# Submission to the Joint Committee on Law Enforcement

*regarding the inquiry into*

# Law enforcement capabilities in relation to child exploitation

25 January 2022

DIGITAL RIGHTS WATCH

Digital Rights Watch is a charity organisation founded in 2016 whose mission is to ensure that people in Australia are equipped, empowered and enabled to uphold their digital rights. We stand for Privacy, Democracy, Fairness & Freedom in a digital age. We believe that digital rights are human rights which see their expression online. We educate, campaign, and advocate for a digital environment where individuals have the power to maintain their human rights.[1]

[1] Learn more about our work on our website: https://digitalrightswatch.org.au/

## General remarks

Digital Rights Watch (DRW) welcomes the opportunity to provide a submission to the Joint Committee on Law Enforcement (the Committee) for consideration as part of the inquiry into law enforcement capabilities in relation to child exploitation (the Inquiry).

As a leading Australian organisation advocating for human rights as realised in the digital age, we are primarily concerned with ensuring an appropriate balance is struck with regard to the impact upon individuals' rights (including childrens' rights) and any adverse impacts on privacy, digital security, freedom of speech and expression, and democratic participation.

We recognise the severity of harm caused by child exploitation, and that there are genuine challenges regarding the safety of children and the distribution of unlawful material online, including the creation, storage and distribution of Child Sexual Abuse Material (CSAM).

We also wish to highlight that due to the sensitive and emotive nature of this area, it can sometimes obscure the dangers of wide-ranging policing and surveillance powers without adequate oversight and accountability. DRW is contributing to this Inquiry in the spirit of seeking to ensure that we don't end up *undermining* safety in the quest to secure it.

We wish to emphasise that while technology can and should play a role in tackling the issue of CSAM and child exploitation more broadly, centralised techno-solutionism in response to complex social problems often results in shortsighted or sometimes actively harmful tech policy and legislation. Complex social problems require holistic responses, in which technology can play a part but should not be the only mechanism. We urge the Committee to consider the range of possible negative consequences that can arise from good-intentioned techno-centric proposals.

**Specifically, we wish to emphasise that the protection of privacy and digital security are essential to the safety of children and young people in the digital age.**

Disagreements about end-to-end encryption have created a perceived dichotomy between a the right to privacy and a child's right to protection from sexual abuse and exploitation. However, the goal of ensuring that children's rights are safeguarded in the digital age involves fulfilment of their rights to both privacy *and* protection from sexual abuse and exploitation.[2]

The right to privacy is not absolute and can be limited. However, the limitation must be proportionate. Undermining the right to privacy *for everyone* can have dire consequences on both an individual and societal level, including inhibiting freedom of speech and expression, preventing people and organisations from holding those in power accountable and shining

[2] 'Encryption, Privacy and Children's Right to Protection from Harm,' *Unicef,* October 2020. Available at: https://www.unicef-irc.org/publications/pdf/Encryption_privacy_and_children%E2%80%99s_right_to_protection_from_harm.pdf

the light on abusive, unethical or unlawful behaviour (such as whistleblowers, political organisers, human rights lawyers and journalists), and in turn, weakening democracy. On an individual level, protecting privacy (often by way of robust digital security mechanisms), is an essential factor in enhancing safety — both online and physical — for people of all ages.

In the remainder of this submission we have outlined some of the challenges related to regularly-proposed technological options for tackling CSAM online. We do this not to dismiss the severity of the issue at hand, but to encourage a necessary dialogue about the limitations and risks of certain approaches such that informed and reasonable debate can ensue. This submission focuses on the following areas:

1. The importance of end-to-end encryption
2. The challenges associated with client-side scanning and automated detection of first generation material
3. The need for more robust safeguards regarding law enforcement surveillance powers, including access to databases

DRW regularly participates in government and industry consultation processes regarding technology policy and legislation. Previous submissions we have made that are relevant to this inquiry include:

- Online Safety draft Industry Codes[3]
- Inquiry into Social Media and Online Safety[4]
- Online Safety (Basic Online Safety Expectations) Determination[5]

## Encryption

We note that the Inquiry's Terms of Reference include remit to consider:

> *"the use by offenders of encryption, encryption devices and anonymising technologies, and Remote Access Trojans to facilitate their criminality, along with the resources of law enforcement to address their use."*

It is important to highlight that encryption, including end-to-end encryption, is widely used by everyday Australians, businesses, and government agencies in order to facilitate a wide range of services and communications. Framing encryption as merely an enabler of criminal activity obscures the essential role that encryption plays in Australia's modern digital society, economy and security.

[3] Submission to the steering group of industry associations on the draft Consolidated Industry Codes of Practice for the Online Industry, Phase 1, *Digital Rights Watch,* October 2022. Available at: https://digitalrightswatch.org.au/2022/10/11/submission-online-safety-draft-industry-codes/

[4] Submission to the Select Committee on Social Media and Online Safety regarding the inquiry into Social Media and Online Safety, *Digital Rights Watch,* January 2022. Available at: https://digitalrightswatch.org.au/2022/01/13/submission-inquiry-into-social-media-and-online-safety/

[5] Submission to the Department of Infrastructure, Transport, Regional Development and Communications regarding the draft Basic Online Safety Expectations, *Digital Rights Watch and Global Partners Digital,* November 2021. Available at: https://digitalrightswatch.org.au/wp-content/uploads/2021/11/Global-Partners-Digital-Digital-Rights-Watch-Joint-Submission.pdf

**Encryption is essential for safety, national and individual security, the economy, and the protection of human rights.**[6]

Encryption facilitates the security of our online activities; protecting data from potential criminals, enabling secure online transactions, and maintaining the privacy and security of online communications, including those of children.

Encryption is critical to ensure children's safety. Their digital devices and communications contain personal information that could compromise their privacy and safety (both physically and digitally) if accessed by malicious actors, such as location data (address, their route to school), behavioural data (routines, regular activities and other patterns of behaviour), and contact information. With an increasing number of home devices being connected to the internet, encryption plays an essential role to ensure that malicious actors do not gain access to devices that children may interact with. For example, encryption is crucial to prevent access to networked devices, including tapping into users' webcams, microphones or other devices, such as baby monitors or smart toys.

Strong encryption serves Australia's national interests by protecting governments, critical services, communities, and the economy from criminal, terrorist, and state-sponsored attacks. Encryption will only become more important for protecting Australian interests as technology advances in the coming decades.

Encryption is essential to our economy. Australians now make the majority of purchases with a card, and are increasingly paying online, with mobile phones, or using online automatic systems.[7] Encryption is central to card payments both online and through point-of sale machines to protect sensitive customer data. The government's *Digital Economy Strategy* estimates that digital technologies could add up to $250 billion to Australia's GDP by 2025.[8] In order for the Australian economy to flourish, the technologies that facilitate online business must be secure and align with global standards. Poor cyber security practices like weak encryption leave businesses and consumers vulnerable to costly criminal attacks and stifle entrepreneurship in Australia's digital economy.

Encryption supports fundamental human rights. End-to-end encryption is particularly essential to uphold the right to privacy, but in addition to this, it facilitates freedom of speech and expression, freedom of assembly, and the right to protest. Research has shown that end-to-end encryption is a vital safety tool for protecting human rights, and that the downsides of its implementation do not outweigh the benefits.[9]

[6] See, for example, 'The Role of Encryption in Australia', *Access Now,* January 2018, Available at: https://digitalrightswatch.org.au/2018/01/19/the-role-of-encryption-in-australia/

[7] 'How Australians Pay', *Reserve Bank of Australia,* 2019 Consumer Payments Survey. Available at: https://www.rba.gov.au/snapshots/how-australians-pay-snapshot/

[8] https://www.industry.gov.au/science-technology-and-innovation/technology

[9] 'Independent Assessment: Expanding End-to-End Encryption Protects Fundamental Human Rights', *Meta,* April 2022. Available at: https://about.fb.com/news/2022/04/expanding-end-to-end-encryption-protects-fundamental-human-rights/

The UN Special Rapporteur on Freedom of Expression has referred to end-to-end encryption as “the most basic building block” for digital security on messaging apps. Because of its critical role, the Special Rapporteur further notes that: “the responsibility to safeguard freedom of expression and privacy may require companies to establish end-to-end encryption as a default setting in their messaging products”. And, the Rapporteur also suggests that companies that offer messaging apps “should seek to provide the highest user privacy settings by default”.[10]

As a jurisdiction without a formal bill of rights, people in Australia are left with minimal protection of their human rights. There are very few limitations regarding law enforcement and intelligence powers regarding proportionality. This includes the breadth and intrusiveness associated with how searches are facilitated through various devices, networks, or infrastructure that are accessed during the execution of a computer access warrant. We urge the Committee to consider how compelling services to weaken, undermine, or otherwise bypass encryption threatens the digital security of Australians at an individual, community, and national security level.

While we appreciate the extent to which encrypted communications can sometimes introduce friction into criminal and intelligence investigations, we encourage much greater evidence-based discussion on these matters as a vital starting point for any consideration of legislative reform.

## Responsibilities of technology providers

We note that the Inquiry Terms of Reference include remit to consider:

> *“the role technology providers have in assisting law enforcement agencies to combat child exploitation, including but not limited to the policies of social media providers and the classification of material on streaming services”*

**Client-side scanning**

Client-side scanning broadly refers to systems that scan message contents (text, images, videos, files etc) for content matches or similarities within a database before the message is sent to the intended recipient.[11] Client-side scanning has received increased attention in recent years as an approach to detect CSAM images, whereby hashed images are compared to hash databases of known CSAM imagery to find matches. For example, the Online Safety Industry Codes considered the extent to which technology companies should

[10] ‘Encryption and Anonymity follow-up report,’ *United Nations Human Rights Special Procedures,* June 2018. Available at: https://www.ohchr.org/Documents/Issues/Opinion/EncryptionAnonymityFollowUpReport.pdf
[11] Fact Sheet: Client-Side Scanning, *The Internet Society,* 2020. Available at: https://www.internetsociety.org/resources/doc/2020/fact-sheet-client-side-scanning/

be required to do proactive scanning of user communications and devices.[12] Client-side scanning is sometimes presented as a way to scan content without breaking end-to-end encryption. While in some cases this may be technically true, this claim has been refuted.[13]

We suggest the Committee consider the following when deliberating client-side scanning:

- it undermines the promise of private and secure communications
- it facilitates the ability to monitor communications at scale, far beyond the detection of CSAM
- it creates vulnerabilities for criminals to exploit by creating additional ways to interfere with communications (increasing the 'attack surface')

Critically, it's not possible to build a client-side scanning system that can *only* be used for CSAM. Even a well-intentioned effort to build such a system can open the door to abuse. A 2021 technical research paper on the risks of client-side scanning highlighted how it provides the ability to pre-emptively scan for any type of content on any device, for any purpose, without a warrant or suspicion.[14] Compelling tech companies to implement such systems creates immense risk of communications surveillance at scale.

For example, local hash matching can check images against a database of known CSAM content. But if the system contains a complete mechanism to block any image content, all it takes is the ability to add items to the hash database to extend the scope of target material. Because of the nature of hash databases, it is difficult to determine the contents of a database just by inspecting it, making the contents effectively un-auditable to journalists, academics, and civil society. This makes it incredibly challenging to ensure that client-side scanning systems remain only used in certain, limited circumstances. Without the ability to build technical safeguards into the design, we are forced to rely purely on legal use restrictions. However, there are examples of scope creep or inappropriate or unlawful use of systems in Australia's recent history. For example, the scope of the metadata retention scheme slowly expanded after it passed in 2015, despite assurances that its use would be strictly limited.[15]

[12] Industry Explanatory Paper, *Online Safety Industry Codes,* 2022. Available at: https://onlinesafety.org.au/

[13] See, for example: 'Why adding client-side scanning breaks end-to-end encryption', *Electronic Frontier Foundation,* 2019. Available at: https://www.eff.org/deeplinks/2019/11/why-adding-client-side-scanning-breaks-end-end-encryption and 'Bugs in our Pockets: The Risks of Client-Side Scanning', October 2021. Available at: https://arxiv.org/abs/2110.07450

[14] 'Bugs in our Pockets: The Risks of Client-Side Scanning', October 2021. Available at: https://arxiv.org/abs/2110.07450

[15] See, for example: 'Australia's surveillance rabbit hole grows deeper,' *Access Now,* April 2020, available at: https://www.accessnow.org/australias-surveillance-rabbit-hole-grows-deeper/, 'Data retention scope creep: Border Force granted metadata access,' *Cnet,* May 2015, available at: https://www.cnet.com/tech/services-and-software/scope-creep-australian-border-force-granted-metadata-access/, 'Data retention: Telco group says councils, Australia Post, among organisations accessing 'metadata', *Computerworld,* November 2018, available at: https://www.computerworld.com/article/3461045/data-retention-telco-group-says-councils-australia-post-among-organisations-accessing-metadata.html

### Automated content moderation and detection of first generation content

While established processes exist to detect *known* CSAM — content that has been identified as CSAM and is contained within one of the relevant databases — the technology to detect material that has not previously been identified, or 'first generation' content, currently presents significant challenges and risks in relation to accuracy, over- or under- capture, privacy and digital security.

High quality automated detection of harmful content is extremely difficult, especially when the scope of target content is broad or not strictly defined.

Machine learning classifiers which seek to automatically flag possibly harmful content (as opposed to matching content to known, or previously identified harmful content) may be improving, but they remain seriously flawed when it comes to classifying complex material at scale. One issue is accuracy and the risk of both over- and under- capture of content. Over-capture can lead to significant negative consequences, in the worst cases leading to wrongful accusation of criminal activity. It can also result in the censorship of legitimate material, including sexual education and health material, the wrongful loss of access to online accounts (and thus, often income, support networks and community).

In a recent widely reported case, a father took images of his son's genitals, at the request of a nurse, and sent them to his doctor. The image was uploaded automatically to Google, which classified it as abuse and suspended his accounts, which he did not get back.[16] Law enforcement and intelligence agencies, as well as tech companies themselves, generally have very little liability for the cost of false positives, despite the immense harm that can follow as a result.

Another issue is that due to the lack of training data, classifier models are more likely to make mistakes related to marginalised groups, and in doing so further entrench existing inequality. Further, there remain ongoing challenges regarding the explainability of machine or deep learning classifiers. While this is an area of ongoing technical research and development, at this stage it may not be possible to explain or justify why some content is flagged by an automated machine learning content classification system.

[16] 'A Dad Took Photos of His Naked Toddler for the Doctor. Google Flagged Him as a Criminal,' *New York Times,* 21 August 2022. Available at: https://www.nytimes.com/2022/08/21/technology/google-surveillance-toddler-photo.html

## Law enforcement surveillance powers

We note the Inquiry Terms of Reference include remit to consider:

> *"reviewing the efficacy of and any gaps in the legislative tools and tactics of law enforcement used to investigate and prosecute offenders"*

Law enforcement already have a significant range of powers available to them to investigate and prosecute offenders, including those most recently passed under the *Surveillance Legislation Amendment (Identify and Disrupt) Act 2020* which included alarming new and intrusive surveillance powers for law enforcement.[17] This is on the back of a swathe of increases to law enforcement and intelligence powers, including:

- *Telecommunications (Interception and Access) Amendment (Data Retention) Act 2015*
- *Telecommunications and Other Legislation Amendment (Assistance and Access) Act 2018 (TOLA)*
- *International Production Orders Act 2021*

In late 2021 the Department of Home Affairs released a discussion paper on *Reform of Australia's Digital Surveillance Framework* which includes a range of proposals to reform Australia's laws governing electronic surveillance to be "clearer, more coherent, and better adapted to the modern world."[18] DRW made a submission[19] providing feedback, emphasising that the surveillance framework should:

1. confine authorisations and warrants to the minimum necessary number of agencies and only add to this via a clear and transparent process;
2. avoid an overly general definition for communications that does not permit nuance, respect individual rights and the capacity to tailor different powers to a demonstrated need by surveillance agencies;
3. simplify warrants but not at the expense of the individuals' rights, and subject them to a necessary and proportionate test;
4. raise the threshold in the definition of a "serious criminal offence";
5. introduce a double lock system as standard;
6. require that the relevant Minister should provide an annual (or more frequent) report that details the number of warrants issued under national security legislation and publish judicial records of these decisions; and

[17] 'Australia's new mass surveillance mandate,' *Digital Rights Watch,* September 2021. Available at: https://digitalrightswatch.org.au/2021/09/02/australias-new-mass-surveillance-mandate/

[18] Reform of Australia's electronic surveillance framework discussion paper, *Department of Home Affairs,* Available at: https://www.homeaffairs.gov.au/reports-and-publications/submissions-and-discussion-papers/reform-of-australias-electronic-surveillance-framework-discussion-paper

[19] Submission to the Department of Home Affairs regarding the Electronic Surveillance Reform Discussion Paper, *Digital Rights Watch,* February 2022. Available at: https://digitalrightswatch.org.au/2022/02/25/submission-electronic-surveillance-reform-discussion-paper/

7. prescribe an independent public interest advocate to make submissions on any warrant application.

We urge the Committee to consider the broad range of powers already available to law enforcement and intelligence agencies that compel companies and organisations to collect, retain and provide access to data, undermine encryption, and conduct networked surveillance without necessary need for a warrant. These powers are deeply invasive and often operate without appropriate oversight or accountability. Before any new powers for law enforcement are considered, a comprehensive and independent review of the current powers available to law enforcement and their efficacy should be conducted.

Providing law enforcement additional powers without requisite oversight and accountability risks those powers being misused. In the past few years there are a wealth of examples of inappropriate, unethical and sometimes blatantly unlawful use of police powers and access to data, which should be taken into consideration as part of any discussion regarding the potential additional powers in relation to child exploitation. Recent examples include but are not limited to:

- Misuse of Victoria Police's sensitive Law Enforcement Assistance Program (LEAP) database[20]
- Ombudsman report found law enforcement routinely break the law in handing private data, and are getting worse rather than better[21]
- NSW Police gave a domestic violence survivor's phone data to her perpetrator[22]
- Queensland police officer accessed a confidential computer system, gave the address of domestic violence survivor to her former partner, and then had his conviction overturned[23]

[20] Victoria Police not doing enough to prevent database misuse, privacy experts say,' *ABC News,* 16 December 2022. Available at: https://www.abc.net.au/news/2022-12-16/victoria-police-privacy-personal-information-criticism-experts/101751500

[21] 'Pretty creepy: Agencies illegally obtained emails, voicemails and texts,' *The Saturday Paper,* 17 September 2022. Available at: https://www.thesaturdaypaper.com.au/news/politics/2022/09/17/pretty-creepy-agencies-illegally-obtained-emails-voicemails-and-texts#hrd

[22] 'Police accidentally gave domestic violence victim's phone data to the attacker,' *ABC News,* 2 June 2021. Available at: https://www.abc.net.au/news/2021-06-02/police-gave-domestic-violence-victim-data-to-attacker/100173270

[23] 'Queensland police officer who leaked address of domestic violence victim has conviction overturned,' *The Guardian,* 1 September 2022. Available at: https://www.theguardian.com/australia-news/2020/sep/01/queensland-police-officer-who-leaked-address-of-domestic-violence-victim-has-conviction-overturned/

## More research needed to support evidence-based approaches

To conclude, we wish to highlight the need for more robust research into the area of CSAM and evidence-informed interventions such that effective and rights-respecting approaches can be established.

A systematic review of criminal justice response to CSAM conducted by the Australian Institute of Criminology revealed a "scarce evaluation research which limited the ability to holistically address CSAM."[24] The report also notes:

> *"Overall, the existing intervention literature in the area of CSAM is largely descriptive, with potentially promising interventions evaluated with low-quality research designs that do not reliably establish effectiveness (see also Gallo 2020; Perkins et al. 2018)."*
>
> *"Without a rigorous evidence base, policymakers and practitioners are unable to make reliable decisions about what criminal justice responses are effective in addressing CSAM offending and, potentially, what may be harmful."*

Given the intrusive nature of regularly-proposed technological "solutions" to CSAM and the significant risks they pose to human rights of all Australians (including children), it would be unreasonable to consider the introduction of any new or expanded technical capabilities or surveillance powers before a rigorous evidence-base is established. A determination as to whether limitations on certain rights — such as the right to privacy — are indeed reasonable or proportionate cannot be demonstrated in the absence of robust, independent research. Otherwise, interventions may stand to undermine human rights and risk creating *additional* harm without substantial assurance off efficacy at actually reducing the harm of CSAM.

## Contact

**Samantha Floreani** | Program Lead | Digital Rights Watch | samantha@digitalrightswatch.org.au

[24] 'Criminal justice response to child sexual abuse material offending: a systematic review and evidence and gap map,' *Australian Institute of Criminology,* April 2021. Available at: https://www.aic.gov.au/sites/default/files/2021-03/ti623_criminal_justice_responses_to_csam_offending.pdf