Australian Government | eSafetyCommissioner

12 February 2021
Senator James Paterson
Chair
Parliamentary Joint Committee on Intelligence and Security
PO Box 6021
Parliament House
Canberra ACT 2600

Dear Mr Paterson,

As Australia's eSafety Commissioner, I welcome the opportunity to provide a submission to the inquiry into extremist movements and radicalism in Australia (inquiry). The inquiry raises important issues that require careful and thorough consideration.

I am writing to you to outline the powers and functions I have as eSafety Commissioner that relate to the inquiry's Terms of Reference. I want to highlight how my broad powers and functions relating to online harms, spanning across education, prevention, proactive change and regulation, connect with the complex area of counter terrorism and countering violent extremism. I also want to provide insights into some of the factors that facilitate and promote extremist movements and radicalism online – and just as importantly, the measures that can disrupt and prevent these activities.

Importantly, I want to underscore for the Committee how broader online safety efforts sit alongside, and ultimately support, specific efforts to counter terrorism and violent extremism.

eSafety

eSafety is Australia's national independent regulator for online safety. Our core objective is to minimise harm to Australians online.

eSafety is the first government agency in the world dedicated specifically to online safety. We lead, coordinate, educate and advise on online safety issues and aim to empower all Australians to have safer, more positive online experiences.

As you may know, on 23 December 2020, the Australian Government commenced a consultation on a Bill for a new Online Safety Act. The proposed new Online Safety Act will expand eSafety's regulatory remit and my functions and powers as eSafety Commissioner, improving the effectiveness, reach and impact of eSafety's work.

AVM

Of my functions and powers as eSafety Commissioner, one relates to Abhorrent Violent Material (AVM).

As you may know, in response to the live-streamed attack in Christchurch, New Zealand on 15 March 2019, the Australian Parliament passed the *Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act 2019*.

AVM is defined as audio and/or visual material that records or streams a terrorist act involving serious physical harm or death, the murder or attempted murder of another person, the torture of another person, the rape of another person, or the kidnapping of another person involving violence. It is material produced by a perpetrator or an accomplice – it does not include bystander coverage. There are also a number of defences for use of AVM, including by journalists, law enforcement agencies, public officials and for research, advocacy or artistic purposes.

The AVM regime gives the eSafety Commissioner the power to issue a notice to any website publishing AVM and/or the service that hosts that website. Rather than require the eSafety Commissioner to monitor the internet for AVM, it is predominantly a complaints-based regime.

Importantly, this is not a power to take down material. Further, the notices do not require the AVM to be removed. However, if a service is later prosecuted for failing to remove or cease hosting AVM, the notice can be used in legal proceedings to show recklessness regarding the AVM.

We have found these notices to be a very effective tool for alerting services to the presence of AVM so they can take appropriate action to protect their users from harm. To date, the eSafety Commissioner has issued 23 notices in relation to content depicting beheadings, shootings and other murders, with the content taken down or restricted for Australian users in 93% of the cases.

Internet Service Provider (ISP) Blocking in an Online Crisis Event

Another of my functions and powers as eSafety Commissioner relates to internet service provider (ISP) blocking in an online crisis event.

Under subsection 581(2A) of the *Telecommunications Act 1997,* the eSafety Commissioner is empowered to give written directions to a service provider in connection with any of the Commissioner's functions and powers. This includes directing ISPs to block Australians from exposure to material that promotes, incites or instructs in, terrorist acts or violent crimes, consistent with a new function conferred on the Commissioner in July 2019.

I exercised this power in September 2019 to formalise blocking action already taken by ISPs against websites providing access to the manifesto and/or video produced by the perpetrator of the Christchurch attack. This was a temporary, six-month direction that was put in place as an interim measure to stem the viral spread of the material and prevent it from causing harm to Australians, while government considered longer-term options for addressing misuse of online platforms by perpetrators of terrorism and violent extremism.

In December 2019, in conjunction with government and industry and in line with a recommendation of the Report of the Australian Taskforce to combat terrorist and extreme violent material online (Taskforce Report), eSafety finalised Protocol (No. 2) Governing ISP Blocking in an Online Crisis Event (Protocol). The Protocol sets out when the eSafety Commissioner will use these powers in relation to potential future 'online crisis events', defined as incidents involving terrorist or extreme violent material being shared widely online in a manner likely to cause significant harm to the Australian community. The Taskforce Report emphasised that such events would require a rapid, coordinated and decisive response by industry and relevant government agencies to contain the viral spread of the material.

The Protocol establishes detailed criteria, high thresholds and checks and balances to ensure the eSafety Commissioner's powers are used only in very limited and very serious circumstances. Any blocking direction made under the Protocol would only be in place for a limited time, to be determined on a case-by-case basis. Following the initial blocking period, the eSafety Commissioner could take further action to address the relevant material, in consultation with the ISPs and affected websites.

Online Safety Act Reforms

As outlined above, the Australian Government is currently consulting on a draft Online Safety Bill to consolidate and update eSafety's regulatory arrangements to ensure Australia's online safety legislation is fit for purpose.

Under the draft Bill, it is proposed that the eSafety Commissioner may give a voluntary blocking request or mandatory blocking notice to an ISP to disable access to AVM or material that promotes, incites or instructs in abhorrent violent conduct for up to three months.

In addition, there is a proposed online content scheme for two classes of material. The draft Bill defines class 1 material as material that is likely to be Refused Classification (RC) under the *Classification (Publications, Films and Computer Games) Act 1995*. This includes publications, films or computer games that directly or indirectly counsel, promote, encourage or urge the doing of a terrorist act; directly or indirectly provide instruction on the doing of a terrorist act; or directly praise the doing of a terrorist act in circumstances where there is a substantial risk that such praise might have the effect of leading a person to engage in a terrorist act.[1]

The draft Bill empowers the eSafety Commissioner to give a removal notice to a social media service, relevant electronic service, designated internet service or hosting service requiring them to take down class 1 material within 24 hours, regardless of whether the service is provided from Australia or overseas.

If a website systemically ignores take down notices for class 1 material, the eSafety Commissioner may give an internet search engine service a link deletion notice to require the provider to cease providing a link. Similarly, if an app systemically ignores take down notices for class 1 material, the Commissioner may give an app removal notice to an app distribution service provider, to cease enabling end-users in Australia to download an app.

The draft Bill also proposes to set Basic Online Safety Expectations (BOSE) for social media services, relevant electronic services and designated internet services. The proposed BOSE include the expectation that the provider of the service will take reasonable steps to minimise the extent to which class 1 content, AVM and material that promotes, incites or instructs in abhorrent violent conduct is provided on the service. It also includes the expectation that the service has clear and readily identifiable mechanisms that enable end-users to report such material.

The draft Bill establishes mandatory reporting requirements that will allow the eSafety Commissioner to require online services to provide specific information about their compliance with the BOSE. This could include requiring services to explain how they are working to minimise terrorist and abhorrent violent material on their service. Reporting requirements could integrate or build on a number of other transparency initiatives. For example, domestically, the Taskforce to Combat Terrorist and Extreme Violent Material Online called on digital platforms to report against a range of metrics in relation to their efforts to detect and remove terrorist and extreme violent material. Internationally, the Department of Home Affairs is leading Australia's engagement on the development of a Voluntary Transparency Reporting Framework spearheaded by the Organisation for Economic Cooperation and Development to promote more thorough and consistent reporting

[1] *Classification (Publications, Films and Computer Games) Act 1995* s 9A

about terrorist and violent extremist content among services. eSafety has participated in and supported both of these initiatives.

Proactive measures

I do not shy away from acknowledging the consequences of harmful material online: the impacts can be profoundly damaging and have long term consequences.

But it is also important to acknowledge not only the immense benefits of being online, but that there are measures than can prevent and mitigate these harms. eSafety has a strong focus on proactive and preventative measures.

eSafety's research and experience points to the fact that online harms can disproportionately impact at-risk and diverse groups. This includes, but is not limited to, Aboriginal and Torres Strait Islander people, people from culturally and linguistically diverse communities, people with disability and people who identify as LGBTQI+, as well as, depending on the circumstances, women, older people and children and young people.

eSafety's research on online hate speech shows that religion, political views, race and gender were the most common reasons cited in both Australia and New Zealand for experiencing hate speech.[2]

eSafety's research on young people and social cohesion showed that 33% of young people have seen videos or images promoting terrorism online and over 50% of young people had seen real violence that disturbed them, racist comments and hateful comments about cultural or religious groups.[3]

Especially in the context of this inquiry, it is important to consider the structural, systemic and social factors that may lead someone to be attracted to, and engage in, negative or dangerous activity online. A whole of community approach and systems approach is therefore needed to understand and address the underlying drivers of this behaviour, as well as provide diversion and alternative pathways to support and assistance.

eSafety also focuses on digital capacity building: giving individuals the skills and strategies to prevent and respond to harmful experiences online and engage online in ways likely to promote safe and positive online experiences.

Capacity building should be a lifelong process that begins at the earliest age possible. It should occur at the individual and community level and at a societal and cultural level. In other words, it needs to focus on building the capacity of the individual, but also of communities and society to understand, recognise and respond to harm online, including extremism and radicalism, and promote safer and more positive experiences.

eSafety has an extensive education and outreach program to support this stream of work. The four Rs of online safety — respect, responsibility, resilience and reasoning — are a basis for examining online information and making an informed judgement on an issue. eSafety has also developed a Best Practice Framework for Online Safety Education based on an evidence review. The review found that a sound online safety education should cover the full range of potential issues, risks and harms that children may encounter and should be delivered in supportive school systems with strong partnerships with other agencies. It is

[2] eSafety, January 2020, Online hate speech. Findings from Australia, New Zealand and Europe, https://www.esafety.gov.au/sites/default/files/2020-01/Hate%20speech-Report.pdf

[3] eSafety and Department of Education and Training, 2017, Young people and Social Cohesion, https://www.esafety.gov.au/about-us/research/young-people-social-cohesion

encouraging that eSafety's most recent research report on the *Digital Lives of Aussie Teens* shows that young people are proactively building positive online relationships, with 9 in 10 teens saying they had engaged in at least one positive online behaviour over the previous 6 months.[4]

eSafety also has a world-leading Safety by Design (SbD) initiative, which aims to drive up standards of user safety in the technology community. SbD emphasises the need to address online harms, alongside user safety and rights, in the design, development and deployment of technological solutions.

Digital platforms and services can take preventative steps to guard against their services being used to facilitate, inflame or encourage illegal and inappropriate behaviours, including violence, extremism, terrorism and hateful or offensive material. Additionally, these platforms can be designed to enable, optimise and support users to have greater control over their activities online. These platforms can also utilise technical measures and tools to assist and help manage exposure to, as well as the amplification of, violent, sexualised or age-inappropriate content and extremist or radical content. In summary, SbD places and embeds safety at the forefront of all stages of the product development lifecycle, in the culture and operations of organisations and in the wider governance, legislative and regulatory systems in which technological solutions exist.

Counter terrorism and countering violent extremism are complex issues with a range of social, cultural and behavioural underpinnings and drivers. As with all our work, eSafety supports a multi-faceted approach that explores how technology can both be a tool for, and extend upon, these underlying issues, while also serving as a means for positive and safe online experiences.

I commend the Committee for undertaking this important inquiry, which I will be monitoring with great interest.

My office and I are happy to provide any further information that would be of assistance to the Committee.

Yours sincerely,

[redacted]

Julie Inman Grant
eSafety Commissioner

[4] eSafety, February 2021, The Digital Lives of Aussie Teens, https://www.esafety.gov.au/sites/default/files/2021-02/The%20digital%20lives%20of%20Aussie%20teens.pdf