DOI: https://doi.org/10.53116/pgaflr.8643

Preliminary Considerations on Artificial Intelligence and Democratic Cyber Safety for the Protection of Children

Michael M. Losavio*

* Associate Professor, Department of Criminal Justice, University of Louisville, Louisville, Kentucky, USA, e-mail: michael.losavio@louisville.edu

Submitted: 15 December 2025 | Accepted: 16 March 2026 | Published online: 30 April 2026

Abstract: Pervasive computing has abused information and communication technologies and created new kinds of systemic risks and societal vulnerabilities. With the rise of artificial intelligence, deep learning, machine learning, and pattern recognition systems, all those petabytes of information can be analysed and used to seriously interfere with the rights and liberties of others, and, in particular, it can be used to reveal the most intimate aspects of the lives of others. This, in turn, has offered new and unheard means of criminality, ranging from financial offenses to child exploitation. Therefore, there is a social responsibility for developing systems capable of mitigating the risks that the development of information and communications technologies poses to society. Given that this particular threat involves expression, efforts to mitigate expression-related misconduct must attend to rules that protect expression from government regulation. This article considers one particular area of expression-related misconduct, namely the online abuse of children.

Keywords: safety, security, AI, child protection, free speech

1. Introduction

Information and computing technologies have become pervasive in our world. Cloud, fog, and edge computing expanded by social media and the Internet of Things collect immense amounts of data in every widespread corner of the information world. Machine learning, deep learning, and "artificial intelligence" provide powerful means to search, analyse, and recognise knowledge and meaning in that immense data corpus. Effectively, almost nothing is hidden, and there is almost nothing that is not discoverable and accessible. Unfortunately, that is matched by increasing power to breach the security of information in many ways and many forms. As such, due to the pervasiveness of this encroachment, the home is no longer a sanctuary. And due to its power of manipulation,

Corresponding author: Michael M. Losavio

fabrication, and analysis, privacy in the physical world is greatly diminished. But a compromise to privacy inevitably brings along a compromise to the protections it offers as well.

Probably the most grievous interference with the right to privacy is related to the expanding phenomenon of "deepfake" technology having immense societal impact through its mass communication of false, humiliating, and manufactured images and footage. Modern information and communications technologies offer a broad ability to disseminate such content practically anywhere and to anyone, harshly interfering with personal autonomy and privacy. What is more, the consumerisation of advanced AI technologies, such as deep learning, pattern recognition, and machine learning, provides the means of information manipulation on a vast scale, capable of encroaching on the right to the freedom of thought, the freedom of expression, and personal safety *en masse*. One example of such criminality is the implementation of AI systems to automate and empower the coercion and exploitation of children for perverse purposes.

That AI systems present such great risks to children gives computer scientists and engineers developing and deploying these systems a choice. They can simply focus on the details of the technical system and leave issues relating to societal impact for others to address. Or they can embrace a moral and ethical position to engineer such exploitation out of their systems, or at least to minimise the risks of exploitation, and provide opportunities for the use of traditional public safety mechanisms to deter wrongdoers, and stop their criminality. We consider this in the context of the online sexual exploitation of children, a global problem for every nation.

2. Protecting children in the cyberspace

2.1. The technical requirements for public safety

The regulation and control of information and communications technologies and AI systems are concerns globally, particularly with the rise of generative pretrained transformers and large language models (LLM). Nations seek to control these amazing systems while not strangling the benefits they produce nor the innovation required to expand their abilities. The European Union is a leader in regulation of these technologies through the comprehensive EU AI Act[1] and its General Data Protection Regulation (GDPR).[2] The People's Republic of China also leads in this process of regulation, including with its Interim Measures for the Management of Generative Artificial Intelligence.[3] But regulation is just a first step in creating a regime of public safety. There

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), OJ L 2024/168.

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC.

[3] Interim Measures for the Management of Generative Artificial Intelligence Services, (effective 15 August 2023) of the Peoples' Republic of China.

must be a process by which the regulations can be enforced, and wrongful conduct investigated, terminated, and punished. This requires the integration of the legal regime with the technical regime to provide for public safety and security in cyberspace. This is especially important as the information and communications technologies regime deals with *information,* with speech and expression for various purposes. Its significance consists in being related to one of the most important of fundamental rights, as the freedom of expression is believed to be "essential to secure all others" (O'Connell, 2020, p. 87). Speech and expression receive, thus, extensive protection from those nations seeking to uphold the very fundamentals of human rights and dignity. The challenge is how to harvest the innovations cyberspace is offering while maintaining the protection of that very speech and expression and constraining possible illegal conduct through the use of that information.

A rather special area where the protection of these freedoms is paramount is the information related to children. The protection of children is an important issue for families, communities, and the state, if not one of the most important of all social issues. The law of the U.S., one of the most solicitous of free expression anywhere, permits regulation generally where there is a *compelling* need to do so, and an effective means for such regulation. See e.g. Reed v. Town of Gilbert, Arizona.[4] See also Ashcroft, Attorney General v. ACLU et al.;[5] Boos v. Barry.[6] Through its judicial case law, it has found there is a compelling legal protectable state interest in protecting children. That interest is deemed legally *compelling,* sufficient to justify state intervention where such intervention can be effective. Among the worst injuries from information technology are those that facilitate and promote child sexual abuse. Among the greatest challenges in law are devising effective and legal systems to regulate the use of information technologies capable of harming children.

2.2. Technical facts and legal regulation

Regulatory tests of limits on free expression under the U.S. standard of "strict scrutiny" are not purely legal. The identification of a compelling state interest in such limits requires an analysis of the technical practicality of protecting that interest through limitations on expression. They require an examination of the facts of the means by which information is distributed to people and the impact on those people, culminating in three issues in particular.

First, the phenomenon of internet connectivity has grown as to become a technology of central significance as noted in *ACLU v. Reno,* that U.S. Supreme Court case observed that the ability of people to connect to all the information in the world has great benefits. The Court further observed that regulation of such means of information distribution must weigh its benefits against possible detriments and the means by which the detriments might be mitigated.

---

[4] *Reed v. Town of Gilbert,* 576 US 155 (2015).

[5] *Ashcroft, Attorney General v. American Civil Liberties Union et al.,* 535 US 564 (2002).

[6] *Boos v. Barry,* 485 U.S. 312 (1988).

The second problem is posed by the Internet of Things phenomenon and ubiquitous data where the data corpus of all information collected on everyone, everything they do, and everywhere they are continues to grow. This growth is exponential. Furthermore, the expansion and distribution of nearly ubiquitous data sensing systems and data distribution and collection technologies, such as cloud computing, continue. This means that nothing is lost and nothing is forgotten, which is a legally significant problem in itself. The European Union, to address this issue of eternal memory, provide for a "right to be forgotten", formally a right of erasure of data, as part of their regulation of data systems.[7]

Finally, AI technologies are prone to algorithmic biases, manifesting themselves either in profiling or dishing out content. This is particularly relevant with a view to child protection as AI systems are capable of doing probabilistic analysis on a particular subject and providing effective responses relevant to that subject and their interests in an unprecedented way. The generative pre-trained transformer architectures for content analytics by AI systems can produce remarkably accurate and directed information that can be framed to appeal to a particular person. And it is this very technology and its scope and scale of focus on an individual, their interests and their desires, that may as a factual matter change the balance regarding whether or not there is a compelling need to regulate such information dissemination systems in a variety of areas.

From this brief overview, it follows that AI technologies are very potent tools capable of harmfully interfering with the protection of children. Their power has already been demonstrated through risks presented across other domains of human activity, such as the AI-driven recommender systems, used widely in commerce and marketing. These are risks that must be recognised and mitigated for child protection.

3. Protecting children and free expression in the new cyber world

3.1. Foundations for technical engagement

The vast reach of the Internet, especially via social media systems, guarantees that users of those systems are "virtually" placed "next" to children, as to enable their exploitation (Wachs et al., 2012). A connection to a child opens the way for the information seduction and exploitation of that child.

The information seduction and exploitation process has been described as "grooming" to prepare a child for what the abuser wishes next. Standard patterns for such grooming have been found, although clear patterns of psychometrics have not yet been established (Bennett & O'Donohue, 2014; see Pollack & MacIver, 2015). The elements and processes of such exploitation are gaining further definition as to permit information analysis. A frequent if not standard practice and protocol has the abuser "gaining access to a child, gaining the child's compliance, maintaining secrecy and avoiding disclosure" (Craven et al., 2006). The practices of grooming behaviours have been collected,

[7] GDPR Regulation (EU) 2016/679, Article 17.

categorised, ordered, and then applied to observe the responses (Winters & Jeglic, 2017). Their analysis produces a structured sequence of grooming behaviours that escalate the process of abuse.

First, there is a selection of the target child. The target exhibits a lack of confidence, low self-esteem and insecurity; the target is in a family with issues of discord, domestic violence, substance abuse, health concerns; there is a perceived ease of access, or the child is vulnerable in some way, such as under low levels of adult engagement and supervision; the target is attractive or appealing to the perpetrator. Next, there must be a connection and access to the target child, that is inevitably followed by isolation of the target child in various ways, physically and psychologically. Third, there is a process of trust creation with the target child, in which trust is being created in their relationship through inquiries as to the target child's interests, encouragement, aid, and offering gifts and secrets to the target child. Finally, at the end stage, there follow activities that lead to physical contact between the target child and the perpetrator.

The successful investigation and prosecution of such cases involving child sexual abuse is supported by understanding these behaviours and seeking evidence of them (Pollack & MacIver, 2015); hence, technical engagement and legal action are intrinsically interdependent. With the definition and listing of such behaviours, with associated text discussions, imagery, and audio, a structure for "grooming" can be built. Black et al. (2015) noted that the linguistic processes for online grooming behaviour were similar to offline behaviour and shared common language patterns, albeit in a different order from that in offline activity. Their analysis highlighted the frequency distribution for words/actions, with flattery, parental presence at work, travel, and "inappropriate behaviour". This linguistic analysis could easily be ported to the training of an analytical language model to engage broadly with targets. Advanced LLMs, like ChatGPT, and specialised small language models (SLMs) can pose and respond to text questions and statements through natural language. This capability offers the potential for efficient and broadly available systems for the detection of misconduct.

Such LLMs can be trained against large sets of data to build their probabilistic model of responses. Similarly, SLMs can be trained against specialised sets of data to create similar models of possible misconduct. These may be found in social media systems generally, as well as those dealing with making social connections, including those on the Dark Web. Lorenzo-Dus et al. (2020) have detailed their lexical and collocation analysis of online grooming transcripts that shows recurring patterns and linguistic structures in grooming language. They suggest the need to develop powerful algorithms to drive detection software of such behaviours, as well as hone the understanding of the *modus operandi* of the offenders.

Unfortunately, such developments may also support the development of "grooming" systems that automate the exploitation process. The semi-supervised and unsupervised training of LLMs and SLMs for child sexual exploitation may be matched by fully supervised and programmed models based on the patterns found in grooming behaviours. The use of SLMs may build effective systems in the child protection domain, given the specialised knowledge, expertise, and language of the target offenders (O'Keeffe & McWhirter, 2023). The use of SLMs for child protection benefits from the psychological

and forensic work addressing online child sexual exploitation. This may permit balancing of performance and ability through design and data choices and lead to more rapid systems of response and protection (Kramer, 2024).

The danger of such automated systems is that those machines working night and day in a globally connected world can advance these evils. They may evolve to detect law enforcement stings and AI detection models to avoid their own detection and alerts by protective measures in place on host systems, such as social media, as well as deliver precisely the illegal grooming behaviour sought to be stopped.

3.2. Models for online child protection

Online child protection in the United States has focused first on promoting safe online practices for children, usually under parental oversight (United States Federal Trade Commission, s. a.; Federal Bureau of Investigation, 2025). For traditional models of law enforcement "patrolling" and "stings" in online areas of risk for misconduct are used (Fowler et al., 2020). These entail law enforcement personnel monitoring online sites, such as social media, that may have child predators using those sites to find targets (Kyrik, 2005). With the advent of AI systems comes the opportunity to automate the process of searching online for child predatory actors. In particular, Language Models have potential for more efficient, cost-effective and timely systems for the online patrol for predators.

The existing data permits a start at building a language model to detect word patterns for "grooming" behaviour as part of the seduction of a target child. Studies show linguistic processes for online grooming behaviour were similar to offline behaviour and shared common language patterns; the primary deviation between the behaviours is that there is a different order of behaviours between online and offline misconduct (Black et al., 2015; Plaisance, 2024). These studies enable the development of preventive measures, making use of hand-coded, supervised, semi-supervised, and unsupervised training of LLMs to accomplish the online protection of children. As such, the Office of the District Attorney for the Parish of Orleans, Louisiana, for instance, has implemented an open-source use of AI systems to track criminal activity (O'Keeffe & McWhirter, 2023). The state Attorney General's office seeks to expand that statewide (Kramer, 2024). This is paralleled by a project helmed by Tulane University researchers to use AI to evaluate the criminal justice system itself (Plaisance, 2024).

Another approach would be the use of SLMs directed at the more limited and stylised practices and language used in child exploitation (Sartain, 2024). The use of SLM systems in the stylised space of online child sexual exploitation might be used for reasoning models that can develop predictive and protective systems against more limited data in a particular jurisdiction with more confined sets of misconduct (Fu et al., 2023).

Studying, creating, and understanding such systems will, unfortunately, be necessary to combat the criminal use of LLM or SLM "grooming" to build relentless means of child exploitation. Working night and day in a globally connected world eases and expands these types of abuse. They can serve to locate and alert on adversarial AI detection models.

They may support investigations into misconduct despite efforts to bypass protective systems that may be deployed, such as content screening for misconduct by large social media services. The U.S. social media service Meta Facebook uses such automated tools to search for malicious material, such as images of child sexual exploitation (Petroff, 2018). This model expands detection to text and voice patterns related to improper conduct with children. The development and use of generative adversarial networks to circumvent image manipulation detection shows how quickly offenders will act to neutralise the effectiveness of protective systems (Mavali et al., 2024). The compelling interest in protecting children may require further regulatory/technical support to build effective protections.[8]

3.3. Expanded boundaries of responsibility

Public safety is not simply a task of state officers. It is a much broader social enterprise for protection generally and for the protection of children specifically. The provision of protection, such as capable guardians watching for threats, embraces more than legal authorities, such as the police. It includes friends, good citizens, neighbours, and parents whose moral suasion alone may help deter misconduct. Such guardians may support the recovery from misconduct and assist legal authorities in the prosecution of those committing misconduct. For online activity, capable guardians must include the online services themselves.

Online services in the U.S. have been shielded from responsibility for the actions of their users; the U.S. Communications Decency Act Section 230 gave immunity from liability for the actions of others using their systems.[9] Yet that immunity has been limited or removed by legislative action in relation to particular types of content, including child sexual exploitation via these systems pursuant to the FOSTA-SESTA amendments to Section 230, addressing that and other illegal exploitation (National Institute of Justice, 2024).[10] Those amendments were the first to begin to limit the safe-harbour immunities for online service providers where those systems were used to promote sex trafficking.

Furthermore, the expansion of services by online services may have also moved them out from under the liability shield of Section 230 immunity. Those expanded services move from passive tender of third-party originated content to actions by the online service provider itself. Those activities and liability for the injuries caused may fall outside of the Section 230 immunity. The hallmark online services include algorithmic systems that keep and heighten the users' engagement in the system by directing the users to particular content, especially through recommender systems. Arguably, these systems may have the

[8] *Osborne v. Ohio* 495 U.S. 103, 111, 109 L. Ed. 2d 98, 110 S. Ct. 1691 (1990).

[9] 47 U.S.C. Section 230, part of the Communications Decency Act of 1996 providing a safe harbour immunity to providers of an "interactive computer service" that may republish content from third parties; see *Gonzalez v. Google LLC* 598 U.S. 617 (2023); *Twitter v. Taamneh* 598 U.S. 471 (2023).

[10] FOSTA-SESTA Acts: The Allow States and Victims to Fight Online Sex Trafficking Act (FOSTA) and the Stop Enabling Sex Traffickers Act (SESTA). There is some questioning of the effectiveness of the FOSTA-SESTA Acts and indicating negative impacts from this legislation.

capacity to block such content from a particular user group, especially populations susceptible to exploitation, like children, the elderly, or those with cognitive dysfunction. The lack of intentionality by the algorithmic system does not change the injury caused by its conduct. The deaths of innocent people were the focus of the litigation in *Taamneh*[11] and *Gonzales,*[12] but the lack of a clear causal connection to the actions of the online services meant liability could not be established.

However, the federal appellate court was persuaded that such an online service's causal and liability connection existed in *Anderson v. TikTok.*[13] The U.S. Court of Appeals for the Third Circuit reversed a district court's dismissal pursuant to Section 230 immunity, and remanded the case back for trial on the issues of liability for the death of a child using the systems. The District Court dismissed the case pursuant to the immunity of Section 230 of the Communications Decency Act for Internet interactive computer services from suit due to "information provided by another information content provider".[14] The Court of Appeals rejected the district court's reasoning by finding that Section 230 did *not* immunise from suit interactive computer services "if they are sued for their own expressive activity or content (i.e. first-party speech)". In doing so, the appellate court cited the 2024 U.S. Supreme Court's ruling in *Moody v. NetChoice LLC*[15] that curation of third-party content to present via algorithm by an interactive computer service is the "expressive product" of the online service itself, subject to First Amendment protection.[16] The *Moody* Court noted that a platform's standards or preferences, encoded in the algorithm's choices, might determine whether or not the algorithmic output, such as recommended content, was itself expressive content, and not simply third-party speech passed on to others with no expressive action by the interactive computer service.

In the *Anderson* case, the ten-year old child's asphyxiation death followed her participation in the "Blackout Challenge" asphyxiation activity she learned of from a video directed to the child Nylah's "For You Page" by TikTok's algorithmic recommender system. That video detailed the Blackout Challenge, which "encourages users to choke themselves with belts, purse string, or any similar until passing out", record it on video and post it. Allegedly TikTok "was aware of the Blackout Challenge", let users participate in it by posting videos of their own self-strangulation and "recommended and promoted Blackout Challenge videos to minors' 'for you pages' through the algorithm." Nylah was one recipient of the directed content video to her "for you page", "which resulted in her death". Section 230 was meant to protect online services from liability for material posted by others, but not the expressive activities of the online services themselves. The curation and direction by algorithmic recommender systems used to keep users engaged with the system were expressive activities by TikTok for which they may be held accountable.

---

[11] *Twitter v. Taamneh* 598 U.S. 471 (2023).

[12] *Gonzalez v. Google LLC* 598 U.S. 617 (2023).

[13] *Anderson v. TikTok, Inc. and Bytedance, Inc.,* ___F3d____(3rd Cir. 2024).

[14] Communications Decency Act, Section 239, 47 U.S.C. §230.

[15] *Moody v. NetChoice LLC* 144 S. Ct. 2383, 2393 (2024).

[16] *Moody* did not address the issue of the impact of different algorithms as to be expressive activity, including those that "respond solely to how users act online", id. at 2404, n.5.

Anderson establishes that the advances in commerce and marketing by online services lead to their services and information falling outside of Section 230 immunity, making them liable for injuries for which they are responsible. This encourages, if not demands, that online service providers protect their users by limiting the harmful content they may provide to them. The online providers have a choice: they can effectively bar vulnerable groups like children from using their systems and being possibly harmed by them, as is required for pornographic materials, or effectively act to protect those vulnerable groups from exploitation.

Responsibility for online injuries may itself be seen as a subset of software liability generally. The European Union's Product Liability Directive seeks to expand liability for software and injuries it may cause (see Shackelford et al., 2025). Acknowledging the inadequacy of liability *post-hoc* measures in some domains to protect others, Shackelford et al. (2025) propose a revised system of liability to promote public safety in systems such as those online. Those revisions expressly extending product liability to include cybersecurity failings require implementation of "secure-by-design" for systems and require transparency and accountability of systems through regulatory frameworks (Shackelford et al., 2025). These may encourage, if not mandate, online systems to implement safety measures for their users.

Another perspective on online systems affecting safety for users can be seen in the example of routine activities theory in crime control as posited by Cohen and Felson (1979). Under the routine activities theory, three elements promote a criminal act: a motivated offender (the criminal), a suitable target (the child), and the absence of a capable guardian. With this comes the greater likelihood of crime. The "motivated offender" class of child sex offenders expands with the pervasiveness and invasiveness of online systems. The online environment expands the set of "suitable targets" that become objects of opportunity for the offender and the calculus of success that an offender makes. That calculus in the online world includes the ease of access to the target, the reward it offers, and features that help avoid detection and promote "ease of escape".

The presence of "capable guardians" refers to constant, present individuals and systems that deter misconduct or sanction offenders. The presence of an effective, capable guardian can mitigate the threat. This approach can be mapped to child protection in the online information space to suggest alternative approaches for the protection of children. This may begin by giving children tools to resist online exploitation through childhood education and using law enforcement investigative activity *pre-* and *post-hoc* to deter and apprehend offenders.

But more is needed, especially as the vast scope of online activity makes automated systems for providing guardians essential. Routine activities theory suggests enhanced protection by development and deployment of the "capable guardian" for online activities, a role made difficult by the engineering and design of online systems.

Major online services provider Meta was found responsible by a New Mexico jury for not protecting minors from solicitation and sexual content via its Facebook and Instagram services (Mulvaney & Bobrowsky, 2026). The next day Meta and Google were found by a California jury to be negligent in their design and operation of their social media systems as to "addict" children to those systems and cause them mental health injuries

(Electronic Privacy Information Center, 2026). If these judgements are appealed there may be greater clarity and legal guidance on online responsibility and the legal liability faced by online service providers for the impact of their services on children.

4. Conclusion

Information and communication technologies, including AI technologies, the data on which AI systems are trained, and the Internet, are all systems of expression. For the U.S., these systems, their input and output, and their expressions of results are protected by the First Amendment to the U.S. Constitution. Although conflicting rationales may be seen in the U.S. case law for validating or preventing regulation of such expression, the majority support free "expression" by such technologies. But those judicial rationales are constrained by the flexibility of analysis that changing facts can lead to changing judicial outcomes. And where speech protections rest on statutory safe harbours for the technology, those may be changed by later statutory amendment.

The compelling interest in protecting children so much at issue in earlier Supreme Court jurisprudence was outweighed by the limited effectiveness of earlier legal and technical systems to protect those children. That, in turn, was weighed against the interests of adults in free expression and free access to information. But times do indeed change, and nothing seems to change more swiftly than information and communication technologies in the modern world. These changes are often for the better, though they may offer new tools for criminality and the victimisation of innocent people. And none are better targets for victimisation than children who are just beginning to develop their skills for critical thinking and analysis. Yet those skills are developing with a limited collection of knowledge and experiences that leave them vulnerable to manipulation and grooming, just as a GPT, trained against a limited language model of partial information, is vulnerable to errors and "hallucinations".

The need for AI-based systems of threat recognition and child protection with online systems is great, and it is needed without delay. Arguably, these systems can be programmed and used to respect freedom of expression in all its forms, though their introduction doubtless requires considerable societal investment and a tailored legal regime.

After all, misconduct exploiting children is banned in all nations, at least in formal law. Yet, additional regulation should be considered that can facilitate machine support for child protection. The information and communications technologies tools that can lead to injuring children can be designed to protect them from human and machine attacks. Even under the broad speech protections of the laws of the U.S., it is possible to regulate means for detecting and policing online child abuse while leaving those systems effective for child protection. If it can be successfully implemented in the U.S., it is available in every nation. Regulation keyed to linguistic analysis of such misconduct may be drawn to narrowly apply and not unnecessarily infringe on the speech rights of adults. Such carefully and narrowly drawn limits and effective technical systems must be used to promote or mandate such protective systems to protect children from online exploitation.

Acknowledgements

The author thanks Antonio M. Losavio, Maura H. Losavio, and Amy L. Losavio for their recommendations and proofreading assistance with this article.

References

Bennett, N. & O'Donohue, W. (2014). The Construct of Grooming in Child Sexual Abuse: Conceptual and Measurement Issues. *Journal of Child Sexual Abuse, 23*(8), 957–976. Online: https://doi.org/10.1080/10538712.2014.960632

Black, P. J., Wollis, M., Woodworth, M. & Hancock, J. T. (2015). A Linguistic Analysis of Grooming Strategies of Online Child Sex Offenders: Implications for Our Understanding of Predatory Sexual Behavior in an Increasingly Computer-Mediated World. *Child Abuse & Neglect, 44*, 140–149. Online: https://doi.org/10.1016/j.chiabu.2014.12.004

Cohen, L. E. & Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review, 44*(4), 588–608. Online: https://doi.org/10.2307/2094589

Craven, S., Brown, S. & Gilchrist, E. (2006). Sexual Grooming of Children: Review of Literature and Theoretical Considerations. *Journal of Sexual Aggression, 12*(3), 287–299. Online: https://doi.org/10.1080/13552600601069414

Electronic Privacy Information Center (2026, March 25). *Jury Finds Meta and Google Negligent in Landmark Social Media Addiction Case.* Online: https://epic.org/jury-finds-meta-and-google-negligent-in-landmark-social-media-addiction-case/

Federal Bureau of Investigation (2025, August 19). *Staying One Step Ahead: How to Protect Kids from Emerging Online Threats.* Online: https://www.fbi.gov/contact-us/field-offices/jacksonville/news/staying-one-step-ahead-how-to-protect-kids-from-emerging-online-threats

Fowler, M. J., Lybert, K. A., M.A., Owens, J. N., M.A. & Waterfield, J. M. (2020, August 6). *Undercover Chatting with Child Sex Offenders.* United States Federal Bureau of Investigation, Featured Articles. Online: https://leb.fbi.gov/articles/featured-articles/undercover-chatting-with-child-sex-offenders

Fu, Y., Peng, H., Ou, L., Sabharwal, A. & Khot, T. (2023). Specializing Smaller Language Models Towards Multi-Step Reasoning. *Proceedings of the 40th International Conference on Machine Learning, PMLR 202*, 10421–10430. Online: https://proceedings.mlr.press/v202/fu23d/fu23d.pdf

Kramer, J. (2024, November 5). *Louisiana Signs $3M Deal for High-Tech Investigations.* The Times-Picayune. Online: https://tinyurl.com/43bpjnr8

Kyrik, Kelly (2005). Trolling for Predators: More and More Law Enforcement Officers Are Actively Working the Internet to Track, Apprehend, and Prosecute Pedophiles. *Police: The Law Enforcement Magazine,* 29(10), 32–40.

Lorenzo-Dus, N., Kinzel, A. & Di Cristofaro, M. (2020). The Communicative *Modus Operandi* of Online Child Sexual Groomers: Recurring Patterns in their Language Use. *Journal of Pragmatics, 155*, 15–27. Online: https://doi.org/10.1016/j.pragma.2019.09.010

Mavali, S., Ricker, J., Pape, D., Fischer, A. & Schönherr, L. (2024). Fake It until You Break It: On the Adversarial Robustness of AI-generated Image Detectors. *arXiv*. Online: https://doi.org/10.48550/arXiv.2410.01574

Mulvaney, E. & Bobrowsky, M. (2026, March 24). Landmark Verdict Says Meta Harmed Children, Allowing Adults to Prey on Them. *The Wall Street Journal*. Online: https://www.wsj.com/tech/landmark-verdict-says-meta-harmed-children-allowing-adults-to-prey-on-them-cb3ad674

National Institute of Justice (2024, May 8). *Unconventional Wisdom: Research Shakes up Assumptions About Sex Trafficking Clues in Online Escort Ads.* Online: https://tinyurl.com/pakavj8x

O'Connell, R. (2020). Freedom of Expression. In *Law, Democracy and the European Court of Human Rights* (pp. 84–119). Cambridge University Press. Online: https://doi.org/10.1017/9781139547246.007

O'Keeffe, K. & McWhirter, C. (2023, September 17). *New Orleans DA Fights "Terrorism" on Streets with AI Spycraft.* The Wall Street Journal. Online: https://tinyurl.com/5xsv9khd

Petroff, A. (2018, April 5). *Yes, Facebook Is Scanning Your Messages for Abuse.* CNN Business. Online: https://tinyurl.com/4cbm6crj

Plaisance, S. (2024, November 18). *Tulane Researchers Partner with Court Watch NOLA to Boost Criminal Justice Transparency.* Tulane University of Science and Engineering. Online: https://tinyurl.com/54xwd4jk

Pollack, D. & MacIver, A. (2015). Understanding Sexual Grooming in Child Abuse Cases. *Child Law Practice, 34*(11), 166–168.

Sartain, M. (2024, June 24). *Large Language Models (LLMs) vs. Small Language Models (SLMs)*. Rackspace Technology. Online: https://tinyurl.com/3cv6vk8s

Shackelford, S., Hiller, J., Makridis, C., Nash, I., Kisska-Schulze, K. & Travis, H. (2025). Moving Slow and Fixing Things. *Indiana Law Journal, 100*(4), 1611–1671. Online: https://www.repository.law.indiana.edu/ilj/vol100/iss4/8/

United States Federal Trade Commission (s. a.). *Protecting Kids Online.* Online: https://consumer.ftc.gov/identity-theft-and-online-security/protecting-kids-online

Wachs, S., Wolf, K. D. & Pan, C. (2012). Cybergrooming: Risk Factors, Coping Strategies and Associations with Cyberbullying. *Psicothema*, *24*(4), 628–633. Online: https://www.psicothema.com/pdf/4064.pdf

Winters, G. M. & Jeglic, E. L. (2017). Stages of Sexual Grooming: Recognizing Potentially Predatory Behaviors of Child Molesters. *Deviant Behavior, 38*(6), 724–733. Online: https://doi.org/10.1080/01639625.2016.1197656