Online Safety Codes

A submission by the
Alannah & Madeline Foundation

September 2022

Contents

Executive summary

At the Alannah & Madeline Foundation (the Foundation), we advocate for the rights of children and young people to be upheld online and offline, including by governments, industry bodies and digital platforms. We also empower children and young people to be positive digital citizens who can participate in society online responsibly, respectfully, and ethically.

We welcome the opportunity to comment on the Head Terms and eight draft codes developed by industry leaders to create improved industry standards to address Class 1A and 1B materials.* These codes were drafted under the terms of the Online Safety Act 2021 in response to guidance by the Office of the eSafety Commissioner in their Position Paper.

These are substantial documents of great public importance. As such, we urge that the consultation process be expanded to include more proactive and in-depth engagement with relevant public bodies, the not-for-profit community services sector, those with professional expertise in online safety and the expertise and insight of children and young people.

Our own submission focuses largely on the need to prevent and end child sexual exploitation material (CSEM), a subcategory of Class 1A material. This issue has particular importance to the Foundation's priorities of supporting children to recover and heal from trauma and bringing to life children's rights.

The Foundation recognises the strengths of the draft codes, including their close engagement with the Online Safety Act, the consistency of themes and approaches across the codes, and the clear guidance provided to help digital participants assess the risk levels of specific elements of their products and services.

However, we believe there are several ways in which the codes should be strengthened to help build a digital world that upholds the rights of children and young people. Specifically, we encourage:

- clear alignment of these codes with the work of relevant public entities
- greater clarity and rigor in risk assessment for industry participants
- stronger design measures to protect children, which are in line with international best practice
- clear measures of progress, outcomes and impacts and an undertaking to communicate these measures appropriately with the community
- alignment of industry requirements for reporting child sexual exploitation material with relevant Australian child safety legislation
- stronger framing of an ‘end goal’ for preventing and ending CSEM and upholding children’s right to live free from CSEM
- partnering with public entities towards an agreed outcome for appropriately detecting and removing first-generation CSEM† and contact between users that could facilitate the production of CSEM

* Class 1A is any material which: promotes or provides instruction of paedophile activity (‘child sexual exploitation’); advocates the doing of a terrorist act, including terrorist manifestos (‘pro-terror’); and/or describes, depicts, promotes, instructs or otherwise deals with matters of extreme crime, cruelty or violence (including sexual violence) without justification (for example, murder, suicide, torture and rape), (‘extreme crime and violence’). Class 1B is any material which: describes, depicts, expresses or otherwise deals with matters of crime, cruelty or violence without justification (‘crime and violence’); promotes, incites or instructs in matters of crime or violence (‘crime and violence’); and/or describes, depicts, promotes, instructs or otherwise deals with matters of drug misuse or addiction without justification (‘drug-related content’).

† ‘First-generation’ refers to material not previously identified and stored in an appropriately maintained NGO database. (See Explanatory Memorandum, p.8)

- reflecting the positive and proactive steps being taking internationally to protect children from online harms, including but not limited to, the requirements for reporting CSEM in the United Kingdom; the private default settings on the social media and gaming accounts of young people under 18 years of age; placing additional requirements on all services that are used by children, not just those that "allow children to have an account"; and stronger protective requirements around children's precise geographic location as in the United Kingdom, Ireland and California.

The drafting of the industry codes is a positive opportunity which we are keen to maximise to get the best outcomes for young Australians. Governments have an obligation to ensure businesses meet their responsibilities to respect and uphold children's rights and to prevent and remedy abuses of children's rights in relation to the digital environment.[1] As such, we hope to see a set of industry codes registered with eSafety which will have the greatest possible impact in helping create a digital world where children's rights are upheld.

About us

The Foundation was established the year after the Port Arthur tragedy, by Walter Mikac AM in memory of his two young daughters, Alannah and Madeline. Our vision is that all children and young people are safe, inspired and have freedom to flourish.

Over the last 25 years our work has grown and evolved but our purpose remains the same. We have three program streams:

- Safe and Strong: recovering and healing from trauma. Linked to our origin story, we have a specialist trauma recovery and therapy service for children who have experienced significant trauma. This has grown in recent years to include working with early childcare providers, kindergartens, and now primary schools to help them build their trauma informed capability and practices. Most of our work in trauma healing and recovery is Victorian based, with our therapists and consultants working from our clients' homes and places of work.

- Safe and Strong: building positive digital citizens. The Foundation works with schools, families and communities nationally to help children build the digital intelligence, skills and competencies they need to stay safe online and to be active, positive digital citizens. With over 10 years' experience working in the cyber bullying and wellbeing space, as technology has become ubiquitous, our work has developed into building digital intelligence, digital ethics and media literacy for all children aged 3-18.

- Safe and Strong: bringing children's rights to life. As a rights-based organisation, this is our policy and advocacy work. Since inception, we have advocated for firearms safety, and we convene the Australian Gun Safety Alliance. In other key policy matters related to our programs, we work closely with the Office of the eSafety Commissioner, the Prime Minister's National Office for Child Safety and other major agencies such as the Australian Federal Police.

In 2018, we partnered with Kate and Tick Everett, after the tragic suicide of their daughter, Dolly. With them we worked to establish Dolly's Dream.

- Safe and Strong: Dolly's Dream, changing the culture of bullying. The purpose is the same, but the programs and services (Parent Hub, telephone help line, school, and community workshops etc.) are specifically designed for remote, rural, and regional families and communities, to meet their unique needs and contexts.

Recommendations

Engage with experts to strengthen child safety measures

1. Extend the consultation period to include more proactive, targeted engagement with the not-for-profit sector, especially services which work with children. eSafety's position paper appears to provide for an extension of the consultation period beyond the minimum 30 days. Sector peak bodies such as the Australian Council of Social Service and the Centre for Excellence in Child & Family Welfare would be key points of engagement. **(eSafety position paper, pp.58-59; Head Terms, p.6)**

2. Ensure industry codes are informed by meaningful engagement with the National Office for Child Safety, National Children's Commissioner, and Human Rights Commissioner, if possible. These bodies could advise about options for aligning digital services accessed by children with the National Principles for Child Safe Organisations.

3. Confirm with the National Children's Commissioner (and/or the state and territory children's commissioners) that the industry codes' requirements about reporting identified child sexual exploitation material to law enforcement align appropriately with relevant child safety laws in Australian jurisdictions. For example, some Australian jurisdictions require that any adult who has formed a reasonable belief that another adult has sexually abused a child must notify police. Consider making reference in the codes to industry participants' obligations under child safety legislation, alongside the current reference to industry participants' obligations under privacy legislation. **(Schedules 1, 2, 3, Objective 1, Outcome 1)**

4. Amend the code review process to extend the minimum public consultation period to six weeks and include targeted, proactive engagement with the not-for-profit sector, especially organisations that work with children. **(Head Terms, p.16)**

Communicate alignment with the work of other key stakeholders

1. Represent clearly how industry codes align with the roles and legislative frameworks of relevant public entities (including but not limited to eSafety[‡]) and how the codes will better enable industry participants to work alongside public bodies to achieve agreed beneficial outcomes. Greater clarity about how the codes align with the work of the Australian Centre to Counter Child Exploitation would be particularly welcome.

Clarify and strengthen risk assessment requirements for industry participants

1. Provide greater detail about how industry participants should accurately assess their *overall* level of risk. This is in addition to the current guidance about assessing individual risk factors. **(Schedules 1-8)**

2. Amend the Head Terms to state that where a single electronic service could fall within the scope of more than one industry code, the industry participant will adopt whichever code provides the highest level of protection against the risks identified for that service. (**Head Terms, p.4**)

3. Clarify the expectations about risk assessment and compliance for an industry participant when that

[‡] For example, the Office of the Australian Information Commissioner, the Australian Communication and Media Authority, and the Australian Competition and Consumer Commission.

participant offers more than one digital product or service. We submit that each product or service should be assessed and reported upon at a tier level that is high enough to address the risks attached to that particular product or service. (**Head Terms, p.4**)

4. Incorporate 'age of end-users' or some equivalent measure into the risk assessment tables of the relevant Schedules, in order to better recognise the risks attached to products and services which are aimed at and/or widely used by children. **(e.g., Schedule 1, p.4; Schedule 2, p.6)**

Clarify and strengthen requirements for reporting and reviewing of codes

1. Work with eSafety to develop standardised, consistent reporting requirements under the codes, at least for Tier 1 industry participants. These requirements should be adapted to the eight industry Schedules. They should not be complex or onerous, but they should enable code reviewers and eSafety to identify and track key trends and changes over time.

2. Articulate measures or indicators of change to sit beneath the nine topics listed for minimum consideration at each code review, so that code reviewers and eSafety can make an accurate assessment of how industry outcomes are changing over time. **(Head Terms, p.16)**

3. Support partnerships between industry, eSafety, law enforcement, relevant NGOs and leading researchers to develop agreed measures or indicators of change in the safety of Australia's online environment. This should enable a meaningful estimate of the impact of the codes on public safety over time, to support continuous improvement.

Strengthen design measures to prevent harm to children

1. When requiring industry participants to adopt design measures intended to reduce children's risk of exposure, refer to international best practice. Demonstrate that the approach taken to protecting Australian children is as high, or higher, than that taken in other jurisdictions, within the limits of Australian legislation. For example, we suggest that the United Kingdom's Children's Code (Age-Appropriate Design Code) sets out more rigorous measures of 'default privacy settings' than those proposed in these draft industry codes. **(Schedules 1, 2 and 7, Objective 1 Outcome 1; Schedules 3 and 5, Objective 1 Outcome 2; Schedule 8, Objective 2 Outcome 7)**

2. Require that information provided to Australian end-users about child safety measures and risks, parental controls, reporting mechanisms, and the role of eSafety should be prominent, timely, concise, up-to-date, and appropriate to different ages and literacy levels. Ideally, such published terms would be developed via engagement with children, young people, parents, and carers. (**Eg. Schedules 1-8, Objective 2, Outcome 7)**

Partner to identify agreed directions for proactive detection of child sexual exploitation material

1. Undertake that industry participants will work with eSafety and the community towards an agreed outcome for appropriate detection and actioning of first-generation CSEM and contact between users that could facilitate the production of CSEM. We suggest a preferred approach would involve CSEM being identified and actioned via suitable, effective technology and appropriately qualified and skilled personnel, supported by adequate infrastructure. Such interventions should have the oversight of a trusted public entity, working within Australian law and in line with international best practice. The interventions should be resourced through partnerships between government and private industry.

Clarify and strengthen the framing of expectations within the codes

1. Consider creating separate items in the codes to address reporting obligations for CSEM and for pro-terror material, in recognition of the different ways Australian law treats these materials. **(Schedules 1, 2, 3, Objective 1, Outcome 1)**

2. Amend the codes to recognise the need to create and maintain a safe online environment for children, whether or not those children are Australian end-users of the specific digital platform. (Additional framing may be needed to recognise that industry participants will act reasonably and within the relevant legislation.) **(Head Terms and industry codes, Objective 1)**

3. Amend the codes to recognise that ending (not limiting) the hosting of CSEM should be the ultimate goal for industry participants. (Additional framing may be needed to recognise that industry participants will act reasonably and within the relevant legislation.) **(Head Terms and industry codes, Objective 1, Outcome 4)**

4. In dialogue with eSafety, consider updating Objective 2 of the codes to articulate an end goal of empowering individuals to avoid, report, and be supported to recover from any exposure to CSEM – not merely to 'manage' their own access and exposure. **(Head Terms and industry codes, Objective 2)**

5. Amend the codes to recognise that industry participants should ensure that any concerned individual should be able to access their reporting mechanisms, information, and tools about Class 1A and 1B material, and referrals to eSafety – whether or not that individual is an Australian accountholder or owner of the digital product. For example, parents should be able to utilise reporting mechanisms to raise concerns about their child's experience on a digital platform, without needing to open an account themselves. **(Head Terms and industry codes, Objective 2, Outcomes 7, 8, 9)**

Engagement with public bodies and the not-for-profit sector

While we welcome the opportunity to respond to the draft codes, we suggest that the current consultation period is not sufficient to enable meaningful engagement by the not-for-profit community services sector. We would welcome an expansion of the consultation to include more targeted, proactive engagement with this sector, which delivers a wide range of universal and targeted services and is estimated to invest $50 billion each year in the course of performing its work.[2] The sector operates diverse digital services, including online forums and online counselling services, and plays a unique role in supporting vulnerable Australians. The insights of the sector would help to enhance the clarity, utility, and ethical approaches of these codes.

Meanwhile, engagement with relevant public entities is also crucial. We welcomed eSafety's statement that their policy positions on industry codes were shaped by consultation with the Australian Communication and Media Authority, the Office of the Australian Information Commissioner, the Australian Competition and Consumer Commission, and the Department of Home Affairs.[3]

However, we submit that community confidence in the industry codes would be strengthened if more detail were provided about how the codes will enable industry participants to work more effectively alongside public bodies to achieve agreed beneficial outcomes – e.g. a reduction in CSEM on digital platforms.

This might be communicated through some clear representation of how the codes align with the work of relevant public entities (including but not limited to eSafety) and the legislative frameworks these entities operate within. For example, greater clarity around the relationship the Industry Codes under the terms of the Online Safety Act 2021, and their compliance with the Commonwealth Classification Act 1995, and the National Classifications Scheme.

Of particular concern is that the codes have the same restricted interpretation of ‘publication’ as Australia’s ‘National Classification Scheme’, not extending to include any and all text-based or image-based material, or other forms of material that might appear to fall within the scope of a plain reading of the definition of ‘publication’, including film, computer games, or advertisements for a publication, film or computer game. Australia's approach to classification needs to be reviewed to ensure that it is aligned to community expectations and can remain relevant in the face of rapidly changing technology.

It would also be useful to learn more about how the codes were informed by input from these public entities. In particular, it would be valuable to have more insight into how industry codes are informed by, and align with, the work of the Australian Centre to Counter Child Exploitation (ACCCE). We recognise that eSafety works closely with the ACCCE and that their consultation about industry codes with the Department of Home Affairs presumably addressed the ACCCE’s work. However, we would welcome more detail about how the industry codes will function in relation to this work.

This issue is particularly pertinent given that the draft codes address how social media services, relevant electronic services, designated internet services, and internet search engine services will work with law enforcement – e.g. in relation to notifications, information sharing, takedowns, collaboration in expert groups, and assisting end-users to report.[4] It would be beneficial for the community to be able to place these expectations within a broader picture of how industry participants (guided by the codes) work in a coordinated way with law enforcement and other public bodies.

In addition, we believe that community confidence in the industry codes would be strengthened if the code development process included meaningful engagement with the National Office for Child Safety, the National Children’s Commissioner and Australia’s Human Rights Commissioner. These public bodies could advise as to the possibilities for aligning digital spaces accessed by children with the National Principles for Child Safe Organisations. This set of voluntary principles was designed to be used (amongst other things) by businesses employing staff who provide services to and work with children and young people. [5]

Enhancing precision in risk assessment

As eSafety recognises in their position paper, the industry codes must acknowledge the diversity of size, maturity, capacity and capability of industry participants. To reach the agreed objectives and outcomes, industry participants are expected to take steps which are reasonable and proportionate, ‘based an assessment of the risk an industry participant's services and/or devices present in respect of class 1 and class 2 material’.[6]

However, we felt there was a certain lack of clarity in the advice provided to industry participants about assessing which code to work within. The Head Terms state:

- ‘For each online activity that they undertake,[§] each participant in the online industry must identify and comply with the industry code that applies to that online activity. Where a single electronic service could fall within the scope of more than one industry code, the relevant industry participant will only be required to comply with one code for that electronic service. The code that will apply in this situation is the code that is most closely aligned with the predominant purpose of the single electronic service.’ **(p.4)**

- no industry participant will have to comply with multiple codes in relation to the same electronic service. **(p.4)**

[§] An online activity is defined according to the Online Safety Act, section 134: an activity that consists of providing a social media service, relevant electronic service, designated internet service, internet search engine, app distribution service, hosting service, and/or internet carriage service to Australian end-users, and/or manufacturing, supplying, maintaining or installing of equipment in relation to the above services.

- where an industry participant must submit a code report to eSafety for multiple online activities, the industry participant may submit a consolidated code report that covers all those activities. **(p.15)**

We submit that 'predominant purpose' can be a rather subjective assessment and does not necessarily indicate the level of risk. Where a single electronic service could fall within the scope of more than one code, we would rather the industry participant adopt whichever code provides the highest level of protection against the risks identified for that service.

Greater clarity would also be welcome regarding how an industry participant should identify their code if they provide more than one service or product. We would hope that each service or product be accurately assessed and reported on according to its own risk level.

Moreover, we would welcome greater clarity as to how industry participants will identify their overall level of risk. The Schedules provide industry participants with guiding tables aligning risk factors with indicators of Tiers 1, 2 and 3. The measures on these charts are specific, clear, and useful. However, no guidance is provided as to how an industry participant should assess the *overall* risk level of their service or product. While we appreciate the difficulty of prescribing an overall measure, we also submit that the current gap in guidance poses a risk. Given the higher compliance levels required from higher risk industry participants, there would seem to be a disincentive for participants to grade their risk at Tier 1, without greater clarity about why they should do so.

In addition, while several of the codes encourage industry participants to consider the age of their end-users when assessing risk,[7] the age group of end-users is not listed with the risk assessment categories in the guiding tables.[**] We would prefer to see age of end-users incorporated into the risk assessment tables, especially given the recognition in eSafety's position paper that the nature of a user base (including whether it is targeted at children) is one indicator of level of risk.[8]

Measuring progress and impacts

We would welcome greater detail about how the compliance of industry participants with the codes will be measured, assessed, reported on publicly, and utilised to inform regular reviews of the codes.

The Head Terms (Objective 3, Outcome 11) require that 'Industry participants publish annual reports about class 1A and 1B material and their compliance with this Code'. For example, it is anticipated that Tier 1 social media services will submit code reports to eSafety with details of their risk assessment, steps taken to comply with minimum compliance measures, and explanation of why these measures are appropriate. **(Schedule 1, Social Media Services, p.17).**

However, in order for the progress of the codes to be recognised and strengthened over time, it is important to have standardised and consistent reporting requirements, at least for higher-risk industry participants. Presumably these should be adapted appropriately to the eight schedules. Reporting requirements need not be onerous, but they should enable code reviewers and eSafety to identify and assess trends in areas such as reporting, classification, actions taken, and complaints received, resolved and/or referred.[††]

We would also welcome greater detail about the approach which will guide the regular review of the codes. According to the Head Terms, each review will be coordinated by industry representatives responsible for developing the codes and 'will be based on the input of industry participants, eSafety, and other interested

[**] For example, social media services are encouraged to consider risk in relation to functionality, purpose, number of active monthly end-users, format of materials, and discoverability, while relevant electronic services are encouraged to consider potential for virality, intended audience, number of active end-users, and discoverability of users. (See Schedule 1, pp.45; Schedule 2, pp.6-7)

[††] When drafting reporting requirements, it might be useful to examine the measures published in eSafety's own annual reports and the 'annual highlights' document of the ACCCE – eg. complaints received, material removed, referrals provided, arrests, charges laid, reach of public communications.

stakeholders.' Drafts of revised codes will be published for comment, with the public invited to make submissions within a consultation period of at least 30 days. **(p.16)**

The current consultation process has been challenging for many not-for-profit community services to engage in effectively, given their limited resources and the number and complexity of the draft codes. Therefore, we suggest that the public consultation period for any revised codes should be longer and should include targeted, proactive engagement with the not-for-profit sector, especially organisations that work with children.

Meanwhile, the review process will also need clear, consistent, meaningful measures of change, so that code reviewers and eSafety can identify how industry has improved their practice by working within the codes. The Head Terms **(p.16)** lists nine topics that each code review will consider at a minimum. We submit that each of these topics should have measures to enable reviewers to assess how the codes have changed industry practice.

Ideally, though, measures of change would delve deeper. At present, the draft industry codes provide high-level measures of how industry participants' own processes will change. But no measures have been proposed to assess how these changes actually affect the safety of Australians online. We would like to see investment in partnerships between industry leads, eSafety, law enforcement, relevant NGOs and researchers (e.g. the Centre for Excellence for the Digital Child) to develop agreed measures or indicators of material impact.

Of course, this would be a complex and challenging task, and it would go beyond the steps articulated in eSafety's position paper. Nonetheless, we believe that community support for the codes would be stronger with clearer measures or indicators of the impact that the codes have had on the safety of Australia's online environment.

Furthermore, we suggest that certain measures of impact are needed to fully demonstrate industry participants' compliance with particular items in the industry codes – e.g. Objective 2, Outcome 9 of the Head Terms: 'Industry participants *effectively* respond to reports and complaints about Class 1A and 1B material' (p.11, our emphasis). Measures of impact would also seem to be needed for code reviewers to accurately assess the 7th item listed for examination in regular code reviews: 'How successful or unsuccessful this Code has been in preventing and mitigating harm'. **(Head Terms, p.16)**

It may not be viable to assess the impact of codes at the level of individual industry participants. But with appropriate investment, cross-industry data might perhaps be synthesised to inform the code review process, to help understand the overall progress that has been made and what changes may be needed.

Designing for child safety

In their position paper, eSafety articulates that, as far as practicable, the codes should address the facilitation of Class 1A material. 'The codes should include an outcome requiring industry participants to have scalable and effective policies, procedures, systems and technologies in place to proactively detect and prevent or mitigate contact between users which could facilitate the production of class 1 - 1A material. For example, contact or messaging involving the grooming of a child to facilitate the production of CSEM.'[9]

The industry codes make relatively few mentions of interventions to proactively detect contact which could facilitate the production of CSEM, although we do note the following:

- Use of technological tools to detect behavioural signals associated with CSEM – optional measure for relevant electronic services.[10]

- Deployment of technological tools designed to detect, flag and/or remove instances of known child sexual abuse material (CSAM)‡‡ from a service, including tools that identify phrases or words commonly linked to CSEM – minimum compliance measure for Tier 1 designated internet services.[11]

- Use of best efforts to ensure that search results seeking images of known CSAM are accompanied by deterrent messaging about risk and criminality and links to reporting mechanisms and support services – compliance measure for all providers of search engines.[12]

We recognise that proactive detection by industry participants of contact between users which could facilitate the production of CSEM is a contested topic, with legal, technological and ethical complexities. We do not feel well placed to comment in depth on this matter within the limited timeframe of this consultation. However, we would hope to see industry and government working together to identify a shared outcome for appropriately identifying and actioning these behaviours through effective technology, qualified and skilled personnel, and appropriate oversight.

In the meantime, we recognise that the draft industry codes require or encourage some design features intended to reduce the risk of children's exposure to CSEM. These involve:

- measures to ensure that only registered account holders can upload or distribute material on social media services and designated internet services[13]

- a minimum age for holding a social media account or designated internet service account, stated in the published terms[14]

- reasonable steps to prevent an underage child from holding a social media account or designated internet service account – approaches are not prescribed and could range from self-declaration to age estimation technologies[15]

- settings designed to prevent social media account holders from unwanted contact with other end-users[16]

- default settings for social media accounts that are designed to prevent children from unwanted contact from unknown end-users, including settings which prevent the child's location from being shared with other accounts by default[17]

- settings for relevant electronic services that enable blocking of messages, enable end-users to be hidden or appear offline, and, in the case of children's accounts, settings which make the accounts of under-16s private by default and prevent the child's location from being shared with any other accounts other than approved ones[18]

‡‡ The Head Terms document (p.8) defines known CSAM as 'material consisting of images (either still images or video images) that has been verified as child sexual abuse material and has been recorded on a database managed by a recognised child protection organisation that: (i) is designed to facilitate the identification of CSAM online; (ii) and which an industry participant is permitted to use for the purpose of utilising technological means to proactively detect such material on its service.' This sits within the broader category of CSEM, which the Head Terms document (p.23) defines as 'Class 1 material that: includes or contains the promotion or provision of instruction in paedophile activity; includes or contains descriptions or depictions of child sexual abuse or any other exploitative or offensive descriptions or depictions involving a person who is, or appears to be, a child under 18 years; or describes or depicts in a way that is likely to cause offence to a reasonable adult, a person who is, or appears to be, a child under 18 (whether or not the person is engaged in sexual activity).'

- age and/or content rating about third-party apps available on app distribution services[19]
- measure to ensure children cannot obtain an internet carriage service without parent / guardian consent[20]
- operating service providers for children's interactive devices set default safety settings to the highest level and make tools available to end-users to restrict unauthorised access to and operation of an adult's interactive device by a child.[21]

While these steps are positive, most are quite modest. We would welcome a reference to benchmarks of international good practice, to demonstrate that the approach taken to protecting Australian children is as high, or higher, than that taken in other jurisdictions.

For example, we note the much higher expectations about default privacy measures articulated in the United Kingdom's Children's Code (or Age-Appropriate Design Code). These state that default high privacy settings include that:

- services do not make children's personal data visible or accessible to other users unless the child amends their settings to allow this
- children are provided with age-appropriate explanations if they attempt to change their privacy settings, to mitigate risk
- children who do change privacy settings have the option of doing so temporarily, so that their account reverts to high privacy after the session
- children are not 'nudged' towards choosing a lower privacy option
- high privacy defaults are retained when software is updated
- different users can choose their own privacy settings on shared devices (e.g. children can have high privacy settings when using a parent's device)
- geolocation options are switched off by default – if tracking is active, it should provide an obvious sign to children and revert to default 'off' after each use.[22]

We welcomed the inclusion in the draft industry codes of requirements for industry participants to provide clear, easily accessible information for end-users (including parents and guardians) about child safety measures and risks, reporting mechanisms, and the role of eSafety.[23] Ideally, we would also like to see a commitment that these communications will be prominent, timely, concise, up-to-date, and appropriate for different ages and literacy levels.[24] Digital platforms' published terms can also be greatly enhanced when developed in consultation with children, young people, parents and carers.

Alignment with Australian legislation concerning reporting of child sexual abuse

We encourage industry leads drafting the codes to engage with the National Children's Commissioner (and, if she recommends it, with the state and territory children's commissioners) to confirm that the codes' requirements about reporting CSEM to law enforcement align appropriately with relevant child protection laws.

At present, the draft codes state that certain industry participants must notify 'appropriate entities' (foreign or local law enforcement or organisations acting in the public interest against child sexual abuse) if the digital provider has identified CSEM on its service and has formed a good faith belief that CSEM is evidence of a serious and immediate threat to the life or physical health or safety of an Australian adult or child. The provider is expected to ensure that such a report is compliant with their obligations under the Privacy Act.

These obligations are set out as a minimum compliance measure for all social media services and providers of search engines, for Tiers 1 and 2 electronic services, and for Tier 1 designated internet services.[25]

However, we would query whether these approaches align adequately with Australian legislation regarding reporting of child sexual abuse. We are concerned that under the codes, CSEM needs to reach a higher threshold before being reported.

Legislation varies between jurisdictions, but some states, like Victoria and New South Wales, require that any adult who forms a reasonable belief that an adult has sexually abused a child must report this to police. These laws do not require the adult to make their own judgement about the seriousness of the threat to the child's life or health. Nor does the adult's obligation vary according to the risk level of the environment where the abuse was identified.[26] However under the codes, rather than a suspicion of abuse, a service would need to believe this material also constituted 'a serious and immediate threat' to a child, in effect, weakening protections.

We recognise that aligning the obligations of industry participants with Australian laws regarding the reporting of child sexual abuse is a complex undertaking. Nonetheless, we believe it is an important one. At present, the draft industry codes specify industry participants' legal requirements under the Privacy Act. It would seem reasonable to also reference their obligations under child protection legislation.[§§]

Note: we recognise that in order to make reporting requirements workable, strong public investment is needed to ensure there is an expert, trusted workforce in place to appropriately assess and respond to such reports. We continue to advocate for such investment.

Finally, we recognise that the abovementioned requirements in the draft industry codes cover pro-terror material as well as CSEM. We are not well-placed to comment on pro-terror content. However, given that Australian jurisdictions have specific requirements about reporting child sexual abuse, we suggest it might be beneficial to create separate items in the codes for these two topics.

Framing interventions to prevent and end CSEM

There are several points in the draft codes where we suggest the framing of interventions in relation to CSEM could be strengthened. In order to do this, the codes may need to identify CSEM more clearly as distinct from other types of Class 1A and 1B material. Specifically:

- Objective 1 of the Head Terms and eight draft codes states 'Industry participants will take reasonable and proactive steps to create and maintain a safe online environment for Australian end-users.' We submit that children do not need to be end-users of a digital product or service in order to suffer harm as a result of CSEM (nor do they need to be Australian.) Naturally, there are practical and legal limitations to what industry participants can undertake to deliver, but the framing of 'reasonable steps' already implies recognition of these limits. We would welcome a clearer undertaking to create and maintain a safe online environment for children especially.

- Objective 1, Outcome 4 of the Head Terms and eight draft codes states 'Industry participants take reasonable and proactive steps to prevent or limit hosting of class 1A and 1B material in Australia.' This differs from eSafety's proposed outcome: 'Industry participants do not host class 1 and class 2 – 2A

[§§] Note: in this section, we refer to industry participants' obligations in relation to CSEM which has been identified – eg. via an end-user report. This section does not discuss any obligation of industry participants to proactively seek out first-generation CSEM for identification. We regard this as a different issue and address it elsewhere in this submission.

material in Australia.' Again, we recognise the practical and legal limitations to what industry participants can undertake to deliver. However, we suggest additional framing is needed to recognise that the ultimate goal in relation to CSEM, must be to end the hosting of this material, not merely to 'limit' it.

- Objective 2 of the Head Terms and eight industry codes states 'Industry participants will empower Australian end-users to manage access and exposure to class 1A and class 1B material.' We submit that 'managing access and exposure' is an inappropriate framing in relation to CSEM, given the illegality of the material and the fact that some end-users seek to 'manage' their engagement by seeking out, storing, and sharing this material. A more appropriate framing is needed in relation to CSEM – e.g. industry participants will empower individuals to avoid, report, and be referred to support to help them recover from any exposure to this material.

- Objective 2 also covers various mechanisms used by industry participants to provide pathways for reporting to platforms and to eSafety. We suggest these items would be strengthened by a guarantee that reporting / complaints mechanisms should be available to any person who has a concern about Class 1A or 1B material on the digital platform, whether or not that person is an Australian accountholder or owner of the product. (An example might be a parent who observes CSEM on the social media feed on their child's phone.)

Detection of first-generation CSEM

eSafety's position paper (Objective 1) states that industry participants will 'proactively detect' Class 1A material, including CSEM. The draft industry codes make more modest commitments. For example:

- Tier 1 social media services will deploy technological tools designed to detect, flag and/or remove from the service instances of known child sexual abuse material (CSAM). Tier 1 services will also make ongoing investments in tools and personnel that support the capacity of the provider to detect Class 1A material.[27]

- relevant electronic services may take the optional measure of using technological tools to detect and remove known CSAM, and/or to detect behavioural signals associated with CSEM.[28]

- Tier 1 designated internet services will use technological tools to detect and remove known CSAM. Tier 1 services will also invest in tools and personnel that support their capacity to detect known Class 1A material.[29]

- all internet search engines will delist search results that surface known CSAM.[30]

We recognise that use of technological tools by private industry to proactively detect first-generation CSEM is a controversial topic for legal, technological, and ethical reasons. Within the timeframe of this consultation, we do not feel well placed to comment in depth on this matter.

Ideally, we would like to see first-generation CSEM identified and actioned via appropriate, effective technology and appropriately qualified, skilled, and resourced personnel, supported by adequate infrastructure. Such interventions should be deployed, or at least regulated, by a trusted public entity working in line with international best practice and Australian law. These interventions should be resourced through partnerships between government and private industry.

Potentially the industry codes could be updated to include some undertaking to work with eSafety and the community towards such an outcome. This might perhaps sit beneath the undertaking in the Head Terms (p.16) that regular reviews of the codes will consider ‘any developments, including technological, that may impact the effective detection of material covered under this Code’.

We would welcome the opportunity to discuss any of these matters further.

Please contact:

Dr Jessie Mitchell,
Manager, Advocacy
jessie.mitchell@amf.org.au

Sarah Davies AM
CEO
sarah.davies@amf.org.au

Ariana Kurzeme
Director, Policy & Prevention
ariana.kurzeme@amf.org.au

[1] United Nations Convention on the Rights of the Child, *General comment No. 25 (2021) on children's rights in relation to the digital environment*, p.6

[2] Australian Council of Social Service, *The profile and pulse of the sector: Findings from the 2019 Australian Community Sector Survey,* 2020, https://www.acoss.org.au/the-profile-and-pulse-of-the-sector-findings-from-the-2019-australian-community-sector-survey/#endnote-002-backlink

[3] eSafety Commissioner, *Development of industry codes under the Online Safety Act: position paper*, September 2021, p.12

[4] Schedule 1 (Social Media Services) Objective 1, Outcomes 1 and 5; Schedule 2 (Relevant Electronic Services) Objective 1, Outcome 1; Schedule 3 (Designated Internet Services) Objective 1, Outcome 5; Schedule 4 (Internet Search Engine Services) Objective 1, Outcomes 1 and 5

[5] Australian Human Rights Commission, *National Principles for Child Safe Organisations*, Sydney, 2018

[6] eSafety Commissioner, *Development of industry codes under the Online Safety Act: position paper*, pp.45, 46, 49

[7] Schedule 1 (Social Media Services), p.4; Schedule 2 (Relevant Electronic Services), p.5

[8] eSafety Commissioner, *Development of industry codes under the Online Safety Act: position paper*, pp.50-51

[9] eSafety Commissioner, *Development of industry codes under the Online Safety Act: position paper*, pp.43, 74

[10] Schedule 2 (Relevant Electronic Services), Objective 1 Outcome 1, p.14

[11] Schedule 3 (Designated Internet Services), Objective 1 Outcome 1, p.9

[12] Schedule 4 (Internet Search Engine), Objective 1 Outcome 1, p.5

[13] Schedule 1 (Social Media), Objective 1, Outcome 1, p.10; Schedule 3 (Designated Internet Services), Objective 1 Outcome 2, p.11

[14] Schedule 1 (Social Media), Objective 1, Outcome 1, p.10; Schedule 3 (Designated Internet Services), Objective 1 Outcome 2, p.11

[15] Schedule 1 (Social Media), Objective 1, Outcome 1, p.10; Schedule 3 (Designated Internet Services), Objective 1 Outcome 2, p.11

[16] Schedule 1 (Social Media), Objective 1, Outcome 1, p.10

[17] Schedule 1 (Social Media), Objective 1, Outcome 1, p.9

[18] Schedule 2 (Relevant Electronic Services), Objective 1 Outcome 1, p.13

[19] Schedule 5 (App Distribution), Objective 1 Outcome 2, p.6

[20] Schedule 7 (Internet Carriage Services), Objective 1 Outcome 1, p.4

[21] Schedule 8 (Equipment), Objective 2 Outcome 7, p.9

[22] UK Children's Code or Age Appropriate Design Code, https://ico.org.uk/for-organisations/childrens-code-hub/

[23] For example, Schedule 1 (Social Media), Objective 2 Outcome 7 and 8, pp.14-15; Schedule 2 (Relevant Electronic Services, Objective 2 Outcome 7, Objective 3 Outcome 10, pp.17, 20; Schedule 3 (Designated Internet Services, Objective 2 Outcome 7, p.13; Schedule 4 (Internet Search Engine) Objective 2 Outcome 7, p.8; Schedule 5 (App Distribution Services) Objective 2 Outcome 7, p.7; Schedule 6 (Hosting Services) Objective 2 Outcome 7, p.6; Schedule 7 (Internet Carriage Services) Objective 2 Outcome 7, p.6; Schedule 8 (Equipment) Objective 2 Outcome 7, pp.8-9

[24] See for example 5Rights Foundation, *Tick to Agree: Age appropriate presentation of published terms* https://5rightsfoundation.com/in-action/tick-to-agree---age-appropriate-presentation-of-published-terms.html , and the UK Age Appropriate Design Code.

[25] Schedule 1 (Social Media Services) Objective 1, Outcome 1; Schedule 2 (Relevant Electronic Services) Objective 1, Outcome 1; Schedule 3 (Designated Internet Services) Objective 1, Outcome 1; Schedule 4 (Internet Search Engines) Objective 2, Outcome 9

[26] For example, NSW Government, 'New legislation to strengthen child sexual abuse laws: Factsheet summary,' 2018, https://www.justice.nsw.gov.au/Documents/Media%20Releases/2018/new-legislation-to-strengthen-child-abuse-laws-summary.pdf; Victorian Government, Department of Justice and Community Safety, 'Failure to disclose offence,' https://www.justice.vic.gov.au/safer-communities/protecting-children-and-families/failure-to-disclose-offence

[27] Schedule 1, Objective 1 Outcome 1

[28] Schedule 2, Objective 1, Outcome 1

[29] Schedule 3, Objective 1, Outcome 1

[30] Schedule 4, Objective 1, Outcome 1