merillot

Submission

to the

Statutory Review of the Online Safety Act 2021

Version 1

Issued June 21st, 2024

Executive Summary

This submission addresses key regulatory and operational concerns surrounding the *Online Safety Act 2021* (Cth) and its application by the eSafety Commissioner. Despite the Act's aim to enhance online safety for Australians, significant issues have arisen, including substantial financial costs, questionable effectiveness of measures, potential harm to Australia's international reputation, threats to online freedom of speech, and concerns about specific censorship actions by the Commissioner. This submission summarises responses to the questions posed in the issues paper and expands on these points in more detail.

The eSafety Commissioner's efforts, particularly in addressing cyberbullying, have been insufficient despite a budget of $53.7 million in FY 2022-2023. The number of formal notices issued is low compared to the prevalence of cyberbullying, indicating a need for more proactive and effective strategies. The Commissioner's attempts to censor content threaten Australians' rights to freedom of speech, especially political communication. Actions such as the pursuit of the Billboard Chris case and censorship of the Wakeley Church stabbing video have raised concerns about overreach and ideological bias, negatively impacting Australia's international reputation.

There is a significant lack of transparency in the eSafety Commissioner's decision-making processes, particularly regarding the criteria for specific censorship actions. This lack of transparency undermines trust and calls for enhanced accountability measures and clearer guidelines to prevent misuse of powers.

The current objects of the Act are generally sufficient but could benefit from an emphasis on proactive measures rather than reactive ones, particularly in addressing emerging online harms and enhancing preventive strategies. Strengthening Basic Online Safety Expectations could enhance accountability and safety, provided they are implemented with clear guidelines to avoid overreach. Greater flexibility around industry codes, involving a broader range of stakeholders, could improve the code drafting process.

The thresholds set for each complaints scheme are appropriate, but the Commissioner should strictly act only when these thresholds are met. The current approach often leads to actions on cases that do not meet the statutory thresholds, representing a misuse of resources. Empowering bystanders to report illegal or seriously harmful material could be beneficial, provided there are safeguards to prevent misuse and ensure reports are handled appropriately. The Commissioner's current powers are adequate but need clearer enforcement guidelines to avoid misuse and overreach. Establishing an independent review body would assist in ensuring that enforcement is done within the guidelines of the law.

The penalties under the Act are generally adequate, but their effectiveness depends on consistent and fair enforcement. Increasing penalties or introducing new powers to impose sanctions could lead to overreach and unintended negative consequences. Strengthening international cooperation and agreements could enhance the enforcement of actions against non-compliant overseas providers within the limits of what can be realistically enforced. However, it is important to explicitly legislate that the eSafety Commissioner does not have the power to take down content globally and is limited to taking down content that us regularly accessible in Australia. The Act should remain technology-neutral to stay adaptable. Introducing a statutory duty of care or Safety by Design obligations should be avoided to prevent stifling innovation and infringing on privacy rights.

The effectiveness of the Act in achieving its object of improving and promoting online safety for Australians is questionable, especially considering the significant resources spent fielding complaints that do not meet the statutory threshold. Fine-tuning the existing framework could ensure it is more effective at a lower cost. Transparency in the eSafety Commissioner's decision-making processes and

establishing independent oversight mechanisms are critical to maintain public trust and ensure actions are justified, proportionate, and aligned with the statutory objectives of the Act.

Expanding the eSafety Commissioner's powers without resolving existing issues of overreach could lead to unintended consequences, such as stifling legitimate discourse and infringing on privacy rights. A balanced approach that prioritises transparency, accountability, and collaboration with global platforms is essential to effectively protect users without compromising fundamental freedoms. Introducing a cost recovery mechanism or additional statutory duties on online services must be approached with caution to avoid placing undue financial burdens on service providers and potentially stifling innovation. This submission underscores the need for a more measured and refined approach to online safety regulation, ensuring that the eSafety Commissioner's actions are firmly grounded in the statutory objectives and aligned with the principles of free speech and democratic accountability.

TABLE OF CONTENTS

1.0 Introduction

This submission addresses regulatory, operational, and concerns surrounding the *Online Safety Act* 2021 (Cth). Despite the Act's objectives to improve and promote online safety for Australians, the application of these provisions by the eSafety Commissioner has raised several significant issues. These issues include substantial financial costs, the effectiveness of measures generally, potential harm to Australia's international reputation, threats to online freedom of speech, and concerns about the Commissioner's specific censorship actions. The submission is structured to first summarise answers to the questions posed in the issues paper and then expands on these points in more detail.

The effectiveness of the eSafety Commissioner's efforts, particularly in addressing cyberbullying, remains questionable. Despite a substantial budget, spending $53.7 million in financial year 2022-2023, the number of formal notices issued is disproportionately low compared to the prevalence of cyberbullying incidents. This reactive approach has proven insufficient in mitigating the issue, indicating a need for more proactive and effective strategies.

The eSafety Commissioner's attempts to censor content threaten Australians' rights to freedom of speech, particularly the freedom of political communication. The Commissioner's actions appear to conflict with the fundamental principles of free speech, raising serious concerns about the overreach of regulatory powers. These censorship activities have led to negative perceptions of Australia on the international stage. These actions have painted Australia as ideologically repressive, drawing widespread criticism and adverse publicity. Such actions undermine Australia's image as a proponent of free speech and democratic values. For example, the pursuit of the Billboard Chris case is troubling. Chris Elston, known as "Billboard Chris," advocates against certain aspects of transgender ideology affecting children. His activities constitute legally protected political expression. The eSafety Commission's actions in this case appear to unjustly interfere with free speech. Similarly, the eSafety Commissioner's efforts to censor the video of the Wakeley Church stabbing are concerning. The censorship of this video can be seen as an attempt to limit public discussions about such incidents, contrary to democratic principles. The video is undeniably of public interest, and its removal contradicts democratic principles of transparency and open discourse.

The lack of transparency about eSafety operations and staff raises serious concerns. In particular, the decision-making criteria on specific decisions is notable absence. Transparency and fairness are critical in ensuring accountability in any future role of eSafety.

2.0 Summary of Questions

This section includes brief answers to the specific questions posed by the review. A more comprehensive discussion and analysis is included in prior sections.

2.1 Part 2 – Australia's regulatory approach to online services, systems and processes

1. Are the current objects of the Act to improve and promote online safety for Australians sufficient or should they be expanded?

The current objects are generally sufficient but could benefit from an emphasis on proactive measures rather than reactive ones, particularly in addressing emerging online harms and enhancing preventive strategies.

2. Does the Act capture and define the right sections of the online industry?

The Act broadly captures the necessary sections of the online industry.

3. Does the Act regulate things (such as tools or services) that do not need to be regulated, or fail to regulate things that should be regulated?

The Act broadly captures the necessary things.

4. Should the Act have strengthened and enforceable Basic Online Safety Expectations?

Yes, strengthening these expectations and making them enforceable could enhance accountability and safety, provided they are implemented with clear guidelines to avoid overreach.

5. Should the Act provide greater flexibility around industry codes, including who can draft codes and the harms that can be addressed? How can the codes drafting process be improved?

Greater flexibility could be beneficial, allowing for industry-specific expertise in drafting codes. The process could be improved by involving a broader range of stakeholders, including consumer advocacy groups, technology experts and groups such as the Electronic Frontier Foundation.

6. To what extent should online safety be managed through a service providers' terms of use?

While terms of use are important, relying solely on them is insufficient. Regulatory oversight is necessary to ensure compliance and address issues that terms of use alone cannot manage. However, the Act or regulations should be amended to explicitly prevent the eSafety Commissioner from making requests to platforms that are merely in violation of terms of service and do not meet the applicable statutory threshold as outlined in section 4.3.

7. Should regulatory obligations depend on a service providers' risk or reach?

Yes, obligations should be proportionate to the risk and reach of the service provider, ensuring that larger, higher-risk platforms have more stringent requirements. Caution should be taken however to not place such obligations on service providers that the Australian market becomes unattractive.

2.2 Part 3 – Protecting those who have experienced or encountered online harms

8. Are the thresholds that are set for each complaints scheme appropriate?

The thresholds are appropriate, but the Commissioner should strictly act only when these thresholds are met. Currently, actions on cases not meeting the statutory thresholds are a misuse of resources and need to be curtailed.

9. Are the complaints schemes accessible, easy to understand and effective for complainants?

Generally, yes. Given the large volume of complaints that do not meet the statutory threshold, it seems that there are no issues accessing the complaint schemes.

10. Does more need to be done to make sure vulnerable Australians at the highest risk of abuse have access to corrective action through the Act?

It is unclear whether more needs to be done for vulnerable Australians specifically on this front. It would be more effective to consider broader activities that assist vulnerable Australians more generally.

11. Does the Commissioner have the right powers to address access to violent pornography?

The current powers are adequate but need clearer guidelines to ensure appropriate and effective use without overreach.

12. What role should the Act play in helping to restrict children's access to age inappropriate content (including through the application of age assurance)?

The Act should support and encourage the implementation of age assurance mechanisms by service providers to restrict children's access to age-inappropriate content, while balancing privacy concerns and ensuring that such measures are not overly invasive. Compulsory participation in digital identification schemes should be avoided. Encouraging best practices and providing guidelines for age verification can help protect children without imposing overly stringent controls.

13. Does the Commissioner have sufficient powers to address social media posts that boast about crimes or is something more needed?

Existing powers are generally sufficient, but clearer criteria and streamlined processes could enhance effectiveness. Various state legislation already covers this, and it is important not to duplicate that legislation.

14. Should the Act empower 'bystanders', or members of the general public who may not be directly affected by illegal or seriously harmful material, to report this material to the Commissioner?

Empowering bystanders could be beneficial, provided there are safeguards to prevent misuse and ensure reports are handled appropriately. Specifically, it is important that merely being offended does not constitute cause for removal.

15. Does the Commissioner have sufficient powers to address harmful material that depicts abhorrent violent conduct? Other than blocking access, what measures could eSafety take to reduce access to this material?

Current powers are adequate but need clearer enforcement guidelines to avoid misuse and overreach. An independent review body would assist in ensuring that enforcement is done within the guidelines of the law. There should further be a public interest exemption, where the eSafety Commissioner may not

remove abhorrent violent content if it is of public interest. This would prevent the use of removal notices to frame certain narratives and keep information from the public, having a chilling effect on open discourse.

16. What more could be done to promote the safety of Australians online, including through research, educational resources and awareness raising?

Increased research, comprehensive educational programs, and public awareness campaigns could enhance online safety.

2.3 Part 4 – Penalties, and investigation and information gathering powers

17. Does the Act need stronger investigation, information gathering and enforcement powers?

No, the current powers are sufficient; however, their application needs stricter controls and clearer guidelines to prevent overreach.

18. Are Australia's penalties adequate and if not, what forms should they take?

The penalties are generally adequate, but their effectiveness depends on consistent and fair enforcement. Higher penalties should not be the focus; rather, ensuring current penalties are applied appropriately, without bias and in line with legislation is key.

19. What more could be done to enforce action against service providers who do not comply, especially those based overseas?

The Act should explicitly state that the eSafety Commissioner cannot enforce the removal of content that is provided to users in other countries. This limitation should be clearly outlined to prevent wasted resources on actions that are beyond Australian jurisdiction. Strengthening international cooperation and agreements could enhance the enforcement of actions against non-compliant overseas providers within the limits of what can be realistically enforced.

20. Should the Commissioner have powers to impose sanctions such as business disruption sanctions?

No, such powers could lead to overreach and unintended negative consequences. Focus should remain on existing penalties and enforcement measures. New, overzealous penalty schemes could make Australia an unattractive environment in which to conduct business.

2.4 Part 5 – International approaches to address online harms

21. Should the Act incorporate any of the international approaches identified above? If so, what should this look like?

Incorporating best practices from international approaches, such as the UK's robust support mechanisms for victims, could enhance the Act's effectiveness. However, the act should not incorporate the duty of care or other complex arrangements found in the UK legislation. The act should adopt an independent ombudsperson or other independent review scheme.

22. Should Australia place additional statutory duties on online services to make online services safer and minimise online harms?

No, Australia should not place additional statutory duties on online services. Introducing such duties could lead to significant additional costs for platforms, which may be passed on to consumers or result in reduced services. Furthermore, these duties could be weaponised by the eSafety Commissioner or various interest groups to suppress dissenting opinions or target specific individuals or groups unfairly, thereby stifling free speech and legitimate discourse.

Online platforms often act merely as intermediaries, hosting user-generated content without directly controlling it. Placing additional regulatory burdens on these platforms holds them unfairly accountable for content they do not create or control. Instead, local law enforcement should be empowered and adequately resourced to take action against individuals who post harmful content in violation of existing state legislation. This approach ensures that those directly responsible for online harms are held accountable, rather than penalising platforms for the actions of their users.

Clear delineation of responsibilities between the eSafety Commissioner and law enforcement will help the eSafety Commissioner remain focused and maintain a balanced approach that protects users without overburdening platforms or infringing on free speech.

23. Is the current level of transparency around decision-making by industry and the Commissioner appropriate? If not, what improvements are needed?

Transparency levels need improvement. Public reporting on the basis for Commissioner decisions and outcomes can ensure accountability and build public trust.

24. Should there be a mechanism in place to provide researchers and eSafety with access to data? Are there other things they should be allowed access to?

Controlled access to data for the public generally and eSafety could improve transparency, understanding of online harms and inform better policies, provided there are strict privacy safeguards to anonymise information Any privacy safeguards need to stop short of going so far as to remove the reasoning used by eSafety in reaching a decision as discussed in section 4.3.

25. To what extent do industry's current dispute resolution processes support Australians to have a safe online experience? Is an alternative dispute resolution mechanism such as an Ombuds scheme required? If so, how should the roles of the Ombuds and Commissioner interact?

Current processes are insufficient. An Ombuds scheme could provide an independent, fair resolution mechanism, complementing the Commissioner's role. The Commissioner should be bound by the independent Ombudsperson decisions.

26. Are additional safeguards needed to ensure the Act upholds fundamental human rights and supporting principles?

Yes, incorporating explicit safeguards for privacy, freedom of speech, and due process is essential to balance regulation with fundamental rights. Specifically, the Act should prevent the Commissioner from infringing the right of free speech unless the statutory threshold is met. This means prohibiting the Commissioner from acting against content which does not meet the threshold. As discussed elsewhere it also means that the Commissioner should be prohibited from acting against content that is in the public interest.

2.5 Part 6 – Regulating the online environment, technology and environmental changes

27. Should the Commissioner have powers to act against content targeting groups as well as individuals? What type of content would be regulated and how would this interact with the adult cyber-abuse and cyberbullying schemes?

No. This does not appear to be required. If implemented, exceedingly clear guidelines must be put in place to prevent such powers being weaponised against freedom of speech as seen in the Billboard Chris case as outlined elsewhere in this submission.

28. What considerations are important in balancing innovation, privacy, security, and safety?

Ensuring regulatory measures do not stifle innovation or infringe on privacy rights is crucial. Transparent, balanced policies that prioritise the open nature of decision making over user safety without compromising other values are needed.

29. Should the Act address risks raised by specific technologies or remain technology neutral? How would the introduction of a statutory duty of care or Safety by Design obligations change your response?

The Act should remain technology neutral to stay adaptable. Introducing a statutory duty of care or Safety by Design obligations should be avoided.

30. To what extent is the Act achieving its object of improving and promoting online safety for Australians?

It is questionable whether the objective has been achieved, and certainly questionable as to whether the objective has been achieved in a cost-effective manner. Vast amounts of money have been spent fielding complaints, many of which do not meet the statutory threshold. Specifically, cyber-bullying of children has seen over 7,000 complaints received by the Commissioner between July 2015 and 30 September 2023, yet only 13 end-user notices requiring removal of cyberbullying material were issued[1]. Similarly for adult cyber abuse, although 710 complaints were received from 1 July 2023 to 30 September 2032, only 6% of those were assessed as meeting the statutory threshold for adult cyber abuse[2]. It is clear that, at least in some form, the eSafety Commissioner is spending a significant portion of time dealing with complaints that do not meet the statutory threshold, and this needs to be resolved.

31. What features of the Act are working well, or should be expanded?

Although it is clear that in some ways the Act has allowed eSafety to remove harmful content such as CSEM, the cost to do so is astronomical. eSafety spend $53.7 million in financial year 2022-2023. In this time only 76 referrals for CSEM were made to the Australian Federal Police. In addition, only 14,975 notifications were sent to INHOPE. This demonstrates, although an effective outcome was achieved, the cost to provide the outcome is significant. It is important that, rather than simply expanding the Act, the existing framework is fine-tuned to ensure it can be more effective at a lower cost. This could be accomplished by explicitly limiting the eSafety Commissioner from acting where the statutory threshold is not met for things such as cyberbullying and allowing a more detailed focus on things such as CSEM.

[1] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.
[2] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.

32. Does Australia have the appropriate governance structures in place to administer Australia's online safety laws?

The current governance structures are lacking sufficient oversight. In particular the lack of an independent review of the eSafety Commissioners decisions and activities.

33. Should Australia consider introducing a cost recovery mechanism on online service providers for regulating online safety functions? If so, what could this look like?

No, introducing a cost recovery mechanism could place undue financial burdens on service providers, particularly smaller ones, and may lead to platforms deciding that Australia is a small and not economically viable market.

3.0 The overarching objects in section 3 of the Act

The overarching objects of the *Online Safety Act 2021* (Cth) are fundamentally sound, aiming to improve and promote online safety for Australians. These objectives align with the government's intent to create a safer online environment, addressing both the prevention of harm and the promotion of safe practices. However, recent actions by the eSafety Office have raised concerns about overreach and misalignment with these core objectives.

The eSafety Office has increasingly engaged in actions that would seem to extend beyond the intended scope of the Act. While the Act focuses on protecting Australians from harmful online content, the Office has taken up cases that involve the removal of content that may offend individuals but does not necessarily meet the statutory threshold of harm outlined in the Act. This indicates an overuse of its powers, addressing content that does not meet the legal standards set by the Act.

Such actions not only strain resources but also deviate from the primary objectives of improving and promoting online safety. The focus should remain on addressing severe and harmful content rather than engaging in selective enforcement that targets material based on subjective criteria of offensiveness. This overreach can lead to a slippery slope where the Office's actions are perceived as policing speech and stifling free expression, rather than protecting users from genuine harm.

The emphasis on removing content that might offend rather than content that is demonstrably harmful undermines the credibility of the regulatory framework. The eSafety Office's actions should be firmly grounded in the statutory objectives, ensuring that interventions are justified, necessary, and proportionate. By aligning its activities more closely with the core goals of the Act, the eSafety Office can better fulfill its mandate without overstepping its regulatory boundaries.

4.0 Effectiveness of Statutory Schemes

4.1 Cyber-bullying Material Targeted at an Australian Child

The statutory scheme addressing cyber-bullying material targeted at Australian children is designed to protect young individuals from online harassment and bullying. Despite significant funding and resources allocated to the eSafety Commissioner's office, the incidence of cyberbullying among children has not significantly decreased. Reports indicate a troubling rise in cases, with a notable increase in serious cyberbullying incidents among younger children, which have tripled since 2019[3].

The current statutory scheme's reactive approach has proven insufficient in addressing the persistent issue of cyberbullying. The eSafety Commissioner's office typically intervenes post-incident, which does little to prevent initial occurrences. For instance, the Australian Institute of Health and Welfare reports that 1 out of 5 of young people aged 15–19 experienced cyberbullying in the past 12 months, indicating a need for more proactive measures[4]. This indicates that, despite the efforts of the eSafety Commissioner, cyberbullying in Australia is occurring at a higher level than in Europe[5].

Evidence suggests that cyberbullying can have severe mental health implications, with many incidents going unnoticed by adults capable of intervening[6]. The reactive nature of current interventions fails to provide a robust safeguard against these harms.

Internationally, countries have implemented various strategies to combat cyberbullying, emphasising the importance of preventive measures and comprehensive support systems. For example, in Denmark, the Health Behaviour in School-aged Children (HBSC) study found that, despite levels bullying generally remaining steady, cyberbullying has increased significantly, prompting the need for educational programs that promote digital literacy and empathy among students[7].

Denmark has adopted a proactive approach to cyberbullying through comprehensive educational programs that emphasise digital literacy and empathy among students. These initiatives are integrated into the school curriculum, aiming to equip children with the skills to manage online interactions responsibly and foster a supportive school environment to prevent bullying before it starts. Similarly, the United States (US) has developed resources and tools through platforms like StopBullying.gov, which provide guidelines for parents, educators, and policymakers on preventing and responding to cyberbullying[8].

---

[3] eSafety Commissioner. "Protecting our (increasingly younger) children from cyberbullying." Accessed June 18, 2024. https://www.esafety.gov.au/newsroom/media-releases/protecting-our-increasingly-younger-children-from-cyberbullying.

[4] Australian Institute of Health and Welfare. "Australia's youth: Bullying and negative online experiences." Accessed June 18, 2024. https://www.aihw.gov.au/reports/children-youth/negative-online-experiences.

[5] World Health Organization. "One in six school-aged children experiences cyberbullying, finds new WHO/Europe study." Accessed June 18, 2024. https://www.who.int/europe/news/item/27-03-2024-one-in-six-school-aged-children-experiences-cyberbullying-finds-new-who-europe-study.

[6] eSafety Commissioner. "Strength in numbers to stop cyberbullying." Accessed June 18, 2024. https://www.esafety.gov.au/newsroom/blog-posts/strength-in-numbers-to-stop-cyberbullying.

[7] World Health Organization. "One in six school-aged children experiences cyberbullying, finds new WHO/Europe study." Accessed June 18, 2024. https://www.who.int/europe/news/item/27-03-2024-one-in-six-school-aged-children-experiences-cyberbullying-finds-new-who-europe-study.

[8] StopBullying.gov. "Cyberbullying Guide." Accessed June 18, 2024. https://www.stopbullying.gov/sites/default/files/documents/Cyberbullying%20Guide%20Final%20508.pdf.

In contrast, the eSafety Commissioner in Australia has been criticised for a predominantly reactive approach[9]. The office typically intervenes only after incidents have occurred, focusing on the removal of harmful content and providing support to victims' post-incident. Despite significant funding and resources, there has been an increase in cyberbullying cases, particularly among younger children, indicating that current measures are not effectively preventing initial occurrences of cyberbullying. Critics argue that this reactive model burdens victims to report content after harm has occurred, rather than implementing preventive strategies to mitigate risks beforehand.

Recommendations:

1. *Emphasise ongoing education and awareness to prevent instances of cyberbullying.*
2. *Develop comprehensive resources and tools for parents to monitor and guide their children's online activities.*
3. *Encourage active parental involvement in setting boundaries and monitoring technology use, as highlighted by educational professionals and law enforcement*
4. *Promote community-wide efforts to support students, involving educators, law enforcement, and community leaders to create a safe digital environment.*
5. *Ensure schools have clear policies and guidelines to address cyberbullying proactively.*
6. *Encourage collaboration between schools, parents, and the community to build a robust support system for students.*

[9] Information Law and Policy Centre. "Regulating Online Safety: Lessons from Australia." Accessed June 18, 2024. https://infolawcentre.blogs.sas.ac.uk/2022/07/19/regulating-online-safety-lessons-from-australia/.

4.2 Non-consensual Sharing of Intimate Images

The statutory scheme addressing the non-consensual sharing of intimate images, commonly known as "revenge porn," in Australia is somewhat unique. Criminal offences exist under state and territory legislation. These laws criminalise the act and provide for prosecution and penalties. However, the *Online Safety Act 2021* (Cth) complements these efforts by empowering the eSafety Commissioner to issue removal notices and civil penalties for non-compliance, but it does not establish criminal offences for these actions. This approach contrasts with international practices, particularly in the United Kingdom (UK) and the US, where more comprehensive legal frameworks exist.

The UK has implemented robust measures under the *Criminal Justice and Courts Act 2015* (UK), which criminalises the non-consensual sharing of intimate images. Recent amendments have broadened the scope to include deepfakes and other digitally altered images. The UK's approach includes severe penalties but also more proactive support mechanisms like the Revenge Porn Helpline, which assists victims in removing content and provides emotional support. The UK government has also integrated these protections into broader legislation, ensuring that the law keeps pace with technological advancements[10].

In the US, both federal and state laws address the non-consensual sharing of intimate images. Many states have specific laws criminalising this behaviour, and platforms like Facebook and Google have strict policies to quickly remove such content. The StopBullying.gov initiative and similar programs offer resources and support for victims, emphasising the importance of preventive measures and rapid response.

In Australia, the *Online Safety Act 2021* (Cth) enables the eSafety Commissioner to act to remove non-consensual intimate images from online platforms. However, the criminal prosecution of offenders remains within the jurisdiction of state laws, such as the *Crimes Act 1900* (NSW). This division of responsibilities can lead to gaps in enforcement and support for victims, who often face delays in content removal and inconsistent responses from online platforms and law enforcement agencies.

Recently, the was proposed by Federal Attorney General, Mark Dreyfus. This legislation would criminalise the non-consensual sharing of intimate images at a federal level.

Attorney-General Mark Dreyfus recently introduced the *Criminal Code Amendment (Deepfake Sexual Material) Bill 2024* (Cth)[11], aiming to address the non-consensual sharing of deepfake porn. Despite its introduction, existing federal laws, such as sections 474.22 and 474.17 of the *Criminal Code Act 1995* (Cth), already criminalise actions related to child abuse material and using a carriage service to menace, harass, or offend. State laws, including section 91H of the *Crimes Act 1900* (NSW) and sections 51C and 53S of the *Crimes Act 1958* (Vic), also cover these offences, highlighting the complexity and potential overlap of legal frameworks.

The new bill proposes new offences with penalties of up to seven years in prison for sharing digitally created sexually explicit material without consent, aiming to address the inadequacies in current laws.

[10] GOV.UK. "New Laws to Better Protect Victims from Abuse of Intimate Images." Accessed June 18, 2024. https://www.gov.uk/government/news/new-laws-to-better-protect-victims-from-abuse-of-intimate-images.
[11] Parliament of Australia. "Criminal Code Amendment (Deepfake Sexual Material) Bill 2024." Accessed June 18, 2024. https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=r7205.

However, critics argue that existing laws already cover much of this conduct, making the new legislation seem redundant[12].

Recommendations

7. *Establish clearer and more stringent enforcement protocols to ensure the rapid removal of non-consensual intimate images.*
8. *Develop comprehensive support services for victims, including legal assistance, counselling, and dedicated helplines.*
9. *Improve coordination between federal and state laws to avoid redundancy and ensure that existing legal provisions are effectively enforced.*
10. *Encourage federal legislation to focus on support mechanisms and enforcement protocols rather than creating overlapping criminal offences.*

[12] Sydney Criminal Lawyers. "The Proposed Federal Offence of Sharing Deepfake Adult Porn in Australia." Accessed June 18, 2024. https://www.sydneycriminallawyers.com.au/blog/the-proposed-federal-offence-of-sharing-deepfake-adult-porn-in-australia/

4.3 Cyber-abuse Material Targeted at an Australian Adult

Although cyber-abuse targeting adults remains an issue, the number of reports to the eSafety Commissioner is low. In the period from 1 July 2023 to 30 September 2023 the eSafety Commissioner received only 710 complaints about cyber-abuse material targeted at an Australian adult[13] Of those only 6%, or 43 complaints, met the threshold for cyber abuse[14] Despite this, eSafety made requests in 103 cases for material to be removed that was not in violation of the law but only in violation of platform terms of service[15]. This could represent a misuse of resources, given that the Commissioner is therefore spending time making requests to various parties for removal of content which it has already identified does not reach the statutory threshold.

Unlike children, adults are generally expected to have more resilience and the ability to manage their online interactions. Indeed, adults should be capable of handling non-violent commentary and criticism online, distinguishing between harmful abuse and mere disagreements or harsh opinions. However, the current approach by the eSafety Commissioner often blurs the line between protecting individuals from genuine abuse and infringing on free speech. It could be argued that, at times, the eSafety Commissioner has overreached, issuing takedown notices for content that may not strictly fall under the intended scope but instead falls into moral policing. moral or subjective categories. This is evident in cases where the Commissioner has acted against material that does not meet the statutory threshold but chooses to act against that content regardless. Additionally, there are situations where the eSafety Commissioner has issued removal notices where respondents and third parties have argued that the content was not cyber abuse material but instead was legitimate commentary or opinion that a member of the public chose to take offence to. Importantly, the fact that a person was offended or upset by content does not mean that the content should be considered offensive.

There is a crucial balance to be maintained between protecting individuals from harm and preserving the broader public's right to free speech and anonymity online. This balance is essential to maintaining a free and open internet while ensuring that genuinely abusive behaviours are effectively curbed.

A notable example of alleged overreach involves twitter user Chris Elston, known as "Billboard Chris," who received a takedown notice from the eSafety Commissioner for a post criticising a transgender activist appointed by the World Health Organization[16]. The post was considered offensive by the eSafety Commissioner, leading to its removal and a subsequent legal challenge by X to protect its users' right to free speech. This incident highlights concerns that the eSafety Commissioner's actions can sometimes stifle legitimate political discourse and criticism under the guise of protecting against cyber-abuse. Unhelpfully, the eSafety Commissioner's office has not released the rationale behind the decision that the material was intended to cause serious harm[17]. Instead, the response to FOI request redacted approximately one and a quarter page of this rationale under section 47E(d) and 47F of the *Freedom of*

[13] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.
[14] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.
[15] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.
[16] Ground News. "X to Take Legal Action Against eSafety Commission Over WHO Trans Activist Post." Accessed June 15, 2024. https://ground.news/article/x-to-take-legal-action-against-esafety-commission-over-who-trans-activist-post.
[17] eSafety Commissioner. "Log 55: Documents Relating to Cyber-Abuse Notice." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-06/LOG_55_Document_Set_0.pdf. p27-28

Information Act 1982 (Cth)[18]. This is peculiar because of the large redaction, in part because section 47E(d) allows exemptions due to "substantial adverse effect on the proper and efficient conduct of the operations of an agency". Without transparency as to the decision making on this type of decision, it leaves room for questions as to whether decisions are made in compliance with the law, or influenced by subjective judgments. X has confirmed it intends to file a legal challenge to this order in aid of protecting free speech[19], further increasing the cost to taxpayers. There is further suggestion that the eSafety Commissioner has issued similar removal notices against those who have spoken out against transgender ideologies, including a 2023 order that expressed views on biological men breastfeeding[20].

Recommendations:

11. *Amend the Act to prevent the eSafety Commissioner from acting against content which does not meet the statutory threshold for adult cyber abuse.*
12. *Publish detailed reports on takedown notices for transparency.*
13. *Amend legislation to distinguish abuse from non-violent criticism.*
14. *Establish consistent content moderation guidelines with platforms.*
15. *Expand mental health support services for victims, including people who feel victimised regardless of intent.*
16. *Conduct regular independent reviews of eSafety Commissioner's decisions.*
17. *Launch educational campaigns to improve online resilience and literacy.*

[18] eSafety Commissioner. "Log 55: Documents Relating to Cyber-Abuse Notice." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-06/LOG_55_Document_Set_0.pdf. p27-28

[19] Rex Widerstrom, "X to Take Legal Action Against eSafety Commission Over Post About Trans Activist at WHO," The Epoch Times, March 31, 2024, https://www.theepochtimes.com/world/x-to-take-legal-action-against-esafety-commission-over-who-trans-activist-post-5617697?utm_source=ground.news&utm_medium=referral.

[20] Rachel Baxendale. "eSafety Commissioner's War on 'Misgendering' Tweet." The Australian, May 2, 2024. Accessed May 3, 2024. https://www.theaustralian.com.au/nation/esafety-commissioners-war-on-misgendering-tweet/news-story/92a74e73508e0443471a945811c2111f?amp&nk=712b333adcbe64316a94c2681a3bd3ac-1714706939.

4.4 Online Content Scheme

The Online Content Scheme aims to regulate harmful online content, including material depicting abhorrent violent conduct. However, the effectiveness of this scheme has been questioned, particularly in light of high-profile incidents where harmful content remained accessible despite regulatory efforts. For example, the case involving the Bishop Mar Mari Emmanuel stabbing video highlighted significant challenges in enforcing content removal and ensuring public safety[21].

The scheme grants the eSafety Commissioner the power to issue removal notices and take action against illegal and offensive content. However, a fundamental problem is the extraterritorial nature of much online content. Many harmful materials are hosted on servers outside Australia, beyond the jurisdiction of the eSafety Commissioner. This creates significant enforcement challenges, as the eSafety Commissioner's powers do not extend internationally. For instance, the eSafety Commissioner issued a takedown notice to X (formerly Twitter) to remove footage of the Sydney church stabbing, but X only geoblocked the material for Australian users. The eSafety Commissioner expressed dissatisfaction with this outcome, pushing for the content to be removed globally. It would seem the eSafety Commissioner decided to explore “novel regulatory powers … to protect Australians from online harm,”[22] in pushing for global takedown of content beyond Australian jurisdiction. This approach not only oversteps the Commissioner's legal boundaries but also sets a concerning precedent for international internet governance, potentially undermining global standards for freedom of expression and digital rights. The attempts to enforce such extraterritorial measures have faced significant pushback from both domestic and international observers[23], highlighting the need for a more balanced and legally sound approach to regulating online content.

It is crucial for the eSafety Commissioner to recognise these jurisdictional limitations and refrain from overstepping international boundaries. There is also the concern that such attempts to enforce content removal internationally could set a dangerous precedent. Free speech advocates argue that issuing global takedown orders not only oversteps jurisdictional boundaries but also threatens free speech everywhere. These actions can potentially legitimise similar practices by authoritarian regimes, thus undermining global standards for freedom of expression. These efforts have drawn international criticism, potentially harming Australia's image as a defender of free speech and digital rights. Such actions could be perceived as overreach and could lead to negative perceptions about Australia's commitment to upholding democratic principles and freedoms. There are broad concerns from industry participants that the eSafety Commissioner is growing or enabling an environment of mass surveillance, risking security and privacy of Australians[24].

---

[21] 9News. "Australia drops case against Elon Musk's X over church stabbing videos,", Accessed June 18, 2024, https://www.9news.com.au/national/esafety-commissioner-australia-drops-case-against-elon-musk-x-over-church-stabbing-videos.

[22] AOL. "Australia drops legal fight against X over church stabbing video." Accessed June 18, 2024. https://www.aol.com.

[23] Ange Lavoipierre. "Federal Court Chooses Not to Extend Temporary Order Blocking Terrorist Attack Vision on Social Media Platform X." ABC News. Accessed June 18, 2024. https://www.abc.net.au/news/2024-05-13/court-chooses-to-end-ban-on-wakeley-stabbing-video-on-x-twitter/103829790.

[24] Digital Rights Watch. "Local and international organisations urge Australia’s eSafety Commissioner against requiring the tech industry to scan users’ personal files and messages." Accessed June 18, 2024. https://digitalrightswatch.org.au/2023/12/20/esafety-joint-letter/.; Trevor Long. "Mass Surveillance Warning from Apple Over New Proposed eSafety Guidelines." 9News. Accessed June 18, 2024. https://www.9news.com.au/technology/apple-news-mass-surveillance-warning-from-apple-over-new-proposed-esafety-guidelines/6206064f-f385-4015-892d-7e3cd6530ea2.

Recommendations:

18. *Ensure that there are adequate appeals mechanisms and procedural fairness in the issuance of removal and blocking notices.*
19. *Amend the Act so that it is clear that the eSafety Commissioner does not have jurisdiction to remove content for persons outside Australia.*
20. *Establish safeguards to prevent the misuse of takedown notices that could infringe on freedom of speech and political communication.*
21. *Increase transparency in the decision-making process for content removal, ensuring public trust and accountability.*
22. *Address concerns regarding mass surveillance and ensure that measures taken to protect online safety do not compromise the privacy and security of Australians.*
23. *Clearly define the limits of the eSafety Commissioner's powers.*

.

4.5 Material Depicting Abhorrent Violent Conduct

The statutory scheme targeting material depicting abhorrent violent conduct has faced significant challenges in enforcement. The eSafety Commissioner's legal battle with X over the removal of the Bishop Mar Mari Emmanuel stabbing video underscores the difficulties in managing such content. The case revealed gaps in the regulatory framework and public dissatisfaction with the Commissioner's handling of the situation[25].

A fundamental issue is the inconsistency in enforcement. The eSafety Commissioner's aggressive pursuit of the takedown of the Mar Mari Emmanuel video stands in stark contrast to the numerous other violent videos that remain accessible online. This selective enforcement raises questions about the criteria used and the effectiveness of current measures.

Critics have pointed out that while the Commissioner pursued legal action against X (formerly Twitter), for this specific content, violent content still proliferates across the internet. This selective enforcement has led to public dissatisfaction and questions about the consistency and criteria used by the Commissioner's office in deciding which content to target for removal (AMI) (eSafety Commissioner).

Given the intent to censor the video globally, the Commissioner's actions were seen as potentially setting a dangerous precedent for global censorship, complicating the balance between protecting the community from harmful content and upholding principles of free speech. The Commissioner noted that Facebook and other major platforms complied with the removal requests, but X did not. When combined with the fact that the eSafety Commissioner is testing "novel powers"[26], it could be considered that they are overreaching and putting pressure on social media platforms to remove content that even the eSafety Commissioner questions the validity of. This raises concerns about the further potential for overreach and the uneven application of regulatory authority.

Although the act allows for the removal of abhorrent violent conduct with a very broad definition, the selective removal of content that fits certain narratives, while allowing other violent content to remain online, undermines the credibility and effectiveness of the regulatory framework. This inconsistency calls for a more balanced and clear approach to content removal, ensuring that all forms of harmful material are addressed uniformly.

In this specific circumstance, although the content certainly depicts violence, there is a level of public interest that dictates the content should remain available. This interest is underscored by the need for transparency in reporting significant events and ensuring public awareness of violent incidents that have broader societal implications. Furthermore, the selective enforcement actions by the eSafety Commissioner raise questions about the consistency and fairness of content removal practices, suggesting a need for a more balanced approach that considers both the potential harms and the public's right to access information. The act could and should prohibit the eSafety Commissioner issuing a removal notice where the content is clearly in the public interest.

---

[25] JURIST. "Social Media Platform X to Challenge Government Order Demanding Removal of Sydney Church Stabbing Posts." Accessed June 18, 2024. https://www.jurist.org/news/2024/04/social-media-platform-x-to-challenge-government-order-demanding-removal-of-sydney-church-stabbing-posts/.; 9News. "Federal Court Overturns eSafety Commissioner's Injunction Against X Over Church Stabbing Video." Accessed June 18, 2024. https://www.9news.com.au/national/federal-court-overturns-esafety-commissioners-injunction-against-x-over-church-stabbing-video/.

[26] eSafety Commissioner. "Statement from the eSafety Commissioner re: Federal Court Proceedings." June 5, 2024. Accessed June 21, 2024. https://www.esafety.gov.au/newsroom/media-releases/statement-from-the-esafety-commissioner-re-federal-court-proceedings.

To ensure the effectiveness of the statutory scheme, it is essential to refine the enforcement mechanisms and establish clear, consistent criteria for content removal. This approach will help mitigate the perception of selective enforcement and ensure a more comprehensive protection for the Australian community from online harms.

Recommendations:

24. *Update the Online Safety Act 2021 (Cth) to have a public interest exclusion, whereby the eSafety Office cannot force removal of content that is of public interest.*
25. *Establish clear, consistent criteria for content removal to ensure uniform enforcement across all platforms.*
26. *Implement a balanced approach that considers both the potential harms and the public's right to access information.*
27. *Introduce regular audits of enforcement actions to identify and address inconsistencies – the content either meets the hurdle or it does not.*

5.0 The operation and effectiveness of the Basic Online Safety Expectations regime in the Act.

The Basic Online Safety Expectations (BOSE) under the *Online Safety Act 2021* (Cth) aim to set minimum safety benchmarks for online service providers to protect the Australian community from harmful content. These expectations include steps to ensure user safety, minimise harmful material, and provide mechanisms for reporting and complaints.

Recent incidents highlight challenges in enforcing these expectations effectively. A notable example involves Instagram's algorithm, which has been found to recommend sexually explicit content to accounts of minors as young as 13 years old. Tests conducted by The Wall Street Journal and researchers from Northeastern University revealed that Instagram's recommendation systems could direct young users to inappropriate content, despite measures intended to protect them[27]. This situation underscores the need for more stringent oversight and better implementation of safety mechanisms by social media platforms[28].

These incidents demonstrate the need for amendments to the BOSE. The reforms proposed provide a good base, ensuring that generative artificial intelligence capabilities and recommender systems are designed with user safety in mind, particularly to prevent the amplification of harmful content to minors.

[27] Jeff, Horowitz. 2024. “Instagram Recommends Sexual Videos to Accounts for 13-Year-Olds, Tests Show.” Wall Street Journal. June 20, 2024. Accessed June 21, 2024. https://www.wsj.com/tech/instagram-recommends-sexual-videos-to-accounts-for-13-year-olds-tests-show-b6123c65.

[28] Engadget. "Facebook and Instagram's Algorithms Facilitated Child Sexual Harassment, State Lawsuit Claims." January 18, 2024. Accessed June 21, 2024. https://www.engadget.com/facebook-instagram-algorithms-child-sexual-harassment-lawsuit-2024.

6.0 Whether additional arrangements are warranted

Given the current issues with overreach, it is not the right time to expand the powers of the eSafety Commissioner. The recent legal battle between the eSafety Commissioner and X (formerly Twitter) over the removal of the Mar Mari Emmanuel video and the Billboard Chris removal notice both highlight the potential for overreach and the complexities involved in regulating global platforms from a single national jurisdiction. Legal experts have pointed out that these cases underscore the need for a balanced approach to online regulation that does not unduly infringe on free speech or extend beyond national borders[29].

Expanding the eSafety Commissioner's powers to address additional online harms, such as online hate, volumetric attacks, and technology-facilitated abuse, without first addressing these overreach concerns, could lead to unintended consequences. For example, new regulations could result in overly zealous removal requests from the Commissioner, or overly cautious content removal practices by platforms, potentially stifling legitimate discourse and important societal discussions. This concern is particularly relevant for social media, which plays a critical role in political expression and advocacy.

There are significant privacy and security risks associated with implementing broad regulatory measures. The case against age verification systems, which were proposed but ultimately not enforced due to their invasive nature and potential to exacerbate online harms, exemplifies these risks. Such measures often require the collection of extensive personal data, which can create new vulnerabilities and privacy issues[30].

Online hate is a vaguely defined area subject to much discussion. The subjective nature of what constitutes "hate" makes it far too open to abuse or misuse. Labelling content as hateful can easily be weaponised to silence unpopular opinions or dissenting voices, thereby infringing on free speech. Providing more and vaguer methods for the eSafety Commissioner to decide that content warrants removal poses significant risks of overreach and requires careful consideration and precise guidelines to avoid misuse.

Although volumetric attacks can be challenging, each piece of content should be dealt with on its own merit. Grouping such incidents under a broad regulatory framework risks ignoring the context and specifics of each case. Overarching regulations could lead to blanket actions that fail to distinguish between genuinely harmful behaviour and non-malicious actions, such as the public voicing their dissatisfaction with a public figure. Each case must be assessed individually to ensure that actions taken are proportionate and justified.

The inclusion of technology-facilitated gender-based violence into the legislation is concerning, especially given the removal notice issued for Billboard Chris. If such provisions are outright included in legislation, it raises significant concerns about what the eSafety Commissioner might do under these expanded powers. The case of Billboard Chris demonstrates how actions taken under the guise of preventing gender-based violence can potentially infringe on free speech and political discourse, especially when critical commentary is involved.

---

[29] B&T. "Elon Musk vs eSafety: Legal Experts Warn That 'Rogue Operators' Like X Are Unlikely To Win Federal Court Battle." Accessed June 21, 2024. https://www.bandt.com.au/elon-musk-vs-esafety-legal-experts-warn-that-rogue-operators-like-x-are-unlikely-to-win-federal-court-battle/.; The Australian Independent Media Network. "Balancing eSafety and Online Censorship, 2024." Accessed June 21, 2024. https://theaimn.com/balancing-esafety-and-online-censorship-2024.

[30] Digital Rights Watch. "Campaign Win: Australian Government Will Not Force Sites to Implement Age Verification." Accessed June 21, 2024. https://digitalrightswatch.org.au/campaign-win-australian-government-will-not-force-sites-to-implement-age-verification.

The provision addressing online abuse of public figures is also fraught with potential for misuse, particularly in silencing dissent against public officials. There is a concern that such measures could be used to shield public figures from legitimate criticism and public accountability. Public figures have access to the same provisions as others, although by nature of their role, they should expect greater public interaction which may include contrary views, frustration, or criticism. This heightened level of interaction is part and parcel of their public position and is essential for democratic discourse. However, regulating such interactions under the guise of preventing online abuse could lead to the suppression of legitimate criticism and dissenting opinions, which are vital for holding public figures accountable and fostering open debate.

While addressing online harms is crucial, expanding the eSafety Commissioner's powers without resolving existing issues of overreach and ensuring robust safeguards could lead to more harm than good. A more measured approach, focusing on transparency, accountability, and collaboration with global platforms, is necessary to effectively protect users without compromising fundamental freedoms.

Recommendations

28. That no new powers for the eSafety Commissioner be implemented until the issues of overreach have been resolved.

7.0 Whether the regulatory arrangements, tools and powers available to the Commissioner should be amended

The introduction of a duty of care requirement towards users, akin to the United Kingdom's *Online Safety Act 2023* or Australia's work health and safety legislation, is a significant point of consideration in this review. However, there are critical concerns regarding the appropriateness and potential consequences of such a regulatory measure.

The core argument against imposing a duty of care on online platforms is that these platforms primarily function as meeting places where users post content. Imposing a regulatory burden on platforms for user-generated content effectively holds them accountable for actions they do not control. This approach could lead to platforms being overly cautious, potentially stifling free speech and legitimate discourse to avoid regulatory penalties. This issue has been highlighted in discussions about the systemic duty of care (SDOC) model, where proactive content moderation could expose platforms to new liabilities and unintended consequences[31].

In the UK, the implementation of multiple 'duties of care' has introduced incredible complexity to the regulatory framework[32]. The UK's approach necessitates distinguishing between different types of content and associating specific duties with each type. This complexity shifts the focus from systemic regulation to a content-first approach, which can dilute the overall effectiveness of the regulatory framework. Additionally, it complicates compliance for platforms, potentially diverting resources from more impactful safety measures, adding cost and risk to the platforms business model.

Existing state and federal laws already address many of the issues that the duty of care seeks to mitigate, such as hate speech, harassment, and child protection. These laws provide a framework for prosecuting individuals who engage in illegal activities online without placing undue burdens on the platforms that host the content.

While the introduction of a duty of care requirement towards users may theoretically enhance user safety, it is essential to approach its implementation with extreme caution. There is a risk of the duty of care being weaponised, where the Commissioner, individuals or groups might exploit these regulations to silence dissenting opinions or target specific platforms, content creators or narratives unfairly. It is crucial to focus on transparency and accountability for the regulator in order to create effective protections without overburdening the platforms themselves and preventing weaponisation of any duty of care provisions. Ensuring robust safeguards against overreach and maintaining a balanced approach will be vital for the success of any new regulatory measures.

[31] Center for Internet and Society. "Systemic Duties of Care and Intermediary Liability." Accessed June 21, 2024. https://cyberlaw.stanford.edu/blog/2020/05/systemic-duties-care-and-intermediary-liability.

[32] Reset Australia. "A Duty of Care in Australia's Online Safety Act: Policy Briefing." April 2024. Accessed June 21, 2024. https://au.reset.tech/uploads/Duty-of-Care-Report-Reset.Tech.pdf.

Recommendations:

29. *That a duty of care is not implemented, as it places the focus on the wrong party in any wrongdoing.*
30. *If a duty of care is implemented, it ensures robust safeguards against misuse by the eSafety Commissioner.*
31. *If a duty of care is implemented, it ensures regular transparent auditing of enforcement action taken by the eSafety Commissioner*

8.0 Whether penalties should apply to a broader range of circumstances.

There should be no increase in penalties or increase in range of circumstances in which penalties apply. The current penalties and enforcement measures under the *Online Safety Act 2021* (Cth) are already robust, providing a balanced approach to regulation without overreach. The existing framework includes substantial financial penalties for non-compliance, reaching up to $156,500 for individuals and $782,500 for corporations. These penalties serve as a significant deterrent and can escalate daily until compliance is achieved, ensuring that offenders face continued financial consequences. Any suggestion that these penalties are insufficient is dishonest.

The Act currently imposes similar penalties for failing to remove both illegal and harmful but not unlawful content, maintaining a balanced approach that does not disproportionately penalise different types of non-compliance. This parity ensures that all forms of harmful content are addressed seriously without creating excessive punitive measures for specific offences. Introducing higher penalties could disrupt this balance and lead to over-punitive regulations. Combined with the selective enforcement seen in the Billboard Chris and Mar Mari Emmanual cases, where certain content is targeted more aggressively than others, raising penalties might exacerbate issues of fairness and consistency in enforcement and could assist in the weaponisation of the eSafety Commissioner role. This approach could undermine public trust in the regulatory system and lead to accusations of bias or unfair treatment of certain content creators or platforms.

Enforcing penalties on overseas platforms presents a significant challenge due to jurisdictional limitations. While the Act formally extends its enforceability to acts, omissions, matters, and things outside Australia, it cannot control what is published to persons in other jurisdictions. This limitation means that efforts to enforce penalties on platforms based overseas can be impractical and resource-intensive, as seen in the Mar Mari Emmanuel case. To prevent such wasted resources, the Act or its regulations should explicitly address these jurisdictional boundaries. By clearly defining the scope and limits of its applicability, the Act can ensure that enforcement efforts are focused on issues within Australia and relevant to Australians. This promotes more achievable and effective measures, rather than on pursuing actions that are unlikely to succeed due to international legal constraints. Increasing penalties or powers will not effectively address this challenge.

Expanding the powers of the eSafety Commissioner risks regulatory overreach, potentially stifling innovation and free expression. The potential for increased penalties and increased powers to penalise to be weaponised against platforms to silence dissenting voices is a serious concern. Maintaining a balanced regulatory approach that protects users while preserving a free and open internet is crucial to avoid these risks.

Rather than increasing penalties, the focus should be on enhancing the effectiveness of existing tools and measures. Improving transparency, accountability, and collaboration with global platforms can address the challenges of enforcing penalties on overseas entities more sustainably and effectively. Strengthening international cooperation mechanisms is essential for addressing online harm without overburdening the platforms.

Recommendations:

32. That the existing penalty regime be maintained in its current form.

9.0 Whether the current information gathering powers, investigative powers, enforcement powers, civil penalties or disclosure of information provisions should be amended

As discussed elsewhere in this submission, the eSafety Commissioner has taken action against a range of content that did not meet the statutory threshold. Particularly in its actions against content that does not meet the statutory threshold for adult cyber abuse between 1 July 2023 and 30 September 2023. In this period only 46 cases of adult cyberbullying were assessed as meeting the threshold under the *Online Safety Act 2021* (Cth)[33]. However, the eSafety Office made requests to platforms for 103 cases based merely on terms of service breaches[34]. This indicates a significant overuse of its investigative and enforcement powers beyond what is stipulated or intended by law.

This overreach is problematic as it not only burdens platforms with requests that exceed statutory requirements but also potentially infringes on the rights of users. When the eSafety Office acts against content that does not meet the statutory threshold, it undermines the credibility and fairness of the regulatory framework. The current provisions allow the Commissioner to investigate and summon individuals, obtain information, and enforce penalties, but these powers must be exercised within the legal boundaries set by the Act.

To address this, there needs to be stricter controls and clearer guidelines on how these powers are used. The Act should explicitly state the conditions under which the Commissioner can make requests to platforms, ensuring that actions are only taken against content that meets the statutory threshold. This would prevent the misuse of powers and ensure that the eSafety Office operates within its legal mandate.

Enhancing transparency in the Office's decision-making process can help in maintaining public trust. Detailed public reporting on the basis for each request and the outcomes can ensure accountability and provide a check against overreach. Establishing an independent oversight body to review the Office's actions could help in curbing excessive or inappropriate use of its powers.

Recommendations:

33. That the powers of the eSafety Commissioner to make requests to platforms where the content does not meet the statutory threshold be explicitly prohibited.
34. That no further additional powers are given to the eSafety Commissioner.

[33] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.
[34] eSafety Commissioner. "Disclosure Log 49: Documents Relating to Senate Estimates." Accessed June 18, 2024. https://www.esafety.gov.au/sites/default/files/2024-01/LOG_49_Document_Set.pdf. p. 2.

10.0 The Commissioner's functions and governance arrangements

The eSafety Commissioner's roles and responsibilities under the *Online Safety Act* 2021 (Cth) establish the office as an independent statutory authority supported by the Australian Communications and Media Authority (ACMA). This governance structure is designed to ensure that the Commissioner can effectively carry out the mandate of promoting and enhancing online safety for Australians.

There are concerns about the current governance arrangements and whether they are sufficient to allow the Commissioner to fulfill this mandate effectively. While the UK's Ofcom has a Board providing strategic direction and the Ofcom Executive managing day-to-day operations, the eSafety Commissioner operates with a single officeholder supported by ACMA staff. This structure may limit oversight and decision-making capacity compared to multi-member commissions seen in other jurisdictions, such as Canada's proposed Digital Safety Commission and Digital Safety Ombudsperson.

One of the critical issues is whether the current functions and powers in the Act are adequate. The eSafety Commissioner has significant investigative and enforcement powers, including the ability to issue removal notices, summon individuals for questioning, and obtain user information. However, there have been instances of overreach, where actions were taken against content that did not meet the statutory threshold. This misuse of powers not only burdens platforms unnecessarily but also raises concerns about fairness and consistency in enforcement. It is essential to ensure that the Commissioner's actions remain within the legal boundaries to maintain public trust and the effectiveness of the regulatory framework.

Another consideration is the proposal to introduce a cost recovery mechanism, where online service providers would bear the cost of regulation. While this approach is used in the UK and proposed in Canada, it may not be suitable for Australia. Implementing cost recovery could place an additional financial burden on service providers, particularly smaller platforms, stifling innovation and reducing the diversity of online services available to Australians. Moreover, the administrative complexity of managing such a system could outweigh the benefits, diverting resources from more critical regulatory functions.

Instead of expanding the Commissioner's powers or introducing cost recovery, the focus should be on improving the current governance structures. This includes enhancing transparency, accountability, and oversight mechanisms to ensure that the eSafety Commissioner operates effectively and within the legal framework. Establishing an independent oversight body, similar to Canada's proposed Digital Safety Ombudsperson, could provide additional checks and balances, ensuring that the Commissioner's actions are justified, proportionate and within the boundaries of the Act and regulations.

Recommendations:

35. Establish an independent oversight body to review the actions of the eSafety Commissioner and ensure they are justified and within the legal framework.

11.0 Conclusion

The regulatory framework established by the *Online Safety Act 2021* (Cth) aims to create a safer online environment for Australians. However, the application of the Act's provisions by the eSafety Commissioner has highlighted several significant issues that need to be addressed to ensure the effectiveness and fairness of these measures.

The substantial financial expenditure, with $53.7 million spent in the financial year 2022-2023, has not corresponded with a proportionate decrease in online harms. This indicates a need for more proactive and efficient strategies rather than reactive approaches that have proven insufficient.

Furthermore, the eSafety Commissioner's attempts at content censorship raise serious concerns about overreach and the potential infringement on Australians' rights to freedom of speech, especially political communication. These actions have not only led to negative perceptions of Australia internationally but also undermine the nation's commitment to democratic values and free expression. The controversial cases involving Billboard Chris and the Wakeley Church stabbing video exemplify the troubling trend of using regulatory powers to stifle public discourse and political expression.

The lack of transparency in the eSafety Commissioner's decision-making processes further exacerbates these concerns, calling for enhanced accountability and clearer guidelines to prevent misuse of powers. Establishing independent oversight mechanisms could help ensure that the Commissioner's actions are justified, proportionate, and aligned with the statutory objectives of the Act.

Expanding the Commissioner's powers to address additional online harms without first resolving existing issues of overreach could lead to unintended consequences, such as stifling legitimate discourse and infringing on privacy rights. A balanced approach that prioritises transparency, accountability, and collaboration with global platforms is essential to effectively protect users without compromising fundamental freedoms.

Any consideration of introducing a cost recovery mechanism or additional statutory duties on online services must be approached with caution. Such measures could place undue financial burdens on service providers, potentially stifling innovation and reducing the diversity of online services available to Australians.

This submission underscores the need for a more measured and refined approach to online safety regulation, ensuring that the eSafety Commissioner's actions are firmly grounded in the statutory objectives, and aligned with the principles of free speech and democratic accountability.

Appendix A – List of Recommendations

This section lists the recommendations made throughout the document.

Number	Recommendation
1	Emphasise ongoing education and awareness to prevent instances of cyberbullying.
2	Develop comprehensive resources and tools for parents to monitor and guide their children's online activities.
3	Encourage active parental involvement in setting boundaries and monitoring technology use, as highlighted by educational professionals and law enforcement
4	Promote community-wide efforts to support students, involving educators, law enforcement, and community leaders to create a safe digital environment.
5	Ensure schools have clear policies and guidelines to address cyberbullying proactively.
6	Encourage collaboration between schools, parents, and the community to build a robust support system for students.
7	Establish clearer and more stringent enforcement protocols to ensure the rapid removal of non-consensual intimate images.
8	Develop comprehensive support services for victims, including legal assistance, counselling, and dedicated helplines.
9	Improve coordination between federal and state laws to avoid redundancy and ensure that existing legal provisions are effectively enforced.
10	Encourage federal legislation to focus on support mechanisms and enforcement protocols rather than creating overlapping criminal offences
11	Amend the Act to prevent the eSafety Commissioner from acting against content which does not meet the statutory threshold for adult cyber abuse.
12	Publish detailed reports on takedown notices for transparency.
13	Amend legislation to distinguish abuse from non-violent criticism.
14	Establish consistent content moderation guidelines with platforms.
15	Expand mental health support services for victims, including people who feel victimised regardless of intent.
16	Conduct regular independent reviews of eSafety Commissioner’s decisions.
17	Launch educational campaigns to improve online resilience and literacy
18	Ensure that there are adequate appeals mechanisms and procedural fairness in the issuance of removal and blocking notices.
19	Amend the Act so that it is clear that the eSafety Commissioner does not have jurisdiction to remove content for persons outside Australia.
20	Establish safeguards to prevent the misuse of takedown notices that could infringe on freedom of speech and political communication.
21	Increase transparency in the decision-making process for content removal, ensuring

	public trust and accountability.
22	Address concerns regarding mass surveillance and ensure that measures taken to protect online safety do not compromise the privacy and security of Australians.
23	Clearly define the limits of the eSafety Commissioner's powers.
24	Update the Online Safety Act 2021 (Cth) to have a public interest exclusion, whereby the eSafety Office cannot force removal of content that is of public interest.
25	Establish clear, consistent criteria for content removal to ensure uniform enforcement across all platforms.
26	Implement a balanced approach that considers both the potential harms and the public's right to access information.
27	Introduce regular audits of enforcement actions to identify and address inconsistencies – the content either meets the hurdle or it does not.
28	That no new powers for the eSafety Commissioner be implemented until the issues of overreach have been resolved
29	That a duty of care is not implemented, as it places the focus on the wrong party in any wrongdoing.
30	If a duty of care is implemented, it ensures robust safeguards against misuse by the eSafety Commissioner.
31	If a duty of care is implemented, it ensures regular transparent auditing of enforcement action taken by the eSafety Commissioner
32	That the existing penalty regime be maintained in its current form.
33	That the powers of the eSafety Commissioner to make requests to platforms where the content does not meet the statutory threshold be explicitly prohibited.
34	That no further additional powers are given to the eSafety Commissioner.
35	Establish an independent oversight body to review the actions of the eSafety Commissioner and ensure they are justified and within the legal framework.