ‘nudify apps’ as designed to sexually harass, abuse and extort victims, and to ensure these apps are banned and their removal from all app stores.
- Technology companies must be held responsible for ensuring safety by prioritising initiatives like ‘Safety-by-Design’ (<https://www.esafety.gov.au/industry/safety-by-design>)
- To develop a robust research agenda to ensure an informed response to risks of developing AI.
- Clear reporting options, including alternative reporting options, and specialised support services. As identified in an internal report completed for ICMEC Australia (Suttipron, et.al., 2025), there is a pressing need to enhance the awareness, quality, and accessibility of reporting mechanisms and support services related to online abuse.

## References

- Australian Centre to Counter Child Exploitation (2025). *AFP warns parents over rise in AI-generated child abuse material*. Retrieved from: <https://www.afp.gov.au/news-centre/media-release/afp-warns-parents-over-rise-ai-generated-child-abuse-material>
- Bracket Foundation (2024). *Generative AI: A new threat for online child sexual exploitation and abuse*. Los Angeles [CA]: Bracket Foundation.
- e-Safety (2025). *Online Safety Advisory: AI chatbots and companions – risks to children and young people*. Retrieved from: <https://www.esafety.gov.au/newsroom/blogs/ai-chatbots-and-companions-risks-to-children-and-young-people>
- Internet Watch Foundation (2024). *What has changed in the AI CSAM landscape*. Cambridge [UK]: Internet Watch Foundation.
- Suttipron, R., Pongpisan S., Buck, K., Lorita, A., & Hussain, A. (2025). The social, policy and legislative considerations of AI-generated Child Sexual Abuse Material. Internal ICMEC Australia/University of Queensland student report: unpublished.
- Thorn (2022). *Online grooming: Examining risky encounters amid everyday digital socialization*. Retrieved from [https://info.thorn.org/hubfs/Research/2022\\_Online\\_Grooming\\_Report.pdf](https://info.thorn.org/hubfs/Research/2022_Online_Grooming_Report.pdf)

Bravehearts Foundation Limited  
 ABN: 41 496 913 890 ACN: 607 315 917  
 PO Box 575, Arundel BC, Qld 4214  
 Phone 07 5552 3000 Email [research@bravehearts.org.au](mailto:research@bravehearts.org.au)  
 Information & Support Line 1800 272 831  
[bravehearts.org.au](https://bravehearts.org.au)