Global
Online Safety Regulators
Network

Position Statement:
AI-generated Child Sexual Abuse Material and Non-Consensual Intimate Imagery

Introduction

The rapid advancement of generative AI has created new possibilities for online sexual abuse and exploitation, including the production and dissemination of both AI-generated Child Sexual Abuse Material (CSAM) and Non-Consensual Intimate Imagery (NCII). These technologies enable the creation of abusive content at scale, with low cost and limited technical expertise, with CSAM consisting of sexual abuse and exploitation of children, and NCII consisting of non-consensual sexual content related to adults. As a result, sexually exploitative material can now be generated instantly. This shift necessitates a considered, coordinated and risk-based response that centres safety in design, development, deployment and use of AI-enabled technologies, and to mitigate the scale and severity of these harms.

The Global Online Safety Regulators Network (GOSRN) provides an important forum to strengthen regulatory coherence, accelerate collective learning and ensure that regulatory responses keep pace with technological change. Across jurisdictions, the misuse of AI tools in online environments means that regulators are confronting escalating harms and increasing public concern. As a result, this issue is a high priority for GOSRN. While respective frameworks differ, there is a shared recognition that generative AI challenges traditional models of detection, attribution, enforcement and cross-border cooperation.

This Statement outlines how the Network understands AI-generated CSAM and NCII ; explains why we are prioritising this topic in our work; and sets out GOSRN's shared position and collective response as well as where to find relevant expertise and resources.

What do we mean by AI-Generated CSAM and AI Generated NCII

We recognise that member jurisdictions may define and classify these terms differently within their respective legal frameworks. As such, these terms are not legal definitions and do not seek to harmonise domestic legal standards but rather provide a shared understanding of the issues.

This statement will use the **term AI-generated CSAM** to mean images, videos, audio, or other synthetic or manipulated media that reflects the sexual abuse of children under 18, created using generative AI tools. When we refer to **AI-generated NCII**, we mean intimate images, videos or other synthetic or manipulated media of an adult over 18 years, created using generative AI tools . This includes AI-fabricated or manipulated depictions of identifiable individuals shown nude or engaged in sexual activity, including, but not limited to, depictions of genitals, underwear, or (in the case of women) breasts.

AI-generated CSAM and NCII are often created using **nudification** tools. These services, including chatbots, websites and applications built on generative AI models digitally remove clothing from real photographs or fabricate sexually explicit depictions. These tools often lack safeguards, can be accessed by children, and are associated with the creation of abusive and illegal material, including NCII and CSAM. . These tools can be integrated into popular online platforms, which can lead to widespread access and dissemination. AI-generated content is also being distributed through a growing range of channels, including encrypted messaging platforms, dedicated forums and features embedded within otherwise mainstream services.

Part of our shared understanding is an acknowledgement that AI-generated content can cause devastating harm. In the case of CSAM, this harm is independent of whether a real child is depicted,'.-AI-generated content constitutes illegal CSAM in many jurisdictions in the same way non-AI-generated content does.

This material risks normalising child abuse, driving further contact offences traumatises victims whose likeness is used even where the abuse depicted did not physically take place. AI tools used to manipulate existing CSAM re-victimises victims and survivors of real-world abuse and the production of this material using image datasets derived from real children means that synthetic content is not categorically separate from real word abuse. That complicates law enforcement investigations and undermines efforts to protect children in danger.

AI-generated NCII cause serious-psychological, reputational, professional - and social harm, and is a growing form of image-based sexual abuse. The gendered nature of this content should be noted: both AI-generated CSAM and NCII disproportionately target women and girls.

AI-Generated CSAM and NCII are a priority for GOSRN

Generative AI tools provide perpetrators with a scalable way to rapidly produce, amplify and distribute CSAM and NCII. At GOSRN we are treating AI-Generated CSAM and NCII as a priority area for our collective focus.

Scale, speed and low barriers to offending

Generative models enable the creation of vast quantities of abuse material at pace that challenges existing detection and moderation capabilities. Individuals with limited technical skills can generate large volumes of content in minutes. AI-generated CSAM and NCII can then be rapidly distributed across interconnected platforms, including messaging services, forums and social media, where it can be reshared, repurposed and recontextualised. The growing scale, speed and accessibility of AI-generated harmful content may also place significant pressure on moderation systems, reporting mechanisms and victim support pathways. Emerging evidence suggests these pressures are already increasing: the National Center for Missing & Exploited Children (NCMEC) reported more than 1.5 million CyberTipline reports with a generative AI nexus in 2025, alongside more than 158,000 AI-generated images and videos categorised as CSAM since 2023.

Challenges to detection and enforcement

Lack of reliable watermarking, weak provenance standards, model opacity and cross-platform distribution can impede efforts to identify the source of content, assess risk pathways and attribute wrongdoing. While hash-matching and other proactive technologies remain effective at detecting known abusive imagery at scale, AI-generated content may create additional challenges where imagery is novel, modified, rapidly reproduced, or distributed across multiple services.

One of the key challenges posed by artificially generated NCII is that it is less likely than non-synthetic images to be proactively detected and relies almost entirely on victims finding and reporting it. By contrast non-synthetic images, may be identified through known datasets, prior reports, or platform detection mechanisms.

These dynamics reinforce the importance of safety-by-design approaches focused on systems, features and functionalities, including risk assessment, friction points, content provenance, reporting pathways. However, where moderation and reporting systems become strained by volume and complexity, victims may experience delayed detection and removal, reduced visibility of reports, and diminished access to effective redress. The time lag between content creation, detection and intervention presents particular challenges, especially where content can proliferate rapidly across platforms and services.

Regulatory gaps and jurisdictional asymmetry

Existing online safety frameworks and legal mechanisms for addressing CSAM and NCII in some jurisdiction's pre-date the current generation of generative AI tools, creating gaps that offenders can exploit. This is particularly the case for NCII, for example, some jurisdictions only regulate the sharing of NCII, while others regulated the generation and possession as well as the sharing of this content.

Moreover, depending on how AI tools are deployed, for example whether they are embedded within existing online platforms or as standalone platforms, they may be subject to different online safety regulation. Some regimes regulate the developers and deployers of AI models, including open-source models, while others only regulate the deployment of these models.

As a result, key areas of regulatory asymmetry can be said to include: the treatment of AI systems embedded within existing platforms versus standalone AI tools; the classification and liability of open-weight models; and cross-border enforcement challenges where services are established in jurisdictions with less developed frameworks and differing definitions of what constitutes illegal content.

In a number of jurisdictions, governments and regulators are increasingly seeking to address these gaps through ongoing legislative reviews, consultations, and emerging regulatory proposals. These developments reinforce the need for co-ordinated, cross-sectoral cooperation and information-sharing across regulators, lawmakers and law enforcement. Where multiple regulatory mandates intersect, clear public communication about regulatory responsibilities can support public understanding and effective recourse.

Approaches and priorities to addressing AI-generated CSAM and NCII

Despite differing mandates and legal frameworks, regulators across the Network agree that there are promising practices and priority actions emerging across jurisdictions. These include:

- Promoting **Safety by Design platform obligations for AI systems**, which includes requirements for platforms to undertake mandatory risk assessments, and implement safeguards that prevent the generation of CSAM, NCII and other illegal content, including guardrails at data selection, training, refinement, input, output and maintenance stages.

- Requiring **transparency and accountability**, including by providing clear information about how AI systems operate, the use of **provenance or labelling mechanisms** where appropriate, meaningful transparency regarding moderation and safety processes and internal governance measures to support oversight, auditability and risk management.

- Ensuring **user empowerment through reporting and redress,** by providing clear, accessible and effective mechanisms for users to report harmful or illegal content, seek redress and exercise meaningful control over their interactions, including contesting moderation outcomes and requesting corrective action.

Underpinning these practices are key principles that guide GOSRN, including:

- Responses that are grounded in the promotion of human rights, including dignity, privacy, freedom from exploitation and the rights of the child. Victim and survivor voices should inform regulatory design and enforcement priorities.
- Prioritising the protection of children, including by promoting age assurance, which can prevent children from accessing nudification tools that can generate harmful content.

How GOSRN is addressing the issue

AI-generated CSAM and NCII are inherently global problems that require a globally coherent response. Taken together, greater regulatory coordination can enhance the effectiveness, consistency and legitimacy of responses to AI-generated harms in a rapidly evolving environment.

GOSRN's work programme focuses on sharing of information and best practice to promote regulatory coherence, and we have a number of work areas that are relevant to promoting common understanding and approaches on AI-generated CSAM and NCII. This includes a recently launched special project involving members and observers on CSAM.

Alongside this project, other areas of focus include:

- Engaging with law enforcement and civility society, including considering approaches to cross-border coordination and information sharing, where possible.

- Building shared technical capacity, including by participating in joint expert groups, developing shared testing frameworks and mutual information sharing, helping regulators keep pace with fast-evolving AI technologies and the abuse patterns they enable.

- GOSRN will also use its unique position to highlight key resources and information from members, observers and the technical community, civil society, academia and law enforcement.

Key resources

The following resources provide supporting context and evidence for this joint statement. Regulators are encouraged to draw on these sources or suggest others.

- National Center for Missing and Exploited Children (NCMEC), CyberTipline 2024 Annual Report
- Internet Watch Foundation (IWF), Harm Without Limits: AI Child Sexual Abuse Material Through the Eyes of Our Analysts (March 2026)
- AI Generated Child Sexual Abuse Material: What's the Harm? Ó Ciardha, C., Buckley, J. & Portnoff, R.S. (2025)
- GOSRN Age Assurance Position Statement (2026)
- UNICEF Innocenti Guidance on AI and Children (2025)
- Georgetown Institute for Women, Peace and Security, Technology-Facilitated Gender-Based Violence (2024)
- United Nations Convention Against Cybercrime (2024)
- Safety by Design CSEA Toolkit (2026)
- Deepfakes | What are deepfakes? | Position statement | eSafety Commissioner (2026)
- Deepfake damage in schools: How AI-generated abuse is disrupting students, families and school communities | eSafety Commissioner (2026)
- Joint Statement on AI-Generated Imagery and the Protection of Privacy (coordinated by Global Privacy Assembly) (2026)

This appendix will be updated as the evidence base develops. Member regulators are encouraged to submit additional resources for inclusion.

Contact current Chair:

gosrn@cnam.ie https://www.cnam.ie/about/what-we-do/international/global-online-safety-regulators-network/

November 2026