Child Sexual Exploitation Material: Investigative and Legal Challenges with Generative Artificial Intelligence

Chad M.S. Steel [a]

[a] George Mason University, Fairfax, Virginia, US

Corresponding Author:

Chad M.S. Steel, csteel@gmu.edu +1-610-639-3884. MS 2B5, George Mason University, Fairfax, VA 22030.

Abstract

Generative Artificial Intelligence (GenAI) has become ubiquitous in the past few years based on advances in both algorithms and computing power. As the technology has proliferated, it has begun to be used to create Child Sexual Exploitation Material (CSEM), and its use has created new categories of offenders and offenses; resulted in new, emerging categories of victims; and presented both theoretical and conceptual legal challenges. This work explores the investigative, digital forensics and legal challenges posed by the proliferation of GenAI related to CSEM offenses. The GenAI technology ecosystem is detailed through a psychological lens, and the current research related to GenAI in CSEM offending is summarized. The impact of this ecosystem is then examined from the perspective of the United States as it directly relates to the challenges posed to investigations and digital forensics, the legal challenges, and the impact to victimology. Current trends are examined, including ongoing investigative and prosecutorial efforts, and underexplored areas for further research are highlighted.

Keywords: Generative Artificial Intelligence, Online Child Sexual Exploitation, Child Sexual Exploitation Material, Cybercrime investigations

In May 2024, the United States Department of Justice charged Steven Anderegg with using Stable Diffusion, a popular generative AI (GenAI) tool, in creating over 13,000 images of child sexual exploitation material (CSEM) (Verma & Harwell, 2024). This case represents the first major investigation where an individual was charged with creating images offline (on a standalone machine), and where no real children were *directly* exploited (though the images were shared over the Internet). With the proliferation of tools like Stable Diffusion, which came under scrutiny after the inclusion of known CSEM in its training data (Levine, 2023), the landscape for CSEM offending is undergoing a dramatic shift. The implications for investigative, digital forensics and prosecutorial efforts is a rapidly evolving area of both current research and practical interest (Moritz, 2023; Pfefferkorn, 2024).

GenAI a subset of artificial intelligence dedicated to the creation of content, relies on training models using deep learning, a form of multilayer neural network technology[1]. To generate text, large language models (LLMs) such as ChatGPT are trained on millions or billions of pages of existing content in order to first understand and then generate new text-based content (Johri, 2023). Text-to-image GenAI such as DALL-E and Stable Diffusion are further trained on large numbers of labeled images, allowing them to take text prompts and generate new images based on those prompts (*DALL·E 3*, 2024; Lee et al., 2024). This has further been extended to video, with tools such as Sora and Runway, that are able to create very short (for now), high resolution video clips (Schwartz, 2024). In addition to the creation of novel content, many image generation tools have image-to-image capabilities, which allow for the transformation of existing content (e.g., changing poses) as well as the insertion of selected content (e.g., the face of another individual) seamlessly into other content, generated or pre-existing, through a technique known as inpainting (Jam et al., 2021).

The amount of AI-generated CSEM is growing - the National Center for Missing and Exploited Children (NCMEC) received 4,700 reports in 2023, and warned of the possibility of the criminal justice reporting system being overwhelmed if their projections hold (Shehan, 2024). The current avenues of GenAI usage by offenders ranges from the creation of completely new CSEM to the modification of existing CSEM to the alteration of SEM images using innocent images of minors. Additionally, GenAI developers face their own challenges - creating tools that are resistant to the generation of CSEM, testing tools (red teaming) in a legal manner, and ensuring their training data don't include already known CSEM (Levine, 2023).

Because of the growth in usage and novel affordances created by GenAI, new investigative, forensic, and prosecutorial challenges are emerging. From an investigative perspective, immense amounts of high-resolution CSEM can be locally generated, offenders are more likely to include minors, and new modalities of offending (e.g., sextortion using deepfakes) are occurring. Forensically, transient content (including that using Augmented Reality), victim identification (discerning real victims from deepfakes), and the lack of support in forensic tools for AI content are present. Prosecutorially, existing laws have gaps in areas ranging from victim compensation to the coverage of new offending methods; the increasing number of underage offenders are creating prosecutorial discretion issues; current statutes don't address issues

[1] For a general overview of GenAI technologies, including large language models, see (Kalota, 2024).

such as inadvertent generation of CSEM during testing; and extant case law is sparse but rapidly evolving.

This paper reviews the current affordances offered by GenAI as it relates to CSEM offending based on the limited existing research, early caselaw and public investigative reports. Based on these affordances, the challenges present throughout the investigative and legal lifecycle are presented. Finally, a series of potential near-term recommendations to address these challenges are offered.

Current GenAI Affordances

This work focuses on usage of GenAI by CSEM offenders on the generation of content, both image and video-based. Other uses, including the generation of text-based content and the malicious usage of chat models that can facilitate visual CSEM offenses are outside the scope of this paper but addressed in other, recent work (Steel, 2024).

1. **Creation of new CSEM-focused datasets.** Perhaps the largest overall risk in GenAI related to CSEM is the creation of a model trained on a large CSEM dataset. To train a model from scratch, most current datasets use tens of millions of images. Individuals have been arrested previously with sufficient images to train a CSEM-specific model in this fashion (Spocchia, 2021)), and once the model is trained it could be used by anyone to create new content. Additionally, the desire to train a model may cause individuals to collect large amounts of CSEM (or communities of offenders to pool resources). Finally, other AI models could be subverted to generate sufficient CSEM to train a custom model based on synthetic training data (e.g., models trained with innocuous child images and adult pornography).
2. **Modification of existing datasets.** While training a model requires tens of millions of images, *boosting* an existing model can be done with orders of magnitude fewer images. Using techniques such as Low-Rate Adaptation (LoRA), existing parameters can be fine-tuned with smaller datasets (e.g., thousands of images) to generate custom content, and techniques such as Dreambooth can even hyperfit existing images generation models to allow for the generation of new images based on a single subject (Hu et al., 2021; Ruiz et al., 2022). These techniques can permit offenders to create a custom CSEM-generation model based on smaller, narrowly targeted datasets for individual use (based on freely downloadable models) or for distribution to specialized CSEM communities.
3. **Modification of models to remove safeguards.** The more responsible industry leaders providing GenAI services have signed on to a series of design principles to reduce the risk of CSEM offenders misusing their tools (Thorn, 2024). Included in these principles are the implementation of guardrails that filter out offending content, either by blocking queries used by offenders or by detecting offending images before providing them to users (or, ideally, both). While this is a significant barrier for GenAI-as-a-service as well as closed source tools, offenders can remove guardrails included in open source AI tools or even subvert them (i.e., only showing images that post-generation AI analysis

flags as offending). This, alongside other benefits including lower risks of detection and the ability to perform custom training, encourages the use of local AI generation by CSEM offenders using modified, open source tools.

4. **Generation of new images from existing series.** One of the key drivers for a small but significant subset of CSEM offenders is a drive to collect all of the content in a specific series (Quayle & Taylor, 2002; Steel et al., 2021). Because legacy series have a fixed amount of abusive content that was produced, these collectors may use GenAI tools to create new content for an existing series. These may be used for individual purposes, or shared with other collectors to establish credibility. Additionally, this creates a new, more damaging form of ongoing victimization beyond the sharing of previously available content.
5. **Generation of novel, hyperspecialized content.** Prior to the Internet explosion of CSEM availability, most content consumption was likely opportunistic instead of preferential based on available materials (Merdian et al., 2013). The advent of GenAI allows for highly customizable, highly targeted content generation. Specific ages, body types, acts, and settings can be created with limitless variations, changing the dynamics of consumption. Additionally, creative prompt engineering may become a core skill set for CSEM offenders as it has for other domains (Marvin et al., 2024), potentially overriding the acquisition of novel content as a key offender knowledge area.
6. **Generation of physically infeasible content**. While the generation of hyperrealistic content is possible, so is the generation of fantastical content. Current genres such as hentai and manga have a history of including minors in them (and in some cases being used as grooming material) (Eelmaa, 2022). GenAI allows for the application of fantastical scenarios previously depicted in certain hentai and manga to be depicted with photorealistic imagery. Additionally, hentai and manga previously required sophisticated artistic abilities to generate - with GenAI, offenders can now create their own content, which was a production avenue previously unavailable. As with the hyperspecialized content, this may result in a dramatic increase in the availability of this content as generation tools become more widespread.
7. **Generation of inpainted images from innocuous images.** Before GenAI, CSEM offenders were able to create content from innocuous images of known minors depicted in still images using tools such as Photoshop, though doing so required significant time and effort. With GenAI, individuals have been found to use images such as yearbook photos to easily create realistic CSEM content of unwitting victims (Tucker, 2024). This has impacted not only direct, sexual exploitation of children, but also enhanced the ability of offenders to commit other crimes such as sextortion and revenge pornography offenses, which no longer require the victim to provide images but can be fully generated by the offender (and used against both minors as well as adults who can be inpainted as offenders) (*Criminals Using A.I. to Alter Images for Sextortion Schemes, State Police Warn*, 2023).

Investigative Implications

GenAI has necessitated significant changes to law enforcement's approach to CSEM investigations. First, the demographics of offending may be shifting, with subclasses of offenders who are also minors on the rise. Second, the increasing capabilities of locally run AI create investigative challenges. Finally, existing risk models may need to be changed based on GenAI-only offenders.

In addition to the general investigative challenges, victim-specific challenges are becoming more prevalent with GenAI. Of primary importance in non-generative AI, victim identification may be meaningless with fully AI-generated images, and complicated with inpainted images. Additionally, transient victims may be present with augmented reality, where there is no permanent record creating ongoing victimization. Finally, the proportion of unaware victims is likely to increase, requiring a change in approach for investigators when interacting with these individuals.

An unanswered question related to AI-generated CSEM is the demographics of the user base. While comprehensive usage data is not currently available (and would be expected to evolve over time), there is reason to believe it may skew younger compared to traditional CSEM production (Steel, 2025). First, the recent predominance of self-produced CSEM highlights a changing trend and potential attitude shift within the under-18 demographic (Finkelhor et al., 2023). Second, initial reports of the use of GenAI tools in school settings are increasing (Cruz, 2024; Shehan, 2024; Sosa, 2024), though it is unknown if this is due to the novelty of the technology or an actual change in overall CSEM increases. These present both investigative prosecutorial challenges.

From an investigative perspective, law enforcement have previously been hesitant to investigate non-consensually shared CSEM offenses where both parties are minors and the originating content consensually generated, barring aggravating circumstances (Dodge & Spencer, 2018), though there have been more recent instances involving GenAI sharing where investigations and prosecutions have occurred (Jones, 2024). The victims themselves in non-consensual image sharing may prefer to deal with the events at a peer-level (Dodge & Lockhart, 2022), and may not want to cooperate with investigators (Dodge & Spencer, 2018), though in many of these cases the victims were originally consensual participants in the generation of the material. With GenAI, victims of inpainting or nudification may have differing attitudes as there may be no direct involvement in the content creation. Additionally, parental pressure for investigative action may be high from the victim's families, despite a potential lack of prosecutorial merit or victim interest in pursuing.

Legal and ethical requirements for victim notification are likewise unclear, especially when pertaining to unknowing victims. Under United States law, investigators are required to provide notice to victims (Crime Victims' Rights Act, 2004), though the definition of a victim needs clarification for GenAI cases. If a victim's innocuous images are used to train GenAI, or if their likeness is used as a seed image (but not present in the resulting, offending images), the applicability of the law is unclear. The boundaries for investigators in terms of both notification and potential victim engagement (e.g., interviews of the victims and their caregivers) are

likewise unclear, and ethical issues arise if investigative actions may cause unnecessary traumatization. Identifying victims in these cases may not even be possible - once trained, GenAI models cannot be easily reverse-engineered to find source images (unless the training data are found to be present in forensic examinations), and for augmented-reality violations, such as those using nudification software on smartphones, images may only exist for the duration of the viewing.

Detection efforts by investigative agencies need further enhancement as well. Substantial success in detecting CSEM on peer-to-peer, web, dark web, and social media (National Center for Missing and Exploited Children, 2022; Panchenko et al., n.d.; Peersman et al., 2016; *Victim Identification Solutions & Partnership*, 2023) will need to be enhanced and redirected as a shift toward AI-based offending occurs. Identification and detection of prompt engineering terms (which will differ from existing CSEM keyword lists), detection of generated CSEM by providers, and incorporation of monitoring/reporting for new areas of exploitation such as the Telegram-based botnets used to create nudes (Vincent, 2020) are some of the areas in which investigation and detection efforts lag offender adoption.

Changing investigative protocols to respond to the new affordances, and specific training on AI for Internet Crimes Against Children (ICAC) investigators is additionally overdue. Strong baseline research on crossover offending for AI-centric offenders, cognitions of those offenders, and whether or not they are a distinct subgroup are unanswered research questions (Steel, 2024). Understanding these factors will drive investigative approaches to interviews in terms of offense-supportive cognitions and contact offending questions (Paquette & Cortoni, 2019); case prioritization (McManus et al., 2011); and the use of risk assessment tools such as CPORT to identify high risk offenders (Seto & Eke, 2015).

In addition to enhancing the understanding of AI offender cognitions to drive investigative interviewing, statutory requirements for AI-generated CSEM may require additional investigative interview questions or actions. One potential avenue for charging in the United States is under the concept of “attempt”, which does not require the images to be real, only that the subject believed them to be real. As such, if the subject is a consumer and/or distributor (as opposed to a producer) of content, showing that the subject believed the images to be real, either through their statements or actions, can solidify later prosecutorial efforts.

The local nature of CSEM crimes could feasibly lead to unchargeable circumstances, even with current technologies. Local GenAI implementations could, in theory, become a perverse Nozick machine[2] for offenders, able to generate infinite amounts of content on-demand, facilitating continuous novelty-seeking. Additionally, if persistence is needed, offenders could save queries alongside of seed values in lieu of actual content (or have tools do so), creating a scenario

[2] A Noizick machine is a thought expericise put forth by Robert Nozich where individuals are presented a choice between living in a simulated environment that provides pleasure or everyday reality.

where the content does not exist until the “link” is clicked[3]. These present both detection problems (the content stays local or is transient) and a lack of a statutory violation to investigation (see below - there is no “interstate commerce” involved).

While direct CSEM investigations generally involve the content itself, newer crime modalities such as sextortion do not necessarily require the same elements. For GenAI-based sextortion, particularly that involving images of adults, two additional non-CSEM options are available, depending on the context. If anything of value is solicited, federal extortion laws can be used, which only require investigators to show that money (or a thing of value) was demanded over the Internet, and that the AI content would have caused a reputation loss (Interstate Communications, 1948). For obscene GenAI content sent to minors, there is no requirement that the images be indistinguishable from a minor, only that they be obscene. In these cases, investigators only need to show that the image was obscene and that it was sent via the Internet to an individual under the age of 16 (Transfer of Obscene Material to Minors, 1998a).

Digital Forensics Implications

Traditional digital forensics in CSEM cases involves searching for offending images; determining how the individual obtained those images (or produced them) and whether or not they further distributed them; and identifying possible victims that may have been previously unknown. While many of the basic digital forensics actions are the same, there are significant, additional steps needed with the advent of GenAI which require changes to both methodology and tools. Identifying likely AI images can allow for appropriate prioritization of victim identification efforts, and can potentially indicate production offenses. The following digital forensics activities now need to be conducted on all CSEM cases:

1. **Perform EXIF (Extensible Image File Format) scanning.** Many of the GenAI tools add metadata to the EXIF information present in common image formats, including JPEG (in the “User Comment” field) and PNG (in the “Parameters” field) files (Thiel et al., 2023). Support in forensics suites to look explicitly for these tags and highlight them is in process, but exporting all EXIF information using tools such as Exiftool (Prior, 2023) and manually scanning for the names of common AI tool suites can be done while the integrated forensics suites catch up. Not all AI tool suites add metadata (and not all filetypes will have associated EXIF data), and both individuals and websites may strip out metadata (Tanner et al., 2013), but if present it provides a strong indicator that an image was AI generated. Similarly, specific AI tools may use default file naming conventions (similar to digital cameras), though comprehensive lists of these flags are not yet readily available in AI suites (Pasquini et al., 2021).
2. **Identify local AI software or access to web-based AI tools.** Detecting AI generated content from the content itself is difficult. Current techniques, including both spatial and frequency domain approaches, have shown poor detection capabilities, ranging from

[3] This creates a legal challenge as well if cloud-based - do a series of prompt/seed links fed into a commercial GenAI tool constitute a violation as the content doesn’t exist until law enforcement (or an offender) clicks the link.

50% to 90% detection rates (with constrained data), which is insufficient for large forensics datasets, but can be used to risk-rank potential AI images (Corvi et al., 2023). Because of this, secondary signals of AI usage can be identified. Web history can be utilized to identify GenAI site visits, as well as AI trading forum visits. Additionally, local AI implementations (e.g., Stable Diffusion) as well as tools with integrated AI (e.g., Photoshop) can be identified which may increase the prior probability that AI-generated CSEM is present.

3. **Utilize facial recognition and age detection approaches.** Traditionally, CSEM digital forensics efforts utilize databases of mathematical signatures known as hashes generated from previously identified CSEM. These hashes either match exact file content (e.g., SHA-256), or are resilient and match content that may have been resized, cropped, or had a format change (Microsoft, 2009). Institutions such as the National Center for Missing and Exploited Children (NCMEC) and the Internet Watch Foundation (IWF) maintain databases for law enforcement use based on previously seen CSEM. With the ability to generate AI on-demand, hash-based approaches are less viable (except to differentiate content with known victims), and more advanced techniques need to be used. Age detection algorithms built into forensic tools (themselves using AI) can be used to identify AI generated CSEM (*Victim Identification Solutions & Partnership*, 2023), but may need to be adapted specific to this problem set - images that detect only facial characteristics may fail when adult faces are placed on the bodies of minors, and those that detect body characteristics may fail when the faces of minors are placed on adult bodies. Holistic, AI-based age detectors will likewise need to be trained on both fully generated and altered CSEM images, creating a legal challenge for researchers. Additionally, tools such as BANE (Westlake et al., 2022), which identify faces and cluster them, can be used to identify AI-modified content using real victims. Because of the ability for AI to iterate on existing images, using facial recognition databases from currently known victim series may be effective as well. Individuals who become fixated on a particular individual or series are now able to create “new” content using the same victims, allowing for indefinite secondary revictimization that would not be detectable using existing hashset approaches.

In addition to the new actions, current forensics techniques need to adapt to generated CSEM. Actions to establish mens rea need to include initial prompt engineering inputs as well as iterations on those inputs. From a psychological perspective, tracking these actions over time may highlight specific preferences previously unidentified. Similarly, current forensics tools will need to incorporate the above approaches, and new tools developed as the technology further evolves.

Finally, digital forensics will need to address new defenses presented through the use of AI. One of the common defenses in computer crimes is the SODDI defense[4], which frequently relies on claims of malware or computer compromise being responsible for criminality (Steel, 2014). GenAI has now created the AI SODDI defense - that offending images of contact offenses are not real, or were generated by third parties (potentially as part of a sextortion claim), or were not

[4] “Some Other Dude Did It”.

children, or were the result of poort programming based on innocuous prompts. CSEM investigators now need to incorporate countermeasures for this defense into their investigative strategies, which may include additional digital forensic examinations to find supporting scienter evidence, comparison of unclothed versions of the actual individuals to the allegedly generated content (necessitating forensic photography), the use of AI detection tools as they become more viable, and including the possibility of the defense into interview questions.

Legal Implications

Current CSEM-specific legislation in the United States was predicated on the presence of underlying physical assault - i.e., an actual child was sexually abused. In New York v. Ferber, the court held that CSEM images were "intrinsically related to the sexual abuse of children", but acknowledged that "harm to the child is exacerbated by their circulation" (*New York v. Ferber*, 1982). In Ashcroft v. Free Speech Coalition, the court struck down the legal prohibition on virtual CSEM (which included the language "is, or appears to be, of a minor"), noting that if the images were not obscene and did not involve actual children, they were protected by the First Amendment, though the court did make a distinction between fully generated virtual images and *morphed* images which involve, at least partially, a real child (*Ashcroft v. Free Speech Coalition*, 2002). The US Congress passed the PROTECT Act of 2003 in response, changing the language to make it illegal to possess (or create or distribute) "a digital image, computer image, or computer-generated image that is, or is indistinguishable from, that of a minor" (Sexual Exploitation and Other Abuse of Children, 1978).

It has been argued that the current US statutes related to child pornography may not withstand legal scrutiny when applied to AI-generated CSEM due to its definitional weaknesses to the same arguments that formed the basis of the Ashcroft decision (Pfefferkorn, 2024). There are inherent differences, however, in the "virtual child pornography" of 2002 and current AI-generated imagery. First, virtual CSEM in 2002 was created with tools such as Photoshop with significant manual intervention by individuals with graphic arts skills, and no real children were generally involved (as distinguished from morphed images). With current GenAI, the tools have all been trained *with actual children*, and while they may not reflect underlying physical abuse, their resemblance to children in the training set may do harm through circulation of the images. Second, the court in Ashcroft intentionally carved out "morphed" images as not being part of their decision (*Ashcroft v. Free Speech Coalition*, 2002), and these are covered by a separate definition in US statute, specifically making images illegal that have "been created, adapted, or modified to appear that an identifiable minor is engaging in sexually explicit conduct" (Sexual Exploitation and Other Abuse of Children, 1978). Inpainted GenAI would fall under this categorization, as would any images from GenAI trained on previously known CSEM. Arguably, GenAI trained on a dataset including innocuous images of real children as well as sexually explicit images of adults could be considered part of this definition as well. Despite these arguments, much of AI-based CSEM is expected to be prosecuted as obscenity, instead of under child pornography statutes, which is a separate exception to free speech.

A final exception to free speech arguments is present under the pandering portion of the child pornography statute. Pandering makes it illegal for anyone who "attempts or conspires to violate" the statute (Certain Activities Relating to Material Constituting or Containing Child Pornography, 1996). This potentially sidesteps the First Amendment issue and the definitional problems of the "indistinguishable from" language, and allows for the charging if the producer/recipient reasonably believes the image to be that of a minor. For AI-generated CSEM, the context in which it was acquired (e.g., a forum dedicated to sex with minors) or produced (e.g., using prompts that would indicate the intent) would allow for charging under this section.

While the child pornography statutes were modified with the phrase "indistinguishable from", the obscenity statute includes even more direct language, stating that "It is not a required element of any offense under this section that the minor depicted actually exist" (Obscene Visual Representations of the Sexual Abuse of Children, 2003). The change in language allowed for the inclusion of both morphed images of real children, as well as wholly new AI-generated images. In the case of morphed images, while there is no primary exploitation there is still a child victim of secondary exploitation. Additionally, it can be argued that there are indirect victims of AI-generated images (those children whose images are used in the dataset), as noted above. A final case, which has no child victimization, are images where adults are de-aged (either consensually, for age-play, or non-consensually). While these images are technically illegal under federal law, there have been no cases yet where these factors have been challenged.

The obscenity statute is distinguished from the child pornography statutes in that it requires the material to be "obscene", making it excluded from First Amendment protections. Obscenity requires content to be subjected to the Miller test, making it obscene if:

1. The average person, applying contemporary community standards' would find that the work, taken as a whole, appeals to the prurient interest;
2. The work depicts or describes, in a patently offensive way, sexual conduct specifically defined by the applicable state law; and
3. The work, taken as a whole, lacks serious literary, artistic, political, or scientific value. ("Miller v. California," 1973)

While there have been instances where simplistic drawings were not considered obscene ("United States v. Arthur," 2022), most AI-generated CSEM, particularly that which is intended to be photorealistic, is likely to qualify, and prior challenges have found that the obscenity statute is neither overbroad nor vague ("US v. Buie," 2019). In particular, showing the *intent* behind the creation, either through investigative or forensic techniques, can generally overcome claims that the content was intended to be non-prurient and created explicitly for its artistic merits.

The obscenity statutes have additional differences from the statutes explicitly targeting child pornography. Repeat offenders, for example, have a tripling of their penalties for convictions under the child pornography statutes, but the same does not apply to obscenity convictions

(Repeat Offenders, 1998), though they do have a mandatory 10 year penalty if the person is a registered sex offender (Penalties for Registered Sex Offenders, 2006).

In terms of sex offender registration, the federal sex offender registration law itself does not explicitly mention the obscenity statute, but has a broad catch-all of “Any conduct that by its nature is a sex offense against a minor” (Relevant Definitions, Including Amie Zyla Expansion of Sex Offender Definition and Expanded Inclusion of Child Predators, 2006), and this has been applied to “morphed” images involving the faces children placed on the bodies of adults engaged in sexual conduct that were charged as obscenity (*People v. Lewis*, 2024). While public support for both sex offender registration and the illegality of virtual CSEM are strong (Steel et al., 2022), there has been no research into support for sex offender registration for virtual CSEM offenses.

Both the child pornography and obscenity statutes have limitations that are relevant to self-produced CSEM. Specifically, there is a need for there to be an interstate commerce nexus. This is generally not an issue with cloud-based AI web services, but may create prosecutorial difficulties with content that is generated using locally installed tools (Pfefferkorn, 2024). Constructivist arguments that the instruments (the AI software or the training sets) were obtained through interstate commerce, as well as conspiracy-based charges against those software providers (Conspiracy to Commit Offense or to Defraud United States, 1948) have yet to be litigated.

AI-generated images will have an unknown impact on victim restitution as well. In the United States, victims are entitled to general restitution for injury or property damage (Mandatory Restitution to Victims of Certain Crimes, 1982). For child pornography offenses, additional restitution is available, covering the full amount of the victim’s losses, including:

> (A)medical services relating to physical, psychiatric, or psychological care;
> (B)physical and occupational therapy or rehabilitation;
> (C)necessary transportation, temporary housing, and child care expenses;
> (D)lost income;
> (E)reasonable attorneys’ fees, as well as other costs incurred; and
> (F)any other relevant losses incurred by the victim.(Mandatory Restitution, 1994)

The victims are additionally entitled to restitution from a defined victim’s fund, providing some compensation when convicted offenders are indigent or otherwise unable to pay. The definition of victim is broad enough to cover minors whose images are used to create modified AI content, but there are several areas not explicitly covered. Adult victims of de-aged images, as well as minors whose innocuous images were used in training a GenAI model for illicit purposes are not clearly covered and there is no extant caselaw to clarify these situations. Additionally, the shift toward charging virtual CSEM offenses under the obscenity statute (Obscene Visual Representations of the Sexual Abuse of Children, 2003) as opposed to the child pornography statutes has an impact on victim restitution. Offenses charged under this statute are not

explicitly mentioned in the child pornography compensation law, limiting the compensation options for victims.

A final complication exists under US law specific to individuals for which it may not be in the best interests of the government to charge[5] based on the new use cases presented through AI-generated CSEM. Self-generated imagery that is not shared and consensually created was noted above, however this is likely to be a rare event. More critically, two current situations are more prevalent that require direct consideration - the charging of minors who engage in AI-related CSEM offenses and the need for tool providers to have a safe harbor to test their applications and.

In general, charging minors with creation of CSEM where both parties are consenting has been avoided, with the Department of Justice highlighting education and prevention as key strategies, however there can be exacerbating circumstances that favor prosecution (Department of Justice, 2023). The amount of teen-to-teen, consensual CSEM generation has already increased as a result of teen sexting, and GenAI is expected to exacerbate that (O'brien & Hadero, 2023). Work by Thorn showed that, in 2023, one in ten minors knew someone who had used AI tools to generate CSEM of other minors. With this increase in prevalence, prosecutorial decisions will need to be made for an increasing number of offenders under the age of 18. While the vast majority of these cases are better served through education and prevention, several factors may weigh toward prosecution in specific cases:

1. **Age of the offender.** Specific obscenity laws already differentiate individuals under 16 with those 16 years of age or older [e.g., (Transfer of Obscene Material to Minors, 1998b)], but consideration of prosecution should be higher for individuals closer to 18 than those younger for other GenAI CSEM violations.
2. **Age differential with the victim.** From both a risk and consent perspective, the greater the age difference between the offender and victim the greater the need for prosecution. Many locales have carve-outs for physical relationships between older teens and adults to avoid nonsensical prosecutions (e.g., prosecuting an individual who is exactly 18 years old having intercourse with another who is 17 years and 364 days old). AI-generated CSEM prosecution decisions would benefit from similar considerations.
3. **Non-consensual generation or sharing.** Even if AI-generated CSEM is consensual, the circumstances of distribution may not be. An individual may use a de-nudification application on themselves, or generate images of another minor (either inpainted or denuded) with their consent. If there is no consent, or if there is non-consensual distribution, these should be considered aggravating factors.
4. **Coercion and/or sextortion.** Coercion is a sliding scale, ranging from a single request to provide images for GenAI CSEM use to full-on extortative behavior. The degree of coercion used should be a factor in charging minors, particularly if there is non-consensual generation of AI images as part of a sextortion offense.

[5] There are and will be, of course, other complications, including novel defenses that have yet to be proposed or litigated.

5. **Financial motives.** Similar to the commercial factors in other criminal violations, a minor that sells AI-generated images of peers or pays individuals for consensual, innocent images that are then modified has a different motivation than age-appropriate sexual expression. The seeking or paying of any monies should factor into proseuctorial decisions.
6. **Presence of malice.** As with many crimes, the intent of the offender can be important. A minor generating AI CSEM of a peer for "curiosity" purposes should be less likely to be charged than an individual doing so as part of a "revenge porn" scheme.
7. **Prior offending.** With many offenses, including traditional CSEM offenses where the offender is a previously registered sex offender, there are enhanced penalties. Similarly, for minors engaged in GenAI CSEM offenses, an initial caution may be warranted, but subsequent offenses should take into account the initial warning for prosecution.

Aside from the challenges of potentially large numbers of minors being engaged in AI-related CSEM offenses, there are other groups that may have issues in the current legal environment. GenAI tool developers are one area in need of special consideration. There is a societal desire for providers to develop tools that are appropriately tested and have guardrails in place to protect against their malicious use (Thiel et al., 2023). Unfortunately, under the current law, there are no legislative safe harbor provisions that would permit the testing of these tools through red-team analysis[6]. These provisions could be incorporated, however, into existing statutory carve-outs for individuals. Both the child pornography and obscenity statutes have similar affirmative defenses that would apply to anyone inadvertently generating AI CSEM (e.g., a person using the prompt "cheese pizza", for example, might generate inappropriate content with a poorly trained tool). Better tool testing, and a requirement for testing, would reduce this risk, but the existing laws already allow an affirmative defense for anyone who:

(1) possessed less than 3 such visual depictions; and
(2) promptly and in good faith, and without retaining or allowing any person, other than a law enforcement agency, to access any such visual depiction—

(A)took reasonable steps to destroy each such visual depiction;(Obscene Visual Representations of the Sexual Abuse of Children, 2003)

Extending similar provisions to legitimate tool testing (either in partnership with law enforcement or independently), while not opening the door to inappropriate "research" or "testing" defenses by individuals ("Townshend Escapes Child Porn Charges," 2003), would be beneficial.

Recommendations

Work done by Stanford University and Thorn, in cooperation with major industry partners, has formed a framework for ethical AI implementations (Thorn, 2024), but it relies on voluntary compliance and doesn't address areas such as open source implementations that, by their nature, cannot include several of the recommended guardrails. Additionally, the research into AI

[6] The Department of Justice actively engages with industry on specific good-faith exceptions from prosecution, however.

technologies by offenders is in its infancy (Singh & Nambiar, 2024), and basic research is needed before applied research can be conducted. Finally, new affordances are arising rapidly as AI technologies become more mainstream, and investigators, forensics analysts, and prosecutors will need to continuously adapt to these new usage models. Based on the current technological and legal environment, however, there are several areas for improvement that are readily apparent. Key recommendations are as follows:

1. **Enhance digital forensics and detection through the development of new, shared capabilities.** Traditional services from sources such as NCMEC and the Internet Crimes Against Children databases allow for hash and fuzzy hash comparisons to identify previously known content (Liberatore et al., 2010; National Center for Missing and Exploited Children, 2020), and providers like the Internet Watch Foundation provide keyword lists targeted at both forensics and monitoring (*Hash List*, 2020). To perform more effective forensics, new key phrase lists related to prompt engineering to create CSEM, lists of known GenAI toolsets (both web-based and local), a repository of EXIF information for GenAI images, and better face recognition databases (to identify inpainted victims) as well as multimodal databases are needed (Westlake et al., 2022). These will require both industry and non-governmental organization action to create and maintain.
2. **Develop targeted training on both investigating and prosecuting AI CSEM offenses.** Training for online CSEM offenses has historically focused on peer-to-peer sharing, victim identification, crossover risks, and understanding the cognitions of online offenders. With the advent of sexting and non-consensually shared, self-generated CSEM (both consensual and non-consensually created), training moved toward messaging platforms and different modus operandi of offenses (e.g., sextortion). As GenAI-based CSEM is likely to become dominant in the near future, both investigators and prosecutors need training on the underlying technologies, recognizing the new modus operandi of offenders, determining if victimization has occurred, how the cognitions of these offenders are different, new defense strategies to avoid conviction, and what factors increase the risk of contact offending (e.g., inpainting of individuals known to the offender may be different than creating GenAI images of underage celebrities).
3. **Develop a new typology of GenAI-based CSEM offenders.** The existing typologies did not consider GenAI-based CSEM offenders, and a new typology of offenders, focused on segmenting by both risk and needed interventions, is required. Potentially distinct types of offending that need to be incorporated include:
    a. *Consumers*. Individuals that only consume GenAI CSEM content that others have created. These may follow individual creators or series, or may be more general consumers based on pre-existing preferences.
    b. *Personal User/Producers*. Users of GenAI tools may create personally consumed content, either locally or through Internet-based services. These individuals will utilize GenAI to create highly specialized or novel content from scratch for personal consumption, and may or may not store that content.

c. *Known Inpainters*. Using inpainting features or pose alteration features of tools, these offenders will utilize innocuous photos of individuals they know or have access to. Subtypes with different risk profiles may include minors modifying the photos of classmates or of adult caregivers modifying the photos of their wards.
d. *Unknown Inpainters.* Similar to the known inpainters, these individuals modify existing, innocuous images of minors to make them explicit or extant CSEM images of existing CSEM victims to create novel content. They potentially have a different risk profile than known inpainters in that they may be less likely to have opportunities for contact offending with their victims.
e. *Prompt Engineering Creators*. Developing content may require specialized prompt engineering skills that create stimulating material for a particular audience. Similar to existing 3D, computer generated adult SEM creators, prompt engineering-based producers may start to develop a following based on their skills at manipulating GenAI tools.
f. *On-demand Producers.* Using high performance computing (or distributed computing), these individuals may supplant current live streaming of abuse-on-demand by providing the same service through AI. Additionally, they may host services that are categorized by highly custom content in fantastical areas not readily available with real imagery.
g. *Content Curators*. One of the expected avenues that may become available with widespread availability of AI-generated CSEM is the curation of that content. Commercial and non-commercial distribution opportunities may arise for the categorization and labeling of the content for highly specialized consumer demands.
h. *Tool Creators.* Individuals may create either commercial or freely available tools (such as the Telegram botnets) that remove the clothing from existing images, create new images from prompts, or allow image inpainting. Tool creators may train specific models, or they may remove the safeguards from existing tools/models.
i. *Dataset Creators.* In addition to individuals who create tools, those with large amounts of pre-existing CSEM may use it to train CSEM-specific models, which provide a new illicit avenue for commercialization or distribution. These individuals may also collect AI images to re-train or boost existing models (e.g., through attended automation).
j. *Sextortionists*. Sextortionists generally have a specific goal, either financial or further sexual exploitation (which may be distinct subtypes of offenders). These offenders will use inpainting to generate offending content, and use it to extort either the pictured victims or secondary targets (e.g., threatening to distribute inpainted images of a real adult with minors).

4. **Create new statutes specific to AI-generated CSEM.** While the existing legal statutes have provisions incorporated that map to AI-generated CSEM, they were not created with the challenges noted above in mind. AI-specific statute(s), covering everything from intentional tool creation (to facilitate CSEM generation) to the use of augmented reality applications (e.g., nudification applications), with appropriate penalties as well as

carve-outs for testing is needed. The legislation could include specific provisions for self-generated CSEM that are age-based (e.g., similar to the transmission of obscenity to a minor statues, which require the offender to be an adult and the victim to be under 16), and could include enhanced compensation for victims of AI-generated CSEM offenses. Inclusion (or exclusion) of AI CSEM offenses from sex offender registration would also need to be addressed, ideally based on risk-based research. Current models like CPORT would need updating specific to AI offenses, but more baseline research is required before this is possible. Additionally, updating the sentencing guidelines to specific address GenAI CSEM would be required.

5. **Focus on prevention and education**. While legislative fixes and investigative enhancements are needed, government spending on prevention and education programs, particularly those targeted at minors, are likely to be the most effective approach. Raising awareness about the risks of using AI tools, and providing reasonable alternatives to minors, is likely to reduce future investigative and prosecutorial demands. The best approach to this prevention and education, specifically related to the use of AI in CSEM offending, is an ongoing research challenge, and programs will need to be evidence-based and avoid both shaming and infeasible approaches (e.g., avoiding AI).
6. **Creating a government-partnered testing model for industry.** The National Institute for Standards and Technology (NIST) creates testing standards and protocols for digital forensics tools, face recognition, and other technologies (Guttman, 2024). NIST should consider partnering with industry to create an approved testing suite for AI tools related to CSEM. Use of these tools by a provider, using approved protocols (e.g., ones that do not require viewing of the resultant content) that make use of automated AI detection, could be an answer to generating safer AI tools, providing safe harbor testing, and reducing liability for AI tool providers that engage in voluntary compliance.

Limitations

This paper addressed the challenges created by use of GenAI by CSEM offenders in the production and distribution of content. Offenders may use GenAI tools for other offensive-supportive behaviors, including facilitating cybergrooming. Additionally, there are defensive uses for AI that were outside the scope of this paper, including the use of GenAI to identify CSEM as well as assist investigators in detecting and responding to grooming behavior. Finally, chatbot uses of GenAI in deterring CSEM offending, either through displacement or treatment, are potential areas of further research (Pearson & Curtis, 2025).

Conclusions

The use of GenAI by CSEM offenders is in its infancy, and there are likely affordances that have yet to come to light. Meanwhile, the technology itself is improving, with the ability to generate high quality video with realistic audio, potentially with video in-painting, on personal devices in the near future. Similarly, the technology is becoming easier to use, increasing the number of offenders likely to engage with it. At the same time, the legal landscape is evolving through

case law as well as likely new legislative fixes, and investigative techniques are similarly adapting. This paper provided an overview of several of the key current, known challenges GenAI has brought to CSEM offending, and provided near-term recommendations to address them. Because this is a point-in-time analysis, it is only intended to start the conversation and provide time-bound guidance. Further research and re-evaluation will be needed commensurate with the march of progress of GenAI technologies. Additionally developers of GenAI technologies should exercise caution and consider potential offending usage related to CSEM when implementing controls on their software.

REFERENCES

Ashcroft v. Free Speech Coalition, 535 U.S. 234 (U.S. 2002). https://scholar.google.com/scholar_case?case=4016009721484982910

Certain Activities Relating to Material Constituting or Containing Child Pornography, 18 U.S.C. § 2252A (1996).

Conspiracy to Commit Offense or to Defraud United States, 18 U.S.C. § 371 (1948).

Corvi, R., Cozzolino, D., Zingarini, G., Poggi, G., Nagano, K., & Verdoliva, L. (2023). On The Detection of Synthetic Images Generated by Diffusion Models. *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.

Crime Victims' Rights Act, 18 U.S.C. § 3771 (2004). https://www.justice.gov/usao/resources/crime-victims-rights-ombudsman/victims-rights-act

Criminals using A.I. to alter images for sextortion schemes, state police warn. (2023, July 17). CBS Pittsburgh. https://www.cbsnews.com/pittsburgh/news/artificial-intelligence-alter-images-sextortion-schemes-warning/

Cruz, I. (2024, April 10). *Fairfax High School investigates inappropriate images shared online*. ABC7 Los Angeles. https://abc7.com/lausd-fairfax-high-school-probes-inappropriate-images-shared-online-bill-targeting-ai-child-porn-moves-through-california-legislature/14642201/

DALL·E 3. (2024). https://openai.com/dall-e-3

Department of Justice. (2023). *Sextortion, Crowdsourcing, Enticement, and Coercion*. https://www.justice.gov/d9/2023-06/sextortion_crowdsourcing_enticement_and_coercion_2.pdf

Dodge, A., & Lockhart, E. (2022). “Young People Just Resolve It in Their Own Group”: Young People’s Perspectives on Responses to Non-Consensual Intimate Image Distribution. *Youth Justice : The Journal of the National Association for Youth Justice*, *22*(3), 304–319.

Dodge, A., & Spencer, D. C. (2018). Online Sexual Violence, Child Pornography or Something Else Entirely? Police Responses to Non-Consensual Intimate Image Sharing among Youth. *Social & Legal Studies*, *27*(5), 636–657.

Eelmaa, S. (2022). SEXUALIZATION OF CHILDREN IN DEEPFAKES AND HENTAI. *TRAMES*, *XXVI*(2), 229–248.

Finkelhor, D., Turner, H., Colburn, D., Mitchell, K., & Mathews, B. (2023). Child sexual abuse images and youth produced images: The varieties of Image-based Sexual Exploitation and Abuse of Children. *Child Abuse & Neglect*, *143*, 106269.

Guttman, B. (2024). *Digital Forensics | NIST*. https://www.nist.gov/programs-projects/digital-forensics

Hash List. (2020). IWF. https://www.iwf.org.uk/our-services/hash-list

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2021). LoRA: Low-Rank Adaptation of Large Language Models. In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/2106.09685

Interstate Communications, 18 U.S.C. § 875 (1948).

Jam, J., Kendrick, C., Walker, K., Drouard, V., Hsu, J. G.-S., & Yap, M. H. (2021). A comprehensive review of past and present image inpainting methods. *Computer Vision and Image Understanding: CVIU*, *203*, 103147.

Johri, S. (2023). *The Making of ChatGPT: From Data to Dialogue*. https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/

Jones, S. (2024, July 9). Spain sentences 15 schoolchildren over AI-generated naked images. *The Guardian*. https://www.theguardian.com/world/article/2024/jul/09/spain-sentences-15-school-children-over-ai-generated-naked-images

Kalota, F. (2024). A Primer on Generative Artificial Intelligence. *Education Sciences*, *14*(2), 172.

Lee, S., Hoover, B., Strobelt, H., Wang, J., Peng, A., Wright, A., Li, K., Park, H., Yang, A., & Chau, P. (2024). *Diffusion Explainer: Stable Diffusion Explained with Visualization*. https://poloclub.github.io/diffusion-explainer/

Levine, A. S. (2023, December 20). Stable Diffusion 1.5 Was Trained On Illegal Child Sexual Abuse Material, Stanford Study Says. *Forbes Magazine*. https://www.forbes.com/sites/alexandralevine/2023/12/20/stable-diffusion-child-sexual-abuse-material-stanford-internet-observatory/

Liberatore, M., Erdely, R., Kerle, T., Levine, B. N., & Shields, C. (2010). Forensic investigation of peer-to-peer file sharing networks. *Digital Investigation*, *7*, S95–S103.

Mandatory Restitution, 18 U.S. Code § 2259 (1994). https://www.law.cornell.edu/uscode/text/18/2259

Mandatory Restitution to Victims of Certain Crimes, 18 U.S. Code § 3663A (1982). https://www.law.cornell.edu/uscode/text/18/3663A

Marvin, G., Hellen, N., Jjingo, D., & Nakatumba-Nabende, J. (2024). Prompt Engineering in Large Language Models. *Data Intelligence and Cognitive Informatics*, 387–402.

McManus, M., Long, M. L., & Alison, L. (2011). Child pornography offenders: towards an evidenced-based approach to prioritizing the investigation of indecent image offences: Michelle McManus, Matthew L. Long and Laurence Alison. In *Professionalizing Offender Profiling* (pp. 195–205). Routledge.

Merdian, H. L., Curtis, C., Thakker, J., Wilson, N., & Boer, D. P. (2013). The three dimensions of online child pornography offending. *Journal of Sexual Aggression*, *19*(1), 121–132.

Microsoft. (2009, December 15). *New Technology Fights Child Porn by Tracking Its "PhotoDNA."* https://news.microsoft.com/2009/12/15/new-technology-fights-child-porn-by-tracking-its-photodna/

Miller v. California. (1973). In *US* (Vol. 413, Issues 70-73, p. 15). Supreme Court.

Moritz, D. (2023). Using Artificial Intelligence to Generate Child Sexual Exploitation Material: Challenges and Implications for Law and Justice. *Algorithmic Justice Symposium*. https://usc.esploro.exlibrisgroup.com/esploro/outputs/99740398702621?institution=61USC_INST&skipUsageReporting=true&recordUsage=false

National Center for Missing and Exploited Children. (2020). *NCMEC - Key Facts*. https://www.missingkids.org/footer/media/keyfacts

National Center for Missing and Exploited Children. (2022). *2021 CyberTipline Reports by Electronic Service Providers (ESP)*. https://www.missingkids.org/content/dam/missingkids/pdfs/2021-reports-by-esp.pdf

New York v. Ferber, 458 U.S. 747 (U.S. 1982). https://scholar.google.com/scholar_case?case=1226851723986989726

O'brien, M., & Hadero, H. (2023, October 24). *AI-generated child sexual abuse images could flood the internet. Now there are calls for action*. AP News. https://apnews.com/article/ai-artificial-intelligence-child-sexual-abuse-c8f17de56d41f05f55286eb6177138d2

Obscene Visual Representations of the Sexual Abuse of Children, 18 U.S. Code § 1466A (2003). https://www.law.cornell.edu/uscode/text/18/1466A

Panchenko, A., Beaufort, R., & Fairon, C. (n.d.). Detection of Child Sexual Abuse Media on P2P Networks: Normalization and Classification of Associated Filenames. *Language Resources for Public Security*. https://core.ac.uk/download/pdf/38625377.pdf#page=32

Paquette, S., & Cortoni, F. (2019). The Development and Validation of the Cognitions of Internet Sexual Offending (C-ISO) Scale. *Sexual Abuse: A Journal of Research and Treatment*,

1079063219862281.

Pasquini, C., Amerini, I., & Boato, G. (2021). Media forensics on social media platforms: a survey. *EURASIP Journal on Information Security*, *2021*(1), 4.

Pearson, S., & Curtis, C. (2025). Erotic AI chatbots offer research opportunities for the behavioral sciences. *Archives of Sexual Behavior*, *54*(3), 855–858.

Peersman, C., Schulze, C., Rashid, A., Brennan, M., & Fischer, C. (2016). iCOP: Live forensics to reveal previously unknown criminal media on P2P networks. *Digital Investigation*, *18*, 50–64.

Penalties for Registered Sex Offenders, 18 U.S. Code § 2260A (2006). https://www.law.cornell.edu/uscode/text/18/2260A

People v. Lewis, 2024 NY Slip Op 248 (Appellate Div. 2024). https://scholar.google.com/scholar_case?case=10608612706325381003

Pfefferkorn, R. (2024). *Addressing computer-generated child sex abuse imagery: Legal framework and policy implications* (The Digital Social Contract: A Lawfare Paper Series). Lawfare. https://s3.documentcloud.org/documents/24403088/adressing-cg-csam-pfefferkorn-1.pdf

Prior, T. (2023, November 12). *Using Exiftool to Extract Metadata from Image Files*. https://www.osintteam.com/using-exiftool-to-extract-metadata-from-image-files/

Quayle, E., & Taylor, M. (2002). Child pornography and the Internet: perpetuating a cycle of abuse. *Deviant Behavior*, *23*(4), 331–361.

Relevant Definitions, Including Amie Zyla Expansion of Sex Offender Definition and Expanded Inclusion of Child Predators, 34 U.S. Code § 20911 (2006). https://www.law.cornell.edu/uscode/text/34/20911

Repeat Offenders, 18 U.S. Code § 2426 (1998). https://www.law.cornell.edu/uscode/text/18/2426

Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., & Aberman, K. (2022). DreamBooth: Fine

tuning text-to-image diffusion models for subject-driven generation. *Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 22500–22510.

Schwartz, E. H. (2024, July 2). *Forget Sora, Runway is the AI video maker coming to blow your mind*. https://www.techradar.com/computing/artificial-intelligence/forget-sora-runway-is-the-ai-video-maker-coming-to-blow-your-mind

Seto, M. C., & Eke, A. W. (2015). Predicting recidivism among adult male child pornography offenders: Development of the Child Pornography Offender Risk Tool (CPORT). *Law and Human Behavior*, *39*(4), 416–429.

Sexual Exploitation and Other Abuse of Children, 18 U.S. Code § 2256 (1978). https://www.law.cornell.edu/uscode/text/18/2256

Shehan, J. (2024, March 12). *"Addressing Real Harm Done by Deepfakes."* United States House Committee on Oversight and Accountability Subcommittee on Cybersecurity, Information Technology, and Government Innovation. https://www.missingkids.org/content/dam/missingkids/pdfs/final-written-testimony-john-shehan-house-oversight-subcommittee-hearing.pdf

Singh, S., & Nambiar, V. (2024). Role of Artificial Intelligence in the Prevention of Online Child Sexual Abuse: A Systematic Review of Literature. *Journal of Applied Security Research*, 1–42.

Sosa, A. (2024, April 15). AI-generated child pornography is circulating. This California prosecutor wants to make it illegal. *Los Angeles Times*. https://www.latimes.com/california/story/2024-04-15/ai-generated-child-pornography-is-circulating-this-california-prosecutor-wants-to-make-it-illegal

Spocchia, G. (2021, September 17). Man with 8.5 million child abuse images jailed for 27 years. *The Independent*.

https://www.independent.co.uk/news/world/americas/crime/eric-eoin-marques-jailed-child-abuse-b1921457.html

Steel, C. M. S. (2014). Technical SODDI defenses: The Trojan Horse defense revisited. *Journal of Digital Forensics, Security and Law*, *9*(4), 4.

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

Steel, C. M. S. (2025). Prevalence of Generative Artificial Intelligence Sexualized Image Usage by Adolescents in the US. *Manuscript Submitted for Publication*.

Steel, C. M. S., Newman, E., O'Rourke, S., & Quayle, E. (2021). Collecting and viewing behaviors of child sexual exploitation material offenders. *Child Abuse & Neglect*, *118*, 105133.

Steel, C. M. S., Newman, E., O'Rourke, S., & Quayle, E. (2022). Public Perceptions of Child Pornography and Child Pornography Consumers. *Archives of Sexual Behavior*. https://doi.org/10.1007/s10508-021-02196-1

Tanner, A., Jefferson, S., & Skelton, G. (2013). Revealing the unseen in social networking sites: Is your metadata protected? *Aquatic Microbial Ecology: International Journal*. https://www.academia.edu/download/31850583/ijmcis01242013.pdf

Thiel, D., Stroebel, M., Portnoff, R., & C. Center. (2023). *Generative ML and CSAM: Implications and Mitigations*. https://stacks.stanford.edu/file/druid:jv206yg3793/20230624-sio-cg-csam-report.pdf

Thorn. (2024). *Safety by Design for Generative AI: Preventing Child Sexual Abuse*. https://doi.org/10.25740/jv206yg3793

Townshend escapes child porn charges. (2003, May 7). *The Guardian*. https://www.theguardian.com/media/2003/may/07/digitalmedia.arts

Transfer of Obscene Material to Minors, 18 U.S.C. § 1470 (1998).

Transfer of Obscene Material to Minors, 18 U.S.C. § 1470 (1998).

https://www.justice.gov/usao/resources/crime-victims-rights-ombudsman/victims-rights-act

Tucker, R. (2024, March 20). *Florida teacher accused of using students' yearbook photos for AI-generated child porn*. https://fox59.com/news/national-world/florida-teacher-accused-of-using-students-yearbook-photos-for-ai-generated-child-porn/

United States v. Arthur. (2022). In *F. 4th* (Vol. 51, Issues 21-50607, p. 560). Court of Appeals, 5th Circuit.

US v. Buie. (2019). In *F. 3d* (Vol. 946, Issues 18-2942, p. 443). Court of Appeals, 8th Circuit.

Verma, P., & Harwell, D. (2024, May 21). In novel case, U.S. charges man with making child sex abuse images with AI. *The Washington Post*. https://www.washingtonpost.com/technology/2024/05/21/doj-arrest-ai-csam-child-sexual-abuse-images/

Victim Identification Solutions & Partnership. (2023, October 11). Thorn. https://www.thorn.org/solutions/victim-identification/

Vincent, J. (2020, October 20). *Deepfake bots on Telegram make the work of creating fake nudes dangerously easy*. https://www.theverge.com/2020/10/20/21519322/deepfake-fake-nudes-telegram-bot-deepnude-sensity-report

Westlake, B., Brewer, R., & Swearingen, T. (2022). Developing automated methods to detect and match face and voice biometrics in child sexual abuse videos. *Trends and Issues in*. https://search.informit.org/doi/abs/10.3316/agispt.20220331064671