Australian Government | eSafetyCommissioner

Submission

Existing and Emerging Sexually Exploitative Practices against Children in the Digital Environment

May 2024

Introduction

Globally we are facing an unprecedented escalation of online child sexual exploitation and abuse (CSEA), with new and emerging risks, as well as established methods being amplified by advanced technologies. We note with particular concern the intersecting types of abuse and harm that children face online, including sexual extortion, peer-on-peer harmful sexual behaviour and impact of extreme pornography, and the alarming rates that online harms are occurring.

There is an urgent need for a collaborative approach to combatting CSEA, which recognises the shared responsibility of governments, communities and industry, and the importance of Safety by Design.

This submission provides information about Australia's online safety framework and the eSafety Commissioner's (eSafety) efforts to prevent and respond to CSEA, with a focus on existing and emerging technologies that are used to facilitate this abuse.

About eSafety

eSafety is Australia's independent regulator for online safety. Established in 2015, our mission is to safeguard Australians from online harms. eSafety's enabling legislation, the *Online Safety Act 2021* (the Act), empowers us to remediate online harms and tackle systemic safety risks at scale.

eSafety champions a human rights-based approach to online safety regulation and harm prevention. We work as part of a cross-sector and multi-jurisdictional online safety ecosystem, across our three pillars of prevention, protection, and proactive and systemic change.

We believe our multifaceted approach developed and refined over nine years provides a model for other jurisdictions to consider.

Prevention

Online harms research

Underpinning eSafety's approach is our research and evaluation program, providing a robust evidence base to support our programs and regulatory functions. This includes research about the online risks encountered by children.

Our 2024 research on the prevalence and predictors of requests for facilitated CSEA on online platforms found that, of 4,011 Australian adults surveyed, 2.8% had received requests to facilitate sexual exploitation of children they had access to. Sharing photos of or information regarding children online was associated with a significantly increased likelihood of being asked questions of a sexual nature about children, as well as being asked, pressured or offered payment for sexual images of children. The results highlight the need for increased awareness of the potential harms of posting photos of and information about children publicly online, and place onus on platforms to warn users of these potential harms.

Our research on young people's experiences of online gaming found a significant minority of young gamers surveyed encountered risks online such as requests for nude images and potential grooming. Our survey of over 2000 young people aged 8-17 found 9% of teen gamers (aged 13-17) and 3% of children (aged 8-12) surveyed had received or been asked for nude images or sexual information in the past year while playing games online. This is broadly consistent with our previous research on young people's online lives, which found that 11% of young people aged 14-17 said that they had been pressured online to send sexual pictures or videos of themselves.

Our survey also found 7% of young gamers experienced other players doing or saying something that made them feel uncomfortable, including asking personal questions, being too friendly or asking them to keep secrets. These types of interactions could be indicative of grooming.

The results of our gaming research highlight the important role industry, educators and parents/carers have in preventing harm and equipping young people to manage risks associated with online gaming.

Our report into age verification and complementary measures to prevent and mitigate harms to children from online pornography also explores the impact that young people viewing extreme pornography can have on social norms around gender and sexual violence.

Education and training

Research underpins our education and awareness-raising efforts, where we lead, coordinate, educate and advise on online safety issues to empower all Australians to have safer online experiences.

We draw on experiences from relevant cohorts to ensure resources, training and education materials we develop are informed by robust evidence, fit for purpose and resonate with the target audience. For example, to prioritise perspectives and aspirations of young people to about the issues they face online, in 2022 eSafety established a Youth Council, giving young people aged 13-24 a platform to influence issues, engage in meaningful discussion, and share their knowledge and experiences. Youth Council insights are considered in broader policy and program development, including the position statement on generative AI and the roadmap for age verification. Our website further details our advice, programs and resources for communities.

Protection

Investigations

eSafety also focuses on protection and harm remediation through our investigations and complaints scheme – serving as a safety net when platforms fail to act on reports of abuse.

The safety of children online, including the rapid removal of CSEA material is a regulatory priority for eSafety. We are Australia's hotline to report CSEA material and operate legislated youth cyberbullying, adult cyber abuse, and image-based abuse schemes, where individuals can make complaints about harmful online content or report illegal and restricted online material. We also use broad systemic powers to increase transparency and accountability to shift responsibility back to platforms.

Under the Act, eSafety can conduct investigations into illegal and restricted content including material showing the sexual abuse of children or acts of terrorism, through to content which should not be accessed by children, such as simulated sexual activity, detailed nudity or high impact violence.

If CSEA material is found to be hosted in or provided from Australia, eSafety will notify Australian police first. Once we are certain their investigation and the potential rescue of a child will not be compromised, we will direct the relevant online service to remove the material.

eSafety can give a removal notice to a website, social media service, relevant electronic service or hosting service provider who must remove the material within 24 hours. Where established, eSafety uses informal pathways for the removal of online CSEA material over formal action. Providers that do not comply with a takedown notice issued by the eSafety Commissioner face serious penalties.

The vast majority of CSEA material reported to eSafety is hosted overseas. In these cases, as the Australian member of INHOPE, eSafety works with hotlines in other countries, who have relationships

with industry and law enforcement agencies in their own country, to facilitate the rapid removal of the content. Where CSEA material is hosted in or provided from a non-INHOPE member country, eSafety will consider the use of formal powers to get the material removed, including through the Australian Federal Police (AFP).

As per our 2022-23 Annual Report, over 90% of investigations into illegal and restricted content (including CSEA material) are finalised within 2 business days, with eSafety sending 14,975 notifications to INHOPE and 76 investigations to the AFP. Reports to eSafety also show sexual extortion almost tripled between 2022 and 2023, with the vast majority of sextortion reports coming from young people aged 18-24 years.

Industry codes and standards

In Australia, and globally, there have been significant increases in reports of online child abuse material. While notice and takedown for CSEA material will remain an important regulatory tool, more needs to be done at the systemic level to prevent and address harms arising from illegal and restricted content.

The Act provides for industry bodies to develop codes to regulate illegal and restricted online material, and for eSafety to register the codes if they meet appropriate community safeguards. If a code does not meet the requirements, eSafety can develop an industry standard for that section of the online industry.

In 2023, eSafety registered six codes, placing enforceable requirements on social media, ISPs, equipment providers, app stores, hosting services and search engines. Under these codes, providers are required to take a range of measures, including in the case of social media services, the detection and removal of unlawful online material, such as videos showing the sexual abuse of children.

Two draft codes did not provide appropriate community safeguards – Relevant Electronic Services (covering messaging services, online dating services and gaming) and Designated Internet Services (including file and photo storage services). Accordingly, eSafety has drafted and is consulting on standards for these sections of the online industry. Once registered, these standards will operate alongside the codes, which are available on eSafety's register of industry codes and standards.

eSafety can receive complaints and investigate potential breaches of the codes or standards, enforceable by civil penalties and other enforcement options. A second set of codes and standards will be developed to ensure industry minimises the risk of children's exposure to restricted content, such as online pornography.

Proactive and systemic change

Transparency and accountability

Through the Act, eSafety also has powers to compel industry transparency and promote greater accountability. The Basic Online Safety Expectations (the Expectations) outline the Australian Government's expectations that online service providers will take reasonable steps to keep Australians safe including in relation to illegal and restricted online content. The Expectations aim to provide transparency and accountability around services' safety features, policies and practices.

eSafety can require online service providers to report on how they are meeting the Expectations. The obligation to respond to a reporting requirement is enforceable and backed by civil penalties. eSafety can publish statements about the extent to which services are meeting the Expectations.

Since the establishment of the Expectations, the Commissioner has issued two rounds of transparency notices focused on CSEA. Providers were asked questions about the tools, policies and processes used to address CSEA, such as the proliferation of online material, grooming, use of video calling services to provide livestreamed child abuse, sexual extortion, and the steps taken to avoid the risk of amplifying harmful content through recommender systems.

From the notices we found that none of the major technology companies were doing enough to combat CSEA content, including significant inconsistencies in the timeliness of responses to reports of CSEA as well as protections to prevent the livestreaming of CSEA.

Summaries of our reporting notices are published on our website. By highlighting what we have learned from these notices, eSafety's aim is that the information is used by researchers, academics, the media and the public to scrutinise the efforts of industry, to encourage implementation of the Expectations and to lift safety practices, protections and standards.

As a global community, it is imperative that we set and enforce expectations of industry to prevent and respond to CSEA. We welcome the Special Rapporteur and Member States calling for greater transparency and accountability of the technology industry.

Safety by Design

eSafety's systemic and enforceable regulatory tools are important to prevent harms at scale. We also need to change the culture of tech companies to ensure robust and effective safety protections are embedded to prevent misuse and abuse.

This is where Safety by Design plays a critical role – an initiative we commenced in 2018 – to encourage the tech sector to put user safety and rights at the centre of the design and development of online services, rather than retrofitting safeguards after harm has occurred.

The initiative was developed through in-depth research and consultation with big tech, NGOs, advocates, parents and young people. At its heart are three principles (Service provider responsibility; User empowerment and autonomy; Transparency and accountability) that elevate safety as the third pillar, alongside privacy and security, in the development process for online and digital technologies.

Safety by Design provides voluntary guidance and tools for the tech sector to bolster and improve standards of safety within companies and as a catalyst to build a thriving and responsible technology ecosystem. We encourage governments to motivate tech companies to utilise these freely available resources to improve their online safety practices.

Emerging trends and technologies

Recent advancements in technologies are changing our online experiences, from the proliferation of generative AI to developments in immersive technologies that have the potential to converge and make online experiences more visceral. The expansion of existing technologies like end-to-end-encryption (E2EE) to services popular with children is creating additional risks.

eSafety conducts horizon scanning and engages with domestic and international experts – across academia, civil society, government and industry – through our Tech Trends work program. This allows us to identify the safety risks and benefits of emerging technologies, as well as regulatory opportunities. Through our position statements, we highlight the importance of a Safety by Design approach to ensure potential risks are assessed upfront and preventative measures built in to reduce the likelihood of harm occurring.

Generative AI

In August 2023, eSafety published a statement on generative AI, which provides an overview of the generative AI lifecycle, examples of its use and misuse, considerations of online safety risks and opportunities, as well as regulatory challenges and approaches. The position statement explores emerging risks to children and young people, including the creation of CSEA material, and child safety risks related to chatbots and other forms of conversational AI. For example, perpetrators can exploit the ability of large language models (LLMs) powered by AI to mimic natural human language allowing them to groom children in automated and more targeted ways.

The position statement highlights a range of industry interventions to minimise existing and emerging generative AI harms. This includes interventions across the entire AI product lifecycle– e.g., establishing a business case and selecting data, informed consent measures for data collection and use, routine stress tests before deployment, establishing escalation pathways to engage with law enforcement, support services or illegal content hotlines, providing real-time support and reporting mechanisms and regular evaluation and third-party audits.

Generative AI may also provide opportunities to improve content moderation and detection, including to combat the proliferation of CSEA material. For example, LLMs can be used to identify harmful content or material, which could reduce the need for humans to be exposed to harmful content during review processes. In addition, it may be possible to train AI models to detect harmful text more effectively than existing key word detection tools.

Immersive technologies

Immersive technologies (including augmented reality, virtual reality and haptics) blend the virtual and physical worlds to create highly interactive sensory experiences, including through touch, sound and visual content.

While immersive technologies provide opportunities in entertainment, education, defence, and health sciences, hyper-realistic experiences could increase the impact of negative interactions online. Our statement highlights concerns about how immersive technologies can be used for CSEA, especially in the gaming environment, as well as around the collection of vast amounts of data, including biometric information.

eSafety champions a proactive approach underpinned by Safety by Design, working with industry and users to mitigate risks and ensure the benefits of immersive technologies can be fully enjoyed.

End-to-end encryption

eSafety's statement notes that E2EE presents child safety risks as it can create digital hiding places that enable significant individual and community harm, as well as limiting the ability of services, regulators, and law enforcement to investigate wrongdoing.

Due to the lower risk of detection, perpetrators often prefer private environments, such as E2EE messaging services. In our statement, we noted how E2EE can prevent or limit the detection of online CSEA and the grooming of children by blocking technologies that can identify illegal material and activity. E2EE can conceal the production, storage, exchange, and proliferation of CSEA, enabling abuse to go undetected and allowing offenders to continue connecting with and abusing new victims. It can also expose survivors to ongoing trauma as images of their abuse continue to be circulated.

eSafety's position is that safety, privacy, and security are not mutually exclusive, and each can be maintained through thoughtful and intentional design. eSafety does not expect companies to design

systematic vulnerabilities or weaknesses into E2EE services. However, deployment of E2EE does not absolve platforms and services of responsibility for hosting illegal content and facilitating the sharing of CSEA and other harmful material. The position paper provides specific examples on how, by adopting a Safety by Design approach, E2EE services can make digital environments safer and more inclusive, especially for children.

Collaboration

As technologies evolve, regulators need to coordinate efforts and equip themselves with the necessary skills to address rapid developments. This includes expanding knowledge, enhancing tech testing capability, and developing auditing skills alongside other regulators and experts. eSafety recognises the importance of collaboration in improving online safety and works with stakeholders to raise online safety expectations globally.

In addition to eSafety being a member of INHOPE, the Commissioner is on the board of the WeProtect Global Alliance, which brings together over 280 members from governments, industry, civil society and intergovernmental organisations to develop policies and solutions in response to CSEA.

eSafety is a founding member of the Global Online Safety Regulators Network – the first global forum dedicated to supporting collaboration between online safety regulators. The Network connects regulators and observers with a shared commitment to a human rights-based approach to online safety regulation and harm prevention. The Network is also focused on promoting coherence and coordination to help prevent fragmentation that could hinder the effectiveness of global and local online safety efforts.

Conclusion

Experiences of online child sexual exploitation and abuse, and the services that are misused to facilitate this abuse, are not confined to national borders. CSEA is a global issue that is amplified by technology.

The tech sector plays an important role in shaping online environments and children's safety. Governments have an important role to set expectations for industry, as well as raising the awareness and capacity of communities to deal with risks, and implementing safety nets to minimise harm when risks occur.

To be a modern and effective regulator, we need to be an anticipatory regulator. Recognising this, the Australian Government is conducting a review of the Act, to consider optimal regulatory settings, and the need for flexibility in response to new and emerging technologies and harms. As part of this process, we look forward to continued engagement and consultation with international partners to achieve our shared goal of ending CSEA.

Our experience shows a multifaceted approach will lead to positive outcomes. We welcome other countries looking to our model, and the opportunity to collaborate with the globally to prevent CSEA in the digital environment.