1 July 2024

AFP Submission to the statutory review of the *Online Safety Act 2021*

AFP

afp.gov.au

Table of contents

Introduction

1. The Australian Federal Police (AFP) welcomes the opportunity to make a submission to assist with Ms Delia Rickard's independent review of the *Online Safety Act 2021* (Cth) (the Act).

2. The proliferation of the internet and social media has enabled like-minded individuals to connect, communicate and spread online messaging to an extent not previously possible. While there are many benefits to this, one negative effect is the online environment becoming increasingly saturated with abusive, violent and extremist material. Individuals' intent on sharing illegal and/or radicalising material are easily able to connect, often creating an echo-chamber of dangerous material and viewpoints.

3. Online violent extremist material and online child sexual abuse is becoming more prevalent, commodified, organised and extreme. Advancements in technology, including Artificial Intelligence (AI), anonymising technologies and end-to-end encryption (E2EE) continue to impact on law enforcement ability to effectively investigate, disrupt and prosecute these emerging threats.

4. This submission focuses on the AFP's operational experience and role in combating crime in the online environment, as well as our close working relationship with partners, including the Office of the eSafety Commissioner (eSafety).

Operational trends

Online Radicalisation

5. The internet provides a permissive environment for like-minded individuals to connect, communicate and spread their messaging widely. It can be used to realise ideologies, plan attacks and mobilise individuals and is becoming increasingly saturated with violent extremist content.

6. Extremist groups are deliberately targeting young and vulnerable individuals for radicalisation. The process used by extremists to radicalise youth often differs from adults. Law enforcement experience indicates Australian youth can be more susceptible and vulnerable to radicalisation by extremists, and are influenced by a number of factors. These include social dislocation, peer influence, mental health challenges, neurodiversity factors, active online engagement with extremists, and triggering or traumatic events.

7. When young people distribute extremist content, it is often done without fully understanding the repercussions of their actions. The AFP continues to witness a highly concerning trend of young people being radicalised online, either by other individuals they are engaging with, or through self-radicalisation, with the majority of active Joint Counter Terrorism Teams (JCTT) investigations involving an online element. Many have no offline social connections with extremists, but access and share propaganda with likeminded individuals and groups online.

8. Of concern are the creation of echo chambers where alternative content is lacking. This can reinforce extremist ideologies and detach individuals from other environments. Similarly, the online environment provides opportunity for individuals to amass huge collections of violent extremist material, and offenders have been known to gather videos, images and text to make content libraries.

9. The AFP has observed the ramifications of filter bubbles, including through algorithmic-based preferencing of online content within social media applications. Such algorithms are designed to increase views and produce higher levels of engagement. This persuasive technology can exploit an individual's motivations to connect and engage with both Religiously Motivated Violent Extremist (RMVE) and Ideologically Motivated Violent Extremist (IMVE) groups.

10. The AFP anticipates the risks within the terrorism environment over the next five years will continue to be:

    a. RMVE and IMVE (including National Racist Violent Extremism)

        i. RMVE remains the predominant threat to Australia. Where the AFP used to see planned, coordinated and resource-intensive threats and attacks, the AFP are now seeing small-cell actors and self-radicalised individuals who may engage in less sophisticated attacks.

        ii. IMVE propaganda with extremist narratives influences a broad audience, and comes in many forms including by spreading disinformation, conspiracy theories, and in some cases in order to incite violence.

    b. Youth radicalisation (including self-radicalisation)

    c. The continued advancement of online technologies (including anonymising technology and E2EE), and;

    d. International instability, of which the implications are transcending international borders.

11. The use of social media by extremists to distribute propaganda has extended to gaming and gaming adjacent platforms. Together with the 'gamification' of mass casualty incidents and the language surrounding them, young people are introduced to extremist concepts in a more relatable way. The AFP has witnessed the chat function on gaming platforms being used to target and radicalise vulnerable young people.

Online Child Sexual Abuse

12. Online child sexual abuse is becoming more prevalent, commodified, organised and extreme, and continues to be complex for police to track and investigate. This is compounded by wide-scale adoption of E2EE, anonymising technologies, streaming services and pay-per-view models, and use of virtual currencies to obfuscate law enforcement detection.

13. The hosting, sharing and distribution of child abuse material (CAM) occurs on dark web hidden services, which require specialised browsers and other anonymising software to access. These technologies are free or low cost for perpetrators to use, yet make a significant impact on the ability for law enforcement to detect and access.

14. The scale of offending on such platforms is significant and many services have tens of thousands of users across the world. Offenders using such platforms are typically concerned about law enforcement detection. Law enforcement have observed offenders producing and sharing 'how to' guides to assist fellow perpetrators on avoiding detection alongside instructional guides for producing CAM. Traditional law enforcement techniques struggle to address the scale of this problem.

15. Online child sex offenders also operate on the clear net, using consumer grade messaging applications many of which already employ E2EE. The challenges in this environment are detailed below under 'End-to-end-encryption'.

16. The AFP has responsibility for the investigation of online child sexual exploitation, including online grooming, live streaming, producing and consuming CAM, sexual extortion of minors and on Australians engaged in the sexual abuse of children offshore. The AFP combats child sexual exploitation in partnership with state, territory and international law enforcement agencies, government organisations and industry partners.

17. Operational since 2018, the AFP-led Australian Centre to Counter Child Exploitation (ACCCE) is a world-class collaborative hub, bringing together law enforcement, public and private sectors and civil society, to drive a national response to deter, disrupt and prevent child exploitation, with a specific focus on countering online child sexual exploitation.

18. The ACCCE has a range of capabilities within its structure to help deliver on its mission. The ACCCE receives reports of child exploitation from a range of sources, including investigative authorities, Commonwealth agencies (including eSafety, victims of crime, members of the public, non-government organisations and private sector organisations). CyberTipline reports via the National Centre for Missing and Exploited Children (NCMEC) based in the US contribute to a significant proportion of total reports; however increased community awareness and emerging trends such as sexual extortion is resulting in an increase in member of public reports to the ACCCE.

19. In response, reports and related material are assessed to determine the most appropriate course of action (24/7 triage). This may include an immediate referral to the relevant investigative authority or requirement for intelligence input, victim identification support or covert online engagement.

20. Reports of online child sexual exploitation to the ACCCE have nearly tripled from 14,285 in the 2018-19 financial year to over 40,000 reports of online child abuse material in the 2022-23 financial year. The spike in reports reflects increasing levels of online child sexual abuse identified, alongside greater awareness in the Australian community of the issue and methods of reporting.

21. The AFP's Child Protection Operations works collaboratively with domestic and international partners to prevent, detect, investigate and disrupt online child sexual exploitation, and sexual abuse of children offshore. This model includes:

    - Joint Anti-Child Exploitation Teams (JACETs) consisting of AFP and state and territory police child protection teams in all Australian jurisdictions (with the exception of NSW, where the AFP maintains a Child Protection Team working closely with its NSW counterpart). The teams were implemented in response to the high volume and velocity of child sexual exploitation information received by the AFP.

Sextortion

22. The AFP is aware of children and young people being targeted by online child sex offenders through social networking, image, video and instant messaging applications to self-produce CAM. Children and young people may self-generate CAM for a number of reasons, including but not limited to, consensual sharing of images, feeling pressured or coerced, grooming, financial gain and in some instances, as a result of sexual extortion (sextortion).

23. As part of sextortion offenders employ fear, coercion and manipulation tactics, including threatening to share material, to force the victim to pay money and/or produce more CAM. Offenders exploit young victims' feelings that they have done something wrong and will be reprimanded by parents or carers and even prosecuted by the law if their actions are discovered. Many victims will feel there is no way out of the situation.

24. On 1 July 2022, the ACCCE in partnership with AUSTRAC, announced Operation HUNTSMAN, a national operational strategy to counter the dramatic increase in online sexual extortion of Australian children. Phase 1 was a financial disruption strategy enacted against over 1,500 Australians, to prevent offshore offenders accessing any funds garnered from the online sexual exploitation of Australian children. This ACCCE-initiated financial action has resulted in the closure of several thousand Australian bank accounts used for moving funds offshore, significantly reducing the number of Australian child victims being targeted.

25. Phase 2 of Operation HUNTSMAN involves taking action against domestic and offshore elements involved in facilitating or profiteering from sextortion. ACCCE are currently working with international partners to disrupt offshore organised crime syndicates targeting Australian children. Following Phase 2 operational activity, there was a 20% reduction in sextortion reports received by the ACCCE.

26. Tackling the sexual extortion of children remains a priority, noting it has been linked to multiple domestic child suicides. The ACCCE, ThinkUKnow, in collaboration with partners such as eSafety, commenced awareness and education programs with the aim to remove the stigma associated with sextortion, educate victims on what to do if someone attempts to sextort them and where to go for help.

End-to-End Encryption

27. E2EE is a technology that prevents third parties from accessing the content of communication between two or more parties. It has been widely adopted by major communication platforms, but poses a challenge for law enforcement agencies who need to access such data for investigation purposes, under a lawful warrant. In 2022-2023, 96.1% of the AFP's lawfully intercepted content was unintelligible due to encryption. E2EE also prevents communications service providers from identifying illegal content on their own platforms and reporting it to law enforcement.

28. Encryption has a legitimate role in protecting information, including banking and identity data, from unlawful access. However, the AFP holds serious concerns about how E2EE has not only diminished law enforcements ability to investigate serious crime, but also the ability of industry to maintain a safe online environment and protect users of their services.

29. The AFP would welcome the independent review considering whether increased regulation in the use of E2EE in online services being used by children would assist in reducing the exploitation of children online as well as assist law enforcement in detecting, preventing and responding to national security and public safety threats.

Deep Fakes and AI-Generated Imagery

30. Deep fakes are digital images, videos or sound files of a real person that have been edited to create an extremely realistic but false depiction of them doing or saying something that they did not actually do or say. They have been utilised to discredit public figures and spread

dis-information, extort funds, and influence democratic processes and can pose significant threat to trust, reputation, and public safety of individuals.

31. The sophistication of deepfakes make it difficult to distinguish 'fact from fiction' and the 'real from virtual', and poses a real challenge for both law enforcement and broader community.

32. The AFP notes the introduction of the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024 into Parliament on 5 June 2024. The Bill will strengthen offences targeting the creation and non-consensual sharing of sexually explicit material depicting adults online, including material created or altered using technology such as deepfakes. These amendments will ensure the legislation is capable of responding to contemporary and emerging technology.

33. Section 473.1 of the *Criminal Code Act 1995* (Criminal Code) defines CAM to include "material that depicts a person, or a representation of a person" and as such, extends to include AI-generated material, anime and text-based material. The ACCCE has defined both AI Generated Child Abuse Material (AIG-CAM) and AI Modified Child Abuse Material (AIM-CAM) for use within the Australian Victim Identification Database. AIG-CAM creates material without the direct use of media from real child victims, AIM-CAM is modified using media from real child victims. This 'artificial' content is of particular concern to specialists in the ACCCE, who need to establish whether or not the CAM depicts a real victim. By having to devote resources to ascertain whether the CAM depicts real human or AI-generated victims, they have less time available to identify and help victims of child sexual exploitation.

34. On March 2024, a Tasmanian man was jailed for two years for uploading and downloading CAM, which included content generated by AI, identified as part of a Tasmanian JACET (TAS-JACET) investigation. The TAS-JACET, comprising of AFP and Tasmania Police members, began an investigation in November 2022 after receiving reports from the National Centre for Missing and Exploited Children (NCMEC) relating to an Australian-based user saving and downloading CAM from a website and social media platform.

35. An examination of the computer seized by police revealed hundreds of files containing CAM, including a significant amount generated by AI. The man was arrested and charged and pleaded guilty to possessing CAM obtained using a carriage service contrary to section 474.22A of the Criminal Code and using a carriage service to access CAM contrary to section 474.22 of the Criminal Code.

Malicious use of Artificial Intelligence

36. Malicious use of AI is an evolving cybercrime-enabler, allowing for the improvement and modernisation of conventional/traditional cybercrime tradecraft.

37. Malicious use of AI can include use of generative AI models providing a suite of capabilities that can be exploited by cybercrime-as-a-service tools. This can include the crafting of very realistic spear phishing emails, vishing (voice phishing/Hi Mum, Hi Dad scams), Business Email Compromise attacks, generating malware, exploit scanning, and target acquisition. Criminals can leverage open-source AI large-language models to reduce development time and provide bespoke tradecraft. The availability of AI lowers the barriers to conducting large-scale attacks where criminals may otherwise lack resources or technical proficiencies.

38. As a result of the increasing availability of such technologies, it is anticipated the frequency and severity of cybercrime incidents will increase and place new demands on the AFP and partner agencies.

39. Further, the AFP is also aware of malicious use of AI to generate increasingly realistic deepfakes, disinformation content and life-like CAM, with the AFP-ACCCE receiving reports of crime containing AIG-CAM.

ThinkUKnow

40. The AFP is a leader in prevention and education initiatives such as through ThinkUKnow program, and Playing IT Safe. The ACCCE also engages in a significant body of work to spread awareness of online child sexual exploitation in Australia and internationally.

41. The ThinkUKnow program has been delivered across Australia since 2009 educating students, parents, carers and teachers about online child sexual exploitation and how to keep children and young people safe online. Since its inception, the scope of the program has expanded to include a suite of education materials including presentations for students, parents, carers and teachers, resources for parents, carers and teachers including Teacher Toolkits, home learning activities, fact sheets, guides and the children's picture book Jack Changes the Game.

42. Through the ThinkUKnow program, the AFP holds the only capability in Australia that sees intelligence, examples from police investigations and victim reports developed into educational resources that are delivered into classrooms nationally. The Online Child Safety Team works alongside operational teams including the Child Protection Triage Unit, Intelligence Fusion Cell and JACETs to develop this evidence-base and ensure that the program addresses safety issues in a contemporary way.

43. In 2022-23 FY the AFP, state and territory police and industry volunteers delivered 257 presentations to an estimated 17,756 parents, carers, and teachers across Australia, including a mix of face-to-face and virtual sessions. A further 2,515 presentations were delivered to an estimated 209,544 students across Australia. The AFP Online Child Safety Team delivered eight ThinkUKnow presentation training sessions that saw 719 attendees from AFP, state and territory police and industry partners.

AFP and eSafety Partnership

44. The AFP and the eSafety Commissioner (eSafety) have a long-standing relationship working collaboratively on a range of Commonwealth matters relating to online crimes. In 2020, the AFP and eSafety signed a Memorandum of Understanding (MOU), with the intention to facilitate positive outcomes for Australian children who are vulnerable to a range of online harms. The MOU addresses how and under what circumstances eSafety will notify the ACCCE about threats to children. For example, eSafety may notify the ACCCE where a matter is reported to eSafety involving online grooming or where CAM is through the Online Content Scheme depicts an identifiable child or offender.

45. Where information may lead to the identification of a victim or offender is found as part of eSafety's investigations, they will provide this to the ACCCE for their consideration. The arrangements for sharing information between eSafety and the ACCCE are contained within a letter of exchange, which operationalises the provisions of the MOU. The letter of exchange relates to the prevention of online child sexual exploitation through education and awareness. The AFP achieves this objective through the ACCCE and AFP-led online child safety initiatives and programs, and law enforcement focused crime prevention and deterrence initiatives to counter online child sexual exploitation.

46. Representation from eSafety is also embedded with the AFP Joint Policing Cybercrime Coordination Centre to support a partnered approach and ready access in cybercrime.

Legislative Framework

Takedown powers

47. The AFP remains concerned about the ease in which graphic terrorist material can be shared online, often very quickly after an attack.

48. The widespread sharing of footage online of the recent attack in Wakeley, NSW, has reinforced the ease with which violent material can be broadcast online, as previously seen during the livestreaming and subsequent sharing of footage of the 2019 Christchurch attack. The circulation of such material online can have wide-ranging implications, including risk of promoting and inciting violence, and encouraging further interest in the harmful content. Increased transparency and accountability by digital industry on their algorithmic systems is important to ensure violent or exploitative content is detected and prevented from going viral.

49. The takedown powers under the Act require social media companies to remove violent extremist content posted online, and are a positive step towards addressing online radicalisation. However, the recent court case involving the e-Safety Commissioner and X highlights the challenges surrounding the enforcement of these powers.

50. The takedown powers only scratch the surface when it comes to addressing and preventing radicalisation. Vulnerable young people engage online in a variety of forums, including niche platforms where controls, scrutiny and takedown of offensive or illegal content are limited or where content is protected from oversight and monitoring due to encryption.

51. Since July 2021, the AFP has commenced investigations and conducted operational activity against 27 individuals that were 17 years old or younger, with the youngest being 12 years old (as at 3 June 2024). The ease of access to extremist content online, lack of accountability by internet service providers, and E2EE all create significant challenges for authorities to monitor and take action against violent and extremist content posted online.

Conclusion

52. Criminals are highly adaptable and employ technological advancements as they emerge. Without appropriate and proportionate mechanisms and safeguards in place, malicious use of emerging technologies will negatively impact law enforcement's ability to investigative, disrupt and prosecute these evolving and concerning threats.

53. To address this ever-evolving threat, the AFP and partners must work together at local, national and international levels to maintain an ability to coordinate a full suite of expertise, resources and technology to swiftly combat these threats and protect the community.

54. The Act is an important mechanism in combating the threat posed by online exploitation and the AFP see's utility in the framework continuing in the future.