Australian Government
Attorney-General's Department

June 2024

Submission

To the Statutory Review of the Online Safety Act 2021

Attorney-General's Department

Content warning: This submission includes statistics and other information relating to child sexual exploitation and abuse. If you, or someone you know, has been impacted by child sexual abuse, or is concerned about a child's safety, there are services and resources available to help. Please visit www.childsafety.gov.au for more information.

Contents

1. Introduction

Scope of submission

The Australian Government is committed to strengthening Australia's policy, regulatory and criminal justice frameworks to prevent and respond to harmful and illegal material online. Following the meeting of National Cabinet on 1 May 2024, the Australian Government reiterated that it has zero tolerance for online harms and abuse, and the use of technology to harm women and exploit vulnerable young people.[1]

At the time of its implementation, the *Online Safety Act 2021* (the Act) was fit-for-purpose, world-leading legislation. The online threat landscape has evolved considerably since this time, with the advent and uptake of new technologies such as end-to-end-encryption and generative artificial intelligence (GenAI) presenting new and emerging threats for its users, and new challenges for policy makers and the criminal justice system. Despite efforts by governments to engage companies early in the design of new products and services, digital platforms and online service providers continue to build and deploy new features without appropriate and considered safeguards.

The Attorney-General's Department (the department) leads the Commonwealth's policy response to child sexual abuse, and the criminal justice response to child sexual abuse offences. The *Criminal Code Act 1995* (Cth) (Criminal Code) provides a comprehensive framework of offences relating to child sexual abuse committed online, via postal services or by Australians overseas including in relation to child sexual abuse material. Under Australia's world leading *National Strategy to Prevent and Respond to Child Sexual Abuse 2021-2030* (National Strategy), this offence framework is complemented by a policy response that focuses on prevention by actively engaging the Australian community, international partners and the tech industry to enhance online child protection. This includes promoting systemic safety uplift across industry and working with international partners to ensure a global safety net of online protection. The Act is an important part of this effort, mandating transparency and accountability. The department works closely with the Australian Federal Police (AFP), and the AFP-led Australian Centre to Counter Child Exploitation. The AFP will be providing a standalone submission.

The Attorney-General is responsible for administering offences in the Criminal Code enacted through the *Counter-Terrorism Legislation Amendment (Prohibited Hate Symbols and Other Measures) Act 2023*. This includes offences for the public display of prohibited Nazi or terrorist organisation symbols, performance of an act that is a Nazi salute in a public place, and using a carriage service for violent extremist material and possessing or controlling violent extremist material obtained or accessed using a carriage service.

The Attorney-General also administers cybercrime offences in the Criminal Code, relating to use of a carriage service to menace, harass or cause offence and for the non-consensual sharing of intimate images.

The department is responsible for aspects of family, domestic and sexual violence policy and reform, including initiatives related to criminal justice responses to sexual assault; justice sector training, education

[1] Tackling online harms | Prime Minister of Australia (pm.gov.au)

and awareness on gender-based violence; addressing coercive control; legal assistance funding; and strengthening the family law system.

The department works closely with the Department of Social Services, which has overarching responsibility for the *National Plan to End Violence against Women and Children 2022-2032*; the national policy framework that will guide actions towards ending violence against women and children over the next 10 years.

The department is responsible for Australia's privacy framework, including administration of the *Privacy Act 1988* (Privacy Act), Australia's legislative mandate that protects the privacy of individuals, and sets out how government and industry can collect, use, and disclose individuals' personal information.

The department welcomes the opportunity to contribute to the review, which is an important opportunity to ensure that powers given to the eSafety Commissioner under the Act are enforceable, keep pace with Criminal Code reforms, and balance personal privacy with the need to protect the Australian community from online harms.

2. Voluntary measures to address online harms

A number of studies have attempted to estimate the global prevalence of child sexual abuse. While there is no agreed figure, most large studies have reported rates of child sexual abuse of between one in ten and one in four children.

The Australian Child Maltreatment Study (ACMS) is the first nationally representative study of the prevalence, context and health impacts of the five forms of child maltreatment in Australia. Published in March 2023, the ACMS reported that one in four children experienced sexual abuse at some time during their childhood. The rates of abuse were higher for girls (over one in three) than boys (nearly one in five). The WeProtect Global Childhood study interviewed more than 5300 participants on their experiences of online sexual harm during childhood. The study reported that one in two participants experienced one or more types of online child sexual abuse during childhood.

This research gives insights into the prevalence of online child sexual exploitation and abuse, and demonstrates the need for a strong whole-of-government response to online harms. The National Strategy includes measures which seek to engage industry and international partners, with a core theme is working with the tech industry to support Australia's response to this crime type.

In March 2020, the Australian Government, alongside its Five Eyes partners, launched the *Voluntary Principles to Counter Online Child Sexual Abuse* (the Voluntary Principles), a high-level framework designed to support industry actors to harden their platforms and services to online child sexual exploitation and abuse, and increase the prevention, detection, and reporting of such material and predatory behaviour. The Voluntary Principles were developed in consultation with six key digital industry companies (Facebook, Google, Microsoft, Roblox, Snap, TikTok and Twitter), non-government organisations and academia. Forums such as the Five Country Ministerial, comprising of Australia, the United Kingdom, the United States, Canada and New Zealand, provide a platform to engage industry as a collective, encouraging industry to endorse and implement the Voluntary Principles.

March 2024 will mark five years since the launch of the Voluntary Principles, with a fraction of tech companies having publicly signed up. Reflecting on the public signatories, the department has observed a

limited commitment to fundamental principles such as design decisions based on the threat landscape, targeting of grooming and livestreaming, and measures to proactively detect child abuse material.

In an effort to champion global voluntary transparency, in 2022 Australia worked with the Tech Coalition to develop TRUST: The Tech Coalition's Transparency Framework (TRUST framework). The TRUST framework was an industry driven initiative to provide flexible guidance to tech companies seeking to demonstrate accountability and transparency in reporting on steps they have taken to combat online child sexual abuse. Despite being developed by industry, the TRUST framework is still to be implemented by all Tech coalition members.

Although these initiatives demonstrate willingness from some companies to lead on safety efforts, government continues to experience challenges in engaging industry more broadly, with some companies taking limited or no action to address harmful or unlawful content. Industry can take further steps to place user safety at the forefront of design and business choices, and demonstrate transparency around what platforms are doing or not doing to protect users. Public statements and pressure from governments and civil society have limited impact on industry design choices, most recently with the rollout of default end-to-end encryption across Meta's platforms.

Voluntary frameworks, such as the Voluntary Principles, remain an important tool for countries where regulation is unlikely to be forthcoming. However, voluntary measures, including those developed by industry, often struggle to generate uptake across the sector, and unless developed in cooperation with government and law enforcement, are informed by industry-level intelligence and insights. Experience also shows that voluntary efforts fail to keep pace with online offending and are not sufficient to ensure industry are implementing appropriate protections for Australian users.

As such, regulatory measures are crucial to ensuring industry compliance and transparency. Enforceable regulation that builds in appropriate guardrails is needed, and penalties must not be perceived as a cost of doing business in the global economy.

3. Basic Online Safety Expectations and Industry Codes

The department supports the strengthening of the Basic Online Safety Expectations (BOSE) to further bolster Australia's response to harmful or unlawful content. The intent of the BOSE and its ability to elicit transparency is an important part of the eSafety Commissioner's powers. The BOSE transparency notices can ask precise and targeted questions, the answers to which would not be willingly published by industry otherwise. The utility of the BOSE has been demonstrated in recent US Senate hearings used to question industry on their lack of action to address child sexual abuse.

While the BOSE notices have resulted in short term changes by industry, sustained transparency efforts over time are required to ensure systematic uplift.

Responses provided to BOSE transparency notices should lead to mandatory action when it becomes apparent to the eSafety Commissioner that industry is falling short. Consideration could be given to establishing a mechanism to ensure that findings from the BOSE transparency notices are able to be used to amend the Industry Codes and Standards, which are enforceable.

eSafety should be adequately resourced to implement wider periodic reporting of the BOSE and to monitor and enforce compliance with the Industry Codes and Standards.

4. Statutory Duty of Care

The department supports the inclusion of a duty of care obligation on industry, similar to that in the UK's *Online Safety Act 2023* (UK's Online Safety Act). The UK's Online Safety Act for example, requires that platforms preventatively safeguard their systems through measures such as risk assessments and mitigation measures, and requires systematic efficiency assessments of these measures. A duty of care provision could build on existing mechanisms designed to ensure systematic uplift at the prevention stage, such as safety by design and the industry codes.

The Act's takedown framework remains a critical tool for the eSafety Commissioner to address illegal or restricted content being hosted online. However, this approach relies on a targeted and individual response for each item of illegal or restricted content and cannot proportionately address the continuous and mass proliferation and distribution of harmful and unlawful content across both the dark and clear web. The harm experienced by victims of online abuse does not cease once content is taken down, and impacts can be long-lasting. Placing an overarching positive obligation on industry would have a significant preventative impact, by minimising harm before it occurs.

If a statutory duty of care is progressed, the department welcomes the opportunity to engage further on its operation, alongside reform to establish a statutory tort for serious invasions of privacy.

5. Emerging technologies

The increase in availability and use of GenAI can facilitate criminal behaviours including grooming, allowing perpetrators to better target children and tailor scripts, such as has been seen in the widespread sextortion of teenagers in Australia. Grooming and other preparatory behaviour is being used systematically by offender networks and organised crime gangs to offend against children, and can lead to tragic outcomes.

In the world of immersive technologies such as in gaming, it will become harder to identify unlawful content due to the live nature of the technology; a challenge Australian law enforcement has encountered in combatting the live streaming of online child sexual abuse. Successful prevention efforts have targeted preparatory grooming behaviours, such as those observed in chat logs between a child and perpetrator, and behavioural signals that are linked to online child sexual abuse and are visible in encrypted spaces, including suspicious user activity (such as mass contacting of unknown or underage accounts), and the use of specific emojis and vernacular. It is important that the Act addresses such criminal behaviours alongside content-based abuse, to ensure a strong regulatory response to online child sexual exploitation and abuse.

A duty of care obligation would likely capture the impact of emerging technologies without the need to continually revise regulatory powers.

6. Civil penalties, investigation powers and information gathering powers

Australia's civil penalty regime was the first of its kind in the world. Since it came into force, Ireland, the EU and the UK have introduced similar regimes with higher penalties. The review should reconsider Australia's penalties in light of this. The scale of penalties should be consistent with other domestic and international legislation, to recognise the significant harm caused by online offending such as child sexual abuse. As a point of comparison, the penalty regime provided for under the UK's Online Safety Act enables their regulator to fine companies an amount up to 10% of their global annual turnover[2]. Canada's proposed Bill C-63 *Online Safety Act* would mirror similar penalties to those provided under the UK's Online Safety Act.

The department supports the enhancement of the eSafety Commissioner's information gathering powers, such as eSafety's ability to collect provider data on the prevalence and nature of harms, and the nature and outcomes of user reports and complaints to platforms. Importantly, this data should be disaggregated by gender and age where possible, reflecting that children and young people, women and gender diverse people are disproportionately impacted by all forms of online harms and abuse. The department notes the eSafety Commissioner handles personal information in accordance with obligations under the Privacy Act.

The department notes that two-thirds of reports to eSafety about cyberbullying, image-based abuse and adult cyber abuse are made by women and girls[3] . The lack of gender-disaggregated data collected from providers represents a gap that could be addressed through reform to the Act, noting the gendered nature of complaints to eSafety's regulatory schemes. Transparency of data would support effective policy responses, including as they relate to the eSafety Commissioner's regulatory schemes such as the image-based abuse scheme, and assist government's ability to better respond to emerging areas of concern.

7. Extra-territorial powers

. The extra-territorial powers under the Act are critical to Australia's online safety architecture, and assist in ensuring a company is held responsible for protecting Australian users, regardless of where the company is head quartered. However, noting most companies with the largest number of Australian end-users are invariably based in the US, this poses challenges to the applicability of domestic legislation and ability for law enforcement to investigate criminal activity.

The department recommends the review considers ways in which the eSafety Commissioner can pursue fines and penalties from overseas-based companies. Fines and penalties must be enforceable in practice, noting that much of the compliance regime is designed to enhance prevention efforts and reduce harm to the Australian community before it occurs. Solutions to addressing challenges to enforcing fines and the penalty regime in Australia could be drawn from international regulatory schemes, such as in the UK, that have sought to address this.

[2] A guide to the Online Safety Bill - GOV.UK (www.gov.uk)
[3] eSafety, How the new Online Safety Act supports women, Online Safety Act fact sheet, updated February 2022

8. Privacy reforms

The Australian Government is committed to uplifting privacy protections while encouraging digital innovation. On 16 February 2023 the Privacy Act Review Report was published, and on 28 September 2023 the Government responded to the Report. The Government Response sets out a pathway for reforms, balancing the benefits of privacy enhancements against compliance costs for entities and other impacts. The reforms will uplift privacy standards, protect vulnerable groups from harm, provide certainty to businesses and enhance enforcement mechanisms.

Serious invasions of privacy

The Government's privacy reforms would introduce changes that complement existing powers in the Act, which allow the eSafety Commissioner to issue takedown notices for harmful content. In response to the department's Privacy Act Review, the Government has agreed to introduce a statutory tort for serious invasions of privacy involving intrusion upon seclusion or misuse of private information, which could address online harms such as doxing. A statutory tort would apply to individuals and a broader range of entities than those regulated by the Privacy Act. This would also complement existing offences in the Criminal Code that criminalise the use of a carriage service to menace, harass or cause offence.

The statutory tort will be informed by the model recommended by the Australian Law Reform Commission[4] and proposed in the Privacy Act Review Report contains a range of safeguards. A person bringing a claim would have to demonstrate that:

- the privacy invasion was serious
- they had a reasonable expectation of privacy
- the invasion was committed intentionally or recklessly (not merely negligently), and
- the public interest in privacy outweighs any countervailing public interests.

There would also be a number of defences, including a defence for fair reporting of proceedings of public concern.

The proposed statutory tort for serious invasions of privacy provides an option for individuals to seek redress if they are victims or targets of image-based abuse by enabling them to seek compensation and other remedies if their private information is misused.

The department is considering stakeholder feedback to ensure any reforms which could address online harms such as doxing are appropriate. The department is also reviewing written submissions made by the public during consultation on doxing and privacy reforms, and is carefully considering proposals to strengthen the protection of personal information in the context of broader updates to the *Privacy Act 1988*.

The Australian Government has announced that it will bring forward legislation in early August 2024 to outlaw the malicious release of personal information online and strengthen existing offences in the Criminal

[4] Serious Invasions of Privacy in the Digital Era (Australian Law Reform Commission Report 123), 74.

Code. This will ensure that perpetrators who seek to abuse or degrade people through doxxing, or by abusing their privacy online, will be subject to serious criminal penalties.

Greater transparency and control for users

The Government has also agreed-in-principle to proposals from the Privacy Act Review to provide individuals with greater transparency and control over their personal information, through the creation of new individual rights which would enhance dialogue and cooperation between entities and individuals on how their personal information is being handled (as outlined in the Government's response to proposals made in Chapter 18 of the Privacy Act Review Report). This includes a right to de-indexation, which would enable individuals to request a search engine to remove search results containing their personal information which is:

- sensitive information
- information about a child
- excessively detailed (for example a home address and personal phone number), or
- inaccurate, out-of-date, incomplete, irrelevant or misleading.

Additionally, a right to erasure would give individuals the ability to request that an entity destroy or de-identify personal information that it holds in relation to the individual.

The department notes there are intersections between the Privacy Act Review's consideration of any new additional arrangements to address online harms, and the government's proposed privacy reforms. For example, the Privacy Act Review Report identified harmful and invasive ways digital platforms can collect and use children's data and exploit children's vulnerabilities. The Government has agreed to progress amendments to the Privacy Act which would:

- require entities to have regard to the best interests of the child when handling their personal information
- provide for a Children's Online Privacy Code
- prohibit entities from targeting and direct marketing to children, unless it is in a child's best interests, and prohibit entities from trading in the personal information of children, and
- require privacy policies and notices to be child-friendly – for example, through the use of easy to understand language and graphics or videos.

9. Removal notices for pro-terror, violent extremist material and hate speech

In January 2024, new offences in the Criminal Code enacted through the *Counter-Terrorism Legislation Amendment (Prohibited Hate Symbols and Other Measures) Act 2023* (Prohibited Hate Symbols Act) came into force. These offences are for the public display of prohibited Nazi and terrorist organisation symbols, the making of a gesture that is the Nazi salute, using a carriage service for violent extremist material, and

possessing or controlling violent extremist material obtained or accessed using a carriage service. The prohibited symbols and salute offences apply to conduct in a 'public place', which is defined in the Criminal Code Dictionary. The explanatory memorandum to the Prohibited Hate Symbols Act specifies that the definition of 'public place' is intended to capture both physical and online places.[5]

The offences were enacted to facilitate law enforcement intervention at an earlier stage in individuals' progress to violent radicalisation, and provide greater opportunities for rehabilitation and disruption of violent extremist networks. The offences are also intended to complement the existing framework for regulating online service providers, including offences for hosting abhorrent violent material.

Noting the harm that the public display of prohibited hate symbols and the dissemination of violent extremist material can have, it is important to consider the application of the eSafety Commissioner's powers to such content. The department understands that the eSafety Commissioner is able to issue removal notices in relation to this content should it fall within one of the existing categories in the Act. To ensure a comprehensive and integrated response to this issue, the review may wish to consider how the eSafety Commissioner's existing powers complement the powers provided to police in the Criminal Code to direct people to act to remove prohibited hate symbols from display.

Similarly, the review may wish to consider the eSafety Commissioner's powers in relation to violent extremist material. Violent extremist material causes harm to the community by encouraging and instructing individuals in the commission of violent acts, and by radicalising individuals to violent extremist ideologies. Examples of material that may constitute violent extremist material, for the purposes of the Criminal Code, include images and videos depicting terrorist incidents such as violent extremist manifestos and propaganda; and instructional material covering topics such as how to build a bomb, attack a person, or manufacture harmful chemicals.

Earlier this year, the Attorney-General announced that the Government would bring forward legislation to strengthen current laws that deal with hate speech. The Government is committed to pursuing legislative amendments to create new criminal offences and strengthen existing laws relating to hate speech. The review may wish to consider how the eSafety Commissioner's powers and hate speech laws may interact.

10. Complaint and content-based removal schemes

Complaint and content-based removal schemes complement efforts to ensure the appropriate criminalisation of harmful conduct online and support law enforcement and regulators to respond to and protect Australians across online platforms and forums.

For example, the Act has a take-down regime for image-based abuse, in which intimate images or videos are shared online without consent. This is a damaging type of technology-facilitated abuse that disproportionately affects women and girls and can inflict deep harm on victims[6].

[5] See paragraph 266 of the revised explanatory memorandum to the Counter-Terrorism Legislation Amendment (Prohibited Hate Symbols and Other Measures) Bill 2023.

[6] Tackling online harms | Prime Minister of Australia (pm.gov.au)

On 5 June 2024 the Government introduced the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 to modernise and strengthen offences for the non-consensual sharing of simulated and real sexual material online. The reforms make clear that the online sharing of sexually explicit material without consent, including material generated or altered using technology (like AI-generated deepfakes), will be subject to appropriate criminal penalties. The Bill models some elements of the content -based removal schemes in the Act to ensure a consistent response.

The department notes the complementary relationship between content-based removal regimes and criminal justice responses in ensuring a robust and effective framework against online harms. The department welcomes amendments to the Act to improve efficiency of the schemes.

11. Conclusion

The department welcomes the opportunity to provide this submission and supports the appropriate strengthening and expansion of powers through the Act to further protect the Australian community.

It is important that Australia's regulatory framework remains tech-neutral and adaptive to allow for a timely response as the online threat landscape evolves and new technologies emerge.

The department will continue to work with the Department of Infrastructure, Transport, Regional Development, Communication and the Arts and the eSafety Commissioner to ensure this Review's consideration of additional arrangements to address online harms takes a holistic approach, and complements other policy and legislative reforms underway.