Senate Estimates opening statement
February 2024

Thank you Chair, and Committee members, for the opportunity to present this statement to update the Committee on developments, achievements, and challenges we faced at the eSafety Commissioner over the past three months.

eSafety has been relentless in its pursuit of enhancing online safety for all Australians.

Last month, we embarked on a partnership with the European Parliament Intergroup on Children's Rights, a new milestone in our global efforts to safeguard children online.

This partnership underscores our commitment to fostering international collaboration and prioritising children's online rights.

Through joint commitments and advocacy efforts – including our work co-founding the Global Online Safety Regulators Network and active membership of other important bodies like INHOPE and the WePROTECT Global Alliance– we aim to address common challenges and promote a safer digital environment for children worldwide.

My recent meetings with European Commission officials and fellow regulators have strengthened this collaboration, setting the stage for future collaboration.

Pleasingly, the Safety by Design approach we have advocated for tech companies since 2018 is being actively embraced by companies, advocacy organisations and legislative schemes around the globe.

In 2023, we saw a notable surge in cyberbullying reports to eSafety, prompting us to urge parents and carers to have open discussions with their children as the new school term approached through an earned media awareness campaign.

It is crucial we continue to engage the entire community, especially with the rapid rise of advanced AI magnifying the potential harms and reinforce that eSafety is there with resources, reporting and support.

We are already receiving a small number of complaints about synthetic (AI-generated) child sexual abuse material and deepfaked porn targeting both prominent and everyday Australian women. As these increasingly powerful, consumer-facing image generation tools proliferate – without effective safety guardrails in place – we are concerned that AI-generated abuse will proliferate too.

Deepfakes are covered under our world-leading image-based abuse scheme, which has close to a 90% success rate in helping Australian victims get their non-consensually shared intimate images and videos taken down from the internet.

And we are exercising our remedial powers taking action against those who are weaponising generative AI to create deepfaked porn of Australian women.

eSafety has given formal removal notices, remedial directions and formal warnings to a range of perpetrators and has commenced civil penalty proceedings against one such individual in Australia's Federal Court. The proceedings are ongoing – and we believe that these enforcement actions send a strong deterrent message.

Just this past week, we led Australia's participation in Safer Internet Day which saw thousands of Australian organisations and individual amplifying this year's theme of Connect, Reflect, Protect.

This global event is eSafety's biggest campaign of the year and a cornerstone of our efforts to promote online safety. This year it also provided a platform to present our latest research focused on young people's experiences gaming, together with new eSafety resources designed to empower parents, caregivers, and educators in fostering safer online gaming environments.

Our multi-faceted marketing campaign utilised stakeholder engagement, strategic partnerships, media, social media, paid advertising, emails, and webinars to effectively disseminate online safety messages to millions of Australians. Notably, on Safer Internet Day, we observed a 36% surge in traffic to eSafety.gov.au compared to 2023, and a 5.25% increase compared to last year's campaign, indicating a growing reach and impact.
Additionally, our campaign garnered 204 media mentions, potentially reaching 5.33 million Australians, with 512 mentions on eSafety accounts, from a cross-section of government, corporate, not for profit and education sectors.

And last month, we released our transparency report summarising responses to the legal notice we sent to Twitter/X in June 2023 seeking specific information about what it was doing to meet the Australian Government's Basic Online Safety Expectations and enforce its own policies in relation to online hate.

The report revealed a concerning and deliberate decrease in global trust and safety personnel, safety engineers, content moderators, and public policy staff following the platform's acquisition in October 2022.

The reinstatement of previously banned accounts without placing them under additional scrutiny, including those banned for hateful conduct violations, underscored critical gaps in addressing online harms and safeguarding users.

Response times to user reports of hateful content also slowed, with a 75% increase in the median time to respond to reports about hate via direct messages.

In December, we initiated civil penalty proceedings against X Corp. in the Federal Court related to its alleged non-compliance with a separate transparency notice issued to find out what steps X Corp was doing to deal with child sexual exploitation material. This followed X Corp's decision not to pay eSafety's infringement notice of $610,500 to the company. X Corp. has also sought judicial review of eSafety's infringement notice, and the original transparency notice.

Our aim is for both the judicial review and civil penalty proceedings to be heard together to expedite both proceedings with the aim of deterring non-compliance by X Corp. and other platforms with their statutory obligation to respond to regulatory notices.Equivalent transparency notices were also issued to Google, TikTok, Twitch and Discord, seeking information on how these companies were addressing child sexual exploitation and abuse material on their services.

X Corp. failed to respond accurately to certain questions, and in some cases left whole sections blank, prompting our action.

Google also failed to comply fully with the notice by providing some general answers to specific questions but its shortcomings were found to be less serious and the company was issued with a formal warning. We continue to engage with Google at senior global levels of the company, and locally, on a range of online safety matters.

In October, we released our summary of the responses we received from service providers to these notices, our second major transparency report under the Basic Online Safety Expectations. We have now issued 13 notices covering 27 major technology services and reported on the information we found via these notices.

Concerningly, our second report demonstrated that many companies weren't using readily available tools and technologies to detect child sexual exploitation and abuse material, let alone detecting grooming. We also found gaps regarding the detection of livestreamed child sexual abuse, and limited steps taken by some companies to prevent banned users creating new accounts and reoffending.

eSafety will continue to use these transparency powers to hold platforms to account and raise safety standards across the industry on a range of high-risk, high-harm safety issues.

In December, five new industry codes came into force covering social media, app stores, ISPs, hosting providers, device manufacturers, and suppliers.

Drafted by industry, the codes require online services in these sectors to combat the most harmful forms of online content, including child sexual abuse and pro-terror material.

Australians can now file formal complaints if such services fail to meet requirements, with eSafety empowered to investigate and enforce compliance with code obligations.

While a sixth code covering search engines will commence next month, I rejected two codes drafted and submitted to eSafety by industry because they failed to meet appropriate community safeguards.

These draft codes covered the Relevant Electronic Services (RES) sector, including instant messaging, email and online dating and gaming and the Designated Internet Services (DIS) sector which covers apps and sites not covered by other sectors and also includes file and photo storage services.

They will be replaced by industry standards currently being drafted by my office.

We are aware that file and photo storage services like iCloud, OneDrive, Google Drive and DropBox are routinely used to store and distribute child

sexual abuse material and encrypted messaging services are also used to share this material.

While the standards **will not require** encrypted services to weaken or break end-to-end-encryption, this cannot be an excuse for companies to do nothing. Indeed, as advanced technologies move apace, we see tremendous opportunity for these companies to continue to invest and innovate in solutions and to deter and disrupt the hosting and proliferation of child sexual exploitation material.

Some end-to-end encrypted services are already taking useful steps, like scanning the non-encrypted parts of their service – including profile and group chat names and pictures that might indicate accounts are providing or sharing child sexual abuse material – or looking for behavioural signals to help identify child abuse perpetrators.

These and other interventions, such as making user reporting options readily accessible, are practical examples of measures companies with end-to-end encrypted services can take to reduce the risk of their services being misused to transmit horrific content.

We're taking into account constructive feedback from stakeholders to make amendments. We expect these standards, which are regulatory instruments, will be tabled before Parliament in the next three-to-six months and we welcome conversations with Parliamentarians to discuss the rationale behind these proposals, the operation of the standards and their intended impact.

In December we announced parent engagement with programs offered by our network of Trusted Education Providers nearly tripled over the previous financial year, underlining the importance of a comprehensive approach to online safety within school communities.

Our Trusted Education Providers have significantly boosted online safety education, reaching nearly half of Australia's schools and more than 1.3 million students, parents, and educators in 2022-23. We now have 1.8 million senior BeConnected learners and had more than two million visitors to our esafety.gov.au safety portal.

In November, the Australian Government's $10 million Preventing Tech-based Abuse of Women Grants Program awarded $3 million in funding to seven projects aimed at tackling online abuse against women and children.

Also in November, we launched eSafety Sport, a comprehensive resource to combat online abuse in the Australian sporting community, created with support from over 50 Australian sporting bodies.

I appreciate the Committee's attention and speak on behalf of everyone at eSafety when I say we remain committed to enhancing online safety for all Australians.