

AI and CSAM: Current Standing and a Future Perspective

Analiv Hasert

Andreia Costa Marchante

Terezia Kucarova

Erasmus University Rotterdam

AI and CSAM: Current Standing and a Future Perspective

Child sexual abuse material (CSAM) continues to be a rapidly growing, fast-evolving threat, representing one of the most profoundly harmful types of online exploitation targeting children. CSAM has been defined to include any material visually depicting a child, or what appears to be a child, engaged in real or simulated sexually explicit conduct, depictions of intimate parts of such person for primarily sexual purposes, or material intended to provide advice, guidance, or instructions on committing an offence related to child sexual abuse material (as per Directive 2011/93/EU, and the currently proposed recast COM(2024)60). A recent trend in the development of artificial intelligence (AI) has been recognized as a major driver of both the volume and severity of CSAM, presenting a new and urgent challenge: AI's capacity to generate synthetic content now enables offenders to produce abuse imagery depicting children in increasingly accessible and realistic manner, with broad generative range, as converging tools increasingly combine image generation, audio deepfakes, and chatbot functionality within single platforms (Ciardha et al., 2025; Internet Watch Foundation, 2026; Steel, 2024; WeProtect Global Alliance, 2025). In 2025, IWF reported assessing over 8000 AI-generated, highly realistic CSAM items, while a combined UNICEF, ECPAT, and INTERPOL study across 11 countries found that at least 1.2 million children disclosed having had their image manipulated into a sexually explicit deepfake in the past year alone, representing 1 in 25 children in some countries (UNICEF, 2026; full national reports are to be released throughout 2026).

In light of these developments, this report aims to review the current body of knowledge in the area of AI and CSAM, namely the: AI creation tools; motivations of creators; motivations of consumers; impact on victims; undertaken mitigation and prevention efforts; and legal and

detection challenges; offering a well-rounded overview on the issue while identifying pressing gaps in existing research, with the purpose of providing a foundation for informed and actionable responses to the issue.

Historical Context

The historical developments in the area of CSAM primarily reflect the gradual advancement of the technological capability of our society. Emerging via small-scale, less expansive ways of direct distribution, CSAM gradually became salient as a global problem with the emergence of the internet (European Parliamentary Research Service, 2020; UNODC, 2023). The 2000s gave rise to more sophisticated editing tools, blurring evidentiary boundaries and raising critical definitional and interpretive concerns for courts handling CSAM cases. Early legal debates followed soon after (notably: *Ashcroft v. Free Speech Coalition*, 535 U.S. 234 (2002)); court rules depictions not involving a real child are not automatically criminalized, highlighting the tension between technological innovation and child protection. In parallel with technological advancement, CSAM volume continued to grow with faster internet connectivity and image hosting platforms expanding across cyberspace (European Parliamentary Research Service, 2020).

Both the Council Framework Decision 2004/68/JHA and the later 2011 Directive mark early attempts at mitigation and prevention of CSAM, emphasizing combatting CSAM emerging via new technologies. A shift critical to AI-CSAM occurred in the late 2010s, with the expansion of deep learning-based generative models now being able to create increasingly realistic content. Earlier AI tools, such as nudification or faceswapping (tools which were powered by generative adversarial networks (GANs)), were followed by higher-fidelity open-source diffusion models, meaning that the technical barriers that might have limited accessible image manipulation

before have eroded substantially (Ciardha et al., 2025; Steel, 2024). By the early 2022-23, these tools began appearing at dark web markets and within online offender communities, marking a significant turning point: AI-CSAM could now be created at scale, and without any direct abuse material source (Borah et al., 2025; IWF, 2024b).

The speed at which the nature of AI CSAM developed from 2023 onwards was rapid, mirroring the large-scale, rapid scaling of generative AI overall (Ciardha et al., 2025; Europol, 2025). In its initial report on AI CSAM published in 2023, the IWF identified primarily content characterized by still images involving one subject, often identifiable by artificial elements such as distorted body parts. By mid-2024, the IWF had noted that AI CSAM continued to grow not only in scale, but also realism and extremity, with critical development following the widespread use of Low-Rank Adaptation (LoRA) models, a fine-tuning technique applied to open-source generative models for targeted outputs. By late 2025, AI CSAM featuring full motion video has become common on the internet (IWF, 2026).

Why Action is Needed Now

Both the scale and severity of AI CSAM have been recently reported to grow dramatically. In 2024, the Internet Watch Foundation (IWF) reported a 380% annual increase in AI-generated CSAM from 2023, indicating an exponential growth in this type of content. The National Centre for Missing and Exploited Children reported a 1325% surge in CyberTipline involving generative AI within the same period (NCMEC, 2025). Out of over 8000 images and videos sexually depicting children that were detected by the IWF in 2025, 3443 items consisted of AI-generated motion videos, representing a 26,385% increase from 13 items assessed in 2024. 65% of these videos were classified as Type A content, the most severe

category encompassing offences like rape, sexual torture, and bestiality (IWF, 2026).

Novel forms of material including videos and deepfake audio are becoming increasingly accessible and photorealistic, with highly gendered content (girls accounting for 97% of AI generated CSAM) (IWF, 2026). An increase of this type of content on clear web, including commercial sites, has been noted as well; offenders actively commercialize custom AI generated content, instructional material, and customized models (IWF, 2026).

The harm of AI-CSAM has been found to be extensive, and ranges from potential risks of desensitization (Knack et al., 2020; Seigfried-Spellar & Rogers, 2013), risk of contact offending (Insoll et al., 2022; Johnson, 2020), perpetuating trauma (Gerwirtz-Meydan et al., 2018), to complex content removal issues (Parti & Szabó, 2024).

While the body of evidence reinforces in pointing towards the harms of AI CSAM, both international and national legal frameworks lack consensus, and detection efforts struggle to keep up with the ever-evolving pace of AI. With the expiration of the EU interim ePrivacy derogation on April 3rd 2026, negotiations on permanent CSA Regulation remain ongoing (European Parliament, 2026). The proposed recast of Directive 2011/93/EU is similarly still under negotiation. Comprehensive federal framework in the United States remains absent.

These trends make AI-CSAM not merely a technical problem, but also a policy imperative, demanding urgent coordination of all involved stakeholders.

AI creation tools

As generative AI continues to advance rapidly, the tools used to create virtual CSAM are simultaneously becoming more accessible to non-

expert users (Ciardha et al., 2025). These rapid developments prompted efforts to understand and classify the methods involved in the content creation, supporting targeted detection, mitigation strategies, and informed policy-making.

Core Methods for Creation of Virtual CSAM

AI-generated CSAM. This category represents content that is entirely synthetic, that is, depicts minors that, by definition, do not exist. One category of generative system used for this purpose is text-to-image diffusion models, a generative AI system which is able to create images from written descriptions, generally known as *text prompts*. In its initial training stage, the model learns to associate visual content with its written description. Once it has been trained, it is able to generate entirely new imagery by refining random visual noise into specified content guided by the text prompt, allowing for a continuous, on-demand generation of CSAM (Ciardha et al., 2025; IWF, 2024b; Steel, 2024).

A critical distinction particularly relevant to AI CSAM concerns open-source and closed-source models: closed-source models are hosted on company servers and can be accessed remotely. Their provider retains control over the model's code and architecture, into which they may integrate safety filters in order to prevent users from generating harmful or illegal content (*Midjourney*). Open-source models, on the contrary, make their code public, for any private individual to access, download, and run locally offline, with no formal external oversight (*Stable Diffusion*). Because of their offline feature, they are frequently used by AI CSAM offenders (IWF, 2026). Notably, closed-source models, typically higher-fidelity ones, are not immune to exploitation either. In order to override the built in safety filters of the models, offenders have been found to utilize prompt engineering, or so called “jailbreaking techniques” or “surrogate prompts”, such as coded

language, intentional misspellings, or descriptions of medical anatomy (Ciardha et al., 2025; Ba et al., 2024).

Beyond the pure generation of content, certain diffusion models are able to enhance the quality of generated content via latent diffusion upscalers (super resolution models), creating realistic body postures or lighting, or pose/depth conditioned generation (IWF, 2023; Steel, 2024; Thiel et al., 2023). IWF (2026) reports that offenders have been found to utilize these tools in order to increase marketability of older material for recirculation.

A further escalation concerns the offender's attempts to overcome the generic nature of the content produced by open-source models through LoRA adapters (Low-Rank Adaptation), a fine-tuning technique that allows users to specialize an open-source generative model for a particular purpose. By training a small, additional layer of code using a limited set of reference images, a user can attach this layer to the open-source code without modifying the base code itself. This method produces more targeted outputs than the base model alone, and is accessible to non-experts with limited technical effort (Ciardha et al., 2025; IWF, 2026). The new files created will be small, shareable, and downloadable and can be operated even offline, which makes them highly resistant to tracing and removal efforts. The base code of the model will remain unchanged and legal, with the LoRAs attached or detached as needed. LoRA models have been observed circulating on dark web forums, with some trained on images of identified CSAM victims, celebrities, or certain sexual acts. 202 different LoRA models were identified by December 2025 (IWF, 2026).

AI-manipulated CSAM. The second category includes content that is produced using an image of a real person as an input and manipulated, altered, or extended using AI further, for instance by removing clothing, altering facial features, or modifying age. Unlike AI-generated CSAM, this

method exploits features of existing individuals, directly victimizing them (Ciardha et al., 2025).

Manipulation tools are numerous, and different methods are known to be used in creation of CSAM. Face-swapping techniques have been employed to replace faces of adults in pornographic material with children's faces. Nudification tools have been observed to be used to generate nude content from ordinary clothed photographs of children; research indicates that nude deepfakes account for up to 98% of all deepfake content online (Internet Matters, 2024). AI-assisted upscaling has been similarly used to enhance resolution and detail of older, low-quality abuse imagery, to increase its marketability.

In addition to the above-mentioned approaches, a number of potential applications of current AI technologies, including age regression or body morphing, have caused alarm regarding their further implementation as tools for extending AI manipulation methods (CameraForensics, 2024; Thiel et al., 2023).

According to IWF analysts, criminals have reportedly used artificial manipulation technology not only to create material, but to obscure the existence of CSAM images by altering the appearance of children's faces or other identifiable features in order to evade detection, posing a critical challenge for law enforcement (IWF, 2026).

The Underlying Technology

Diffusion models constitute the primary architecture powering the above described methods, and account for much of the current scale of AI-CSAM production (Steel, 2024; Thiel et al., 2023). Unlike the earlier generative adversarial networks (GANs, which are, in comparison to diffusion models, less flexible and stable), diffusion networks generate images by progressively de-noising random static into coherent imagery,

that is, they are able to create content guided by text, prompts, or sketches (Ciardha et al., 2025, Karaagac, 2025).

A fundamental concern lies in the training process of widely used diffusion models: because these models utilize learning based on text-visual association, the content they digest to train directly shapes the content that is produced. Recent findings suggest that latent diffusion models (e.g., *Stable Diffusion*) have been trained on highly problematic content, including racial/ethnic/gender stereotypes, pornography, as well as illegal content such as sexual abuse and CSAM (Downing, 2024; IWF, 2024b; Steel, 2024; Thiel, 2023).

Newly Emerging Techniques

Synthetic Motion Video. Moving imagery generated by AI can be created using various techniques, including superimposing the face of a child onto existing video material (face-swap deepfakes), creating completely new videos based on text prompts (text-to-video models), or animating still pictures to produce moving imagery (image-to-video).

According to the IWF (2026), face-swap deepfakes marked the earliest example of AI-generated CSAM video, whereas the use of entirely synthetic videos and image-to-video approaches has become more common since then. What had previously been characterized as an emerging issue has become increasingly accessible; researchers have reported existence of videos that mimic the movement of the body in natural settings with photorealistic lighting and camera motion, with quality reaching a point where the content can deliberately imitate amateur videography, depicting multiple subjects and increasingly extreme forms of abuse. The so far identified motion techniques are evolving rapidly, with researchers warning that more advanced, feature-length content may become increasingly

common (European Parliamentary Research Service, 2025; IWF 2026; Steel, 2024).

Deepfake Audio. Voice-cloning or speech-synthesis technologies can create realistic audio tracks using just small amounts of speech, mimicking not only the voice's tone but its accent or emotional affect. Previous research has already identified possible implications for abuse-related activities when combining the described ability with synthetic video generation, suggesting that the two approaches could result in photorealistic multimodal abuse materials: extending the range of abuse beyond visual content (European Parliamentary Research Service, 2025; Steel, 2024). According to the IWFs 2026 report, the combination of the mentioned technologies has been observed: AI systems have been used to combine image, audio, and text inputs into videos with synthesized voices. Although AI-generated CSAM featuring audio elements alongside other types of media is not largely widespread, these developments signal a trajectory towards increasingly multimodal and photorealistic abuse material (IWF, 2026).

Conclusion and Knowledge Gaps

The realism of AI-generated content poses a serious issue in the matter of differentiating it from the authentic material (Ciardha et al., 2025; IWF, 2025; IWF, 2026). It is plausible that the amount of AI CSAM cases that are currently registered is underdocumented, due to the fact that the content is considered to be synthetic only when explicitly traced from the source or via the metadata (IWF, 2026). Moreover, the advances in the architecture of generative models made it possible to reduce technological barriers to both access to and creation of illicit content, with tools such as LoRA introducing unprecedented capacity for targeted output. While the

evidence base on how AI tools are exploited to create CSAM strengthened considerably, critical gaps remain underexplored.

In terms of detection and prevention of AI-enabled CSAM, currently used strategies are becoming insufficiently effective. Traditional perceptual hashing systems, which have been useful in locating known CSAM, cannot be used to locate new AI-generated material, as the image will be unique every time, avoiding existing hash databases (Steel, 2024; Thorn & All Tech Is Human, 2024). Watermarking technology, advocated as an alternative option, is similarly insufficient, as offenders may manipulate images by removing or adding a false watermark to complicate detection efforts (IWF, 2026; Steel, 2024). Robust research is urgently needed to explore options of defensive use of AI, especially in terms of capabilities of identification of AI-CSAM (Ciardha et al., 2025; Steel, 2024).

From a technological perspective, a number of emerging areas require further examination. The increase in the use of AI chatbots enabling both simulation of sexual conversations with minors as well as an environment where AI-CSAM can be hosted, marking an intersection between technologies used for text creation and image generation, warrants additional research (IWF, 2026). The convergence of capabilities of integration of image generation, chatbot conversations, and audio deepfakes into a single platform, lowering technical barriers and broadening the modalities involved in offending, has similarly received limited attention in research (IWF, 2026). Additionally, an area requiring critical attention concerns content embedded in training datasets, particularly, exploration of existing or possible safeguards preventing CSAM from entering large-scale training datasets, their effectiveness, and the burden of responsibility when the safeguards fail (Downing, 2024; Steel, 2024; Thiel, 2023).

Another research gap concerns exploration of the possible implications of agentic AI, a system that is able to operate independently,

with minimum human input, to pursue certain goals. With the rise of the first generation of agentic AI technologies, the concern regarding proliferation of CSAM has emerged; offenders were found to openly discuss using such systems to automate the creation and distribution of CSAM (Ciardha et al., 2025; IWF, 2026). While these risks are currently only emerging, the pace of development across the AI arena emphasizes the need for proactive research before such capabilities become widely accessible.

Although considerable progress has been made on documenting how the existing tools have been misused, policy recommendations calling for establishment of industry-wide standards for AI development and its maintenance remain largely aspirational (Thorn & All Tech Is Human, 2024). The proliferation of nudify apps, tools reported as having no single legitimate use, further highlights the gap between the pace at which technology develops and spreads, compared to the speed of regulatory responses (IWF, Child Helpline International, INHOPE, NCMEC, Offlimits, Safe Online, & WeProtect Global Alliance, 2026).

Motivation: Creators

Beyond the technological tools of creating AI-CSAM, it is important to understand the motivations of those who create it. These motives can range from individual sexual gratification to organized financial exploitation, and may involve both adults and young people. Although these domains have been studied in online communities where CSAM is posted, the use of AI alters the conditions, lowers barriers, and changes the pool of potential creators. Overall, this enables new forms of offending that do not rely on traditional access pathways.

Before AI, live-streamed CSAM typically required offenders to have direct physical access to the child (Huikuri, 2023). However, AI can reduce

or remove the need for direct access to a child in some forms of offending. Instead, facilitation of offending includes access to technological devices with models like Low-Rank Adaptation (LoRA) that can be trained to create AI CSAM within 15 minutes using 20 to 30 images (Internet Watch Foundation, 2026). Due to this shift, it is important to study what drives the creation of specifically AI CSAM, to allow for more tailored responses.

Monetary gain can be a motivation for the creation of AI CSAM. TRM Labs (2025) estimates that a cryptocurrency transaction associated with CSAM takes place around every two minutes, highlighting the market that AI CSAM could entail. Emerging evidence further suggests that this market is becoming increasingly structured, with over half of individuals involved in AI-CSAM (54%) reporting that they had commissioned or produced such material for financial gain (Protect Children, 2026). AI-CSAM circulates across dark-web, clear-web, and social platforms, with monetization occurring through marketplaces, subscriptions, and paid access structures. (Crawford & Smith, 2023; Internet Watch Foundation, 2023; TRM Labs, 2025). Dark web forums are often either structured as admin-owned markets (profits for site operators) or contributory platforms that share revenue with uploaders (TRM Labs, 2025). In these markets, certain groups advertise on-demand or custom AI CSAM, which allows buyers to specify certain targets and scenarios (TRM Labs, 2025). However, even in the clear web, platforms like Patreon (in the USA) can function as monetized clear-web ecosystems that may connect users to further sexual-content sites (Crawford & Smith, 2023). Therefore, creators can profit financially without even entering the dark web. Furthermore, the rise of AI-generated videos suggests product diversification (Internet Watch Foundation, 2024a). This diversification accelerated in 2025, with Internet Watch Foundation assessing 3,443 AI-CSAM videos (65% Category A - most severe), suggesting that more advanced video production is increasingly present

(Internet Watch Foundation, 2026). In addition to direct sales of CSAM, the wider AI-CSAM ecosystem also includes peer-to-peer guides in underground communities (WeProtect Global Alliance, 2024) and YouTube tutorials on the clear web (Davy & Lundrigan, 2024) that facilitate creation practices and lower technical barriers. There is also evidence that generative models are developed, exchanged, and sold specifically to produce sexualized images of minors (Internet Watch Foundation, 2026; European Child Sexual Abuse Legislation Advocacy Group & DOT Europe, 2025). These systems may reduce production costs for users and enable a single individual to create and distribute large volumes of content (Kaul, 2024). Overall, these findings suggest that AI CSAM is viewed as a low-cost and high-reward market for mass production.

There are other motivations beyond financial gain, including status, belonging, and access to exclusive channels. Mink et al. (2026) found that broader AI-generated sexual content creators were motivated by technical prestige (mastering jailbreaks/LoRAs), community belonging, and status within creator hierarchies. The Internet Watch Foundation (2026) similarly describes dark-web and clear-web communities in which sharing models and "teaching" workflows establishes reputation, although Internet Watch Foundation analysts note that prestige appears to be an important motive alongside financial incentives. Nevertheless, these reputation-based incentives likely interact with financial motives, creating hybrid creator subtypes motivated by both profit and prestige in parts of the broader AI sexual-content ecosystem (Mink et al., 2026).

Sextortion is a type of threat where an offender uses intimate images of the victim to make a demand. These demands can consist of more explicit images, monetary payments, sexual favors, and more (Ray & Henry, 2025). AI has been used to create these images, with the use of a face photo alone, indicating that offenders no longer need a real intimate

photo to initiate coercion (Aronashvili, 2024; Internet Watch Foundation, 2024a). Furthermore, some guides explicitly encourage blackmail of minors using nudification tools (Internet Watch Foundation, 2024a). Regarding this, the FBI reports an increase in cases where offenders transform social media photos into explicit images and demand money (often gift cards or crypto) or sexual content (Internet Crime Complaint Center, 2023). Overall, AI is not only seen as a tool to satisfy sexual interest but also as a mechanism of domination, financial extraction, or humiliation.

Some creators mistakenly believe that fully synthetic images fall outside criminal law or are “less serious” than real CSAM (Lucy Faithfull Foundation, 2024). Recent 2026 research has also found that these beliefs persist (Mink et al., 2026) and this perception can encourage creation. However, it remains unknown whether these perceptions form from deliberate misinformation within forums, observation of weak enforcement, or a genuine cognitive distinction between synthetic and real harm.

One relevant subgroup is youths who create AI-generated nude or sexual images. In a Belgian survey of 15 to 25-year-olds (N = 2,819), 41.9% had heard of “deepnudes,” 23.1% had seen one, 12.8% knew specific apps, and among those who knew an app, 60.5% had tried it (Szyf et al., 2023). Another study based in Czech Republic primary and secondary schools, found that 2.77% of respondents had created nude images with AI (Kopecký & Voráč, 2025). The male population was over-represented across the measures of familiarity, creation, and possession of this material (Kopecký & Voráč, 2025; Szyf et al., 2023). Reported motives behind creating this content include curiosity or fun, popularity or impressing friends, “pranks” or dares, and revenge/retaliation or humiliation, with distribution primarily via Snapchat, X/Twitter, and Instagram (also TikTok and Telegram; Szyf et al., 2023). Overall, a reinforcing loop may exist, where this content is easy to create, has a high payoff for sharing,

attracting more attention and group approval for the creator, thereby may normalize these images and escalate the harm (Internet Watch Foundation, 2024a; Szyf et al., 2023). These patterns suggest that what may begin as experimentation can carry real and lasting consequences for young people and their peers. Although research investigates the prevalence and motivations of this subgroup, it would be important to further examine the early warning signs of adolescents/children who create this content. Furthermore, longitudinal data does not exist examining whether youth creators remain situational actors or develop sustained offending patterns, and if they do, which factors contribute to this development.

Last but not least, creators of AI-CSAM can be consumers as well, when the motivation behind creating this content is for their own sexual gratification. The ease of creating this content, in addition to the realism linked to it, allows interested individuals to create for their own paraphilic exploration (Crawford & Smith, 2023). The motivation behind the consumers of such content will be elaborated in a section below.

Overall, research now documents a wide array of AI-CSAM creator motivations, from financial gain and social prestige to sextortion and personal gratification. The Internet Watch Foundation (2026) confirms commercial ecosystems (custom sales, creation courses) alongside community status incentives, while Mink et al. (2026) documents similar patterns among broader AI sexual content creators. Although motivations are not mutually exclusive, research is only beginning to systematically measure prevalence or examine how mixed motives interact and evolve over time.

Motivation: Consumers

Besides the motivations for creating this type of content, it is also key to understand the motivations of consumers who indirectly contribute to

this urgent issue by engaging with CSAM, particularly AI-generated or manipulated CSAM. CSAM consumers are typically heterogeneous, representing a wide variety of backgrounds. However, some studies have focused on identifying common patterns among individuals who watch child sexual abuse material and reported some interesting findings, which have allowed the establishment of likely motivators. CSAM consumers have been found to be predominantly male, single (Barroso et al., 2025; Brown, 2023; Napier et al., 2024; Prat & Jonas, 2012) and of younger age (18–34 year old) despite a few studies with contradicting results that evidenced that CSAM consumers are more likely to be older than other offenders (Brown et al., 2018; Brown, 2023; Napier et al., 2024). Further research also investigated distinctions between individuals who had been formally charged and those who had not, revealing that adult sexual offenders with an antisocial background, a sexual interest in CSAM, and frequent consumption of adult pornography displayed markedly different personal, motivational, and behavioural profiles compared to non-charged CSAM consumers, who commonly reported that their use began with viewing adult pornography (Lahtinen et al., 2025). Understanding the target public and common motivators is key to applying appropriate and effective prevention initiatives (awareness campaigns, (psycho)therapeutic treatment, (psycho)education, helplines, among others) as studies have suggested that, even though initiatives directed at contact child sexual abuse offenders are often not effective in reducing offending, some of these preventive interventions can encourage self help-seeking behaviours among other target groups (Gannoni et al., 2023), possibly CSAM consumers.

Building on the discussion of motivators, one of the motives research has identified, and the most obvious, is sexual gratification (Internet Watch Foundation, 2023). A study examining the determining factors of initial exposure to CSAM and subsequent deliberate engagement, conducted on

5512 internet users, reported 742 individuals having viewed CSAM at least once, and a subsequent set of 459 intentional CSAM users (Napier et al., 2024). In the subset sample, 47.5% of the individuals reported viewing CSAM again after first exposure by choice, showing that even accidental exposure can lead to future intentional viewing among a significant proportion of individuals (Napier et al., 2024). Furthermore, this study also demonstrated that while people of all genders can be exposed to CSAM, men experiencing loneliness and visiting pedophilic chat forums online. are more likely to come across it and deliberately view it again (Napier et al., 2024). Nonetheless, approximately one-quarter of participant female viewers also reported intentional viewing after their first exposure (Napier et al., 2024).

Despite all the limitations of the research that has been done on this topic due to methodological issues and diverse (individual) symptomatology, another possible motivation is addiction. Multiple studies have demonstrated that people can get addicted to pornography, exhibiting similar basic mechanisms and underlying neural processes to substance addiction, so that the DSM-5 in 2014 added the diagnosis of hypersexual disorder, which includes compulsive and problematic pornography consumption, and symptoms such as sexual dysfunction and psychosexual dissatisfaction (De Alarcón et al., 2019; Hilton & Watts, 2011; Love et al., 2015). On the other hand, research also suggests a link between compulsive behaviours and engaging more in CSAM use (Kim & Banks, 2024). Repeated CSAM use can show similar characteristics to substance addiction. Both users exhibit compulsive behaviour and report that their consumption is highly rewarding (Christensen et al., 2021; Kim & Banks, 2024). This addictive nature increases the risk of escalation, with individuals potentially seeking out more extreme and disturbing content and may be more likely to engage in more severe forms of abuse (a study showed that

52% to 58% of CSAM viewers in their sample reported fearing that viewing CSAM might lead to sexual acts against a child) (Christensen et al., 2021; Huikuri, 2023; Insoll et al., 2022). Indeed, comparable to porn addiction, the thrill of novelty provided by CSAM decreases over time, and the continued engagement with similar content decreases dopamine in the synapses, contributing to the need to seek more severe material to create a similar level of dopamine (Christensen et al., 2021). Altogether, these similarities suggest that there could be a link between addiction and CSAM consumption, like there is between addiction and pornography consumption, that is not fully understood yet. This urges further research on the topic to enable the development of appropriate prevention initiatives, especially because it has also been shown that individuals often progress from passive consumption to active collection, as possessing such material facilitates interaction with other offenders and the acquisition of more extreme content (Christensen et al., 2021).

Moreover, paraphilic sexual interests, particularly pedophilia, are strong predictors of CSAM consumption. This is particularly true when hypersexuality is present, since possessing CSAM can show specific sexual interests (Lahtinen et al., 2025; Reichel et al., 2024). Research shows that people convicted of CSAM offences have higher levels of atypical sexual interests than those convicted of contact child sexual abuse (Lätth et al., 2025). These offenders often have cognitive distortions that help them rationalise and justify illegal behaviour (Barroso et al., 2025). As a result, they may not fully understand concepts like re-traumatisation, because they rely on false ideas about victims, society, and themselves (Huikuri, 2023). However, it is important to note that not all cases of CSAM use can be attributed to paraphilic or pedophilic disorders, as various explanatory models, including the motivation-facilitation model, explain CSAM consumption and use with factors such as sexual desire and antisocial

tendencies besides paraphilias (Seto, 2017). A considerable proportion of CSAM consumers (82%) reported regularly viewing adult pornography prior to seeking out CSAM, with over half engaging in daily use (Lahtinen et al., 2025). This pattern may reflect a heightened sexual drive, identified as a key motivational factor in the model of sexual offending (Lahtinen et al., 2025; Seto, 2017). Additionally, researchers have also argued that experiencing abuse as a child, namely sexual abuse, also constitutes an enormous risk factor as it may normalise such behaviours through repetitive exposure or past experiences (Christensen et al., 2021; Huikuri, 2023; Johnson, 2025).

Normalisation may also contribute to all motivations. CSAM consumers may foster impunity, escalation, and a lack of empathy for child victims through their engagement in forums. Some believe viewing CSAM decreases the likelihood of acting on atypical sexual desires. However, research shows this is not true. Viewing, possessing, or making CSAM strengthens deviant thoughts, fantasies and urges, and can increase the risk of escalation (Christensen et al., 2021; Internet Watch Foundation, 2026; Johnson, 2020). Two different studies have shown that a large proportion of CSAM consumers report having thoughts about seeking direct contact with children and fear that their consumption could lead them to do so, which evidences a link between CSAM consumption, desensitisation and hands on, or hands off, sexual offending (Insoll et al., 2022; Internet Watch Foundation, 2026; Protect Children, 2021). Lastly, since there is already the problematic assumption among consumers that CSAM is a victimless crime because there is no contact with the victim (Johnson, 2025), it would be predictable that this belief is stronger in AI-fabricated or manipulated material and would, therefore, reduce the sense of culpability as a consequence. Thus, these offenders might also see this as a legal loophole, thus potentially increasing the demand for synthetic CSAM. This also

desensitises CSAM, for example, in cases of fully fabricated images, the guilt or empathy internal mitigating factors no longer have the same impact since no real child is being harmed, dehumanizing children, entrenching objectification, and reinforcing offender cognitive distortions (Christensen et al., 2021; Internet Watch Foundation, 2026). Plus, the aggravating factor is that generative AI is likely to exacerbate this problem by facilitating infinite creative potential that enables offenders to satisfy their insatiable demand and further fuel their addiction.

In summary, while substantial research has examined the motivations and influential factors of CSAM consumption, individual and external motivators vary widely by person and over time. This variability makes it challenging to generalise or group consumers according to similar profiles and motivations, complicating the design of preventive interventions.

Psychological Impact and Harm: Victims

Emotional Distress

AI-driven online sexual exploitation of children has been linked to severe psychological harm (American Academy of Pediatrics (AAP), 2025; Page et al., 2025), combining both immediate and chronic emotional disturbance. Victimization by AI-CSAM has been associated with experiences of fear, shame, humiliation, violation, anger, and self-blame, causing feelings of helplessness and loss of control over identity, extending to fear of public exposure and stigmatization (AAP, 2025; Page et al., 2025). The specific nature of AI sexual exploitation affects well-being profoundly. Victims report difficulties in interpersonal relationships and persisting feelings of self-blame. Research indicates that experience of AI-CSAM abuse may trigger symptoms consistent with clinical expressions of anxiety, depression, and post-traumatic stress disorder (PTSD; i.e., intrusive thoughts, nightmares, and hypervigilance) (Page et al., 2025).

Perpetuation and Amplification of Trauma

The nature of the trauma specific to experiences lived by victims of AI-CSAM is underexplored and damaging. Characteristics of AI models and dynamics of cyberspace (e.g., content embedded in training algorithms, infinite replication, spread on social media) are significantly challenging in terms of effective content removal, which in turn enables repeated violation of both dignity and privacy of the victims (Ciardha et al., 2025; Thorn, 2025a). In addition, secondary victims, i.e., friends and family members, have been also found to suffer considerable emotional distress linked to social isolation (IICSA, 2023).

Barriers to Justice

Victims of AI-CSAM consistently report serious obstacles in seeking effective redress and prompt content removal (National Center for Victims of Crime, n.d.). Platforms often tend to delay, or they refuse requests for content removal, intensifying the already debilitating impact that online sexual exploitation has on its victims (Thorn, 2025a). Legal and procedural challenges further restrict effective access to remedies. This phenomenon gets even more complicated by fragmented jurisdictional authorities, the tension between national and international criminal justice, and a lack of oversight about who specifically is responsible for the content governance (Wolbers et al., 2025). Furthermore, victims (i.e., disproportionately young males) report experiencing fear of disbelief or stigmatization within their environment, which may prevent them from effective access to support systems (AAP, 2025).

Knowledge Gaps

The nature of AI-CSAM exploitation is fundamentally unique and increasingly prevalent, yet the research dedicated specifically to its impacts on victims remains limited (Ciardha et al., 2025). Impact on secondary victims, especially close family members, remains under-addressed despite

reports of severe emotional distress and social consequences linked to such cases of abuse (National Center for Victims of Crime, n.d.). Research that addresses psychological harm caused specifically by victimization by CSAM strongly suggests that these individuals are more prone to experiencing heightened vulnerability to chronic psychiatric conditions. This finding highlights the importance of long-term, holistic care, i.e., addressing the harm systematically as opposed to acutely (Page et al., 2025).

Interventions that exist at the moment continue to be challenged in light of the impacts of generative AI abuse material (AAP, 2025). AI-CSAM contains complex data, often rooted deeply within AI algorithms, fueling repeated revictimization (IWF, 2024; Steel, 2024). The current evidence-based intervention models lack sufficient capability to provide holistic and effective victim support (Ciardha et al., 2025).

Mitigation and Prevention Efforts

EU Efforts

Despite the current urgency of this problem, it is not a recent issue and different governments and political unions have tried to address it. The European Union (EU) adopted a law named Directive 2011/93/EU to combat the sexual abuse and exploitation of children. Given the key role of the internet in the spread of CSAM, this directive replaced the earlier Council Framework Decision 2004/68/JHA and aimed to combine strong criminal measures with victim-centred protection, prevention, and international cooperation to ensure that children are safeguarded from sexual abuse and exploitation offline, but also online (European Parliament & Council of the European Union, 2011). The main goal of this directive is to ensure that a broad range of child-related offences, such as sexual abuse, online grooming, and the creation or distribution of CSAM, are consistently prevented, investigated, and punished across the EU (European Parliament

& Council of the European Union, 2011). The directive also places strong emphasis on victim protection, including child-friendly judicial procedures and safeguards to reduce further trauma (European Parliament & Council of the European Union, 2011). Additionally, this law also calls for prevention measures such as public awareness efforts, education for parents and caregivers, training for professionals, and programs to reduce reoffending through treatment and monitoring (European Parliament & Council of the European Union, 2011). The directive expands the EU's jurisdiction by allowing its members to prosecute crimes committed abroad if the offender or victim is an EU citizen, resident, or national. This change helps prevent offenders from avoiding justice because of legal loopholes (European Parliament & Council of the European Union, 2011). Although these legal improvements were made, member states did not uniformly apply the rules and new challenges also appeared, especially related to digital technologies, like the spread of child sexual abuse material through encrypted and AI-assisted platforms (European Commission, 2024). This led to a proposal (introduced in 2024), which takes the form of a recast directive, consolidating and updating the existing law which calls for stronger cooperation between national authorities, Europol, and the planned EU Centre to Prevent and Counter Child Sexual Abuse, and would play a coordinating role, facilitate data sharing, and help EU members identify, remove, and prevent the spread of illegal content (European Commission, 2024). For example, through widening the scope of criminal law to include new and emerging forms of abuse, like the creation and distribution of AI-CSAM. This proposal passed by unanimity in 2025 and is currently being negotiated among the Parliament, the Council of the EU, and the European Commission to determine the final shape of the law.

In 2024, the European Union approved the AI Act to ensure that AI systems in Europe are safe, transparent, and respect fundamental rights

such as privacy, equality, and human dignity (European Parliament & Council of the European Union, 2024). This regulation classifies AI systems according to four risk levels and has strict standards for each of those levels, as well as specific obligations for all generative AI systems (European Parliament & Council of the European Union, 2024). The four risk levels are unacceptable risk (includes content that threatens safety or fundamental rights and is banned), high risk (encompasses content with the potential of seriously impacting health, safety or fundamental rights and is strictly regulated), limited risk (includes content that poses moderate risk due to deception issues but has no direct threat to safety or fundamental rights and requires compliance with transparency obligations) and minimal or no risk (comprehends content that poses no threat to fundamental rights, safety, or privacy and is unregulated) (European Parliament & Council of the European Union, 2024). For example, high-risk AI systems need to pass conformity checks before entering the market. Breaking the law can result in heavy penalties (European Parliament & Council of the European Union, 2024).

UK Efforts

In the United Kingdom, the production or manipulation of child sexual abuse material (CSAM) using artificial intelligence is already prohibited under existing laws. The Protection of Children Act 1978 criminalises the creation, possession, and distribution of indecent photographs or pseudophotographs of children, while the Coroners and Justice Act 2009 extends this prohibition to non-photographic images, including drawings, cartoons, and animations (UK Government, 1978; UK Government, 2009). However, the rise of AI-generated content has created new enforcement challenges. As synthetic images become increasingly indistinguishable from real images, content moderators and law enforcement struggle to determine whether the material is criminal (Internet Watch Foundation,

2023). This can slow down removal and investigation compared to traditional CSAM (Internet Watch Foundation, 2023). To address these challenges, the UK government introduced reforms through the Crime and Policing Bill, specifically under the “Measures to Combat Child Sexual Abuse” section (Papaleontiu et al., 2025). This legislation establishes new criminal offences targeting the creation, possession, or distribution of digital models or files used to generate photorealistic CSAM, as well as the provision or maintenance of online services, including dark web forums, that facilitate such activity (UK Government, 2025). By doing so, the law closes a loophole that offenders previously exploited through AI technologies (UK Government, 2025). Additionally, this reform also enhances the supervision of registered sex offenders, strengthens police powers, tightens notification requirements, and restricts name changes to prevent offenders from concealing their identities (UK Government, 2025).

US Efforts

In the United States, the PROTECT Act (Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today Act) was implemented. This act represents a comprehensive federal effort to prevent and punish child sexual exploitation by establishing mandatory minimum sentences for the production, distribution, or possession of CSAM and, similarly to the EU Directive, expands federal jurisdiction to include offences involving U.S. citizens or residents, enabling cross-border investigations and strengthening international cooperation (United States Congress, 2003). In addition, closely to what is being done in the UK and EU, this act enhances preventive and support measures as it also mandates the creation of a national sex offender registry, integrated with state systems, requiring offenders to maintain up-to-date information, and the allocation of federal resources to support victims, providing counseling, legal assistance, and trauma-informed care (United States Congress, 2003).

Moreover, the US is also making federal efforts to implement the STOP CSAM Act, which restructures the federal framework to broaden protections for victims who testify in court, to require providers to report CSAM to a national reporting system within 60 days and prohibit these providers from storing or facilitating CSAM, and to expand related criminal and civil penalties for who fails to comply with these requirements (passed Senate Judiciary 2025, pending; United States Congress, 2025b). Furthermore, regarding AI-CSAM specifically, the US is also focused on implementing the ENFORCE Act, which aims to strengthen legal protections for children victims and hold offenders accountable for creating or distributing this type of content by ensuring that AI-CSAM creators are prosecuted and penalised under existing child sexual abuse laws (Senate-passed Dec 2025, House pending; United States Congress, 2025a).

General Prevention Efforts

On a different note, AI companies are also collaborating with governments and researchers to help address these challenges. Stanford AI researchers recommend that classifications also include whether images were AI-generated to help law enforcement identify real victims among millions of fake images (Krishna et al., 2024). They also urge AI companies to train their models not to create CSAM and bake watermarks into AI-generated images to aid the justice system in finding and prosecuting predators (Krishna et al., 2024). Promising training data models that attempt to identify and eliminate CSAM, such as the LAION-5B, are starting to emerge and constantly being reviewed and updated (Thiel, 2023). LAION-5B (used to train diffusion models) does not contain the images themselves, but metadata (URLs, captions, hashes, and embeddings) for approximately 5 billion image-text pairs (Thiel, 2023).

On another note, different tech groups that represent companies such as OpenAI, TikTok, Snapchat and Meta have made independent and joint

statements in which they say they are joining forces with governments for the first time to tackle the spread of CSAM and the danger posed by AI for child sexual abuse. One good example of this being in practice is the lawsuit that Meta took against a company that ran advertisements on its platforms due to them promoting an app that allows people to create nonconsensual and sexualized images of others using AI technology, avoiding their regulation, which prohibits this (Gerken, 2025).

Furthermore, on a more individual level, research has focused on the efficacy of different available treatments and therapies in managing paraphilias or sexual-related disorders and problematic pornography use and related behaviours that could stem from addiction or hypersexuality, revealing that cognitive-behavioural therapy and pharmacological treatment demonstrate positive outcomes in some cases, especially when combined (Garcia & Thibaut, 2011; Kaplan & Krueger, 2012; Minarcik, 2016; Roza et al., 2023). However, it is key to highlight that these claims have multiple limitations, such as methodological limitations, underreporting of sexual offences, sample variability, inconsistencies across treatment interventions and poor evidence of (long-term) effective treatment (Kaplan & Krueger, 2012; McManus et al., 2013). Nonetheless, any therapies or treatments are only used to complement the strong legal action, victim safeguarding, and public health response measures mentioned before.

Legal Gaps and Detection Challenges

Despite the progress that different legislations have made in criminalizing different forms of CSAM, there are still significant legal and practical gaps when it comes to AI-generated material. These gaps weaken the effectiveness of existing legal frameworks, which make certain enforcement and prosecution processes difficult. Furthermore, as legislation

is often reactive, new loopholes and gaps will likely continue to emerge as technology, especially AI, evolves.

General Legal Challenges

There are general challenges that all countries still face, including questions about who is culpable for such a complex offense and certain definitional challenges. AI CSAM does not always include an identifiable “real” child, which has historically led some to view it as less harmful. However, research suggests that it can reinforce abusive fantasies, it relies on training datasets with real CSAM, and contributes to normalizing child sexual abuse (Dushi et al., 2025; Parti & Szabó, 2024). Furthermore, this problem can be even less certain for non-photorealistic material (e.g., anime), where it can be unclear whether a character should be qualified as a “child” (Van den Bergh, 2024). Especially given the definitional challenges of assigning age to a non-existent figure.

Criminal intent standards may still need to be adjusted, as multiple actors can be seen as contributing to the crime (Dushi et al., 2025; Van den Bergh, 2024). These include users, model trainers and tuners, and general platforms. Furthermore, liability issues, specifically regarding mens rea, can arise if a user claims not to be fully in control of the output, depending on how a model is trained (e.g., reckless or negligent use of AI could produce CSAM images without the intent to do so). This coincides with the debate about criminal liability, as it is questionable whether AI companies or hosting platforms should carry some liability, as they enable or fail to mitigate the abuse in addition to the end users creating the prompts.

Lastly, procedural gaps persist, which slow prosecutions. Jurisdiction questions arise when AI CSAM is generated in one country, shared in another, and hosted on servers elsewhere. This was seen in Europol’s Operation Cumberland, where 19 countries identified 273 suspects across borders but faced challenging prosecutions due to legislative gaps (Europol,

2025; Internet Watch Foundation, 2026). Furthermore, offenders use VPNs or anonymous tools. Overall, this means that costly digital forensics are necessary, while international cooperation would need to take place more quickly.

EU Legal Gaps

In the EU, there is currently a transitional and interpretive gap, because the Directive 2011/93/EU predates modern generative AI and the recast Directive is still pending. While the current Directive's language suggests coverage of non-actual depictions ("realistic images," "simulated conduct," "any person appearing to be a child"), it does not explicitly mention digitally or AI-generated content, creating interpretational uncertainty (Dushi et al., 2025). Courts may extend "simulated" to cover AI-generated depictions, but this broad interpretation does not have definite legal certainty (Dushi et al., 2025). This leads to inconsistent implementation across EU States, with some having tighter definitions, while others include all AI CSAM (Kokolaki & Fragopoulou, 2025). As of now, the 2024-25 recast proposal explicitly addresses AI CSAM, but until it is implemented, transitional inconsistencies and interpretive gaps persist. In addition to this, the temporary ePrivacy derogation that had allowed platforms to voluntarily detect and report CSAM expired on 3 April 2026, creating a further gap in the EU's detection framework (European Parliament, 2026).

Interaction with the EU AI Act

Although the EU AI Act (phased implementation from 2024) introduces a definition of deepfakes that regulates certain high-risk AI systems, loopholes still remain. Dushi and colleagues (2025) point out a disconnect between the AI Act and recast Directive 2011/93/EU. While the recast Directive 2011/93/EU criminalizes possession of (AI) CSAM, the AI Act does not classify models capable of generating or reproducing CSAM as

“high risk” systems. This AI Act label results in these systems not needing to undergo pre-market conformity assessments or detailed risk documentation (Dushi et al., 2025). This allows providers to comply formally with the AI Act (e.g., labelling outputs as AI-generated) while still producing illegal content. Therefore, law enforcement may face uncertainty when it comes to accountability, questioning whether it lies with the model provider, the platform hosting the model, or the end-user generating the image (Dushi et al., 2025).

Gaps outside of the EU (UK and US)

Regarding other countries, outside of the EU, similar gaps and inconsistencies appear. In the UK, new measures such as the ongoing UK Crime and Policing Bill 2025 (Ministry of Justice; 2025) address AI CSAM, but it may be difficult to draw a line between general-purpose models and those “optimized” for CSAM, making it challenging to determine when development of these models should be criminalized.

In the USA, federal efforts like the STOP CSAM Act (passed Senate Judiciary 2025, pending; United States Congress, 2025b) and the ENFORCE Act (Senate-passed Dec 2025, House pending; United States Congress, 2025a), which address the creation and possession of AI CSAM in the USA. However, legal standards differ across specific states depending on what is classified as obscene (Schidlow et al., 2025). Overall, these inconsistencies create a jurisdictional loophole for offenders, especially as this content is often distributed across borders.

Detection Challenges

Detection of AI CSAM faces further challenges as well. First and foremost, the volume and speed of the generation of AI CSAM that is created overwhelms current detection systems. Furthermore, traditional detection approaches are poorly adjusted to detecting AI-CSAM. Overall,

there is debate over whether encrypted chats should be scanned on the sender's side via client-side scanning (ENR et al., 2025).

Perceptual hashing, a method used to identify known CSAM, does not work with AI content because each output is unique and does not match existing hash databases (WeProtect Global Alliance, 2024). Watermarking tools have been suggested as an alternative, but these also fall short, as offenders can easily crop or recolor watermarks from the image, or conversely insert false watermarks to mislead investigators (Scholten, 2025). Although detectors are lagging behind generators, big corporations are responding, for example, Google is attempting to watermark everything created with AI (SynthID), which will be invisible (Google DeepMind, n.d.).

As AI CSAM becomes more photorealistic, investigators are having increasing difficulty distinguishing synthetic from real abuse imagery. This leads to limited resources being invested in verifying authenticity rather than focusing on rescuing victims (Internet Watch Foundation, 2024a). This increasing realism can be seen in IWF's 2024 analysis of dark-web forums, where 90% of AI-CSAM images were photorealistic enough to be assessed under existing UK legal standards. At the same time, production barriers are low, and offenders can generate and distribute large quantities of AI-CSAM at minimal cost and skill. This combination of amount and realism overwhelms law enforcement.

A Future Perspective: Research Recommendations

Offender Motivation & Psychological Mechanisms

Offender motivations, whether on the part of the creator or the consumer, require further research. This would be important because understanding why people justify or minimize their behaviour helps prevention efforts target specific mechanisms.

First and foremost, research needs to continue to study the cognitive and moral frameworks that shape both creator and consumer behaviour. Some studies have shown that the assumption that 'no real child means no harm' is not corroborated by research as it has been evidenced that it re-victimises children and increases the risk of escalation for consumers of CSAM. However, research still fails to corroborate the professional's presumption that this belief among creators and consumers diminishes guilt emotions, especially in AI-CSAM, possibly leading to more individuals engaging with this content that would not otherwise. It is unclear whether this belief is a deliberate self-justification or a genuine belief that synthetic content does not cause harm. Furthermore, there is no research on how these misperceptions could be emerging or being spread. Understanding these perceptual distortions is important for designing awareness strategies that directly challenge this narrative. Debunking this belief could help people understand the actual harm and legal consequences connected to AI CSAM, and may prevent individuals who might otherwise drift into this material from continuing once those misconceptions are corrected.

Regarding the creators of AI CSAM, research still needs to address the broader social and developmental mechanisms that drive their behaviour, even though online communities, technical prestige, and status-seeking dynamics are now beginning to be documented. While it is clear that forums include tutorial-sharing, feedback loops, and reputational hierarchies, we have only limited understanding of how social learning, peer guidance, and community normalization shape creators' motivations and engagement over time. It would be important to investigate whether these dynamics interact with status-seeking incentives, which have been studied in traditional CSAM contexts, and whether social rewards operate independently or amplify financial and sexual motives. At the same time, youth-driven creation is a new and understudied phenomenon, with no longitudinal evidence on

whether adolescent creators remain situational or transition toward more offending patterns. Understanding how social rewards, peer instruction, and early behavioural markers interact is key to designing prevention strategies that interrupt these pathways before they solidify.

Regarding the consumers of AI CSAM, additional research is necessary to continue to establish patterns that allow the creation of target groups with similar profiles and motivators. This would allow the improvement of preventive interventions as these groups would account for both the individual and external factors that can cause this variability and the problem of generalization. Simultaneously, it is key to further investigate the link between addiction and CSAM consumption. Research has focused on pornography addiction, especially on people with hypersexuality symptoms, however, it fails to address CSAM excessive consumption specifically, focusing only on adult pornography. Therefore, it is key that this link is explored further as it is expected that there will also be a link between addiction and “child pornography”. Thus, most of the existing studies on pornography addiction are focused on hypersexual people, which only represent a part of the population that engages with this content, and existing research does not cover specifically sexual content that is not performed by adults.

Harm & Victimology

The body of research examining the unique victimology of persons targeted by AI-CSAM remains scarce. The consequences of the offence-specific circumstances, that is, the possibility that an AI-CSAM case has no direct identifiable victim, as well as the emerging evidence of CSAM being embedded within AI training datasets, are largely unknown. Future research should focus on the ways AI-manipulation harms existing victims, alongside the perpetuation of trauma stemming from the online presence of the content, vulnerability to psychiatric conditions, long-term psychological

impacts, and secondary victimization of close relatives. Given the strongly gendered nature of AI CSAM (girls comprising 97% of identified content, IWF, 2026), research should examine how these harms manifest across the gender dimension, including the distinct vulnerabilities of boys targeted through sextortion (AAP, 2025). Additionally, research examining the effectiveness of current victim support frameworks, including prompt access to justice, training of law enforcement, and holistic interventions minimizing harm, is needed.

Technology & Model Misuse

Although recent findings have considerably improved our knowledge on how AI technologies are being abused for creation of CSAM content, critical gaps remain. Systematic research beyond currently available observational data is needed on the specific processes, technical pipelines, and offender communities involved in exploitation of open-source diffusion models (Steel, 2024). Apart from this, the integrity of AI training material requires attention. Recent research has confirmed the existence of CSAM and other harmful content embedded in large AI databases, with offenders actively sharing illegal material for purposes of training (IWF, 2026), raising questions about the adequacy of existing safeguards and accountability schemes for when those safeguards fail. Furthermore, additional phenomena requiring immediate attention have been identified: convergence of image generation, chatbot technology, and audio deepfakes into single platforms lowering barriers to offending; the rising popularity of nudify applications and its impact on children; and the likely underreporting of AI-CSAMs true scale (IWF, 2026; Internet Matters, 2024). The foundation for all these phenomena lies in the necessity of developing reliable mechanisms for detection and distinguishing AI-generated content from photographic material, which is crucial for both measuring the scale of the issue and effective victim identification (Ciardha et al., 2025; Steel, 2024).

Additionally, the potential impact of AI agentic models, currently being discussed in offender communities as a means to automate creation and distribution, warrants comprehensive investigation. (Ciardha et al., 2025; IWF, 2026).

Brief conclusion

Altogether, the present report addresses the ever-evolving concern of AI-CSAM, shedding light on the creation tools, motivations of creators and consumers, impact on victims, mitigation efforts, and legal and detection challenges. These findings highlight the importance of developing a policy framework and practical checklist for organizations that structure how AI tools and datasets are used and created. This can be done by critically analyzing existing initiatives and adapting and creating new or improved versions that will add to or complement the current shortcomings. Additionally, prioritizing preventative rather than reactive approaches would be essential.

Furthermore, this paper identified concrete research gaps crucial to guide targeted interdisciplinary efforts in attempts to mitigate harm and inform prevention mechanisms. This serves as a stepping stone for other professionals to have some guidance on what is key to address in future research, as these efforts are foundational for safeguarding children and disrupting the exploitative ecosystems enabling widespread misuse of artificial intelligence.

References

American Academy of Pediatrics. (2025). The impact of deepfakes, synthetic pornography, & virtual child sexual abuse material.

<https://www.aap.org/en/patient-care/media-and-children/center-of-excellence-on-social-media-and-youth-mental-health/qa-portal/qa-portal-library/qa-portal-library-questions/the-impact-of-deepfakes-synthetic-pornography--virtual-child-sexual-abuse-material/>

Aronashvili, R. (2024). *The Evolution Of Sextortion Attacks: How Generative AI Is Taking A Front Seat*. Forbes. Retrieved December 12, 2025, from

<https://www.forbes.com/councils/forbestechcouncil/2024/02/20/the-evolution-of-sextortion-attacks-how-generative-ai-is-taking-a-front-seat/>

Ba, Z., Zhong, J., Lei, J., Cheng, P., Wang, Q., Qin, Z., Wang, Z., & Ren, K. (2024). SurrogatePrompt: Bypassing the safety filter of text-to-image models via substitution. In *Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1166–1180). ACM. <https://doi.org/10.1145/3658644.3690346>

Barroso, R., Silva, S., Fishere, M., Nentzl, J., Vo, T. N., Forero, C. G., Gómez-Durán, E. L., Ferreira, C. B., Gieseler, H., Westfal, V., Franch-Martínez, B., Krejčová, L., & Beier, K. M. (2025). Child sexual abuse material (CSAM): a systematic review of risk profiles, risk factors, and

typologies of users. *Sexual Medicine Reviews*, 14(1).

<https://doi.org/10.1093/sxmrev/qeaf081>

Borah, S. K., Ramaswamy, S., & Seshadri, S. (2025). The online specter: Artificial intelligence and its risks for child sexual abuse and exploitation. *Journal of Indian Association for Child and Adolescent Mental Health*, 21(2), 107–112.

<https://doi.org/10.1177/09731342251334293>

Brown, R. (2023). Prevalence of viewing online child sexual abuse material among Australian adults. *Trends & Issues in Crime and Criminal Justice*, No. 682.

[https://www.aic.gov.au/sites/default/files/2023-](https://www.aic.gov.au/sites/default/files/2023-12/ti682_prevalence_of_viewing_online_csam_among_australian_adults.pdf)

[12/ti682_prevalence_of_viewing_online_csam_among_australian_adults.pdf](https://www.aic.gov.au/sites/default/files/2023-12/ti682_prevalence_of_viewing_online_csam_among_australian_adults.pdf)

Brown, Rick & Bricknell, S.. (2018). What is the profile of child exploitation material offenders?. *Trends and Issues in Crime and Criminal Justice*, No. 564.

<https://www.researchgate.net/publication/329800466>

CameraForensics. (2024, February 8). The dark reality of Stable Diffusion. <https://www.cameraforensics.com/blog/2024/02/08/the-dark-reality-of-stable-diffusion>

Christensen, L. S., Moritz, D., & Pearson, A. (2021). Psychological perspectives of virtual child sexual abuse material. *Sexuality &*

Culture, 25(4), 1353–1365. <https://doi.org/10.1007/s12119-021-09820-1>

Ciardha, C. Ó., Buckley, J., & Portnoff, R. S. (2025). AI generated child sexual abuse material - What's the harm? *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2510.02978>

Crawford, A. & Smith T. (2023). Illegal trade in AI child sex abuse images exposed, In BBC News. <https://www.bbc.co.uk/news/uk-65932372>

Davy, D. & Lundrigan, S. (2024). Artificial intelligence-produced child sexual abuse material: Insights from Dark Web forum posts. International Policing & Public Protection Research Institute (IPPPRI). <https://www.cambridgenetwork.co.uk/sites/default/files/IPPPRI-Insight-No-1-AI-CSAM.pdf>

De Alarcón, R., De La Iglesia, J. I., Casado, N. M., & Montejo, A. L. (2019). Online Porn Addiction: What We Know and What We Don't—A Systematic review. *Journal of Clinical Medicine*, 8(1), 91. <https://doi.org/10.3390/jcm8010091>

Downing, S.,. (2024). *Unveiling the challenges of AI-generated CSAM*. <https://www.cameraforensics.com/blog/2024/09/06/unveiling-the-challenges-of-ai-generated-csam/>

Dushi, D., Berdufi, N., & Karagianni, A. (2025). The legal framework and legal gaps for AI-generated child sexual abuse material.

Computer Law & Security Review, 59, 106205.

<https://doi.org/10.1016/j.clsr.2025.106205>

ENR with AFP, ANSA, CTK, Belga, dpa, EFE, LUSA, PAP, & STA. (2025, October 15). Privacy vs. child protection: EU's "chat control" plans split member states. *European Newsroom*.

<https://europeannewsroom.com/privacy-vs-child-protection-eus-chat-control-plans-split-member-states/>

European Child Sexual Abuse Legislation Advocacy Group & DOT Europe. (2025, June). *Joint ECLAG–DOT Europe position paper on artificial intelligence-generated child sexual abuse material*.

<https://doteurope.eu/wp-content/uploads/2025/06/Joint-ECLAG-DOT-Position-Paper-AI-CSAM.pdf>

European Commission. (2024). *Proposal for a directive of the European Parliament and of the Council on combating the sexual abuse and sexual exploitation of children and child sexual abuse material, and replacing Council Framework Decision 2004/68/JHA (recast)* (COM/2024/60 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2024%3A60%3AFIN>

European Parliament. (2026, March 26). *Child sexual abuse online: Voluntary detection measures will not be extended* [Press release]. <https://www.europarl.europa.eu/news/en/press->

[room/20260325IPR39207/child-sexual-abuse-online-voluntary-detection-measures-will-not-be-extended](#)

European Parliament & Council of the European Union. (2011).

Directive 2011/93/EU of the European Parliament and of the Council of 13 December 2011 on combating the sexual abuse and sexual exploitation of children and child pornography, and replacing Council Framework Decision 2004/68/JHA. Official Journal of the European Union, L 335, 1–14. <http://data.europa.eu/eli/dir/2011/93/oj>

European Parliament & Council of the European Union. (2024).

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Official Journal of the European Union, L. <https://data.europa.eu/eli/reg/2024/1689/oj>

Europol (2025, February 28). *25 arrested in global hit against AI-generated child sexual abuse material*. Europol. Retrieved April 9, 2026, from <https://www.europol.europa.eu/media-press/newsroom/news/25-arrested-in-global-hit-against-ai-generated-child-sexual-abuse-material>

Gannoni, A., Voce, A., Napier, S., Boxall, H., & Thomsen, D. (2023). *Preventing child sexual abuse material offending: An international review of initiatives*. <https://doi.org/10.52922/rr78764>

Garcia, F. D., & Thibaut, F. (2011). Current concepts in the pharmacotherapy of paraphilias. *Drugs*, 71(6), 771–790. <https://doi.org/10.2165/11585490-000000000-00000>

Gerken, T. (2025, June 12). *Meta urged to go further in its crackdown on "nudify" apps*. <https://www.bbc.com/news/articles/cgr58dlne5o>

Gewirtz-Meydan, A., Walsh, W., Wolak, J., & Finkelhor, D. (2018). The complex experience of child pornography survivors. *Child Abuse & Neglect*, 80, 238–248. <https://doi.org/10.1016/j.chiabu.2018.03.031>

Google DeepMind. (n.d.). *SynthID: A tool to watermark and identify content generated through AI*. <https://deepmind.google/models/synthid/>

Hilton, D. L., & Watts, C. (2011). Pornography addiction: A neuroscience perspective. *Surgical Neurology International*, 2(1), 19. <https://doi.org/10.4103/2152-7806.76977>

Huikuri, S. (2023). Users of online child sexual abuse material. *Journal of Police and Criminal Psychology*, 38(4), 904–913. <https://doi.org/10.1007/s11896-023-09611-4>

Insoll, T., Ovaska, A. K., Nurmi, J., Aaltonen, M., & Vaaranen-Valkonen, N. (2022). Risk Factors for Child Sexual Abuse Material

Users Contacting Children Online: Results of an Anonymous Multilingual Survey on the Dark Web. *Journal of Online Trust and Safety*, 1(2). <https://doi.org/10.54501/jots.v1i2.29>

Internet Crime Complaint Center. (2023, June 5). *Malicious actors manipulating photos and videos to create explicit content and sextortion schemes*. <https://www.ic3.gov/PSA/2023/psa230605>

Internet Matters. (2024). The new face of digital abuse: Children's experiences of nude deepfakes. <https://www.internetmatters.org/wp-content/uploads/2024/11/Childrens-experiences-of-nude-depfakes-research.pdf>

Internet Watch Foundation. (2023). *How AI is being abused to create child sexual abuse imagery* [Report]. https://www.iwf.org.uk/media/q4zll2ya/iwf-ai-csam-report_public-oct23v1.pdf

Internet Watch Foundation. (2024a). *What has changed in the AI CSAM landscape?* [Report]. https://www.iwf.org.uk/media/nadlcb1z/iwf-ai-csam-report_update-public-jul24v13.pdf

Internet Watch Foundation. (2024b, July). *How AI is being abused to create child sexual abuse imagery*. <https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/>

Internet Watch Foundation. (2025, July 11). *Full feature-length AI films of child sexual abuse will be inevitable as synthetic videos make huge leaps in sophistication in a year*. <https://www.iwf.org.uk/news-media/news/full-feature-length-ai-films-of-child-sexual-abuse-will-be-inevitable-as-synthetic-videos-make-huge-leaps-in-sophistication-in-a-year>

Internet Watch Foundation. (2026). *Harm without limits: AI child sexual abuse material through the eyes of our Analysts*. AI CSAM Report 2026. <https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/>

Internet Watch Foundation, Child Helpline International, INHOPE, National Center for Missing & Exploited Children, Offlimits, Safe Online, & WeProtect Global Alliance. (2026, February 10). *No to Nudify* [Joint statement]. <https://www.mentalhealtheurope.org/wp-content/uploads/2026/02/no-to-nudify-statement-eng.pdf>

Johnson, S. A. (2020). Child porn users & risk for engaging in contact offenses: Faulty data minimizes offender's risk & puts more children at risk for sexual abuse. *Forensic Research & Criminology International Journal*, 8(2), 93–99.
<https://doi.org/10.15406/frcij.2020.08.00311>

Johnson, S. A. (2025). Child sexual abuse material (CSAM): understanding the problem, the offender, & risk concerns. *Forensic Research & Criminology International Journal*, 13(1), 44–57.

<https://doi.org/10.15406/frcij.2025.13.00436>

Kaplan, M. S., & Krueger, R. B. (2012). Cognitive-behavioral treatment of the paraphilias. *The Israel journal of psychiatry and related sciences*, 49(4), 291–296.

Karaagac, M., (2025, October 23). *Diffusion models: Mechanism, benefits, and types (2025)*.

<https://www.archivinci.com/blogs/diffusion-models-guide>

Kaul, A. (2024, April 22). *Generative AI is supercharging the creation of child exploitation content*. Resolver.

<https://www.resolver.com/blog/generative-ai-supercharging-csam-creation/>

Kim, S., & Banks, J. (2024). Expanding experiences and anxieties: gender-identity and sexual-orientation differences in attitudes toward synthetic pornography. *Porn Studies*, 1–19.

<https://doi.org/10.1080/23268743.2024.2400945>

Knack, N., Holmes, D., & Fedoroff, P. (2020). Motivational pathways underlying the onset and maintenance of viewing child pornography on the Internet. *Behavioral Sciences & the Law*, 38(2), 147–162.

<https://doi.org/10.1002/bsl.2450>

Kokolaki, E., & Fragopoulou, P. (2025). *Unveiling AI's threats to child protection: Regulatory efforts to criminalize AI-generated CSAM and emerging children's rights violations* (arXiv:2503.00433). arXiv.

<https://doi.org/10.48550/arXiv.2503.00433>

Kopecký, K., & Voráč, D. (2025). The phenomenon of deep nudes—A new threat to children and adults. *AI & SOCIETY*.

<https://doi.org/10.1007/s00146-025-02425-4>

Krishna, S., Dubrosa, F., & Milanaik, R. (2024). Rising threats of AI-Driven child sexual abuse material. *Pediatrics*, 153(2).

<https://doi.org/10.1542/peds.2023-063954>

Lahtinen, H., Honkalampi, K., Insoll, T., Nurmi, J., Quayle, E., Ovaska, A. K., & Vaaranen-Valkonen, N. (2025). Investigating the disparities among child sexual abuse material users: Anonymous self-reports from both charged and uncharged individuals. *Child Abuse & Neglect*, 161, 107299. <https://doi.org/10.1016/j.chiabu.2025.107299>

Lätth, J., Joleby, M., McMahan, A., Luke, T. J., & Rahm, C. (2025). Child Sexual abuse material users on the darknet: Psychiatric morbidities related to offence behavior. *Sexual Abuse*.

<https://doi.org/10.1177/10790632251347562>

Love, T., Laier, C., Brand, M., Hatch, L., & Hajela, R. (2015).

Neuroscience of Internet Pornography Addiction: A Review and

update. *Behavioral Sciences*, 5(3), 388–433.

<https://doi.org/10.3390/bs5030388>

Lucy Faithfull Foundation. (2024). A call to end AI-generated child sexual abuse. *Lucy Faithfull Foundation*.

<https://www.lucyfaithfull.org.uk/a-call-to-end-ai-generated-child-sexual-abuse/>

McManus, M. A., Hargreaves, P., Rainbow, L., & Alison, L. J. (2013). Paraphilias: definition, diagnosis and treatment. *F1000Prime Reports*, 5. <https://doi.org/10.12703/p5-36>

Minarcik, Jennifer. (2016). Proposed Treatment of Problematic Pornography Use: A Cognitive-Behavioral Approach. 10.13140/RG.2.2.19375.84649.

Ministry of Justice. (2025). *Equality impact assessment [EIA] – Crime and Policing Bill: Child sexual abuse and violence against women and girls measures*.

[https://assets.publishing.service.gov.uk/media/692f07d9ce50d215cae96385/Crime and Policing Bill EIA -
_Child Sexual Abuse and Violence Against Women and Girls.pdf](https://assets.publishing.service.gov.uk/media/692f07d9ce50d215cae96385/Crime_and_Policing_Bill_EIA_-_Child_Sexual_Abuse_and_Violence_Against_Women_and_Girls.pdf)

Mink, J., Qin, L., & Redmiles, E. M. (2026). “Unlimited Realm of Exploration and Experimentation”: Methods and Motivations of AI-Generated Sexual Content Creators (arXiv:2601.21028). arXiv.

<https://doi.org/10.48550/arXiv.2601.21028>

Napier, S. S., Seto, M. C., Cashmore, J., & Shackel, R. (2024).

Characteristics that predict exposure to and subsequent intentional viewing of child sexual abuse material among a community sample of Internet users. *Child Abuse & Neglect*, 156, 106977.

<https://doi.org/10.1016/j.chiabu.2024.106977>

National Center for Missing & Exploited Children. (2025). *2024 CyberTipline report*.

<https://www.missingkids.org/gethelpnow/cybertipline/cybertiplinedata>

National Center for Victims of Crime. (n.d.). Effects of child sexual abuse on the victim. <https://victimsofcrime.org/effects-of-csa-on-the-victim/>

Page, S., & McFeeters, S. (2025). Psychological and emotional harms of online child sexual exploitation and abuse: A systematic review. *Child Abuse Review*, 34(4), e70048.

<https://doi.org/10.1002/car.70048>

Papaleontiu C., Deputy Director & Tackling Child Sexual Abuse Unit (2025). Measures to Combat Child Sexual Abuse. In *Crime and Policing Bill: Equality Impact Assessment – Child Sexual Abuse and Violence Against Women and Girls Measures*.

<https://assets.publishing.service.gov.uk/media/686fc62381dd8f70f5d>

[e3e11/Crime and Policing Bill EIA -](#)

[_Child Sexual Abuse and Violence Against Women and Girls.pdf](#)

Parti, K., & Szabó, J. (2024). The Legal Challenges of Realistic and AI-Driven Child Sexual Abuse Material: Regulatory and Enforcement Perspectives in Europe. *Laws*, 13(6), 67.

<https://doi.org/10.3390/laws13060067>

Prat, S., & Jonas, C. (2012). Characteristics of French users of child pornography: Description of a cohort subjected to forensic evaluation. *Medico-Legal Journal*, 80(4), 162–166.

<https://doi.org/10.1258/mlj.2012.012029>

Protect Children (2021). CSAM Users in the Dark Web: Protecting Children Through Prevention. Suojellaan Lapsia. Available online at: <https://www.suojellaanlapsia.fi/en/post/csam-users-in-the-dark-web-protecting-children-through-prevention>

Protect Children. (2026). CSAM Perpetrator Research Report: Findings from a Survey of CSAM Perpetrators on Digital Platform Use and Design (Tell Me More About Tech).

<https://www.protectchildren.fi/en/post/tmat-csam-perpetrator-research-report>

Ray, A., & Henry, N. (2025). Sextortion: A Scoping Review. *Trauma, Violence & Abuse*, 26(1), 138–155.

<https://doi.org/10.1177/15248380241277271>

Reichel, R., Daser, A., Gnielka, F. M., Schmidt, A. F., Blokland, A., & Lehmann, R. J. B. (2024). A review of risk factors for online and mixed child sexual abuse material offending: what is being researched? *Journal of Sexual Aggression*, 1–44.

<https://doi.org/10.1080/13552600.2024.2418100>

Roza, T. H., Noronha, L. T., Shintani, A. O., Massuda, R., Lobato, M. I. R., Kessler, F. H. P., & Passos, I. C. (2023). Treatment Approaches for Problematic Pornography Use: A Systematic Review. *Archives of Sexual Behavior : The Official Publication of the International Academy of Sex Research*, 53(2), 645–672.

<https://doi.org/10.1007/s10508-023-02699-z>

Schidlow, J., Gaitis, K. K., Lu, M., Stevenson, J., & Fry, D. (2025). Legal challenges in tackling AI-generated child sexual abuse material within the USA (Report). *Childlight GCSI*.

https://www.pure.ed.ac.uk/ws/portalfiles/portal/535227053/5-EYES-REPORT_USA.pdf

Scholten, C. (2025, June 23). *Semantic watermarks for AI image recognition can be easily manipulated*. Ruhr-Universität Bochum News. <https://news.rub.de/english/press-releases/2025-06-23-computer-science-semantic-watermarks-ai-image-recognition-can-be-easily-manipulated>

Seigfried-Spellar, K. C., & Rogers, M. K. (2013). Does deviant pornography use follow a Guttman-like progression? *Computers in Human Behavior*, 29(5), 1997–2003.

<https://doi.org/10.1016/j.chb.2013.03.012>

Seto M. C. (2017). The Motivation-Facilitation Model of Sexual Offending. *Sexual Abuse*, 31(1), 3–24.

<https://doi.org/10.1177/1079063217720919>

Steel, C. M. S. (2024). Artificial intelligence and CSEM - A research agenda. *Child Protection and Practice*, 2, 100043.

<https://doi.org/10.1016/j.chipro.2024.100043>

Szyf, N. D., Gilen, A., Giacometti, M., Van de Heyning, C., & Walrave, M. (2023). *Deepnudes among young people in Belgium: The numbers, the market, the impact*. University of Antwerp & Institute for the Equality of Women and Men. [https://igvm-iefh.belgium.be/sites/default/files/deepnudes among young people i
n belgium.pdf](https://igvm-iefh.belgium.be/sites/default/files/deepnudes_among_young_people_in_belgium.pdf)

Thiel, D. (2023). Identifying and Eliminating CSAM in Generative ML Training Data and Models. *Stanford Digital Repository*.

<https://doi.org/10.25740/kh752sm9123>.

Thiel, D., Stroebel, M., & Portnoff, R. (2023). Generative ML and CSAM: Implications and mitigations. *Stanford Digital Repository*.

<https://purl.stanford.edu/jv206yg3793>

Thorn. (2025a, March 3). *Deepfake nudes & young people: Navigating a new frontier in technology-facilitated nonconsensual sexual abuse and exploitation*. Thorn.

<https://www.thorn.org/research/library/deepfake-nudes-and-young-people/> <https://www.thorn.org/research/library/deepfake-nudes-and-young-people/>

Thorn. (2025b, July 7). *How Thorn makes the internet safer & helps stop the cycle of abuse*. Thorn. <https://www.thorn.org/blog/how-thorn-makes-the-internet-safer-helps-stop-the-cycle-of-abuse/>

Thorn & All Tech Is Human. (2024). *Safety by design for generative AI: Preventing child sexual abuse*. <https://info.thorn.org/hubfs/thorn-safety-by-design-for-generative-AI.pdf>

TRM Labs. (2025, March 28). *The evolving CSAM landscape: Vendors increasingly leveraging AI as they return to the dark web*. TRM Labs. <https://www.trmlabs.com/resources/blog/the-evolving-csam-landscape-vendors-increasingly-leveraging-ai-as-they-return-to-the-dark-web>

UK Government. (1978). *Protection of Children Act 1978*. <https://www.legislation.gov.uk/ukpga/1978/37>

UK Government. (2009). *Coroners and Justice Act 2009*. <https://www.legislation.gov.uk/ukpga/2009/25/contents>

UNICEF. (2026, February 4). *'Deepfake abuse is abuse'* [Press release]. <https://www.unicef.org/press-releases/deepfake-abuse-is-abuse>

United Nations Office on Drugs and Crime. (2023). *Background paper: Expert Group Meeting on removal of child sexual abuse material from the internet* [PDF].

https://www.unodc.org/pdf/criminal_justice/endVAC/EGM/EGM_CSAM_Removal_Background_Paper.pdf

United States Congress. (2003). *PROTECT Act of 2003*, Pub. L. No. 108-21, 117 Stat. 650.

<https://www.congress.gov/108/plaws/publ21/PLAW-108publ21.pdf>

United States Congress. (2025a). *ENFORCE Act*, H.R. 4831, 119th Congress. <https://www.congress.gov/bill/119th-congress/house-bill/4831>

United States Congress. (2025b). *STOP CSAM Act of 2025*, S. 1829, 119th Congress. <https://www.congress.gov/bill/119th-congress/senate-bill/1829/text>

WeProtect Global Alliance. (2024, September 11). *AI-produced child sexual abuse material: Insights from Dark Web forum discussions*. We Protect Global Alliance. <https://www.weprotect.org/blog/ai-produced-child-sexual-abuse-material-insights-from-dark-web-forum-discussions>

WeProtect Global Alliance. (2025). *Artificial intelligence and 'Gen AI'*.

<https://www.weprotect.org/thematic/artificial-intelligence-and-gen-ai/>

WeProtect Global Alliance & Columbia University. (2025). *Global*

Threat Assessment 2025: Key findings.

[https://www.weprotect.org/global-threat-assessment-25/key-
findings/](https://www.weprotect.org/global-threat-assessment-25/key-findings/)

Wolbers, H., Cubitt, T., & Cahill, M. (2025). Artificial intelligence and
child sexual abuse: A rapid evidence assessment. In *Australian*

Institute of Criminology eBooks. <https://doi.org/10.52922/ti77802>

Van den Bergh, C (2024). *Legislation in the age of innovation:*

*Regulating AI-driven child sexual abuse material in the European
Union. Fact or fiction?*

<https://repository.gchumanrights.org/handle/20.500.11825/2848>