Vol. 95 issue 2, 2024

RIDP

Gert Vermeulen, Nina Peršak
& Stéphanie De Coensel *(Eds.)*

Researching the boundaries of **sexual integrity, gender violence and image-based abuse**

Revue Internationale de Droit Pénal
International Review of Penal Law
Revista internacional de Derecho Penal
Международное обозрение уголовного права
刑事法律国际评论
المجلة الدولية للقانون الجنائي
Revista Internacional de Direito Penal
Rivista internazionale di diritto penale
Internationale Revue für Strafrecht

MAKLU

Researching the boundaries of **sexual integrity, gender violence and image-based abuse**

Edited by

Gert Vermeulen
Nina Peršak
Stéphanie De Coensel

RIDP
Revue Internationale de Droit Pénal
International Review of Penal Law
Revista internacional de Derecho Penal
Международное обозрение уголовного права
国际刑事法律评论
المجلة الدولية للقانون الجنائي
Revista Internacional de Direito Penal
Rivista internazionale di diritto penale
Internationale Revue für Strafrecht

Maklu

Antwerpen | Apeldoorn | Portland

AIDP – Association Internationale de Droit Pénal | The International Association of Penal Law is the oldest association of specialists in penal law in the world. Since 1924, it is dedicated to the scientific study of criminal law and covers: (1) criminal policy and codification of penal law, (2) comparative criminal law, (3) international criminal law (incl. specialization in international criminal justice) and (4) human rights in the administration of criminal justice. The Association's website provides further information (http://www.penal.org).

RIDP – Revue Internationale de Droit Pénal | The International Review of Penal Law is the primary publication medium and core scientific output of the Association. It seeks to contribute to the development of ideas, knowledge, and practices in the field of penal sciences. Combining international and comparative perspectives, the RIDP covers criminal law theory and philosophy, general principles of criminal law, special criminal law, criminal procedure, and international criminal law. The RIDP is published twice a year. Typically, issues are linked to the Association's core scientific activities, ie the AIDP conferences, Young Penalist conferences, world conferences or, every five years, the International Congress of Penal Law. Occasionally, issues will be dedicated to a single, topical scientific theme, validated by the Scientific Committee of the Association, comprising high-quality papers which have been either presented and discussed in small-scale expert colloquia or selected following an open call for papers. The RIDP is published in English only.

Peer review: All contributions are subject to double-layered peer review. The primary scientific and peer review responsibility for all issues lies with the designated Scientific Editor(s). The additional scientific quality control is carried out by the Executive Committee of the Editorial Board, which may turn to the Committee of Reviewers for supplementary peer review.

Disclaimer: The statements and opinions made in the RIDP contributions are solely those of the respective authors and not of the Association or MAKLU Publishers. Neither of them accepts legal responsibility or liability for any errors or omissions in the contributions nor makes any representation, express or implied, with respect to the accuracy of the material.

ISSN – 0223-5404
ISBN 978-90-466-1263-7
D/2024/1997/33
NUR 824
BISAC LAW026000
Theme: LNF, LAR, JBFK2

Maklu Publishers
Somersstraat 13/15, 2018 Antwerpen, Belgium, info@maklu.be
Koninginnelaan 96, 7315 EB Apeldoorn, The Netherlands, info@maklu.nl
www.maklu.eu

USA & Canada
International Specialized Book Services
920 NE 58th Ave., Suite 300, Portland, OR 97213-3786, orders@isbs.com, www.isbs.com

Summary

THE POLICING AND REPORTING OF ONLINE CHILD SEXUAL ABUSE MATERIAL: A SCOPING REVIEW

*Nena Decoster**

Abstract

Child sexual abuse material (CSAM) has exponentially increased in recent years because of technological advances and the rise of the internet. The call for new investigative tools became urgent. As a result, automated technologies became indispensable in the field of detection, filtering, classification and prioritization of CSAM. Private tech companies have slowly but surely been dominating the development of those automated tools. Furthermore, private organizations, NGO's, big tech companies and online service providers hold the reins when it comes to selecting which reports to pass on to law enforcement, aided by automatization and their own corporate policies. At the end of the line, the reports and images the police receive, have gone through a major funnel, significantly impacted by private entities. The field of evidence on the policing and reporting of CSAM remains scattered and is characterized by a lack of coherent and integrated literature. This scoping review aims to meet this knowledge deficit by critically mapping relevant and available literature.

1 Introduction

Online child sexual abuse material (CSAM) saw the light of day in the late 1980's and has been a pressing issue ever since.[1] CSAM refers to any visual image or record of sexually explicit activity that involves minors.[2] Advances in technology through the rise of internet access and social media shape the current landscape concerning CSAM, causing constant growth and adaptation and new, safer and cheaper means for accessing and distributing CSAM.[3] As a consequence of these advanced technological possibilities, CSAM

* Nena Decoster is a PhD researcher (FWO) at the Institute for International Research on Criminal Policy (IRCP) at the Department of Criminology, Criminal Law and Social Law, Faculty of Law and Criminology at Ghent University, Belgium. She conducts research on CSAM and surveillance. For correspondence: <nena.decoster@ugent.be>.

[1] Jon Taylor, 'Policing social networking sites and online grooming', *Internet Child Abuse: Current Research and Policy* (Routledge-Cavendish 2010).

[2] Mayana Pereira and others, 'Metadata-Based Detection of Child Sexual Abuse Material' (2020) 99 IEEE Transactions on Dependable and Secure Computing.

[3] Bracket Foundation, *Artificial Intelligence: Combating Online Sexual Abuse of Children* (2019); Liora Lazarus and others, *Respecting human rights and the rule of law when using automated technology to detect online child sexual exploitation and abuse* (2021); Elie Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet* (2019); Elena Martellozzo, *Online Child Sexual Abuse: Grooming, Policing and Child Protection in a Multi-Media World* (Routledge 2013); David Mount and others, 'Triaging online child abuse material: testing a decision support tool to enhance law enforcement and investigative prioritisation' (2021) 44 Policing: An International Journal; Juliane A Kloess, Anthony R Beech and Leigh Harkins, 'Online child sexual exploitation: prevalence, process, and offender characteristics' (2014) 15 Trauma, Violence & Abuse 126.

has exponentially increased in recent years.[4] Reports of presumed CSAM on platforms of online service-providers (OSP's) such as Facebook or Google are notified through the National Centre for Missing and Exploited Children's (NCMEC) Cybertipline, a US-based private, non-profit organisation. NCMEC received 10 000 reports a year when it started its activities in 1998 and received a record number of reports in 2022 of 32 million.[5] This pressures law enforcement to improve their approach to tackle CSAM since the reviewing process greatly exceeds their manual capacities.[6] Because of these trends and their consequences, the call for new investigative tools and alternative means for proactively detecting CSAM and processing reports is urgent.[7]

Where technology offers more opportunities for offenders, it equally provides opportunities to fight crime and adapt to transitions in cybercrime.[8] More specific, the use of automated technologies has become indispensable in the field of detection, classification and prioritisation of CSAM.[9] Automatic detection tools are essential to track down and remove potential CSAM on platforms to prevent redistribution.[10] In addition to detecting images, classification (ie which images are in fact child abuse and which are not) is a crucial task that is also being automated.[11] Lastly, automated technologies are used for the prioritisation of cases (eg. when the report contains recent or very clear sexual abuse), aiding the decision-making process of law enforcement agencies (LEA).[12]

Private actors have slowly but surely been dominating the development of automated detection tools in the fight against CSAM. Private actors such as Facebook and Google

[4] Mount and others, 'Triaging online child abuse material: testing a decision support tool to enhance law enforcement and investigative prioritisation'; John Carr and Zoe Hilton, 'Combating child abuse images on the Internet: International perspectives', *Internet Child Abuse: Current Research and Policy* (Routledge-Cavendish 2010); Jennifer Martin, 'Child sexual abuse images online: Implications for social work training and practice' (2016) 46 The British Journal of Social Work 372.

[5] Lazarus and others, *Respecting human rights and the rule of law when using automated technology to detect online child sexual exploitation and abuse*; NCMEC, *CyberTipline 2022 Report* (2023).

[6] Hee-Eun Lee and others, 'Detecting child sexual abuse material: A comprehensive survey' (2020) 34 Forensic Science International: Digital Investigation; Laura Sanchez and others, 'A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM)' (2019) 29 Digital Investigation.

[7] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

[8] Bryce Westlake, Martin Bouchard and Richard Frank, 'Assessing the Validity of Automated Webcrawlers as Data Collection Tools to Investigate Online Child Sexual Exploitation' (2015) 29 Sexual Abuse.

[9] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet;* Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

[10] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

[11] Foundation, *Artificial Intelligence: Combating Online Sexual Abuse of Children;* Jillian Richardson, *Three Ways Artificial Intelligence is Good for Society* (2017); Peter Buell Hirsch, 'Tie me to the mast: artificial intelligence & reputation risk management' (2018) 39 Journal of Business Strategy 61.

[12] INHOPE, *Annual Report 2020* (2020); Lazarus and others, *Respecting human rights and the rule of law when using automated technology to detect online child sexual exploitation and abuse.*

have been developing own techniques to detect and remove CSAM from their platforms.[13] Moreover, numerous profit-oriented companies are making similar software and hardware packages available to law enforcement. These private companies, NGO's and non-profits hold the reins when it comes to selecting which reports to pass on to LEA. US-based ESPs (electronic service providers: both social media companies and web engines) are responsible for submitting reports to NGO's like NCMEC.[14] NCMEC is then responsible for making these reports available to law enforcement worldwide, after assessing the reports manually or automatically using privately made technology.[15]

It can be stated that knowledge on the policing and reporting processes of CSAM exists but is scattered throughout academic and grey literature. There is a lack of comprehensive and coherent evidence that ties together the most important facets of the CSAM reporting and policing chain. The aim of this paper will be to critically tie together literature on the most important players active in fighting CSAM such as NGO's, ESPs, law enforcement and the government. On top of that, the operational reporting and policing chain of CSAM, including technological tools such as hashing and AI will be studied. This will be achieved by performing a scoping review. It is important to note that this scoping review generated limited and dated information on the European regulatory framework concerning the policing of CSAM. It remains essential to touch upon the most important legislative developments. Therefore, this paper will commence with an overview that includes relevant general obligations on online service providers as well as a discussion of instruments to tackle CSAM specifically, to adequately sketch the European Union's legal framework. After that, the text will briefly outline the methodology used for the scoping review, followed by the results and a short discussion. The results will firstly touch upon the broader regulatory framework in which the policing of CSAM occurs, followed by an analysis of the policing process in subsequent order: scanning and detection, processing by NGO's and the eventual reception of reports by law enforcement.

1.1 Notice & Takedown

The Digital Services Act (DSA), approved in September 2022, is an ambitious instrument aimed to regulate and harmonize digital services, with a focus on digital harms and aims to build on the E-Commerce directive, with attention to the challenges around online intermediaries and platforms.[16] The DSA holds significant relevance for the scanning and

[13] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet;* Pereira and others, 'Metadata-Based Detection of Child Sexual Abuse Material'.

[14] Mar Negreiro, *Curbing the surge in online child abuse: The dual role of digital technology in fighting and facilitating its proliferation* (2020); Ethel Quayle, 'Prevention, disruption and deterrence of online child sexual exploitation and abuse' (2020) 21 ERA Forum.

[15] NCMEC, *Cybertipline: What Happens to Information in a CyberTip?* (2020).

[16] Buri I and Hoboken Jv, The Digital Services Act (DSA) proposal: a critical overview (2021).

reporting of CSAM.[17] It connects the size of a platform with the potential it poses for harm. The larger and more complex the digital service, the greater the obligations to address illegal content under Union law.[18] Therefore, very large online platforms (VLOPs) and very large online search engines (VLOSEs) are subjected to different and additional obligations than intermediary services, hosting services and online platforms. The status of VLOP or VLOSE is awarded when it has 45 million active users per year or 50 million in annual revenue (Art. 3 & 33(1)).[19] Under the DSA, VLOPs and VLOSEs are subjected to additional obligations to tackle illegal content, on top of those imposed on the other providers. In terms of content moderation, hosting providers and online platforms should allow relevant judicial or administrative authorities to notify illegal material and need to expeditiously remove it.[20] VLOPs and VLOSEs are faced with more stringent and cumulative due diligence duties.[21] They need to perform a risk assessment (Art. 34) to identify and analyze systemic risks stemming from the design and functioning of their services, including their algorithmic systems. Two of these systemic risks are specified in Art. 34(1)(a) and 34(1)(2)(d) as the dissemination of illegal content through their services or any actual or foreseeable negative effects in relation to the protection of minors.[22] It should also be analyzed how these risks are influenced by the amplification and potentially rapid and wide dissemination of illegal content. Once these risks are assessed, they should be mitigated by, for example, adapting content moderation processes in terms of speed and quality of notice and takedown (Art. 35). It is also laid out that VLOPs and VLOSEs should undergo yearly independent audits (Art. 37) and share their data with the Commission, national authorities, and vetted researchers (Art. 40) for DSA-compliance monitoring or research concerning the systemic risks in the EU.[23] Whereas DSA content moderation obligations are strict for hosting providers and online platforms, they are even more stringent for VLOPs and VLOSEs. Although there is no legal obligation to search for illegal content, the DSA introduced several possibilities for such actions,

[17] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277.

[18] Florence G'sell, 'The Digital Services Act: a General Assessment' in Antje von Ungern-Sternberg (ed), *Content Regulation in the European Union - The Digital Services Act*, vol 1 (Trier Studies on Digital Law, 2023).

[19] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277.

[20] Peter Church and Ceyhun Necati Pehlivan, 'The Digital Services Act (DSA): A New Era for Online Harms and Intermediary Liability' (2023) 4 Global Privacy Law review.

[21] Ibid; Elisabetta Stringhi, 'The due diligence obligations of the Digital Services Act: a new take on tackling cyber-violence in the EU?' (2024) International Review of Law, Computers & Technology.

[22] Regulation (EU) 2022/2065 of The European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277.

[23] Ibid.

through which providers are given an active role in influencing the digital environment in which users operate.[24]

Additionally, audiovisual media services like television and broadcasting services have a flagging and reporting system similar to the DSA. The 2018 amendment to the Audiovisual Media Services Directive (AVMD) (Directive 2018/1808)[25] introduced several measures to protect the public from content depicting CSAM. Art. 28b of the consolidated AVMD stipulates possibilities to introduce, amongst others, flagging and reporting mechanisms for CSAM content. It also leaves the possibility to the Member States to take more detailed or stricter measures to ensure prompt removal or blocking of access to CSAM than those provided for in the AVMD, insofar as this does not entail a general monitoring obligation (Art. 28b(6)).[26]

1.2 Data exchange with private parties

Several instruments have been facilitating (or imposing) possibilities for data exchange with private parties in pursuit of tackling online illegal content such as CSAM. The temporary derogation[27] made it possible for service providers to exchange data with law enforcement for the purpose of removing and detecting CSAM.

Art. 26b of the amended Europol Regulation (Regulation (EU) 2022/991)[28] expands the possibilities for data sharing with and data requests by Europol to address the online dissemination of CSAM. Until then, it was generally not possible for Europol to have direct contacts with private parties.[29] Art. 26b includes the possibility for private parties to send over personal-data packages to Europol on a voluntary basis and vice-versa, for Europol to transfer and transmit data to private parties (Art. 26b (1) & (3)). However, it also entails a possibility for Europol to request and obtain personal data from private parties, on an involuntary basis (Art. 26b(6)). However, whereas Europol's powers are

[24] Marcin Rojszczak, 'The Digital Services Act and the Problem of Preventive Blocking of (Clearly) Illegal Content' (2023) 3 Journal of Administrative Sciences.
[25] Directive (EU) 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities [2018] OJ L303.
[26] Ibid.
[27] Regulation (EU) 2021/1232 of the European Parliament and of the Council of 14 July 2021 on a temporary derogation from certain provisions of Directive 2002/58/EC as regards the use of technologies by providers of number-independent interpersonal communications services for the processing of personal and other data for the purpose of combating online child sexual abuse [2021] OJ L274.
[28] Regulation (EU) 2022/991 of the European Parliament and of the Council of 8 June 2022 amending Regulation (EU) 2016/794, as regards Europol's cooperation with private parties, the processing of personal data by Europol in support of criminal investigations, and Europol's role in research and innovation [2022] OJ L169.
[29] Wanqi Lai, Amalia Van Vaerenbergh and Wannes Bellaert, 'Europol and its growing alliance with private parties' (2021) 92 Revue Internationale de Droit Pénal 45.

significantly expanded, this has not been accompanied by strong data protection safeguards.[30]

1.3 The (not so) new CSAM regulation

Lastly, important steps have been taken to establish regulatory instruments aimed specifically at tackling online CSAM. Below, attention will be paid to the interim derogation and the proposed CSAM regulation.

The 2018 recast European Electronic Communications Code Directive (EECC) regulates electronic communication services in the EU. It extends the definition of electronic communication services to include number-independent interpersonal communication services (NI-ICS) and services merely using numbers as an identifier, such as instant messaging. This change in definition affects the e-Privacy Directive[31], which covers the security and confidentiality of electronic communications (both traffic data and content).[32] It sets out the rights, freedoms and protections that users enjoy in relation to their electronic communications. As a consequence of the revised definitions in the EECC, online service providers active in the scanning, removal, and reporting of online CSAM have to adhere particularly to Art. 5 and 6 of the e-Privacy Directive, ie ensure confidentiality of content and traffic data which prevented them from listening, tapping, storage or other kinds of interception or surveillance of communications. Therefore, the voluntary scanning using specific technologies was no longer possible.[33]

To ensure continued voluntary detection, reporting, and removal of CSAM by NI-ICS using specific technologies without breaching provisions from the e-Privacy Directive, a Proposal for a Regulation on a temporary derogation from certain provisions of the e-Privacy Directive for the purpose of combating child sexual abuse online was introduced.[34] Despite the mention by the Commission that it took account of fundamental rights to privacy and data protection and complied with the principle of proportionality, the interim derogation was met with critique. The proposal was amended and certain issues clarified. The scope was narrowed as the scanning of audio communications were excluded and it was determined that service providers had to act in accordance with the

---

[30] European Data Protection Supervisor, *Amended Europol Regulation weakens data protection supervision* (2022).

[31] Jeanne Mifsud Bonnici and others, *Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse* (2021); Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201.

[32] Bonnici and others, *Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse*.

[33] Ibid.

[34] Regulation (EU) 2021/1232 of the European Parliament and of the Council of 14 July 2021 on a temporary derogation from certain provisions of Directive 2002/58/EC as regards the use of technologies by providers of number-independent interpersonal communications services for the processing of personal and other data for the purpose of combating online child sexual abuse [2021] OJ L274.

GDPR when processing personal data.[35] The text was adopted in July 2021.[36] The temporary or interim derogation was set to apply until the 3rd of August 2024, until a permanent legal framework with more safeguards was in place.[37]. However, the amended and adopted version of the derogation still lacks effective safeguards against general and indiscriminate monitoring of private communications.[38] Due to a lack of agreement on the proposed permanent CSAM regulation, the temporary derogation was extended until April 2026, despite some privacy concerns.[39]

Certain provisions of the E-Privacy Directive serve as a *lex specialis* to the General Data Protection Regulation, specifically with respect to the processing of personal data in the electronic communications sector.[40] For example, Art. 5 and 6 of the e-Privacy Directive offer higher protections for the processing of traffic and location data than the GDPR does. The GDPR applies to protection of personal data, regardless of which technology is used. Furthermore, the GDPR[41] stipulates what constitutes a lawful basis for the processing of personal data by for example, online service providers in Art. 6. The six lawful grounds for this are consent, contract, legal obligation, vital interest, public task or legitimate interest. When processing special categories of personal data such as data from racial or ethnic origin or concerning a person's sex life, additional conditions for the processing are laid out in Art. 9.[42] In the context of a study by the European Parliamentary Research Service[43] of these legal grounds in the context of detecting CSAM online, it was determined that processing could only be possible based on the legal ground of vital

[35] Charlotte Somers, 'The proposed CSAM Regulation: trampling privacy in the fight against child sexual abuse?' (2023) <https://www.law.kuleuven.be/citip/blog/the-proposed-csam-regulation-trampling-privacy-in-the-fight-against-child-sexual-abuse/> accessed 2 July 2024.

[36] Lazarus and others, *Respecting human rights and the rule of law when using automated technology to detect online child sexual exploitation and abuse.*

[37] Bonnici and others, *Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse*.

[38] European Data Protection Supervisor, *Opinion 8/2024 on the Proposal for a Regulation amending Regulation (EU) 2021/1232 on a temporary derogation from certain ePrivacy provisions for combating CSAM* (2024).

[39] European Parliament, 'Child sexual abuse online: current rules extended until April 2026' (2024) <https://www.europarl.europa.eu/news/en/press-room/20240408IPR20311/child-sexual-abuse-online-current-rules-extended-until-april-2026> accessed 2 July 2024.

[40] Bonnici and others, *Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse*.

[41] 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 (General Data Protection Regulation) [2016] OJ L119.

[42] European Data Protection Board, 'Process personal data lawfully' (n.d.) <https://www.edpb.europa.eu/sme-data-protection-guide/process-personal-data-lawfully_en#toc-1> accessed 2 June 2024; 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 (General Data Protection Regulation) [2016] OJ L119.

[43] Bonnici and others, *Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse*.

interest or legitimate interest. However, it was mentioned that despite the fact that detecting, reporting and removing of CSAM pursues a legitimate aim, envisaged measures should always be in accordance with other fundamental and human rights affected by those measures.[44]

On May 11 2022 the European Commission proposed a regulation laying down rules to prevent and combat child sexual abuse.[45] Once adopted, this regulation should replace the interim derogation and serve as a permanent legal basis for the scanning of CSAM.[46] The proposal sets rules to target, combat, and prevent online CSAM. Additionally, it includes obligations for relevant hosting service providers and providers of interpersonal communication services to assess and mitigate the risk of CSAM on their platforms. Providers should conduct a risk assessment of their services to identify whether their platforms are being misused for CSAM and should take appropriate measures to mitigate those risks.[47] The proposal also introduces a possibility for judicial or independent administrative authorities to issue detection, removal, or blocking orders to services when the risk assessment provides reasons therefore, entailing a legal obligation for scanning.[48] Despite providing safeguards concerning the scope of the detection order, its timeline, and the protection of personal data, experts strongly oppose the proposal. The proposal emphasizes detecting unknown CSAM, despite the less accurate technology compared to known CSAM, which potentially increases false positive reports and law enforcement's workload. It remains unsure whether the CSAM regulation will include the possibility for detection orders within end-to-end encrypted (hereafter: E2EE) communications.[49] However, should it include E2EE, this will conflict with the purpose of E2EE, namely protected and private communications.[50] Whereas this proposal would positively impact the protection of children, it would infringe on general data retention and general monitoring prohibitions and significantly impact fundamental rights.[51]

---

[44] Ibid.

[45] Proposal for a Regulation of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse [2022] COM(2022) 209 final.

[46] European Parliamentary Research Service, *Proposal for a regulation laying down the rules to prevent and combat child sexual abuse Complementary Impact assessment* (2023).

[47] Proposal for a Regulation of the European Parliament and of the Council laying down rules to prevent and combat child sexual abuse [2022] COM(2022) 209 final.

[48] European Parliamentary Research Service, *Proposal for a regulation laying down the rules to prevent and combat child sexual abuse Complementary Impact assessment*.

[49] Julia Tar, 'EU nears consensus on child abuse draft law, new agency takes lead on privacy preservation' (2024) <https://www.euractiv.com/section/law-enforcement/news/eu-nears-consensus-on-child-abuse-draft-law-new-agency-takes-lead-on-privacy-preservation/> accessed 3 July 2024.

[50] European Parliamentary Research Service, *Proposal for a regulation laying down the rules to prevent and combat child sexual abuse Complementary Impact assessment*.

[51] Ibid.

2 Methodology

As mentioned, the aim of this paper is to map out available evidence concerning the automated policing of CSAM. A scoping review provides the best methodological framework to meet this objective, since the envisaged outcome of this paper was not to offer specific solutions to an identified problem or to answer detailed questions but rather to provide a critical and coherent illustration of existing evidence on this topic.[52] The output of this review will thus attempt to meet some of the knowledge deficits concerning processes of the automated policing of CSAM mentioned above.[53] Methodological literature on scoping reviews specifically tailored to social sciences and the humanities is rather hard to find. The material that is available can be traced back to nursing studies and medicine.[54] In many cases, when a scoping review is executed, the methods or procedures behind it are not clarified.[55] In this review we will rely on the methodological framework of Arksey and O'Malley.[56] It has been developed with the aim of usage within the sphere of social sciences which makes it highly applicable to this research. Moreover, since literature on scoping reviews is quite scattered, this framework serves as a guide often used by scholars thanks to its clarity and conciseness.[57] Following these guidelines, the research question at hand was formulated broadly, to generate plenty of literature. The research question aimed to answer in this research sounds: *'what is known from the existing literature about the automated policing and reporting of CSAM (child sexual abuse material)?'*.

---

[52] Yinthe Feys, 'Criminological scoping reviews as part of a student assignment: methodological considerations' (2022) 57 Quality & Quantity 1; Zachary Munn and others, 'Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach' (2018) 18 BMC Medical Research Methodology; Danielle Levac, Heather Colquhoun and Kelly K O'Brien, 'Scoping studies: advancing the methodology' (2010) 5 Implementation Science 1; Mai T Pham and others, 'A scoping review of scoping reviews: advancing the approach and enhancing the consistency' (2014) 5 Research Synthesis Methods 371; Mark Petticrew and Helen Roberts, *Systematic reviews in the social sciences: A practical guide* (John Wiley & Sons 2008).

[53] Hilary Arksey and Lisa O'Malley, 'Scoping Studies: Towards a Methodological Framework' (2005) 8 International Journal of Social Research Methodology 19.

[54] Jessica Peterson and others, 'Understanding scoping reviews: Definition, purpose, and process' (2016) 29 Journal of the American Association of Nurse Practitioners 12; Micah D J Peters and others, 'Updated methodological guidance for the conduct of scoping reviews' (2020) 18 JBI Evidence Synthesis; Munn and others, 'Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach'; Maria J Grant and Andrew Booth, 'A typology of reviews: an analysis of 14 review types and associated methodologies' (2009) 26 Health Information and Libraries Journal 91.

[55] Arksey and O'Malley, 'Scoping Studies: Towards a Methodological Framework'; Heather L Colquhoun and others, 'Scoping reviews: time for clarity in definition, methods, and reporting' (2014) 67 Journal of Clinical Epidemiology 1291; Levac, Colquhoun and O'Brien, 'Scoping studies: advancing the methodology'.

[56] Arksey and O'Malley, 'Scoping Studies: Towards a Methodological Framework'.

[57] Colquhoun and others, 'Scoping reviews: time for clarity in definition, methods, and reporting'; Monika Kastner and others, 'What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review' (2012) 12 BMC Medical Research Methodology 1; Peterson and others, 'Understanding scoping reviews: Definition, purpose, and process'.

In this study both academic and grey literature was consulted, taking advantage of the flexibility a scoping review offers, to broaden the scope and diversity of the review.[58] Following Gusenbauer and Haddaway[59], a division was made between principal search engines, which meet all necessary quality requirements (Base, ScienceDirect, Scopus and Proquest), and supplementary databases that were chosen because of their unique qualities (IEEE Explorer and Calio for their speciality in technological and child abuse literature). Afterwards, snowball sampling was applied to the selected literature through hand-searching reference lists for more relevant studies.[60] The initial search terms included synonyms, related aspects and various characteristics related to the research question, aiming for full coverage of the topic.[61] After a test of the search terms, they were adjusted and revised when necessary, resulting in following search strings (automation AND 'Child sexual abuse material', automatic AND 'child sexual abuse material', AI AND 'child sexual abuse material', algorithm AND 'child sexual abuse material', hash AND 'child sexual abuse material', database AND 'child sexual abuse material', technology AND 'child sexual abuse material', 'content moderation' AND 'child sexual abuse material', 'digital forensics' AND 'child sexual abuse material', detection AND 'child sexual abuse material' , 'machine learning' AND'child sexual abuse material', 'big data' AND 'child sexual abuse material', 'digital investigations' AND 'child sexual abuse material'). It must be noted that the search string 'child sexual abuse material' was interchanged by commonly used synonyms CSAM, CEM (child exploitation material) and CSEM (child sexual exploitation material) for each search string to generate complete results. When the results were generated, a title-based selection occurred first, followed by an abstract based selection and finally, full text assessment.

Inclusion and exclusion criteria were formulated after the testing of the search strings, when familiarity with the literature was gained. Texts were included from January 2010 - June 2023, guided by the test phase which demonstrated no relevant literature before 2010. Only Dutch or English literature was included, based on the linguistic capacities of the researcher. Texts were included when they dealt with the *operational* side of policing CSAM. Consequently, texts were excluded that were solely describing legislation, written from a victimological/psychosocial perspective or discussed various types of cybercrime, with limited attention for CSAM. For the sake of strictly delineating the scope of this review, the focus was put on child sexual abuse images and not videos (but texts

[58] Arksey and O'Malley, 'Scoping Studies: Towards a Methodological Framework'; Karen M Benzies and others, 'State-of-the-Evidence Reviews: Advantages and Challenges of Including Grey Literature' (2006) 3 Worldviews on Evidence Based Nursing 55; Peterson and others, 'Understanding scoping reviews: Definition, purpose, and process'; Pham and others, 'A scoping review of scoping reviews: advancing the approach and enhancing the consistency'; Richard J. Adams, Palie Smart and Anne Sigismund Huff, 'Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies' (2017) 19 International Journal of Management Reviews 432.

[59] Michael Gusenbauer and Neal R Haddaway, 'Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources' (2020) 11 Research Synthesis Methods.

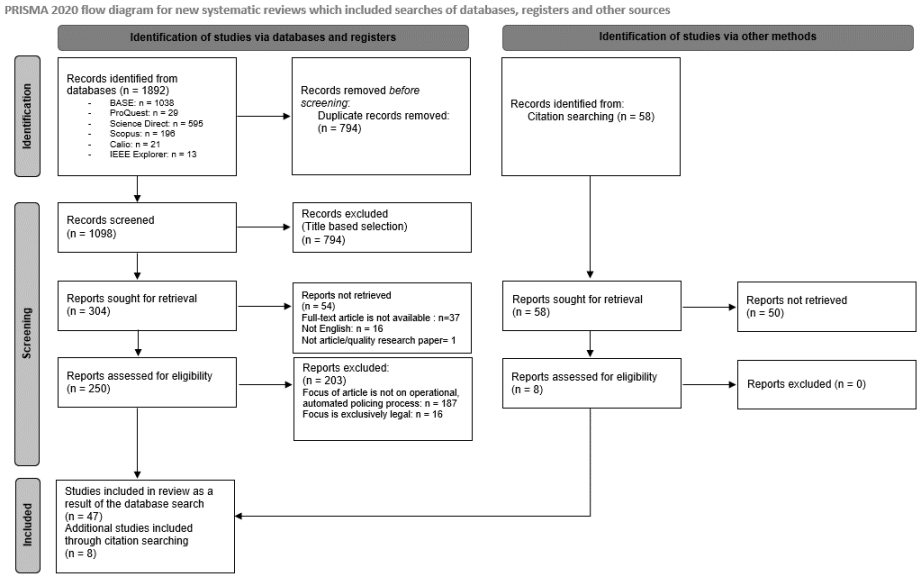
[60] Arksey and O'Malley, 'Scoping Studies: Towards a Methodological Framework'.

[61] Feys, 'Criminological scoping reviews as part of a student assignment: methodological considerations'.

discussing both were included). Texts on grooming were excluded and had to concern CSAM circulation on the surface web, not peer 2 peer groups or the dark net.

3 Results

The PRISMA flow diagram below elaborates on the number of texts that were included in the scoping review. In total 55 texts were analysed following a database search and a citation search. Reports were excluded based on different grounds. In the first title-based selection round 794 texts were excluded for not being relevant, based on the title. In the second round 37 texts were not retrieved because they could not be accessed, due to the university's and/or open source license and subscription limits, sixteen were not English and one was a low quality thesis, which was not included in the review. In the next phase, texts were selected based on the content of their abstracts. When it was unclear whether or not a text was relevant based on the abstract, the full text was assessed. In this round 203 texts were excluded because they did not discuss the automated policing of CSAM and were either looking at the topic from a psychological, victimological, offender-based perspective or an exclusively legal scope. After these selections, 47 texts were selected from the database search for full review and inclusion and eight additional texts were deemed relevant through citation searching.



The table below gives an overview of the texts included in the scoping review and maps out several categories of topics and the methodologies most used. The large majority of studies included were literature reviews or desk reviews. Other methodologies were qualitative interviewing and surveys or experimental designs to test automated tools, but these were in the minority. It can also be stated that a large amount of studies date

from the last five years, suggesting that attention for the topic has grown recently. Furthermore, it must be made clear that defining the territorial scope of the literature was not evident. Therefore in the overview below, those texts that have countries marked with an asterisk (*) means that the exact territorial scope could not be specified, and the geographic affiliation of the authors is presented. When no asterisk is present, the geographic scope of the study was clear and is specified. Most studies that had no specified territorial scope were literature reviews without such geographic limits and generally including global or international literature or experimental designs testing automated tools. Aside from this, either the author's affiliation or the territorial scope of the study were predominantly Western.

Table 1: scoping review literature

Author	Year	Type of publication	Geographic region	Study topic	Methodology
Anton Moser, Marlies Rybnicek and Daniel Haslinger	2015	Conference proceeding	Austria*	Issues paired with CSAM-detection technologies	Experimental design testing for child and nudity detection
Argyro Chatzinikolaou	2020	Journal article	EU	Online policies for CSAM	Literature review
Aurelija Pūraitė and Natalja Prokofjeva	2020	Journal article	EU	Policies around children's safety online	Literature review
Ben Wagner	2016	Journal article	Germany, UK, US	Governance of the internet sector concerning CSAM	Literature review
Benoit Leclerc, Jesse Cale, Thomas Holt and Jacqueline Drew	2022	Journal article	Australia	Police officers perspectives on investigating CSAM	Qualitative interviewing
Berisha Blerim	2019	Journal article	EU	EU guidelines concerning harmful content online	Literature review
Boglárka Meggyesfalvi	2021	Journal article	Hungary*	Content moderation on social media platforms	Literature review
Bryce Garreth Westlake	2020	Book chapter	US*	Challenges experienced by law enforcement in investigating CSAM	Literature review
Bryce Westlake, Martin Bouchard and Richard Frank	2012	Journal article	Canada*	Tools to detect CSAM	Experimental design using web-crawler to detect CSAM
Camila Laranjeira da Silva, Joao Macedo, Sandra Avila and Jefersson dos Santos	2022	Journal article	Brazil	Identifying attributes to use in CSAM detection	Literature review

Chad M.S. Steel	2014	Journal article	Global	Impact of measures to limit CSAM	Analysis of online search engine query statistics
Christina Angelopoulos and Stijn Smet	2016	Journal article	EU	Notice-and-takedown mechanisms	Literature review
Deisy Chaves , Eduardo Fidalgo, Enrique Alegre, Rocío Alaiz-Rodríguez, Francisco Jáñez-Martino and George Azzopardi	2020	Journal article	Spain, the Netherlands*	CSAM detection tools	Experimental design testing face detection tool performance
Ekaterina Markovich	2017	Research paper	Sweden	International measures against CSAM	Literature review
Elena Martellozzo and Jeffrey DeMarco	2020	Journal article	UK	IWF's processes at removing CSAM	Qualitative interviewing
Elena Pilipets and Susanna Paasonen	2022	Journal article	Austria, Finland*	User reactions to Tumblr's NSFW ban	Online content analysis
Elie Bursztein, Travis Bright, Michelle DeLaune, David M. Eliff, Nick Hsu, Lindsey Olson, John Shehan, Madhukar Thakur and Kurt Thomas	2019	Conference proceeding	US	Challenges experienced by law enforcement in investigating CSAM	Analysis of CSAM reports
Felix Anda , Nhien-An Le-Khac and Mark Scanlon	2020	Journal article	Ireland*	DeepUage detection model	Experimental design testing age-estimation tool performance
Genevieve Cameron, Ellie Mendez Sayer, Louisa Thomson and Sophie Wilson	2015	Industry report	Global	International measures against CSAM	Literature review
Hee-Eun Lee, Tatiana Ermakova, Vasilis Ververis and Benjamin Fabian	2020	Journal article	Germany, the Netherlands, Portugal*	Technological tools and accomplishments in investigating CSAM online	Literature review
Ion Rusu	2012	Journal article	EU	Regulations to tackle CSAM on a European level	Literature review
Jasmine V. Eggestein and Kenneth J. Knapp	2014	Journal article	US	Challenges in tackling online CSAM	Literature review

Jesse Cale, Thomas Holt, Benoit Leclerc, Sara Singh and Jacqueline Drew	2021	Journal article	Australia*	CSAM offending: offences, distribution and production	Literature review
John Carr and the International Centre for Missing and Exploited Children	2017	Industry report	Global	The global approach against CSAM	Literature review
Julia Davidson, Adriano Schimmenti, Vincenzo Caretti, Angelo Puccia, Elisa Corbari, Stefan Bogaerts, Janneke D. Schilder, Mia Scally, Antonia Bifulco and Jeffrey Nicholas DeMarco	2020	Book chapter	UK, Italy, the Netherlands, Ireland	Police approach to tackle CSAM	Qualitative interviewing, surveys, case studies
Jyri Rajamäki, Iiro Lahti and Johanna Parviainen	2022	Journal article	Finland*	Exploring the possibility of using OSINT on the dark net	Literature review
Kaspar Rosager Ludvigsen, Shishir Nagaraja, and Angela Daly	2022	Journal article preprint	UK*	Apple's client side scanning to tackle CSAM	Literature review
Kathryn C. Seigfried-Spellar, Gary R. Bertoline and Marcus K. Rogers	2011	Journal article	US, global	Regulation and measures to fight CSAM	Literature review
Kemal Veli Açar	2019	Journal article	Turkey*	CSAM repositories	Literature review
Laura Sanchez, Cinthya Grajeda, Ibrahim Baggili and Cory Hall	2019	Journal article	US	Challenges in investigating online CSAM	Survey
Lee MacLeod, David King and Euan Dempster	2020	Journal article	UK*	Age estimation forensic tools	Literature review
Lukas Struppek, Dominik Hintersdorf, Daniel Neider and Kristian Kersting	2022	Journal article	Germany*	The perceptual hashing components of NeuralHash	Experimental design testing Neuralhash performance
Martine B. Powell, Peter Cassematis, Mairi S. Benson, Stephen Smallbone and Richard Wortley	2013	Journal article	Australia	Police officers' perceptions of challenges of working with CSAM	Qualitative interviewing
Mary Graw Leary	2022	Research report	US	Section 230 of the Communications Decency Act	Literature review

Mateus de Castro Polastro and Pedro Monteiro da Silva Eleuterio	2010	Journal article	Brazil*	Nudetective as forensic tool to detect CSAM	Experimental design for testing Nudetective detection tool performance
Melissa Stroebel and Stacy Jeleniewski	2015	Research report	Global	CSAM hotlines around the world	Survey
Michael Salter and Elly Hanson	2021	Journal article	Australia*	Historical regulation strategies concerning CSAM	Literature review
Michael Salter and Lloyd Richardson	2021	Journal article	Canada, Australia*	The takedown of the Trichan imageboard	Literature review
Michal Grega, Damian Bryk and Bartłomiej Grabowski	2012	Journal article	Czech Republic*	Forensic tools to catalogue CSAM evidence	Experimental design for testing evidence cataloguing tool
Mitali Thakor	2018	Journal article	US*	Bias of online detection tools	Literature review
Mitali Thakor	2017	Journal article	US*	Interplay between police, state and tackling CSAM	Literature review
Olivia Cullen , Keri Zug Ernst , Natalie Dawes , Warren Binford and Gina Dimitropoulos	2020	Journal article	US	Challenges experienced by law enforcement officers in investigating CSAM	Qualitative interviewing, focus groups
Paul Bleakley, Elena Martellozzo, Ruth Spence and Jeffrey DeMarco	2023	Journal article	US, UK*	Challenges experienced by law enforcement in investigating CSAM	Literature review
Rubel Biswas, Deisy Chaves, Franciso Janez-Martino, Pablo Blanco-Medina , Eduardo Fidalgo, Carlos Garcıa-Olalla and George Azzopardi	2021	Journal article	Spain, the Netherlands*	Age estimation forensic tools	Experimental design testing age-estimation tool performance
Sarah Napier and Coen Teunissen	2022	Journal article	Australia*	ESP & CSAM: measures and processes	Literature review
Seny Kamara, Mallory Knodel, Emma Llansó, Greg Nojeim, Lucy Qin, Dhanaraj Thakur and Caitlin Vogus	2021	Industry report	US*	Content moderation in E2EE systems	Literature review
Thomas J. Holt, Jesse Cale, Benoit Leclerc and Jacqueline Drew	2020	Journal article	US, Australia*	Challenges experienced by law enforcement in investigating CSAM	Literature review

Timothy Gernand	2022	Journal article	US	Apple's Neuralhash	Literature review
TJ McIntyre	2010	Journal article	EU	Developments in blocking technologies	Literature review
TJ McIntyre	2013	Journal article	EU, UK, US	CSAM blocking by OSP	Literature review
Virginia N.L. Franqueira, Joanne Bryce, Noora Al Mutawa and Andrew Marrington	2017	Journal article	US, UK	Challenges in investigating online CSAM	Survey
Wanqi Lai, Amalia Van Vaerenbergh and Wannes Bellaert	2021	Journal article	EU	Art. 26 of the 2022 Europol regulation	Legislative analysis
WeProtect Global Alliance	2016	Industry report	Global	An 'ideal' model to tackle CSAM	Literature review
Wesam Al-Nabki, Eduardo Fidalgoa, Enrique Alegrea and Rocío Alaiz-Rodrígueza	2020	Journal article	Spain*	CSAM detection tools	Experimental design testing file classifiers for CSAM
Younes Akbari, Somaya Al-maadeed, Omar Elharrouss, Fouad Khelifi, Ashref Lawgaly and Ahmed Bouridane	2021	Journal article	UK, UAE, Qatar*	Source video identification techniques for forensic analysis	Literature review

3.1 First things first: regulating online regulation

Before diving deeper into the policing of online CSAM, it is useful to touch upon relevant regulation to explain how and why some policing processes occur. As mentioned above, the focus of this research is not on those legal frameworks governing online policing of CSAM but for the sake of completeness of the study, the overarching policy framework cannot be excluded.

As Kamara and colleagues[62] noted, technological solutions cannot singlehandedly solve the issue of CSAM. It is argued that technologies to detect CSAM are only effective *if* a strict policy or legal framework exists and that the absence of these guidelines in several countries may very well be one of the Achilles' heels of policing CSAM.[63] Therefore, political and societal causes of CSAM should also be considered and addressed. Changes in encryption software, image editing or the increasing use of peer-to-peer networks and

[62] Seny Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems* (2021).
[63] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

social media for CSAM-distribution are only a few examples in which technological developments outpace the law.[64] The responsibility to respond to this should not solely lie with the poster of harmful content but equally belongs to the government and ESPs.[65] However, throughout history industry has steadily chosen profits over measures benefitting child safety, which was tacitly allowed by government. The latter has valued growth and tech-profitability over strict CSAM regulation and has long sat on the sidelines and refrained from putting obligations on ESPs to tackle CSAM. Even when decisions made to tackle CSAM do come about, it is at a much slower pace than for-profit solutions to tackle the same problem.[66] Salter & Richardson[67] add to this and remark that calls for technological evolution are more often driven by civil society organizations rather than industry. As a result, many industry-NGO partnerships sprouted to tackle CSAM.[68]

It needs to be said that an evolution has been seen in regulatory stances throughout the years and strongly depends on the country at hand. Salter and Richardson[69] remarked in 2021 that governments are becoming more interventionist. When we look at the regulatory climate in Europe, it is in 2006 that the European Commission took an active stance against CSAM with the Safer Internet Program which would support blocking efforts and hotlines' activities and work more closely with ESPs.[70] Another noteworthy initiative is CIRCAMP (Cospol Internet Related Child Abusive Material Project) Action Plan. This is a network that involves assisting countries in establishing blocking measures through a collaborative effort between 16 countries, Interpol and Europol. This system is heavily police led which means that law enforcement decides on the illegal status of material.[71]

---

[64] Michael Salter and Lloyd Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material' (2021) 13 Policy & Internet 385.

[65] Boglárka Meggyesfalvi, 'Policing harmful content on social media platforms' (2021) 69 Belügyi Szemle.

[66] Jasmine V. Eggestein and Kenneth J. Knapp, 'Fighting Child Pornography: A Review of Legal and Technological Developments' (2014) 9 Journal of Digital Forensics, Security and Law.

[67] Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.

[68] Bryce Garreth Westlake, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection', *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (Palgrave Macmillan 2020); Paul Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?' (2023) 21 European Journal of Criminology.

[69] Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.

[70] T. J. McIntyre, 'Blocking child pornography on the Internet: European Union developments' (2010) 24 International Review of Law, Computers & Technology; Aurelija Pūraitė and Natalja Prokofjeva, 'Policy of the European Union On The Safety of Children In Cyber Space ' (2018) 21 Public Security and Public Order.

[71] McIntyre, 'Blocking child pornography on the Internet: European Union developments'; Ekaterina Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet', Lund University (2017); Genevieve Cameron and others, *Child Sexual Exploitation: A study of international comparisons* (2015).

Mcintyre[72] mentions that this system embraces overblocking, supported by the argument that this can have a deterrent effect. Furthermore, possibilities for direct data-exchange between Europol and private parties such as service providers have been expanded, allowing for more flexibility and direct requests and receipt of information in the field of CSAM.[73] A study by Rusu[74] stated that in Europe, member states need to ensure measures for prompt removal of internet pages containing or disseminating CSAM hosted on their territory. However, those measures should be proportional so that restrictions are limited to what is necessary and should ensure transparent procedures and adequate protections for users. In 2018, the European Commission released guidelines for tackling illegal content online, as part of its Digital Single Market strategy. These guidelines encouraged online platforms to automate flagging of this content on a voluntary basis. They are also stimulated to engage in rapid automated detection and effective removal of harmful content as well as preventing detected material from reappearing. The guidelines do emphasize transparency thereof and the provision of safeguards to users.[75]

When assessing the state of regulation in the United States, it becomes clear that a different approach is adhered to. In the US an industry-based model was most prominent to regulate CSAM. The Communications Decency Act criminalized indecent communication but simultaneously pushed self-regulation by private actors.[76] In line with this, Mcintyre[77] argues that US-based ESPs developed their own in-house blocking lists, fearful of having to adhere to a government-made list. In the UK a similar industry-led approach is taken, albeit heavily government-influenced. It should be mentioned that none of those private actors mentioned have the capacity to regulate the internet on their own. However, this power has been spread across the internet and amongst private parties, which results in a trust on private implementation of content regulation and the emergence of self-regulatory systems.[78]

That being said, Chatzinikolaou[79] studied policies of online service providers and determined that platforms such as Facebook, Instagram, Snapchat and YouTube adhered to quite similar approaches towards images featuring sexual depictions of children. As can be expected, CSAM is restricted on these platforms, but nude representations of children

[72] T.J. McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems' in Ian Brown (ed), *Research Handbook on Governance of the Internet* (2013).
[73] Lai, Vaerenbergh and Bellaert, 'Europol and its growing alliance with private parties'.
[74] Ion Rusu, 'Preventing Offenses of Sexual Abuse, Sexual Exploitation of Children and Child Pornography in the European Union' (2012) 8 Acta Universitatis Danubis Juridica.
[75] Blerim Berisha, 'Platforms Regulation in the Digital Single Market Strategy-Illegal Content and Intellectual Property as a Key Challenge' (2019) 22 European Research Studies Journal.
[76] Ben Wagner, 'The Public Sector and Content Regulation: Focussing on Pornography and Child Sexual Abuse Material', *Global Free Expression - Governing the Boundaries of Internet Content, Law, Governance and Technology* (Springer International Publishing 2016).
[77] McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems'.
[78] Wagner, 'The Public Sector and Content Regulation: Focussing on Pornography and Child Sexual Abuse Material'.
[79] Argyro Chatzinikolaou, 'Sexual images depicting children: the EU legal framework and online platforms' policies' (2020) 11 European Journal of Law and Technology.

taken in an innocent context are as well. The setting in which an image was taken is of little to no relevance when regulating content on these platforms, resulting in the restriction of more than only CSAM. This can be explained by the platforms trying to avoid liability for not removing or blocking potential CSAM. Consequently, it was noted that this causes over-enforcement and signals the need for and importance of clear policy guidelines.[80]

Scholars often voice critique about the inadequacy of these policies and the need for governments and companies to step up, as it is acknowledged that cooperation between the two is very important.[81] To illustrate this, in most EU-states blocking systems are not legislated and rely on voluntary agreements between service providers and law enforcement. While acknowledging the importance of cross-stakeholder collaboration, Salter and Richardson[82] note that this multistakeholder paradigm inherent to tackling CSAM is exactly what makes it so difficult for regulations to have impact and meet their goal. They refer to a takedown incident whereby authorities experienced significant issues in getting CSAM-littered imageboards inaccessible. They note that the government is hesitant on interfering in internet regulation and thus allows tech companies to decide on the scope of censorship, regulatory costs and industry liability. They also mention that there are no clear frameworks or benchmarks that illustrate not only what to do when hosting providers are unresponsive to takedown requests but what is expected from providers in fighting CSAM generally.

But even when guidelines on aforementioned issues are established, they are met with critique questioning how far collaboration between state agencies and service providers should be able to go. Paradoxically, Salter & Hansen[83] note that when certain companies receive critique for not acting upon reports of CSAM, they defend this lack of action by trying to protect their users' privacy, despite their business models being centred upon surveillance and privacy invasion. Likewise, Markovich[84] argues that automated tools may entail censorship and surveillance on citizens' communications and that allowing algorithms to filter out legal from illegal is harmful to our fundamental rights. Moreover,

---

[80] Ibid.

[81] Olivia Cullen and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States' (2020) 9 Laws; Meggyesfalvi, 'Policing harmful content on social media platforms'; Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet; Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.

[82] Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.

[83] Michael Salter and Elly Hanson, '"I Need You All to Understand How Pervasive This Issue Is": User Efforts to Regulate Child Sexual Offending on Social Media' in Jane Bailey, Asher Flynn and Nicola Henry (eds), *The Emerald International Handbook of Technology Facilitated Violence and Abuse* (2021).

[84] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

it is said that putting the responsibility for scanning and monitoring of our communications on online platforms leads to excessive focus on the speed and amount of takedown without adequately assessing the consequences regarding principles like free speech and privacy or determining whether removed content is actually illegal.[85] A possible explanation for allowing this comes from Thakor[86], who argues that usage of these technological tools are justified through the moral obligation and urge to fight CSAM, alluding to a general belief in technological infallibility.[87]

Furthermore, the lack of uniform legislation and understanding of what defines CSAM is seen as problematic when trying to establish an effective, global fight against CSAM. There is a large necessity for policies to be harmonized and unified between countries and companies.[88] This is specified in a study by Stroebel & Jeleniewski[89] and Bleakley and others[90] which mention that legal content in one hosting country may not be referred to a country in which the material is illegal or that a CSAM report could be found illegal by an analyst who sends it to a country in which it is not prohibited and thus, no action is taken upon it. Similarly, Puraite and Prokofjeva[91] note that whereas countries could have a similar criminalization of CSAM, the understanding of what constitutes a child and the age of consent or criminal responsibility could differ, which makes effective cooperation more difficult.[92]

3.2 Scanning, detecting and reporting: where it all begins

CSAM is increasingly detected using digital techniques. This has transformed the policing of CSAM and distributed these tasks to actors such as technology companies and computer scientists.[93] Alleged CSAM is discovered in several ways. Webpages and social media are a popular method for hosting and distribution, signalling the important role

[85] Berisha, 'Platforms Regulation in the Digital Single Market Strategy-Illegal Content and Intellectual Property as a Key Challenge'.
[86] Mitali Thakor, 'Digital Apprehensions: Policing, Child Pornography, and the Algorithmic Management of Innocence' (2018) 4 Catalyst feminism, theory, technoscience.
[87] Mitali Thakor, 'How to Look: Apprehension, Forensic Craft, and the Classification of Child Exploitation Images' (2017) 39 IEEE Annals of the History of Computing.
[88] WePROTECT Global Alliance, *Preventing and Tackling Child Sexual Exploitation and Abuse (CSEA): A Model National Response* (2016).
[89] Melissa Stroebel and Stacy Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges* (2015).
[90] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.
[91] Pūraitė and Prokofjeva, 'Policy of the European Union On The Safety of Children In Cyber Space'.
[92] Julia Davidson and others, 'Exploring policing and industry practice in the prevention of online child sexual abuse' in India Bryce and Wayne Petherick (eds), *Child Sexual Abuse: Forensic Issues in Evidence, Impact, and Management* (Academic Press 2020).
[93] Thakor, 'Digital Apprehensions: Policing, Child Pornography, and the Algorithmic Management of Innocence'.

ESPs have in the proactive policing of CSAM.[94] The most common ways CSAM is discovered, is through user-reporting and upon proactive detection by service providers.[95] Social media companies hire content review teams that can be located in-house or are outsourced. These content reviewers decide whether to report certain material. Moreover, they are training technology through specifying what it can or cannot recognize as abusive material.[96] It is noted that service providers enjoy a unique position to report potential CSAM to law enforcement as they are gatekeepers to the internet.[97] Therefore, content moderators have an essential role in the self-regulation of platforms.[98] Consequently, collaboration between ESPs and law enforcement is necessary, since ESPs have a much more advantageous position to access potentially abusive material or web pages with abuse circulating on it, as opposed to law enforcement. Furthermore, Holt and colleagues[99] mention that these service providers have greater knowledge of their platform's weaknesses and the type of users (and their data) on it. They also have better access to certain data sets, technology, budgets and human resources. Currently, numerous service providers such as Google, Microsoft, Instagram, Meta and Twitter are involved in the protection of children.[100] Important to note is that ESPs are not systematically required to monitor webpages and that these actions remain voluntary.[101] They are not allowed to arrest or prosecute individuals but should material be found, they are required to report the potential CSAM to a designated organization in certain circumstances.[102]

One of the most commonly used technologies to automatically detect CSAM is by using a hashing system.[103] Google, Facebook and Microsoft have produced these sort of tools

---

[94] Jesse Cale and others, 'Crime commission processes in child sexual abuse material production and distribution: A systematic review' (2021) Trends & issues in crime and criminal justice; Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'; Thomas J. Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses' (2020) 55 Agression and Violent Behavior.

[95] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.

[96] Thakor, 'Digital Apprehensions: Policing, Child Pornography, and the Algorithmic Management of Innocence'.

[97] Westlake, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection'; Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

[98] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.

[99] Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'.

[100] Ibid; Timothy Gernand, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge' (2022) 127 Dickinson Law Review.

[101] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

[102] Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'; Kathryn C Seigfried-Spellar, Gary R. Bertoline and Marcus K. Rogers, 'Internet Child Pornography, U.S. Sentencing Guidelines, and the Role of Internet Service Providers' (International Conference on Digital Forensics and Cyber Crime).

[103] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

and use it to detect material on their services.[104] Hashing technology is used to detect known or already discovered material and is popular since it allows for rapid scanning of large volumes of images and enables the verification of images that are copies of those already stored.[105] Previously detected images get assigned a hash, comparable to a digital fingerprint. This hash is generated through a mathematical algorithm and consists of a 24-character numerical code. This code serves as authentication for a certain image. Confirmed hashes of CSAM material are stored in a database which allows for quick hash-matching in case an already detected image surfaces so it can be blocked quickly.[106]

These databases warrant some extra attention. Multiple currently exist that either store verified hashes of confirmed CSAM, URLs of confirmed webpages hosting CSAM, confirmed victims and perpetrators or confirmed keywords used by perpetrators.[107] These are established by involved parties such as law enforcement, child-protection organizations or tech companies and ESPs. A famous example of this is the ICSE database (International Child Sexual Exploitation database) manufactured by Interpol. It provides law enforcement with software that executes analysis of images and connects victims, abusers and locations.[108] NCMEC also operates a large database to which service providers can share information. Many smaller national databases exist as well. It should be remarked that little harmonization between the repositories currently exists. It may thus be possible that a country is a member of Interpol, the NCMEC database and has their own national one, and therefore needs to enter information multiple times. Granted, certain databases have a different focus. ISCE is limited to abuse of actual children whereas the NCMEC database also includes virtual or manufactured CSAM that doesn't involve 'real' children. Still, the lack of standardization and unification between these databases does not strengthen the effectiveness in fighting CSAM.[109]

Furthermore, machine learning or deep learning, which are both versions of AI, are used to detect new, previously undiscovered material since hashing is only capable of detecting material that has already been assigned a digital fingerprint.[110] These techniques are noted by Lee and others[111] to be much more robust against modifications than hashing

[104] Chad M.S. Steel, 'Web-based child pornography: The global impact of deterrence efforts and its consumption on mobile platforms' (2015) 44 Child abuse & Neglect; Eggestein and Knapp, 'Fighting Child Pornography: A Review of Legal and Technological Developments'.

[105] Gernand, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge'.

[106] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'; Eggestein and Knapp, 'Fighting Child Pornography: A Review of Legal and Technological Developments'.

[107] Westlake, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection'; John Carr, *A Joint Report on Online Child Protection Combatting Child Sexual Abuse Material on the Internet* (2017).

[108] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

[109] Kemal Veli Açar, 'Framework for a Single Global Repository of Child Abuse Materials' (2019) 11 Global Policy.

[110] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

[111] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

technologies. Additionally, it can be used to limit the reviewing process and perform facial clustering (to group reports that contain a same child or perpetrator) or to limit the investigation process by prioritizing which material should be looked into first.[112] The latter can be achieved by detecting scene information (to detect objects or landmarks to give geographic indications) or by facial detection, gender and age prediction and nudity recognition.[113] Deep learning approaches have been noted by some scholars to be highly effective for automated face detection with high accuracy and could be used to discern CSAM from adult pornographic imagery or to filter out faces of minors from adult faces.[114] Lee and colleagues[115] and Westlake and others[116] mention that the best detection results are obtained when several techniques are combined.

It needs to be noted that detection tools and software have been developed and manufactured by technology companies and made available to law enforcement and child-protection agencies. This phenomenon is coined 'detection as a service' by Lee and colleagues.[117] Accordingly, Microsoft has offered their hashing technology PhotoDNA for free to law enforcement and is used globally by over 100 companies. Popular companies like Snapchat, Twitter and TikTok all use PhotoDNA.[118] Nevertheless, it took companies like Google almost five years to use PhotoDNA. Subsequently, some NGO's have succeeded in integrating PhotoDNA in their own software for detection, reporting and removal of CSAM.[119] Similarly, Google provides its Content Safety API, a tool for identifying and triaging CSAM by using AI, to companies and NGO's.[120] Facebook also developed technology to detect new CSAM and offers its technology to stakeholders.[121]

3.3 The perils and pitfalls of technological solutionism

The previous paragraph illustrates that parties engaged in tackling CSAM heavily rely on automated technologies to process large volumes of potential CSAM circulating

---

[112] Felix Anda, Nhien-An Le-Khac and Mark Scanlon, 'DeepUAge: Improving Underage Age Estimation Accuracy to Aid CSEM Investigation' (2020) 32 Forensic Science International: Digital Investigation.

[113] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*; Mateus de Castro Polastro and Pedro Monteiro da Silva Eleuterio, *NuDetective: a Forensic Tool to Help Combat Child Pornography through Automatic Nudity Detection* (2010).

[114] Deisy Chaves and others, 'Assessment and Estimation of Face Detection Performance Based on Deep Learning for Forensic Applications' (2020) 20 Sensors.

[115] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

[116] Bryce Westlake, Martin Bouchard and Richard Frank, *Comparing Methods for Detecting Child Exploitation Content Online* (2012).

[117] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

[118] Coen Teunissen and Sarah Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview' (2022) Trends & issues in crime and criminal justice.

[119] Salter and Hanson, '"I Need You All to Understand How Pervasive This Issue Is": User Efforts to Regulate Child Sexual Offending on Social Media'.

[120] Teunissen and Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview'.

[121] Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'; Teunissen and Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview'.

online. Nonetheless, it seems that deployed technologies are not adequate and leave room for improvement.[122]

Criticism was detected in literature concerning the over-blocking of content and referred to risks such as a lack of transparency, legitimacy, accountability, fair procedures, possible mission creep, privacy issues and insufficient effectiveness.[123] In 2018, social media application Tumblr was removed from the app store after it was accused of being a breeding ground for CSAM.[124] Consequently, it started removing all NSWF content ('not suitable for work', including CSAM). However, it soon became clear that the deployed AI algorithm removed more than just potential CSAM as the platform received significant backlash and dissatisfaction. Angelopoulos and Smet[125] argued that filtering out illegal content is impossible without screening the totality of content available and therefore, those systems result in general monitoring, regardless of whether the process is automated or performed by human supervision. Westlake[126] adds that this extensive monitoring may also be seen as financially and labour intensive, aside from giving rise to severe privacy infringements. False positives and corporate-led content-moderation is also linked to interfering with freedom of expression.[127] Moreover, the effectiveness of these tools is often contested and it is argued that they merely serve as an illusion for action and instil a naïve faith in technology which reduces pressure to establish clear policies (also called technological solutionism).[128] This is mirrored in how these technologies are often marketed: as being highly accurate to produce little false positives and being able to process CSAM at a speed and accuracy unattainable by manpower.[129] Salter & Richardson[130] coined the term 'techlash' for referring to criticism surrounding the monopoly that service providers have and their impact on user's privacy paired with the lack of accountability and transparency in their reporting and detection processes. Per illustration, there is little info on the CSAM-definitions that providers use and for some, little is known about the general detection process and tools used.[131]

---

[122] Meggyesfalvi, 'Policing harmful content on social media platforms'.
[123] McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems'.
[124] Elena Pilipets and Susanna Paasonen, 'Nipples, memes, and algorithmic failure: NSFW critique of Tumblr censorship' (2022) 24 New media & society.
[125] Christina Angelopoulos and Stijn Smet, 'Notice-and-fair-balance: how to reach a compromise between fundamental rights in European intermediary liability' (2016) 8 Journal of Media Law.
[126] Westlake, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection'.
[127] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.
[128] McIntyre, 'Blocking child pornography on the Internet: European Union developments'.
[129] Thakor, 'How to Look: Apprehension, Forensic Craft, and the Classification of Child Exploitation Images'.
[130] Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.
[131] Teunissen and Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview'.

Furthermore, it appears that perceptual hashing systems are not robust against detection evasion attacks and seem to be less progressive detection methods.[132] This entails that it is possible to slightly modify images or to divide it into multiple parts which generates a new hash, without needing extensive technological knowledge, resulting in detection evasion.[133] Similarly, an image can be altered to change a hash to a specific value which could deliver false-positive results for normal, non-illegal content. Another method to circumvent detection is the manufacturing of unique images that have no recognized hashes.[134] Nevertheless, the public opinion on hashing techniques is not black or white as some call this method to be a form of mass-surveillance whereas other say it is minimally intrusive.[135] A specific hashing tool that received ample critique was Apple's Neuralhash which was introduced to prevent reputation damage after the company was said to be favouring privacy-respecting measures as opposed to actively fighting CSAM. Neuralhash aimed to install CSAM hashes on user's devices and scan their personal information.[136] This proposed scanning on user's devices (client side scanning) was unprecedented and met with severe backlash to have serious implications for users privacy, democratic principles and human rights since the technique is said to be unable to achieve what it is supposed to and allows mass surveillance.[137] Struppek and colleagues[138] and Kamara and colleagues[139] also note that caution is in order when it comes to hashing technologies since risks of the technology's scope expanding are possible, which makes it sensitive to misuse.

Laranjeira and colleagues[140] mentioned that current technology relevant for detecting new CSAM (being AI) tends to focus on attributes related to age and pornography (ie age estimation tools trying to correctly filter out minors or tools aimed at nudity detection as an attribute for CSAM). However, Lee and colleagues[141] state that both methods are paired with difficulties related to discerning CSAM from pornography and correctly

---

[132] Michał Grega, Damian Bryk and Maciej Napora, 'INACT—INDECT Advanced Image Cataloguing Tool' (2012) 68 Multimedia Tools and Applications.

[133] Lukas Struppek and others, *Learning to Break Deep Perceptual Hashing: The Use Case NeuralHash* (2022); Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*; Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

[134] Eggestein and Knapp, 'Fighting Child Pornography: A Review of Legal and Technological Developments'.

[135] McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems'.

[136] Gernand, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge'.

[137] Ibid; Kaspar Rosager Ludvigsen, Shishir Nagaraja and Angela Daly, 'YASM (Yet Another Surveillance Mechanism)' (2022) Arxiv; Teunissen and Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview'.

[138] Struppek and others, *Learning to Break Deep Perceptual Hashing: The Use Case NeuralHash*.

[139] Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*.

[140] Camila Laranjeira and others, *Seeing without Looking: Analysis Pipeline for Child Sexual Abuse Datasets* (2022).

[141] Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.

estimating the age of minors. Chaves and colleagues[142] also declared that when the object of automated face detection is CSAM or minors in general, it is met with challenges pertaining to tarnished image quality, bad resolution and (too) strict time limitations in which the processing must occur. Also, face mimics and positioning further aggravate this issue and often, a clear distinction between nudity in an abusive setting and nudity in, for example, a holiday setting cannot be made.[143] Another study by Biswas and others[144] and Anda and others[145] noted that many of these age estimation tools are trained with unbalanced data sets and do not perform well for minors or have little ethnic and demographic variety. The issue of unreliable data or databases was touched upon by many authors and seems to be one of the biggest challenges related to the deployment of AI.[146] Thakor[147] specified that databases of abused children hosted by NCMEC contain more images and records of white children than non-white children. This needs to be kept in mind since software is very much dependent on the quality of the data it is trained with and could induce biased tools since an AI algorithm is only as good as the data it is trained with. Something that is found noted by several authors is that no matter how good an AI algorithm is, it will always be essential to have a human in the loop for control, which has practical limitations in it application.[148] Laranjeira and others[149] propose an approach that not only pays attention to age estimation or nudity detection but also assesses visual cues like environment, scene, objects in the frame or child related cues.

Lastly, while efforts to fight CSAM by companies such as Google, Microsoft and Meta should be applauded, it cannot be stressed enough that many service providers and technology companies are not contributing to the fight against CSAM. It is therefore possible that blocking efforts on one medium leads to a displacement of users to platforms not

[142] Chaves and others, 'Assessment and Estimation of Face Detection Performance Based on Deep Learning for Forensic Applications'.

[143] Anton Moser, Marlies Rybnicek and Daniel Haslinger, 'Challenges and Limitations Concerning Automatic Child Pornography Classification' (10th International Conference on Computer Vision Theory and Applications).

[144] Rubel Biswas and others, 'Reinforcement of age estimation in forensic tools to detect Child Sexual Exploitation Material' (2021) 34 Investigación en Ciberseguridad: Actas de las VI Jornadas Nacionales.

[145] Anda, Le-Khac and Scanlon, 'DeepUAge: Improving Underage Age Estimation Accuracy to Aid CSEM Investigation'.

[146] Younes Akbari and others, 'Digital forensic analysis for source video identification: A survey' (2022) 41 Forensic Science International: Digital Investigation; Lee MacLeod, David King and Euan Dempster, *A review of age estimation research to evaluate its inclusion in automated child pornography detection* (2020); Westlake, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection'.

[147] Thakor, 'Digital Apprehensions: Policing, Child Pornography, and the Algorithmic Management of Innocence'.

[148] Jyri Rajamäki, Liro Lahti and Johanna Parviainen, 'OSINT on the Dark Web: Child Abuse Material Investigations' (2022) 53 Information & Security 21; Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*.

[149] Laranjeira and others, *Seeing without Looking: Analysis Pipeline for Child Sexual Abuse Datasets*.

actively fighting CSAM.[150] On top of this, some platforms only act after facing big reputational damage for not engaging in the fight against CSAM. Apple, Twitter, Snapchat and YouTube are only a few examples of this.[151]

3.4 NGO's and non-profits: separating the wheat from the chaff

Once such material is detected through user-reporting or proactive detection by an ESP, it is either made aware to an ESP or a child-protection hotline.[152] In the US, should service providers fail to report to NCMEC's Cybertipline, they violate federal law.[153] These hotlines have played a very important role as clearinghouses for CSAM-reports and are responsible for triaging and sending them to the country in which the material is hosted. They are agencies specialized in CSAM and since they (often) lack official accountability, they are focused on public service tasks and in a powerful position to engage public monitoring mechanisms. These tasks include serving as a knowledge hub and operating a public hotline for the reception of CSAM-reports. So, whereas they cannot arrest or prosecute offenders, they are essential in providing leads to law enforcement.[154] Most of these hotlines are non-profits, government organizations, academic institutions or service provider associations.[155]

Hotlines will have three to four analysts who process reports and are tightly involved in the technological improvement and development of their services. Some hotlines will pass on reports without performing verification checks whereas others will assess the illegal status of a report. If a report is located within the jurisdiction of that hotline, it can directly ask a service provider to take the material down.[156] If it is hosted elsewhere, the hotline reports to INHOPE, an umbrella organization that functions as a connection network for all globally existing CSAM-hotlines. INHOPE is a network of 51 hotlines in 45 countries worldwide, committed to fighting CSAM.[157] These hotlines can also report directly to Interpol or Europol.[158]

---

[150] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

[151] Salter and Hanson, '"I Need You All to Understand How Pervasive This Issue Is": User Efforts to Regulate Child Sexual Offending on Social Media'.

[152] Chatzinikolaou, 'Sexual images depicting children: the EU legal framework and online platforms' policies'; Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

[153] Gernand, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge'.

[154] Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'.

[155] Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.

[156] Seigfried-Spellar, Bertoline and Rogers, 'Internet Child Pornography, U.S. Sentencing Guidelines, and the Role of Internet Service Providers'; Carr, *A Joint Report on Online Child Protection Combatting Child Sexual Abuse Material on the Internet*.

[157] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.

[158] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

The most well-known of those hotlines is the US' National Centre for Missing and Exploited Children's Cybertipline. 2008 legislation in the US requires service providers to notify NCMEC once potential CSAM is found.[159] This obligation and the fact that many service providers are located in the US, may explain the reporting gap between the US and other countries.[160] In 2022 it housed 340 employees and plays a principal role in fighting CSAM.[161] Despite the significant amount of employees, the number of CSAM reports are exceeding even NCMEC's capabilities. Holt and others[162] mention that this incapability to keep up may limit their legitimacy and perceived competency in the public eye. Another such hotline is the UK based Internet Watch Foundation. It is involved in detecting and removing CSAM content in the UK. Like many other hotlines they compose a hash and URL list of confirmed CSAM. The IWF also performs proactive searches for CSAM in order to take it down. Many other countries globally such as Sweden or New-Zealand similarly operate hotlines and have NGO-law enforcement collaborations to detect and report CSAM.[163]

Whereas the role of these organizations is vital in tackling CSAM, it is by no means flawless. Hotline analysts report that they struggle to keep up with the increasing number of CSAM on social network sites and reaching the more hidden places on the internet. The increasing volume of data per report adds to the issue. They also report problems related to a lack of funds and resources, specifically related to technological development.[164] A study by Martellozo and DeMarco[165] demonstrated that the IWF lacks standardised practices when it comes to their notice & takedown procedures and that there is a lack of harmonization in categorization of images. Furthermore, it was discovered that employees have difficulties in maintaining a charitable status about the organization while operating within a business-oriented framework. Other challenges mentioned is non-compliance by employees in using provided software tools that should make their work easier or the fact that they need to juggle several hash lists, since many of these organizations have their own isolated list, as a result of which work and information is duplicated.

[159] Thakor, 'How to Look: Apprehension, Forensic Craft, and the Classification of Child Exploitation Images'.
[160] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.
[161] Gernand, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge'; Cameron and others, *Child Sexual Exploitation: A study of international comparisons;* Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'.
[162] Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'.
[163] Seigfried-Spellar, Bertoline and Rogers, 'Internet Child Pornography, U.S. Sentencing Guidelines, and the Role of Internet Service Providers'.
[164] Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.
[165] Elena Martellozzo and Jeffrey DeMarco, 'Exploring the removal of online child sexual abuse material in the United Kingdom: Processes and practice' (2020) 22 Crime Prevention & Community Safety.

Importantly, stakeholders working with IWF reported that the organization lacks transparency in their work. Stroebel and Jeleniewski[166] pointed out that analysts use around three off-the-shelf and commercial applications and that the types of tools used vary widely. Hotlines that use their own in-house applications, reported using fewer tools than those that exclusively use off-the-shelf tools. Some of these applications used were to establish databases, search URL's or send reports through the INHOPE network. It is also striking to note that respondents in this study perform most operational steps manually. Following steps are ranked from least to most automated: the forwarding of reports, the assessment of content, notice & takedown procedures, tracing a report and prioritizing a report. Furthermore, in this survey two thirds report to *not* use hashing tools. Similarly, Bursztein and others[167] mention that analysts are still required to manually check the large influx of automatically detected reports.

Challenges related to organizational and international collaboration are prevalent. IWF-members reported to be quick in responding to takedown requests but due to their lack of legal power in enforcing them and having to rely on law enforcement for this, the process takes longer.[168] It is important to add that the level of collaboration between service providers and hotlines differs greatly. Respondents in Stroebel & Jeleniewski's[169] study note that the cooperation level between hotlines and service providers is actually very high and that generally, abusive material gets taken down quite quickly. They do mention that personal contact with the providers is essential for this. Nevertheless, Bleakley and others[170] mention that child protection NGO's report unacceptably long delays in the takedown of material and the re-emergence of already reported material. In addition to this, Canadian child-protection project Arachnid reported a small but significant delay in the removal of CSAM, when requesting ESPs. Furthermore, other hotlines analysts acknowledge the difficulty in working together and data sharing due to the absence of uniform operational guidelines and mention that it is a challenge to collaborate with other hotlines due to policy differences. They note that often, work and efforts are duplicated and investigations are slowed down.

3.5 From report to law enforcement: all's well that ends well?

The previous paragraphs have shown that detecting reports from social media platforms and getting them to relevant organizations for processing is not an easy feat. Additionally, one would hope that when a report reaches the designated law enforcement agency, the most difficult steps in the reporting chain are over and police are significantly aided

---

166 Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.

167 Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

168 Martellozzo and DeMarco, 'Exploring the removal of online child sexual abuse material in the United Kingdom: Processes and practice'.

169 Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.

170 Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.

by the efforts made by ESPs and civil organizations. Unfortunately, this idea remains but an illusion. A sidenote needs to be made that in policing CSAM, traditional policing tasks are not solely destined for official, recognized law enforcement agencies anymore. That task is also divided between content moderators, analysts at tech companies and child-protection hotlines. This expansion into a policing network is essential to understand CSAM-policing.

A study by Leclerc and others[171] demonstrated that law enforcement officers feel like they lack knowledge on the technological and media aspects related to CSAM and that they are insufficiently educated and trained on how to investigate the crime.[172] Despite all technological advancements made by law enforcement, it remains a costly and time-consuming task.[173] A lack of human, technological and hardware resources was mentioned several times as being a burden for investigations.[174]

Moreover, technology (and offender's knowledge thereof) is evolving too rapidly in order for them to keep up in a sufficient way alongside the volume of data in cases rising too quickly to deal with in a timely manner.[175] Cale and colleagues[176] demonstrated that investigators are experiencing significant challenges following the rise in popularity of encrypted communications.

Law enforcement officers feel unfamiliar with foreign media (an inherent part of the transnational nature of CSAM) and report challenges in international collaboration. Besides, there is great variability in how international cooperation is experienced. Bear in

[171] Benoit Leclerc and others, 'Child Sexual Abuse Material Online: The Perspective of Online Investigators on Training and Support' (2022) 16 Policing: A Journal of Policy and Practice.
[172] Martine B. Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation' (2013) 37 Policing: An International Journal.
[173] Davidson and others, 'Exploring policing and industry practice in the prevention of online child sexual abuse'; Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.
[174] Cale and others, 'Crime commission processes in child sexual abuse material production and distribution: A systematic review'; Cullen and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States'; Sanchez and others, 'A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM)'; Virginia N.L. Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches' (2018) 24 Digital investigation; Davidson and others, 'Exploring policing and industry practice in the prevention of online child sexual abuse'; Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation'.
[175] Cullen and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States'; Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'; Sanchez and others, 'A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM)'; Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches'.
[176] Cale and others, 'Crime commission processes in child sexual abuse material production and distribution: A systematic review'.

mind, law enforcement is highly dependent and reliant on service providers for receiving reports of potential CSAM and for replying to notice and takedown orders to block CSAM.[177] Still, both parties remain dependent on user-reports. Salter and Richardson[178] declared that in an attempt to take down imageboards circulating CSAM by an official child-protection NGO, certain private actors refused to remove certain material, since this cooperation is based on voluntary principles. This was confirmed by Cullen and others[179] and Powell and colleagues[180] who reported that law enforcement investigators are concerned with the slow or non-response of service providers. Davidson and colleagues[181] report that law enforcement doesn't see industry as a frequent partner, despite welcoming this collaboration. They are concerned that without mandatory reporting regulations, the response rate would be very low since industry still works in accordance with their own agenda. A study by Franquiera and others[182] and Powell and others[183] showed that the lack of uniform investigative guidelines and priorities for law enforcement, judiciary services and service providers further challenge international cooperation. Limits on obligations for service providers to store and process user data in different countries seems to further complicate matters in investigations.[184]

It is important to shed light on the use of technology and software tools by law enforcement investigators investigating CSAM. These tools are not only used by hotline analysts or service providers that detect material. When a report is in the hands of a law enforcement investigator, the same technology can be used. A survey by Sanchez and others[185] demonstrated that most investigators use one or more tools and that the most popular ones are commercial, previously touched upon and labelled as 'detection as a service'. Open-source tools readily available are less popular. The most used ones are skin detection tools, followed by age and gender estimation. Benefits of these systems include

[177] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet; Holt and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses'.
[178] Salter and Richardson, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material'.
[179] Cullen and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States'.
[180] Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation'.
[181] Davidson and others, 'Exploring policing and industry practice in the prevention of online child sexual abuse'.
[182] Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches'.
[183] Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation'.
[184] Markovich, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet'.
[185] Sanchez and others, 'A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM)'.

speeding up investigations and aiding prioritization. However, another study by Franquiera and colleagues[186] notes that investigators want more research into age-classification tools, despite this already being a widely researched avenue. It may be possible that in the case of age-estimation, theory fails to translate into practise. Remarkably, Powell and others[187] reported that law enforcement officers were aware of software to aid CSAM investigations but only few had access to it. They also note several limitations, namely the fact that hashing tools cannot detect new CSAM (but they also admit that this makes it work faster) and the generation of false positives and negatives by certain software. Additionally, it was mentioned that hash lists can sometimes be outdated and that there tends to be an overreliance on them to tackle CSAM.[188]

4 Discussion

When assessing the information the scoping review generated, it can be said that many challenges and issues arise concerning the policing of CSAM. These issues are quite similar regardless of the phase or actor concerned. In this discussion, the most important findings will be highlighted and the observed gaps in the literature will be touched upon, followed by a number of recommendations. Lastly, the limitations of this review will be briefly explained.

This paper initiated with an outline of regulatory developments in the European Union, followed by an overview of trends and challenges concerning the regulation of online CSAM that were found in the scoping review itself. It is interesting to compare the recent legislative developments with those that the scoping review generated. There was an overarching emphasis on the importance of harmonized regulations and guidelines voiced by every actor concerned in the literature. They state that in order to utilize technological solutions and effectively share information in the fight against CSAM, strict policies need to exist.[189] When looking at recent legislative developments, European regulatory bodies have been taking big steps towards the development of those stricter policies. For example, the proposed CSAM regulation is aimed exactly at harmonizing policies amongst member states and imposing strict detection and removal orders on service providers. Similarly, the obligations imposed upon VLOPs and VLOSEs to deal with illegal content in general are considerably far-reaching. So the image painted in the generated literature about a mainly non-interventionist government seems to be more nuanced than that, especially when assessing developments in the EU. Furthermore, litera-

[186] Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches'.
[187] Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation'.
[188] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'; Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches'.
[189] Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*.

ture refers to the voluntary and own-initiative scanning efforts of ESP but with the proposed CSAM regulation for example, a turn is seen towards a legal obligation for detection of CSAM. This shows that it are not always the service providers who decide to deploy automated technologies and infringe on user's privacy but that they are also nudged (or pushed) into that direction by several legislative developments. Those legislative developments in turn, have been criticized for not giving adequate attention to safeguards, suggesting that not only service providers may open the door for infringing user's fundamental rights.

Furthermore, different procedures in judicial and police investigations or NGO and ESP-reporting processes complicate an effective fight against CSAM. The collaboration of aforementioned actors is widely recognized to be the backbone of fighting CSAM despite this being exactly what makes the coherent policing of CSAM such a challenge, since everyone involved has divergent traditions, working routines, objectives and agendas. Therefore, shared guidelines and objectives should be essential in approaching CSAM.[190] The construction of these frameworks should include every relevant party since tackling CSAM demands shared responsibility.

Secondly, the use of technologies in the policing of CSAM was critically inspected, shedding light on many caveats paired with its deployment. The large influx of CSAM reports and prevalence on online platforms necessitates the use of automated tools to aid its processing. Despite the potential of hashing and AI to speed up investigation, important weaknesses are paired with these technologies. Overblocking of content and the generation of false positives was mentioned several times. Issues such as a lack of transparency, legitimacy, accountability, privacy infringements and a risk for mission creep also arose in the literature.[191] On top of that, it seems that both hashing tools and AI - aside from ethical implications - are not technologically infallible. Issues pertaining to the robustness of hashing methods or biased data and incorrect decision-making by AI were mentioned.[192] The importance of automated tools cannot be underestimated but a blind focus on technology as being *the* solution for CSAM, is highly misleading. This technological solutionism stimulates blind faith in and reliance on technological tools to solve delicate and complex issues and distracts from the need for collaborative policy setting and transparency in those tools.[193] Lastly, it needs to be mentioned that a large amount of CSAM

[190] Cullen and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States'; Meggyesfalvi, 'Policing harmful content on social media platforms'.
[191] McIntyre, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems'.
[192] Struppek and others, *Learning to Break Deep Perceptual Hashing: The Use Case NeuralHash*; Kamara and others, *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems;* Lee and others, 'Detecting child sexual abuse material: A comprehensive survey'.
[193] Salter and Hanson, '"I Need You All to Understand How Pervasive This Issue Is": User Efforts to Regulate Child Sexual Offending on Social Media'.

reports come from only a small, albeit highly influential, number of ESPs.[194] This may hint at a lack of effort to tackle CSAM (or to report on it) by smaller, lesser known ESPs.

Child protection NGO's and hotlines are essential clearinghouses for CSAM reports. They send reports to the relevant hotline in the hosting country or to law enforcement. Hotlines struggle to keep up with the number and volume of reports and signal a lack of human, budgetary and technological resources.[195] Furthermore, certain hotlines lack standardized notice & takedown procedures and transparency in their practises. Remarkably, hotline analysts struggle to or do not use provided software tools that should aid their work and still perform simple tasks manually.[196] They report significant challenges in international and cross-organizational collaboration including a lack of legal power in enforcing takedown requests, large waiting times and delays for the removal of this material.[197] However, other evidence shows that nuance is fitting here since it appears that other hotline workers do experience this cooperation in takedown requests to be swift.[198] This makes it clear that different NGOs have different working processes and practices.

The struggles that NGOs experience surrounding a lack of human, technological and budgetary resources are also seen to be tarnishing the good practices of law enforcement. Additionally, law enforcement agencies report a lack of knowledge and training concerning CSAM and its investigation.[199] Similarities are also found in problems surrounding collaboration with ESPs related to slow responses to takedown requests and efficient cooperation in general. It seems that technological developments, whereas they should aid law enforcement, are not entirely fitted to their needs. This can be concluded from the fact that not all officers have access to certain tools or that they note a high rate of false positives and outdated hash lists.[200] It should be taken notice of that despite the necessity of ESPs using their 'gatekeeper' position to the internet for providing reports to law enforcement or the manufacturing of software tools by tech companies, this dependence - and sometimes overreliance - on private actors should be critically assessed.

---

[194] Teunissen and Napier, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview'.

[195] Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.

[196] Bursztein and others, *Rethinking the Detection of Child Sexual Abuse Imagery on the Internet*.

[197] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'.

[198] Stroebel and Jeleniewski, *Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges*.

[199] Powell and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation'; Cale and others, 'Crime commission processes in child sexual abuse material production and distribution: A systematic review'.

[200] Bleakley and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?'; Franqueira and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches'.

Making valuable recommendations concerning abovementioned issues is not an easy task, considering the amalgam of transnational and cross-stakeholder challenges the policing of CSAM is wrapped up in. Nonetheless, certain suggestions can be made. It is irrefutable that there is a pressing need to redraw policies concerning virtually every factor related to the policing of CSAM. Police and NGO's need compatible, uniform guidelines and working practices and the influence of and safeguarding principles for ESPs and industry need to be clearly delineated. It is important that legislative developments are shaped with adequate safeguards for defending human rights and acknowledging what technology can and can not do. Policing CSAM cannot exist in a vacuum whereby decisions are solely made by government, law enforcement or the industry. Therefore, the policy-making process in this field should have the transnational and cross-stakeholder character that is inherent to fighting CSAM. This should permit the application of checks and balances by opposing views put forth by opposing partners, limiting one view or the other from being all-encompassing.

It is essential to facilitate communication between law enforcement, NGO's and industry, since evidence has shown that the level of this communication and collaboration is highly ambivalent. Findings such as the unwillingness or inaccessibility of law enforcement to use certain software or the duplication of work concerning CSAM databases, illustrate the need for cross-fertilization between sectors. Aside from this happening in the context of participative policy setting, academia should equally be at the forefront of these efforts. As a matter of fact, more research is one of the tools for establishing these policies. Existing research tends to focus on one player or another, further perpetuating the existence of these challenges in sector-based silos. Research should comprehensively aim to tie together views by all those involved and critically circle back findings from one sector to the other to allow for growth in knowledge and perspectives.

Secondly, it is needless to say that further research into both hashing tools and AI is needed to increase their efficiency and try to meet the concerns for their weaknesses. Paired with this is the need for balanced and geographically and ethnically diverse datasets to allow AI to be trained as accurately as possible. Even so, research into these tools to aid CSAM investigation should rarely only be purely technological in nature. Since the operation of these systems has implications for - among others - our personal data, the ethics of these measures need to be considered just as much. Transparency on how tools function and reports on their results should be available to all those tackling CSAM *and* the public. This is not the only case that warrants transparency reporting. ESPs and the industry should provide (more detailed and consistent) reports on their actions and their working practises. Important is that a balance needs to be stricken between those service providers that are highly active (and possibly overreporting) in fighting CSAM and those providers that aren't doing anything. All providers need to equally play their part in fighting CSAM.

Aside from that, there is an urgent need for the landscape of databases to be rethought. It is undeniable that the databases currently existing contain unprecedented amounts of information. However, the lack of harmonization and standardization highlights a

missed opportunity and the need for increased collaboration between all those operating databases, to avoid duplicated work and efforts.

Abovementioned research into technologies to aid CSAM investigation should focus on specific needs that NGO's and hotlines have. Evidence shows that hotline analysts don't use certain technologies that are supposed to facilitate their work, which indicates that perhaps the software that is made, is not in tune with realities on the workfloor. Research should be done to get a full grasp on the reality of working in a child-protection NGO and what they truly need to aid their work. When revisiting the concept that certain NGOs experience delays in notice & takedown processes and that others perceive this process to be swift, it may hint at organizational differences between different hotlines. Therefore, future research into the organizational culture at divergent NGO's may be useful to assess what works and what doesn't. The same goes for law enforcement. Research should focus on the needs of law enforcement and how currently existing software tools and reporting processes are being experienced. Lastly, the literature found in this review touched upon law enforcement and NGO's but the exact working processes of industry or ESPs remain under-researched. Solely relying on their own produced (well-marketed communications) is insufficient. This would allow researchers to determine just how much influence ESPs have in these policing processes and what consequences they bring about. Comparably, some sources describe the interplay between law enforcement and NGO's vis à vis industry or ESPs. The interplay between law enforcement and NGOs remains underexposed. For both these cases, objective research is necessary.

4.1 Limitations of the study

In order to correctly interpret displayed results, the limitations of this study should also be reported. First and foremost, since the topic of the policing of CSAM is highly active, it is possible that interesting and relevant literature was published after the cutoff date for the literature selection in this study (June 2023). This leaves the possibility that new and recent finding are not included here. This was particularly the case when looking at legislative developments. For that reason, the introduction of this paper includes a recent overview thereof, since the literature offered an incomplete image. It was shown that certain findings in the literature do not entirely match the current legislative context around CSAM. The same could be possible for other topics. On top of this, in order to guard feasibility of the study, certain exclusion grounds were set. These can also entail a limit of the study. As a consequence of the nature of the databases consulted, book chapters and policy reports were generated less frequently. These could have delivered interesting insights.

Next, it should be noted that in literature concerning the automated fight against CSAM generated in this study, Western countries tend to be overrepresented both in territorial scope or author affiliation. There was not a certain geographical scope adhered to within this review since the main scope demarcation of the study was topic-wise, namely to find

all relevant texts that discussed the operational, automated policing of CSAM and introducing a geographic limit on top of this, would possibly generate even less relevant literature. Certain sources from for example Asia or South-America were found, but they did not discuss the automated policing of CSAM. For that reason they were excluded from the review. It is possible that the findings are therefore skewed in favour of a mostly Western perspective since many technology companies' headquarters are based here and the debate around the automated policing of CSAM is mostly had in these regions. Furthermore, this could be aggravated due to the linguistic capacities of the researcher and the inclusion of English and Dutch texts only. This could also be the case for the ESPs mentioned. Naturally, the service providers mentioned in this paper are the ones most vocal (and typically most popular) about their CSAM-fighting efforts. Logically, those who are not as large are underrepresented in academic literature and therefore possibly also in this study.

5 Conclusion

This scoping review demonstrated that quite some knowledge is available concerning the automated policing and reporting processes of CSAM and that challenges and issues are prevalent on every relevant echelon. Many studies mentioned the necessity of cross-stakeholder collaboration since CSAM is an issue that does not respect geographically nor organizationally established borders. Unfortunately, this cooperation often goes awry. Most of all, it seems that there is a significant lack of communication and standardization between those echelons. ESPs, the industry, NGOs and law enforcement all function in accordance with their own agendas and exist in their own isolated silos, losing sight of an effective approach against CSAM. On top of this, there is a remaining lack of knowledge of the working processes and needs of every actor involved and their reciprocate relationships. Whereas technology has become indispensable for policing CSAM, it should equally be treated with enough scepticism concerning far-reaching consequences and concerns related to issues such as surveillance, transparency and privacy. On top of that, the current software output is not always usable for those sectors it is designed and meant for. Further research into these knowledge gaps should be encouraged to move towards a more balanced, transparent and effective fight against CSAM.

References

Included in scoping review (N=55)

Açar KV, 'Framework for a Single Global Repository of Child Abuse Materials' (2019) 11 Global Policy

Akbari Y and others, 'Digital forensic analysis for source video identification: A survey' (2022) 41 Forensic Science International: Digital Investigation

Alliance WG, Preventing and Tackling Child Sexual Exploitation and Abuse (CSEA): A Model National Response (2016)

Al-Nabkia MW and others, 'Short Text Classification Approach to Identify Child Sexual Exploitation Material' (2020) Journal of Applied Soft Computing Preprint

Anda F, Le-Khac N-A and Scanlon M, 'DeepUAge: Improving Underage Age Estimation Accuracy to Aid CSEM Investigation' (2020) 32 Forensic Science International: Digital Investigation

Angelopoulos C and Smet S, 'Notice-and-fair-balance: how to reach a compromise between fundamental rights in European intermediary liability' (2016) 8 Journal of Media Law

Berisha B, 'Platforms Regulation in the Digital Single Market Strategy-Illegal Content and Intellectual Property as a Key Challenge' (2019) 22 European Research Studies Journal

Biswas R and others, 'Reinforcement of age estimation in forensic tools to detect Child Sexual Exploitation Material' (2021) 34 Investigación en Ciberseguridad: Actas de las VI Jornadas Nacionales

Bleakley P and others, 'Moderating online child sexual abuse material (CSAM): Does self-regulation work, or is greater state regulation needed?' (2023) 21 European Journal of Criminology

Bursztein E and others, Rethinking the Detection of Child Sexual Abuse Imagery on the Internet (2019)

Cale J and others, 'Crime commission processes in child sexual abuse material production and distribution: A systematic review' (2021) Trends & issues in crime and criminal justice

Cameron G and others, Child Sexual Exploitation: A study of international comparisons (2015)

Carr J, A Joint Report on Online Child Protection Combatting Child Sexual Abuse Material on the Internet (2017)

Chatzinikolaou A, 'Sexual images depicting children: the EU legal framework and online platforms' policies' (2020) 11 European Journal of Law and Technology

Chaves D and others, 'Assessment and Estimation of Face Detection Performance Based on Deep Learning for Forensic Applications' (2020) 20 Sensors

Cullen O and others, '"Our Laws Have Not Caught up with the Technology": Understanding Challenges and Facilitators in Investigating and Prosecuting Child Sexual Abuse Materials in the United States' (2020) 9 Laws

Davidson J and others, 'Exploring policing and industry practice in the prevention of online child sexual abuse' in Bryce I and Petherick W (eds), Child Sexual Abuse: Forensic Issues in Evidence, Impact, and Management (Academic Press 2020)

Eggestein JV and Knapp KJ, 'Fighting Child Pornography: A Review of Legal and Technological Developments' (2014) 9 Journal of Digital Forensics, Security and Law

Franqueira VNL and others, 'Investigation of Indecent Images of Children cases: Challenges and suggestions collected from the trenches' (2018) 24 Digital investigation

Gernand T, 'Scanning iPhones to Save Children: Apple's On-Device Hashing Algorithm Should Survive a Fourth Amendment Challenge' (2022) 127 Dickinson Law Review

Grega M, Bryk D and Napora M, 'INACT—INDECT Advanced Image Cataloguing Tool' (2012) 68 Multimedia Tools and Applications

Holt TJ and others, 'Assessing the challenges affecting the investigative methods to combat online child exploitation material offenses' (2020) 55 Agression and Violent Behavior

Kamara S and others, Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems (2021)

Lai W, Vaerenbergh AV and Bellaert W, 'Europol and its growing alliance with private parties' (2021) 92 Revue Internationale de Droit Pénal 45

Laranjeira C and others, Seeing without Looking: Analysis Pipeline for Child Sexual Abuse Datasets (2022)

Leclerc B and others, 'Child Sexual Abuse Material Online: The Perspective of Online Investigators on Training and Support' (2022) 16 Policing: A Journal of Policy and Practice

Lee H-E and others, 'Detecting child sexual abuse material: A comprehensive survey' (2020) 34 Forensic Science International: Digital Investigation

Ludvigsen KR, Nagaraja S and Daly A, 'YASM (Yet Another Surveillance Mechanism)' (2022) Arxiv

MacLeod L, King D and Dempster E, A review of age estimation research to evaluate its inclusion in automated child pornography detection (2020)

Markovich E, 'Two Clicks Away: An analysis of the offence of viewing child sexual abuse materials on the Internet' (Lund University 2017)

Martellozzo E and DeMarco J, 'Exploring the removal of online child sexual abuse material in the United Kingdom: Processes and practice' (2020) 22 Crime Prevention & Community Safety

McIntyre TJ, 'Blocking child pornography on the Internet: European Union developments' (2010) 24 International Review of Law, Computers & Technology

McIntyre TJ, 'Child Abuse Images and Cleanfeeds: Assessing Internet Blocking Systems' in Brown I (ed), Research Handbook on Governance of the Internet (2013)

Meggyesfalvi B, 'Policing harmful content on social media platforms' (2021) 69 Belügyi Szemle

Moser A, Rybnicek M and Haslinger D, 'Challenges and Limitations Concerning Automatic Child Pornography Classification' (2015) (10th International Conference on Computer Vision Theory and Applications)

Mount D and others, 'Triaging online child abuse material: testing a decision support tool to enhance law enforcement and investigative prioritisation' (2021) 44 Policing: An International Journal

Pilipets E and Paasonen S, 'Nipples, memes, and algorithmic failure: NSFW critique of Tumblr censorship' (2022) 24 New media & society

Polastro MdC and Eleuterio PMdS, NuDetective: a Forensic Tool to Help Combat Child Pornography through Automatic Nudity Detection (2010)

Powell MB and others, 'Police officers' perceptions of the challenges involved in Internet Child Exploitation investigation' (2013) 37 Policing: An International Journal

Pūraitė A and Prokofjeva N, 'Policy of the European Union on The Safety of Children In Cyber Space ' (2018) 21 Public Security and Public Order

Rajamäki J, Lahti L and Parviainen J, 'OSINT on the Dark Web: Child Abuse Material Investigations' (2022) 53 Information & Security 21

Rusu I, 'Preventing Offenses of Sexual Abuse, Sexual Exploitation of Children and Child Pornography in the European Union' (2012) 8 Acta Universitatis Danubis Juridica

Salter M and Hanson E, '"I Need You All to Understand How Pervasive This Issue Is": User Efforts to Regulate Child Sexual Offending on Social Media' in Bailey J, Flynn A and Henry N (eds), The Emerald International Handbook of Technology Facilitated Violence and Abuse (2021)

Salter M and Richardson L, 'The Trichan takedown: Lessons in the governance and regulation of child sexual abuse material' (2021) 13 Policy & Internet 385

Sanchez L and others, 'A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM)' (2019) 29 Digital Investigation

Seigfried-Spellar KC, Bertoline GR and Rogers MK, 'Internet Child Pornography, U.S. Sentencing Guidelines, and the Role of Internet Service Providers' (2011) (International Conference on Digital Forensics and Cyber Crime)

Steel CMS, 'Web-based child pornography: The global impact of deterrence efforts and its consumption on mobile platforms' (2015) 44 Child abuse & Neglect

Stroebel M and Jeleniewski S, Global Research Project: A Global Landscape of Hotlines Combating Child Sexual Abuse Material on the Internet and an Assessment of Shared Challenges (2015)

Struppek L and others, Learning to Break Deep Perceptual Hashing: The Use Case NeuralHash (2022)

Thakor M, 'Digital Apprehensions: Policing, Child Pornography, and the Algorithmic Management of Innocence' (2018) 4 Catalyst feminism, theory, technoscience

Thakor M, 'How to Look: Apprehension, Forensic Craft, and the Classification of Child Exploitation Images' (2017) 39 IEEE Annals of the History of Computing

Teunissen C and Napier S, 'Child sexual abuse material and end-to-end encryption on social media platforms: An overview' (2022) Trends & issues in crime and criminal justice

Wagner B, 'The Public Sector and Content Regulation: Focussing on Pornography and Child Sexual Abuse Material', Global Free Expression - Governing the Boundaries of Internet Content, Law, Governance and Technology (Springer International Publishing 2016)

Westlake B, Bouchard M and Frank R, Comparing Methods for Detecting Child Exploitation Content Online (2012)

Westlake BG, 'The Past, Present, and Future of Online Child Sexual Exploitation: Summarizing the Evolution of Production, Distribution, and Detection', *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (Palgrave Macmillan 2020)

Other references

Adams RJ, Smart P and Huff AS, 'Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies' (2017) 19 International Journal of Management Reviews 432

Arksey H and O'Malley L, 'Scoping Studies: Towards a Methodological Framework' (2005) 8 International Journal of Social Research Methodology 19

Benzies KM and others, 'State-of-the-Evidence Reviews: Advantages and Challenges of Including Grey Literature' (2006) 3 Worldviews on Evidence Based Nursing 55

Bonnici JM and others, Commission proposal on the temporary derogation from the e-Privacy Directive for the purpose of fighting online child sexual abuse (2021)

Buri I and Hoboken Jv, The Digital Services Act (DSA) proposal: a critical overview (2021)

Carr J and Hilton Z, 'Combating child abuse images on the Internet: International perspectives', Internet Child Abuse: Current Research and Policy (Routledge-Cavendish 2010)

Church P and Pehlivan CN, 'The Digital Services Act (DSA): A New Era for Online Harms and Intermediary Liability' (2023) 4 Global Privacy Law review

Colquhoun HL and others, 'Scoping reviews: time for clarity in definition, methods, and reporting' (2014) 67 Journal of Clinical Epidemiology 1291

Feys Y, 'Criminological scoping reviews as part of a student assignment: methodological considerations' (2022) 57 Quality & Quantity 1

Bracket Foundation, Artificial Intelligence: Combating Online Sexual Abuse of Children (2019)

G'sell F, 'The Digital Services Act: a General Assessment' in Ungern-Sternberg Av (ed), Content Regulation in the European Union - The Digital Services Act, vol 1 (Trier Studies On Digital Law, 2023)

Grant MJ and Booth A, 'A typology of reviews: an analysis of 14 review types and associated methodologies' (2009) 26 Health Information and Libraries Journal 91

Gusenbauer M and Haddaway NR, 'Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources' (2020) 11 Research Synthesis Methods

Hirsch PB, 'Tie me to the mast: artificial intelligence & reputation risk management' (2018) 39 Journal of Business Strategy 61

INHOPE, Annual Report 2020 (2020)

Kastner M and others, 'What is the most appropriate knowledge synthesis method to conduct a review? Protocol for a scoping review' (2012) 12 BMC Medical Research Methodology 1

Kloess JA, Beech AR and Harkins L, 'Online child sexual exploitation: prevalence, process, and offender characteristics' (2014) 15 Trauma, Violence & Abuse 126

Lai W, Vaerenbergh AV and Bellaert W, 'Europol and its growing alliance with private parties' (2021) 92 Revue Internationale de Droit Pénal 45

Lazarus L and others, Respecting human rights and the rule of law when using automated technology to detect online child sexual exploitation and abuse (2021)

Levac D, Colquhoun H and O'Brien KK, 'Scoping studies: advancing the methodology' (2010) 5 Implementation Science 1

Leary MG, '§230 of the Communications Decency Act: Regarding Child Sexual Abuse Material - The Experiment is Done and it Failed' (2022) 2 The Catholic University of America Columbus School of Law Legal Studies Research Paper Series

Martellozzo E, *Online Child Sexual Abuse: Grooming, Policing and Child Protection in a Multi-Media World* (Routledge 2013)

Martin J, 'Child sexual abuse images online: Implications for social work training and practice' (2016) 46 The British Journal of Social Work 372

Munn Z and others, 'Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach' (2018) 18 BMC Medical Research Methodology

NCMEC, CyberTipline 2022 Report (2023)

NCMEC, Cybertipline: What Happens to Information in a CyberTip? (2020)

Negreiro M, Curbing the surge in online child abuse: The dual role of digital technology in fighting and facilitating its proliferation (2020)

Pereira M and others, 'Metadata-Based Detection of Child Sexual Abuse Material' (2020) 99 IEEE Transactions on Dependable and Secure Computing

Peters MDJ and others, 'Updated methodological guidance for the conduct of scoping reviews' (2020) 18 JBI Evidence Synthesis

Peterson J and others, 'Understanding scoping reviews: Definition, purpose, and process' (2016) 29 Journal of the American Association of Nurse Practitioners 12

Petticrew M and Roberts H, *Systematic reviews in the social sciences: A practical guide* (John Wiley & Sons 2008)

Pham MT and others, 'A scoping review of scoping reviews: advancing the approach and enhancing the consistency' (2014) 5 Research Synthesis Methods 371

Quayle E, 'Prevention, disruption and deterrence of online child sexual exploitation and abuse' (2020) 21 ERA Forum

Richardson J, Three Ways Artificial Intelligence is Good for Society (2017)

Rojszczak M, 'The Digital Services Act and the Problem of Preventive Blocking of (Clearly) Illegal Content' (2023) 3 Journal of Administrative Sciences

Somers C, 'The proposed CSAM Regulation: trampling privacy in the fight against child sexual abuse?' (2023) <https://www.law.kuleuven.be/citip/blog/the-proposed-csam-regulation-trampling-privacy-in-the-fight-against-child-sexual-abuse/> accessed 2 July 2024

Stringhi E, 'The due diligence obligations of the Digital Services Act: a new take on tackling cyber-violence in the EU?' (2024) International Review of Law, Computers & Technology

Taylor J, 'Policing social networking sites and online grooming', Internet Child Abuse: Current Research and Policy (Routledge-Cavendish 2010)

Tar J, 'EU nears consensus on child abuse draft law, new agency takes lead on privacy preservation' (2024) <https://www.euractiv.com/section/law-enforcement/news/eu-nears-consensus-on-child-abuse-draft-law-new-agency-takes-lead-on-privacy-preservation/> accessed 3 July 2024

Westlake B, Bouchard M and Frank R, 'Assessing the Validity of Automated Webcrawlers as Data Collection Tools to Investigate Online Child Sexual Exploitation' (2015) 29 Sexual Abuse

This special issue brings together nineteen topical and innovative papers, researching the boundaries of sexual integrity and affirmative sexual consent, gender violence, and image-based or online sexual abuse, including child sexual abuse material and non-consensual sexual deepfakes. It offers an original and nuanced approach to understanding the important legal elements, various agents and harms of topic-related deviant conduct as well as legislative processes aimed at tackling it. In light of recent societal developments, including changes in societal sensibilities, and recent or on-going legislative amendments at national and supranational levels, research on these topics is timely and much needed.

Gert Vermeulen is Senior Full Professor of European and international criminal law, sexual criminal law and data protection law, Department Chair Criminology, Criminal Law and Social Law, and Director of the Institute for International Research on Criminal Policy (IRCP), all at Ghent University. He is also General Director Publications of the AIDP, and Editor-in-chief of the RIDP.

Nina Peršak is Scientific Director and Senior Research Fellow, Institute for Criminal-Law Ethics and Criminology (Ljubljana), Full Professor of Law, University of Maribor (habilitation), Academic Consultant at the Institute for International Research on Criminal Policy (IRCP), Ghent University, and Co-Editor-in-Chief of the RIDP.

Stéphanie De Coensel is an FWO Postdoctoral Researcher and Visiting Professor in Advanced Criminal Law at the Institute for International Research on Criminal Policy (IRCP), Ghent University, and Editorial Secretary of the RIDP.

www.maklu.be
ISBN 978-90-466-1263-7