Australian Government

Australian Institute of Criminology

Trends & issues in crime and criminal justice

No. 636 September 2021

Abstract | Cyber strategies play a role in combating child sexual abuse material (CSAM). These strategies aim to detect offenders and prevent them from accessing and producing CSAM, or to identify victims. This paper explores five cyber strategies: peer-to-peer network monitoring, automated multi-modal CSAM detection tools, using web crawlers to identify CSAM sites, pop-up warning messages, and facial recognition. This research synthesis captures the background of each strategy, how it works and the evaluative research, along with the benefits, limitations and implementation considerations, offering a practical overview for a broad audience.

Cyber strategies used to combat child sexual abuse material

Graeme Edwards, Larissa S Christensen, Susan Rayment-McHugh and Christian Jones

Introduction

There is no definitive statistic regarding the amount of child sexual abuse material (CSAM) accessible online. The National Crime Agency of the United Kingdom identified 2.88 million accounts registered on darknet CSAM sites in their national strategic assessment (Johnson & Patel 2019). The National Center for Missing and Exploited Children (NCMEC) received more than 18 million reports in 2018 (NCMEC 2020). According to the Australian Federal Police there were 17,905 reports of online CSAM images identified in 2018 alone (Akerman 2019).

This crime presents significant challenges to law enforcement, as new technology continuously creates novel opportunities for CSAM perpetrators (Seto 2013). The anonymity, accessibility and affordability of the internet also offers a unique environment in which to commit offences (Wortley & Smallbone 2006). Research indicates online sharing platforms have been responsible for the acceleration in the pace of CSAM creation and distribution. The NCMEC recorded 23.4 million instances of CSAM between 1998 and 2017, with 9.6 million (40%) being recorded just in 2017 (Bursztein et al. 2019). The proliferation is also evident in Australia, with the Australian Federal Police receiving over 8,000 more CSAM referrals in 2018 than in 2017, with each report potentially relating to thousands of images and videos (Akerman 2019).

Child Sexual Abuse Material Reduction Research Program

The rapid rise in the availability of CSAM has occurred alongside swift advances in technologies such as high-speed internet, increasingly fast central processing units in computers, and virtual reality (Broadhurst 2019). Distribution channels have also followed the development of new forms of online technology. Where email and computer networks were once the main methods of distributing CSAM files, new apps and TOR (The Onion Router) browsers have since been adopted (Bursztein et al. 2019). To protect their anonymity, offenders also now use protected network connections (virtual private networks), encryption and proxy servers (Balfe et al. 2015), along with the darknet.

The internet has also created new forms of CSAM such as live streaming of abuse with payment through various financial instruments (United Nations Office on Drugs and Crime 2015). Live streaming involves viewing child abuse in real time, allowing the CSAM offender to interact with the abuser and to request acts they want committed against the child (Açar 2017; Europol 2019). Often these offenders will make the payments using various anonymous service providers, financial institutions and cryptocurrencies to avoid drawing suspicion (Australian Transaction Reports and Analysis Centre 2019).

CSAM is regularly found on and removed from social media by platform providers. For example, a Northern Territory resident was recently arrested by Northern Territory Police for using social media platforms to transmit CSAM which was identified by NCMEC (Mirage 2020). In 2018, Facebook announced that in one quarter they had removed 8.7 million pieces of content violating their child nudity or sexual exploitation policies identified using artificial intelligence (AI; Hutchinson 2018). Facebook has teams working constantly to identify and remove this content and make reports to NCMEC which are then referred to international law enforcement (Hutchinson 2018). However, these strategies have been critiqued as insufficient, with calls for social media companies to 'do more' to combat CSAM (Hunter 2019). Facebook's recent plans to implement end-to-end encryption, for example, received international criticism given the technology will present a barrier to identifying and reporting CSAM (Hunter 2019).

While therapeutic and educational strategies have long been a key part of child sexual abuse prevention and response (Smallbone, Marshall & Wortley 2008), and many of these have been translated to the prevention of CSAM, cyber strategies used in the prevention of CSAM are relatively new. In essence, this research synthesis captures the current state of practice for cyber strategies that aim to reduce CSAM. It offers to enhance understanding of the strategies for a broad audience of non-cyber experts who are devoted to CSAM reduction, by identifying and breaking down complex technological interventions. Thus, this paper should be a valuable resource for those in other disciplines such as law or counselling, and in non-cyber specialist roles within law enforcement. Innovatively, this paper focuses on the use of technology to reduce CSAM, which is distinct from previous research which aimed to identify instances of CSAM across the many traditional and evolving platforms. In doing so, this synthesis identifies key strategies and explores the background of each, how they work, and the evaluative research, along with the benefits, limitations and implementation considerations. The strategies presented are not recommendations; instead, they represent a review of the research.

Methodology

This research aimed to identify and explain current national and international cyber CSAM reduction strategies. This was achieved through a comprehensive literature search of unpublished (grey) and published (academic) literature. The literature search was designed to be exhaustive in identifying current strategies and was not restricted by the methodological hierarchy of evidence. The search process is displayed in Table 1. Data extraction was undertaken by a primary reviewer and verified by two secondary reviewers.

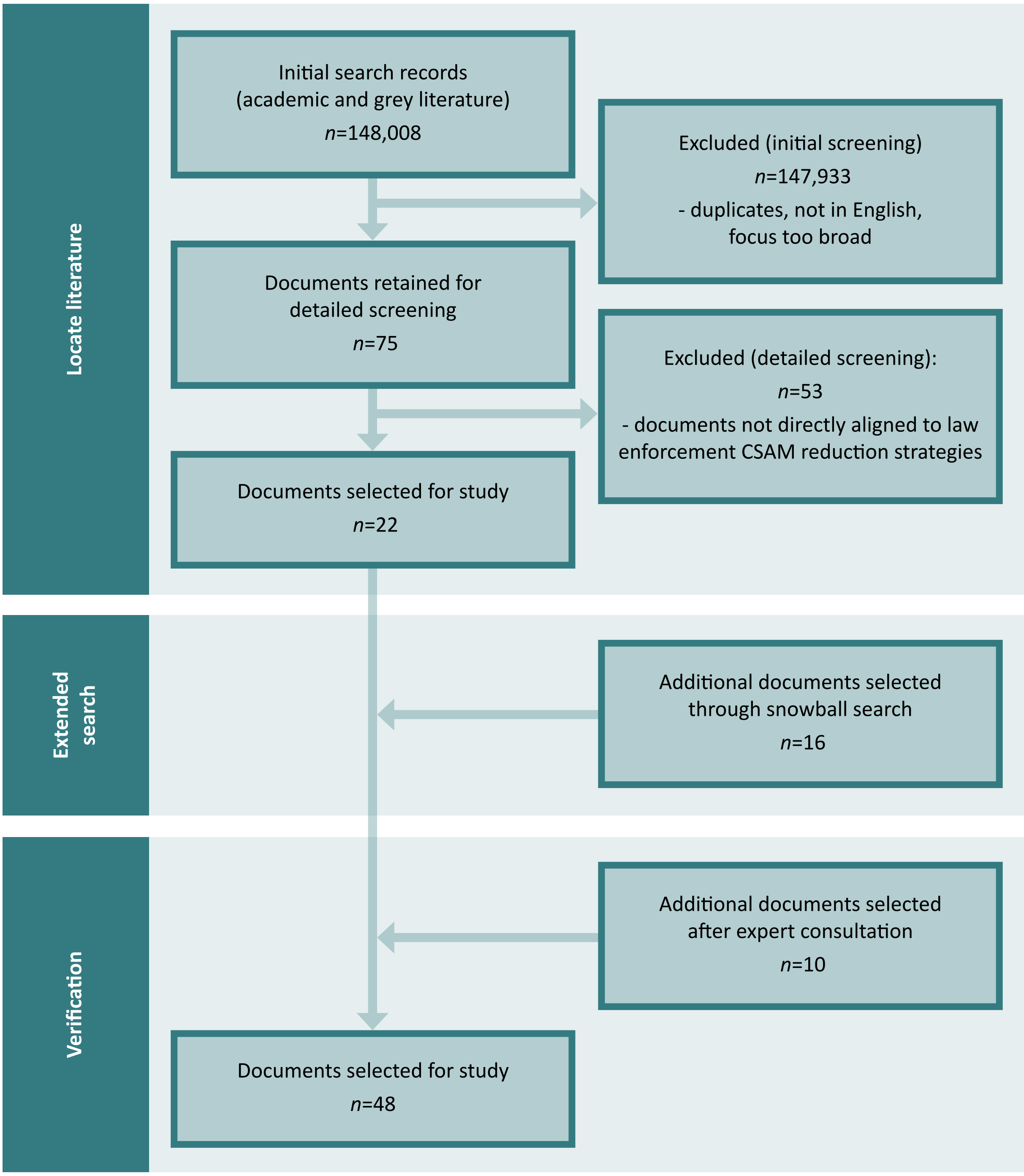
Search results are displayed in Figure 2. Five key strategies emerged across the 48 articles. The research synthesis was completed strategy by strategy using a deductive approach, with documents examined for the background of the strategy, how each strategy is theorised to work, the evaluative research, along with its benefits, limitations and implementation considerations. Data analysis was undertaken by a primary reviewer and verified by two secondary reviewers. Discussion among reviewers was also used to advance interpretation.

Literature sourced is indicated in the reference list with a (*) symbol. Due to publication constraints, representative literature is included in this paper. The authors can be contacted for a complete list. The research team also hosted a roundtable in December 2019 with six key national cyber experts and academics. During the roundtable, the research team summarised each strategy identified from the initial search records, selected snowball documents and literature sourced from expert consultation. These experts verified that the strategies were representative of those employed in the field.

Table 1: Search process

Stage	Description	Steps	Activities undertaken
1	Preparation	• Establish search aims and guidelines	• Academic and grey literature • Not restricted by methodological hierarchy of evidence • Date range: open • Search conducted June/July 2019
2	Locate literature	• Conduct search of selected search platforms • Two-stage document screening	• Online search platforms included: HeinOnline, Google Scholar, Scopus • Broad search terms included: child exploitation, cybercrime, forensics • Initial screening (title) excluded documents not published in English, duplicates, and documents not specific to CSAM • Detailed document screening (abstracts/articles) excluded documents not directly focused on cyber CSAM reduction strategies, documents not available in full text
3	Extended search	• Snowballing search of reference lists and citations	• Reference list and citation searches on selected documents to ensure comprehensive identification of current CSAM reduction strategies
4	Verification	• Contact with experts (academics, practitioners, stakeholders) to source unpublished research literature and practice wisdom	• Two international experts (United States of America) shared published research literature • Subsequent roundtable discussion with 6 national experts (representing 4 organisations) to verify representativeness of identified strategies
5	Documentation	• Search processes and results documented	• See Figure 1

Figure 1: Search results


Locate literature
Initial search records (academic and grey literature) *n*=148,008
Excluded (initial screening) *n*=147,933 - duplicates, not in English, focus too broad
Documents retained for detailed screening *n*=75
Excluded (detailed screening): *n*=53 - documents not directly aligned to law enforcement CSAM reduction strategies
Documents selected for study *n*=22
Extended search
Additional documents selected through snowball search *n*=16
Verification
Additional documents selected after expert consultation *n*=10
Documents selected for study *n*=48

Results

Five cyber strategies were identified: peer-to-peer network monitoring, automated multi-modal CSAM detection tools, using web crawlers to identify CSAM sites, pop-up warning messages, and facial recognition. One of these strategies (facial recognition) was not subject to evaluation and the other four had very limited evaluation (peer-to-peer network monitoring, automated multi-modal CSAM detection tools, using web crawlers to identify CSAM sites, and pop-up warning messages).

The authors acknowledge that law enforcement also use other strategies to combat CSAM, such as proactive investigations. However, this paper focuses on key technological strategies; less cyber-focused strategies are addressed in a separate publication (Christensen et al. 2021). The authors also acknowledge the role of strategies beyond cyber and law enforcement disciplines, including intervention helplines for potential offenders, public awareness campaigns, education for children, and therapeutic programs for CSAM offenders. These strategies are discussed in a separate paper (Rayment-McHugh, McKillop & Christensen forthcoming).

The cyber disruption strategies discussed below may be explained by situational crime prevention (SCP) theory (Clarke 1980). SCP underpins all five disruption strategies identified in this paper. This theory focuses on the specific characteristics of a situation, as opposed to the motive and characteristics of an offender (Huisman & van Erp 2013). Such a perspective is imperative when understanding CSAM offences. SCP is also applied to contact child sexual offending—for example, public awareness campaigns can remove excuses and reinforce standards of behaviour in the community, reducing permissibility (Wortley & Smallbone 2006). However, in contrast with contact child sexual offending, the internet offers a unique criminogenic environment characterised by anonymity, accessibility and affordability (Wortley & Smallbone 2012). This unique environment creates challenges for the prevention of CSAM and, given the role of technology and the internet, the prevention strategies used for contact child sexual offending cannot be seamlessly transferred by cyber specialists.

SCP proposes that crime prevention occurs by altering opportunity structures within a given situation or environment (Huisman & van Erp 2013). In particular, Cornish and Clarke (2003) identified 25 SCP techniques, which come under five primary strategies:

- increasing effort by making it more difficult to commit a crime in that environment;
- increasing the risk of getting caught;
- reducing rewards from crime;
- reducing provocations to offend; and
- removing excuses for crime.

Strategies that disrupt networks (increasing effort and reducing rewards), that identify offenders and victims (increasing risks) and that set standards of behaviour (remove excuses) appear most relevant to CSAM reduction (Wortley & Smallbone 2012).

Peer-to-peer network monitoring

Peer-to-peer (P2P) communication is a common form of file sharing. It was originally used for illegal access to movies and music but has evolved to include the widespread distribution of CSAM. Monitoring P2P network communication assists law enforcement to identify major CSAM distributors, to locate and prosecute persons in possession of CSAM, to identify child victims, and to remove CSAM sites (Schell et al. 2007).

P2P network monitoring involves law enforcement using investigation tools to collect data travelling across a network to identify whether it contains CSAM (Schell et al. 2007). Included in the data captured are the internet protocol addresses (a numerical label assigned to a computer or device) of those computers that are party to the communication of CSAM. Hash values (numeric values created by transforming a dataset such as an image) are also compared to those in law enforcement databases, and positive matches are taken for follow-up investigation (Wolak, Liberatore & Levine 2014). File names are also automatically monitored on the P2P network (Peersman et al. 2014).

This strategy enables law enforcement to identify and remove large libraries of CSAM images, reducing access to this material. Therefore, the strategy disrupts offending by reducing opportunities and by increasing the risk of detection (Wortley & Smallbone 2012). In line with SCP, network monitoring may also increase the perceived risk of offending among CSAM consumers, challenging the perception that the internet is anonymous and safe, but only if network monitoring capabilities and arrests are publicised (Wortley & Smallbone 2012).

Some studies have investigated the dimensions of P2P CSAM networks (eg Wolak, Liberatore & Levine 2014) and used this knowledge to estimate the extent to which overall CSAM availability could be reduced with targeted intervention. Wolak, Liberatore and Levine (2014) estimate that if law enforcement targeted high-contribution computers on just one network (Gnutella), CSAM availability on that network could be reduced by 30 percent. Hurley et al. (2015) further suggest that removing the top 0.01 percent of P2P sites would effectively remove up to 41 percent of available CSAM.

Network monitoring is a valuable tool to be used alongside other technological options. It also has the effect of alerting CSAM users that there is an active law enforcement presence on the networks, which may influence some consumers' decisions about continuing to view this content. However, this strategy has several limitations. It requires intensive labour and investment of law enforcement resources (Wolak, Liberatore & Levine 2014). The quantity of CSAM amplifies this challenge. Technology itself may also limit effectiveness. For example, when new images not previously classified are not identified by hash values, they may go undetected by law enforcement (Peersman et al. 2016). Internet protocol addresses can also be changed at the will of the offender.

While little research has explored the application of network monitoring in Australia, implementation requires the use of relevant software for forensic analysis. It also necessitates action by law enforcement to proactively investigate and serve subpoenas on internet service providers to identify the account to which an internet protocol address is allocated. Further inquiries are required to identify possible persons linked to that address, before seeking and then executing a search warrant.

Automated multi-modal CSAM detection tools

The automatic detection of CSAM can be challenging, particularly in the presence of legal adult pornography. Compared to simply analysing file names and hash values, a multi-modal classification approach to identifying CSAM videos and images has yielded greater discrimination, improving the accuracy of CSAM detection. For example, using a single form of detection (eg skin tone analysis) is not necessarily reliable in differentiating CSAM from adult pornography. Thus using multiple methods to classify material and detect CSAM is more robust and accurate (Schulze et al. 2014). Such automated content classification tools are imperative given the dramatic growth in the number of images and videos online.

One example of automatic detection is the PhotoDNA program developed by Microsoft, which identifies images on publicly available websites (Penna, Clark & Mohay 2005; Westlake, Bouchard & Frank 2012). PhotoDNA compares the hash values of online images against those in the NCMEC database, and any matches are reported to the police for investigation (Microsoft 2020a). The capabilities of PhotoDNA are in the process of being extended to video, using similar identification protocols as for images (Microsoft 2020b).

This technology increases the efficiency and accuracy of CSAM detection. The proliferation of CSAM means digital forensic practitioners spend lengthy periods analysing data, delaying investigations. Automatic detection assists with this workload, providing law enforcement with a time-efficient alternative to visually detecting CSAM. In line with SCP, the risk of detection is a significant consideration in an individual's decision to offend, and the perceived anonymity of the internet is a factor in CSAM offending. Thus any efforts to increase the risk of detection (and reduce anonymity) may reduce this crime (Wortley & Smallbone 2012).

While limited research has explored the impacts of multi-modal classification, in their study of 2,500 CSAM images and 2,500 non-sexual images of children, Gangwar, Fidalgo, Alegre and González-Castro (2017) found CSAM detection methods that use multiple features to identify CSAM (eg shape, text, and colour) perform better than those using a single feature (eg skin colour). Schulze et al. (2014) found that using multiple features improved accuracy in distinguishing between CSAM and adult pornography, reducing the rate of errors from 16 percent to eight percent for videos and from 17 percent to 10 percent for images, in contrast with using only one classification method.

This strategy has many benefits. For example, it has the capacity to identify CSAM that was unknown and whose hash value had not been included in any law enforcement database. Tools such as PhotoDNA provide a valuable resource to law enforcement investigators locating and categorising large volumes of known CSAM in suspects' collections, eliminating the need for investigators to do this work and reducing the potential harm to them. However, several limitations exist, such as the issue of steganography. Steganography is the practice of hiding a message within a seemingly ordinary object or message in plain sight, such that it can only be found by someone who knows where to look. For example, a CSAM image may be embedded within an innocuous photo (Penna, Clark & Mohay 2005). While this multi-modal approach aims to capture these instances where images have been modified, as time goes on, offenders may identify methods to avoid detection. Due to the highly sophisticated nature of this multiple classifications approach, expert training on the use and interpretation of this tool is required.

Using web crawlers to identify CSAM sites

Websites hosting CSAM are often linked. Web crawlers are automated scripts or programs that are used to automatically 'crawl' across many websites. Law enforcement activate a web crawler on a known CSAM site. The web crawler follows the links on each site, identifying the volume of confirmed CSAM (Westlake, Bouchard & Frank 2012). This approach seeks to identify where CSAM is being distributed and presents investigators with an opportunity to remove the central sites, inhibiting CSAM distribution.

The detection of CSAM is enhanced by the automated nature of web crawlers. Web crawlers assist law enforcement by automatically detecting CSAM and mapping relationships between sites. After detecting CSAM sites, law enforcement can disrupt network relationships by removing larger, more popular sites, which in turn starves the smaller sites of the traffic they previously received from them.

This technology works much faster than manual methods, allowing for the quicker identification of CSAM. Knowledge of web crawlers among offenders, particularly those hosting CSAM sites, may prevent crime as, according to SCP (Cornish & Clarke 2003), web crawlers may increase risk of detection and reduce rewards through the resultant site disruption. Eliminating ties between sites may also make it more difficult for offenders to access certain sites (Joffres et al. 2011).

Law enforcement statistics show the effectiveness of web crawlers in identifying CSAM. For example, Project Arachnid, operated by the Canadian Centre for Child Protection, used a web crawler over a six-week period and processed over 230 million web pages, over 5.1 million of which hosted known CSAM images, including over 40,000 unique images (Canadian Centre for Child Protection 2017). This highlights the utility of web crawlers in detecting CSAM. Targeting websites with the highest degree of centrality can be effective in reducing the connectivity between these sites and those linked to them (Joffres et al. 2011). Joffres et al. (2011) showed that removal of the five CSAM sites with the strongest relationships to other CSAM sites reduced site capacity by 36 percent, damaging CSAM networks more than randomly removing websites. If these sites no longer exist, viewers of CSAM cannot access them to view content or use them to directly link to other sites.

Web crawlers are beneficial to law enforcement particularly as the active involvement of investigators is minimal. Taking down these sites creates the potential for suspects to be identified and arrested, as well as child victims to be identified and located. Regarding limitations, web crawlers are inherently restricted to the keywords and hash values used. Along with the training required to operate web crawlers, this strategy requires law enforcement to identify certain keywords, hash values, and major sites. Law enforcement must nominate an initial site and develop a list of CSAM related keywords and hash values for the web crawling tool to search for. Some checking of web crawler results is also required by law enforcement.

Pop-up warning messages

Pop-up warning messages attempt to prevent offenders, particularly first-time offenders, from accessing CSAM. These messages appear when individuals type certain words into a search engine (Prichard, Watters & Spiranovic 2011). This strategy places focus on users understanding that searching for and viewing CSAM is a criminal act, and the unexpected warning elicits concerns about being identified and caught. The message increases in severity with the seriousness of the search terms used.

In line with SCP (Cornish & Clarke 2003), pop-up warning messages are thought to reduce a potential offender's perceived sense of anonymity, increasing the perceived risks of getting caught. Pop-up warning messages can also potentially prevent offending by removing excuses (Prichard, Watters & Spiranovic 2011). Targeting distorted thinking patterns is important, as researchers suggest that cognitive distortions can contribute to individuals commencing and continuing their engagement with CSAM (Prichard, Watters & Spiranovic 2011). For example, warning messages could challenge the distortion that CSAM offenders are 'just viewing' the material and 'not touching' a child.

Recent research (Prichard et al. 2021) examined the effectiveness of warning messages in discouraging individuals from accessing a 'honeypot' (pretend) website claiming to contain barely legal pornography. Results support the effectiveness of warning messages as a CSAM reduction strategy. Deterrence focused warning messages (eg those stating that IP addresses can be traced), significantly reduced the number of individuals who proceeded to access the site. This suggests value in increasing the perceived risk associated with this behaviour. Results, however, did not support the effectiveness of messages about the harm to victims, suggesting the type and format of the messages may impact on their success.

This strategy requires further research and evaluation, including the examination of longer term impacts and potential displacement to other sites (Prichard et al. 2021). Other possible limitations should also be examined. While a warning might deliver a deterrent effect, the warning message alone does not offer longer term support or assistance for the individual (Baines 2018). Including referrals to relevant services as part of the warning message may address this.

Facial recognition

New and evolving facial recognition technology seeks to identify victims and offenders in CSAM. AI facial recognition applications such as Clearview AI have been used by law enforcement internationally. An image of a suspect (or victim) is uploaded to Clearview AI by the investigating officer and algorithms are used to map the face of the suspect (or victim). Unique characteristics of the face (eg distance between eyes, nose width) are used to generate this algorithm as a distinct identifier of the individual. The software then checks this against a facial recognition database to find any instances of the suspect (or victim) where they can be positively identified, notifying law enforcement. One example is the Child Sexual Exploitation Image database, which has the capacity to identify victims, offenders and likely locations through forensic analysis of images and hash values (Broadhurst 2019).

The NCMEC use facial recognition software to search images across the different web surfaces (surface web, deep web and darknet) to match images of missing children (International Centre for Missing and Exploited Children 2018). Such software can provide highly accurate facial analysis, face comparison, and face search capabilities. It can also be used to match potential abductors against law enforcement databases, thus providing a potential lead to the location of the child (Oxford 2019). In line with SCP, facial recognition technology increases the risk of detection, disrupts the production of further CSAM (Schell et al. 2007), and removes victims from harm. By identifying CSAM victims and offenders, law enforcement can extend CSAM investigations beyond existing suspects to their social networks.

While no evaluation research has assessed the effectiveness of facial recognition technology, it is an emerging tool that receives some praise among law enforcement. Detectives in Georgia, Texas and Florida have reported that the Clearview AI application works 'incredibly well' and has helped them solve cases (Hill 2020). In terms of victim identification, Indiana detectives uploaded the images of 21 victims (from the same offender) to Clearview AI and identified 14 of them (Hill & Dance 2020). The Royal Canadian Mounted Police's National Child Exploitation Crime Centre used Clearview AI on 15 cases, resulting in the positive identification of two children (Russell 2020). Clearview AI estimate that 600 law enforcement agencies use their facial recognition software in criminal investigations (Kashmir 2020). Broadhurst (2019) argues the use of AI has the potential to enhance the ability of law enforcement enough to impact CSAM distribution networks. This can be seen as an evolving field of opportunity for law enforcement investigators. However, this strategy can be limited if there is a lack of cooperation among international governments and social media corporations. Without cooperation, there may be limited or no access to image databases. As this is a developing strategy, understanding of further limitations may grow in the future.

Discussion

This paper captures the current state of practice for cyber strategies that aim to reduce CSAM. It was evident that limited evaluation research has been undertaken. This means there is limited information on the effectiveness of these strategies, how often the strategies are deployed and whether they can be effectively used in unison with other strategies. This is one area requiring further research. While it may be years before rigorous evaluation research is carried out on these strategies, the current findings offer useful insights into existing practice and enhance understanding of how the strategies work and the available evaluative research, along with the benefits, limitations and implementation considerations.

Five cyber strategies were identified: P2P network monitoring, automated multi-modal CSAM detection tools, using web crawlers to identify CSAM sites, pop-up warning messages, and facial recognition. The automated nature of these strategies is particularly important given the demands placed on law enforcement by the dramatic growth in CSAM (Australian Federal Police 2017). It was evident in the literature that the reduction of CSAM online is a joint endeavour by both law enforcement and industry. This means that reducing instances of CSAM online is neither the sole responsibility of any one organisation nor a task that law enforcement can undertake alone.

SCP (Clarke 1997) underpins all five strategies. Key to this approach are activities which aim to make offending behaviour difficult (increase the effort required), increase the risk of detection, reduce rewards associated with offending behaviour, remove situational precipitants (reduce provocations), and remove excuses and clarify the offender's role in the behaviour (Cornish & Clarke 2003). Of note is that all five identified cyber strategies, except for pop-up warning messages, are centred on detection and identification. That is, the risk of detection and the effort required to commit CSAM offences are increased, and associated rewards lowered. For example, P2P network monitoring allows law enforcement to remove large libraries of CSAM, thereby increasing the effort involved in offending. In contrast, pop-up warning messages are largely focused on increasing the perceived risk of CSAM offending, by reducing a potential offender's perceived sense of anonymity online, reducing permissibility and setting expected standards of behaviour. Opportunities to further eliminate excuses for CSAM offending, or to design new strategies to target online situational precipitants (prompts, permissibility, pressure, or provocations) are worthy of further exploration.

As many of these strategies are new and emerging, future research is required. Facial recognition technology is evolving and becoming a valid law enforcement tool to identify victims and suspects. There appears to be great potential in researchers working closely with law enforcement to establish the effectiveness of facial recognition. This research could explore what proportion of attempts to use facial recognition to identify suspects and victims are successful. Research with social media providers and web hosting services could establish whether and how often AI products have identified children and led to their rescue. Government agencies that hold facial images of their citizens, such as Australian Border Force (customs) and the Australian Passport Office, could be asked to assist in the comparison of facial recognition mapping against their databases.

Despite the benefits of these strategies, several limitations were identified throughout, particularly concerning technology. For example, due to steganography, law enforcement need to keep track of the keywords and hash values used, particularly when using web crawlers and P2P monitoring. Hash values can be changed when an image is cropped, which may or may not be done deliberately. Moreover, due to the ever-enhancing technology used by offenders, many of these strategies will need to be continually updated to avoid becoming obsolete. These strategies may also be limited by the level of cooperation offered by social media corporations and international government agencies, which requires ongoing consideration in future planning.

Given that law enforcement alone cannot reduce instances of CSAM, professionals in other fields who are devoted to CSAM reduction could use their expertise in various ways to contribute to cyber strategies. For example, the ethical and legal issues with facial recognition require further exploration. In particular, law professionals could carry out research into the legality of using facial recognition software when images are obtained from commercial applications. This research could also explore the legal issues for investigators using facial recognition tools and databases of images, including concerns as to the uploading of images of children being subject to sexual abuse onto commercial computer servers located in foreign countries. To provide another example, practice professionals who deliver therapy to CSAM offenders could significantly contribute to developing the text of pop-up warning messages. These professionals could identify the messages that would be most effective among this heterogeneous offending group, along with ways to build in referrals to relevant services.

Conclusion

The current paper identified five key strategies: P2P network monitoring, automated multi-modal CSAM detection tools, using web crawlers to identify CSAM sites, pop-up warning messages, and facial recognition. Disrupting CSAM across the multitude of online platforms and services requires collaboration between international law enforcement and industry. No organisation, whether technical or law enforcement, can operate in isolation. In the distribution of CSAM, the technical advantages are strongly in the CSAM offender's favour, although it can be argued that emerging technology, particularly in the fields of AI incorporating facial recognition, gives law enforcement a unique opportunity. Law enforcement have the ability to not only identify and rescue child victims, but identify and prosecute the offenders, preventing new CSAM being uploaded and shared. Through further research exploring the number of positive identifications made due to AI, or how to enhance the effectiveness of pop-up warning messages, cyber strategies have the potential to make CSAM offending a riskier and less rewarding proposition for individuals. The expertise of professionals in other fields who are devoted to CSAM reduction can also contribute to cyber strategies.

Acknowledgements

This project was supported by funding from the Australian Institute of Criminology's Child Sexual Abuse Material Reduction Research Program (reference ID P18/320). This research was informed by industry experts who attended a roundtable and/or shared research with the researchers. We acknowledge Andrew Childs for assistance searching the literature.

References

URLs correct as at May 2021

*included in review

Açar KV 2017. Web cam child prostitution: An exploration of current and futuristic methods of detection. *International Journal of Cyber Criminology* 11(1): 98–109. https://doi.org/10.5281/zenodo.495775

Akerman T 2019. Hi-tech damage control. *The Australian*, 22 May. https://www.theaustralian.com.au/inquirer/hitech-damage-control/news-story/cacf9027ba9746600464ba751e251048

Australian Federal Police 2017. *Annual report 2016–17*. Canberra: Australian Federal Police. https://www.afp.gov.au/about-us/publications-and-reports/annual-reports

Australian Transaction Reports and Analysis Centre 2019. *Combating the sexual exploitation of children for financial gain: Activity indicators*. https://www.austrac.gov.au/business/how-comply-guidance-and-resources/guidance-resources/combating-sexual-exploitation-children-financial-gain-activity-indicators-report

*Baines V 2018. Online child sexual exploitation: Towards an optimal international response. *Journal of Cyber Policy* 4(2): 197–215. https://doi.org/10.1080/23738871.2019.1635178

Balfe M, Gallagher B, Masson H, Balfe S, Brugha R & Hackett S 2015. Internet child sex offenders' concerns about online security and their use of identity protection technologies: A review. *Child Abuse Review* 24(6): 427–439. https://doi.org/10.1002/car.2308

Broadhurst R 2019. Child sex abuse images and exploitation materials. In R Leukfeldt & T Holt (eds), *The human factor of cybercrime*. Routledge: 310–36

Bursztein E, Clarke E, DeLaune M, Eliff DM, Hsu N, Olson L, Shehan J, Thakur M, Thomas K & Bright T 2019. Rethinking the detection of child sexual abuse imagery on the internet. Paper presented at the World Wide Web Conference, San Francisco, CA: 2601–2607

*Canadian Centre for Child Protection 2017. Groundbreaking tool to remove online child sexual abuse material. https://protectchildren.ca/en/press-and-media/news-releases/2017/project_arachnid

Christensen LS, Rayment-McHugh S, Prenzler T, Chiu YN & Webster J 2021. The theory and evidence behind law enforcement strategies that combat child sexual abuse material. *International Journal of Police Science and Management*. Advance online publication. https://doi.org/10.1177/14613557211026935

Clarke RV 1980. Situational crime prevention: Theory and practice. *British Journal of Criminology* 20(2): 136–47

Clarke RV 1997. *Situational crime prevention: Successful case studies*. Albany, NY: Harrow and Heston

Cornish DB & Clarke RV 2003. Opportunities, precipitators and criminal decisions: A reply to Wortley's critique of situational crime prevention. In MJ Smith & DB Cornish (eds), *Theory and practice in situational crime prevention*. Monsey, NY: Criminal Justice Press: 41–96

Europol 2019. *Internet organised crime threat assessment 2019*. https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment

*Gangwar A, Fidalgo E, Alegre E & González-Castro V 2017. Pornography and child sexual abuse detection in image and video: A comparative evaluation. Paper presented at the 8th International Conference on Imaging for Crime Detection and Prevention, Madrid, Spain: 37–42

*Hill K 2020. Unmasking a company that wants to unmask us all. *New York Times*, 20 January. https://www.nytimes.com/2020/01/20/reader-center/insider-clearview-ai.html

*Hill K & Dance G 2020. Clearview's facial recognition app is identifying child victims of abuse. *New York Times*, 7 February. https://www.nytimes.com/2020/02/07/business/clearview-facial-recognition-child-sexual-abuse.html

Huisman W & van Erp J 2013. Opportunities for environmental crime: A test of situational crime prevention theory. *British Journal of Criminology* 53(6): 1178–1200. https://doi.org/10.1093/bjc/azt036

Hunter F 2019. Encryption can't put tech giants beyond the reach of the law, Minister says. *Sydney Morning Herald*, 11 December. https://www.smh.com.au/politics/federal/encryption-can-t-put-tech-giants-beyond-the-reach-of-the-law-minister-says-20191211-p53ize.html

*Hurley R, Prusty S, Soroush H, Walls R, Albrecht J, Cecchet E, Levine BN, Liberatore M, Lynn B & Wolak J 2015. *Measurement and analysis of child pornography trafficking on P2P networks*. https://ojjdp.ojp.gov/library/publications/measurement-and-analysis-child-pornography-trafficking-p2p-networks-final

Hutchinson A 2018. Facebook outlines enhanced efforts to remove child exploitation content from its platform. *Social Media Today*, 25 October. https://www.socialmediatoday.com/news/facebook-outlines-enhanced-efforts-to-remove-child-exploitation-content-fro/540505/

International Centre for Missing and Exploited Children 2018. GMCNgine: Revolutionizing the search for missing children. https://plussocialgood.medium.com/gmcngine-revolutionizing-the-search-for-missing-children-d01772b32e49

*Joffres K, Bouchard M, Frank R & Westlake B 2011. Strategies to disrupt online child pornography networks. European Intelligence and Security Informatics Conference, Atlanta, Greece: 163–170

Johnson B & Patel P 2019. Multi-million pound funding to protect child abuse victims and track down offenders. https://www.gov.uk/government/news/multi-million-pound-funding-to-protect-child-abuse-victims-and-track-down-offenders

*Kashmir H 2020. Meet Clearview AI, the secretive company that might end privacy as we know it. *Chicago Tribune*, 18 January. https://www.chicagotribune.com/nation-world/ct-nw-nyt-clearview-facial-recognition-20200119-dkdqz7ypaveb3id42tpz7ymase-story.html

Microsoft 2020a. PhotoDNA. https://www.microsoft.com/en-us/photodna

Microsoft 2020b. PhotoDNA: FAQ. https://www.microsoft.com/en-us/photodna/faq

Mirage News 2020. 19-year-old charged with multiple child abuse offences. *Mirage News*, 30 June. https://www.miragenews.com/19-year-old-charged-with-multiple-child-abuse-offences/

National Center for Missing and Exploited Children 2020. NCMEC Data. https://www.missingkids.org/ourwork/ncmecdata

Oxford T 2019. I can see you. *Linux Format*, June. https://www.linuxformat.com/archives?issue=250

*Peersman C, Schulze C, Rashid A, Brennan M & Fischer C 2016. iCOP: Live forensics to reveal previously unknown criminal media on P2P networks. *Digital Investigation* 18: 50–64. https://doi.org/10.1016/j.diin.2016.07.002

*Peersman C, Schulze C, Rashid A, Brennan M & Fischer C 2014. *iCOP: Automatically identifying new child abuse media in P2P networks*. IEEE Security and Privacy Workshops. San Jose, CA: 124–131. https://ieeexplore.ieee.org/document/6957295

*Penna K, Clark A & Mohay G 2005. *Challenges of automating the detection of paedophile activity on the internet*. First International Workshop on Systematic Approaches to Digital Forensic Engineering. Washington, DC: 206–220

*Prichard J, Krone T, Spiranovic C & Watters P 2019. Transdisciplinary research in virtual space: Can online warning messages reduce engagement with child exploitation material? In R Wortley, A Sidebottom, N Tilley & G Laycock (eds), *Routledge handbook of crime science*. UK: Routledge: 309–19

*Prichard J, Watters P & Spiranovic C 2011. Internet subcultures and pathways to the use of child pornography. *Computer Law & Security Review* 27(6): 585–600

Prichard J, Wortley R, Watters P, Spiranovic C, Hunn C & Krone T 2021. Effects of automated messages on internet users attempting to access "barely legal" pornography. *Sexual Abuse*. Advance online publication. https://doi.org/10.1177/10790632211013809

Rayment-McHugh S, McKillop N & Christensen LS forthcoming. Educational and therapeutic programs that combat child sexual abuse material: A research synthesis

*Russell A 2020. RCMP used Clearview AI facial recognition tool in 15 child exploitation cases, helped rescue 2 kids. *Global News*, 27 February. https://globalnews.ca/news/6605675/rcmp-used-clearview-ai-child-exploitation/

*Schell BH, Martin MV, Hung PCK & Rueda L 2007. Cyber child pornography: A review paper of the social and legal issues and remedies—and a proposed technological solution. *Aggression and Violent Behavior* 12(1): 45–63. https://doi.org/10.1016/j.avb.2006.03.003

*Schulze C, Henter D, Borth D & Dengel A 2014. Automatic detection of CSA media by multi-modal feature fusion for law enforcement support. In Proceedings of International Conference on Multimedia Retrieval. Glasgow, United Kingdom: 353–360

Seto MC 2013. *Internet sex offenders*. Washington, DC: American Psychological Association

Smallbone S, Marshall W & Wortley R 2008. *Preventing child sexual abuse: Evidence, policy and practice*. Devon UK: Willan Publishing

*Smallbone S & Wortley R 2017. Preventing child sexual abuse online. In J Brown (ed), *Online risk to children: Impact, protection and prevention*. London: Wiley: 143–162

United Nations Office on Drugs and Crime (UNODC) 2015. *Study on the effects of new information technologies on the abuse and exploitation of children*. Vienna: UNODC. https://www.unodc.org/documents/organized-crime/cybercrime/Study_on_the_Effects.pdf

*Westlake B, Bouchard M & Frank R 2012. Comparing methods for detecting child exploitation content online. European Intelligence and Security Informatics Conference. Odense, Denmark: 156–63

*Williams KS 2005. Facilitating safer choices: Use of warnings to dissuade viewing of pornography on the internet. *Child Abuse Review* 14(6): 415–429. https://doi.org/10.1002/car.920

*Wolak J, Liberatore M & Levine BN 2014. Measuring a year of child pornography trafficking by U.S. computers on a peer-to-peer network. *Child Abuse & Neglect* 38(2): 347–56. https://doi.org/10.1016/j.chiabu.2013.10.018

Wortley R & Smallbone S 2012. *Internet child pornography: Causes, investigation, and prevention*. Global crime and justice series. Santa Barbara, CA: Praeger

Wortley R & Smallbone S 2006. Applying situational principles to sexual offenses against children. In R Wortley & S Smallbone (eds), *Situational prevention of child sexual abuse*. Monsey, NY: Criminal Justice Press: 7–36

Dr Graeme Edwards is a Senior Lecturer of Cybersecurity and Cybercrime Investigations in the Institute for Cyber Investigations and Forensics at USC Australia.

Dr Larissa S Christensen is Co-Leader of the Sexual Violence Research and Prevention Unit at USC Australia.

Dr Susan Rayment-McHugh is Co-Leader of the Sexual Violence Research and Prevention Unit at USC Australia.

Dr Christian Jones is a Professor of Interactive Digital Media and Leader of the Engage Research Lab at USC Australia.

This project is part of the AIC's Child Sexual Abuse Material Reduction Research Program, funded under section 298 of the Commonwealth *Proceeds of Crime Act 2002*.

General editor, *Trends & issues in crime and criminal justice* series: Dr Rick Brown, Deputy Director, Australian Institute of Criminology. Note: *Trends & issues in crime and criminal justice* papers are peer reviewed. For a complete list and the full text of the papers in the *Trends & issues in crime and criminal justice* series, visit the AIC website at: aic.gov.au

ISSN 1836-2206 ISBN 978 1 922478 31 3 (Online)

https://doi.org/10.52922/ti78313

GPO Box 1936
Canberra ACT 2601, Australia

Tel: 02 6268 7166

Disclaimer: This research paper does not necessarily reflect the policy position of the Australian Government