Department of Police, Fire and Emergency Management

OFFICE OF THE COMMISSIONER AND SECRETARY

GPO Box 308 HOBART TAS 7001
Phone 1800 765 827
Email [redacted]

Our ref: A24/271234

The Director, Strategy and Research
Online Safety, Media and Platforms Division
Department of Infrastructure, Transport, Regional Development, Communications and the Arts

By email – OSAReview@communications.gov.au

Dear Director

TASMANIAN GOVERNMENT RESPONSE TO THE REVIEW OF THE *ONLINE SAFETY ACT 2021*

The Department of Police, Fire and Emergency Management, on behalf of the Tasmanian Government, welcomes the opportunity to comment on the operation and effectiveness of the *Online Safety Act 2021* to protect Australians from online harms. This is particularly necessary recognising the reliance we now have as a global community on the internet for daily life, coupled with the emergence of new technologies intentionally or unintentionally aiding more prevalent and intelligent online criminal activity or antisocial behaviour.

It is noted that the Act's objectives are to improve and promote online safety for Australians. The Act provides for the eSafety Commissioner and Commission, as well as complaints schemes, formalised arrangements for dealing with material depicting abhorrent violent conduct, industry codes of conduct and expectations, amongst other matters.

Tasmania's response is structured around the sections of the Issues Paper:

- Australia's regulatory approach to online services, systems, and processes.
- Protecting those who have experienced or encountered online harms.
- Penalties, and investigation and information gathering powers.
- International approaches to address online harms.
- Regulating the online environment, technology, and environmental changes.

It is requested that this submission is considered confidential and is not published.

Australia's regulatory approach to online services, systems, and processes

The Issues Paper notes that the Act has two schemes that take a systems-focused approach to preventing online harm: "Online Content Scheme" including industry codes and standards, and "Basic Online Safety Expectations". The Issues Paper outlines the services regulated by these codes, standards, and expectations, and seeks feedback on whether these measures are appropriate or should they be expanded.

In response, the Tasmanian Government provides the following advice.

Basic Online Safety Expectations regime

Summary current situation: the Basic Online Safety Expectations set out the Government's minimum safety expectations of online service providers to take proactive steps to protect Australians from online harm i.e. social media services, internet carriage services etc. Service providers are to report against the expectations to boost transparency. There are penalties for failing to comply with expectations outlined in the Minister for Communication's Determinations.

Tasmanian comment: It is considered that the Basic Online Safety Expectations regime could be strengthened by making it legally enforceable with consideration given to establishing mandatory minimum safety expectations for regulated online service providers, as well as others identified by the appropriate Minister. These mandatory minimum expectations could be reviewed in line with latest research regarding online harm on an ongoing basis e.g. every 3 years. Without penalties, the value of the Expectations may have limited impact.

The way industry is currently defined may no longer be fit for purpose or flexible enough for regulatory frameworks to apply to emerging 'services.' There remains a need to be able to expand sections as new technologies are developed and applied in emerging contexts. For example, the increased use and application of various artificial intelligence across online sectors (as defined) will need to be included, with scope for future modifications retained.

Online Content Scheme

Summary current situation: The Act requires that industry bodies or associations representing a regulated industry prepare draft codes to regulate "Class 1" and "Class 2" illegal and restricted online material. The eSafety Commissioner can then register codes developed by industry bodies or associations. Once registered, they are mandatory to comply with.

Tasmanian comment: It is considered appropriate for industry bodies and associations to initially draft industry codes and standards, with the eSafety Commissioner assuming that role in the absence of an appropriate industry body. It is presumed the registration process allows for them to be assessed and modified if necessary.

The focus of the codes/standards is limited to illegal or restricted material, and the National Classification Code (NCC) is referenced in this regard. All states and territories are part of the National Classification Scheme. The co-operative Scheme establishes the National Classification Board, which is responsible for classification decisions regarding films, computer games and certain publications. Classification decisions are made according to the principles set out in the NCC, which were agreed upon by the Commonwealth, state and territory governments.

It is noted that the Australian Government is also currently consulting on reforms to the National Classification Scheme to ensure it is fit-for-purpose for the modern media environment.

Consideration could be given to expand the principles set by the NCC, as not all online harms are currently captured e.g. technology-facilitated gender-based violence, volumetric attacks. These could be included in the Act, but not as an exhaustive list as the legislation needs flexibility to anticipate new forms of abuse emerging are able to be addressed.

There is merit in an enhanced duty of care requirement to all users. Both the United Kingdom model and the Work Health and Safety legislation model are precedence for this, and their intent is supported in principle.

Protecting those who have experienced or encountered online harms

Summary current situation: Under the Act there are four complaints and content-based schemes: the child cyberbullying scheme, the adult cyber-abuse scheme, the non-consensual sharing of intimate images ('image-based abuse') scheme, and the Online Content Scheme. Each focuses on specific types of harmful online material. The Issues Paper seeks clarification as to whether the thresholds set for each complaint scheme is appropriate and whether schemes are accessible and easy to understand for complainants.

Tasmanian response: The accessibility of a complaint scheme in the current Act is considered sufficient, however more could be done by service providers to make complaint pathways more responsive and transparent for complainants.

For example, vulnerable Australians at the highest risk of abuse should be prioritised to have access to corrective action through the Act. Anecdotal evidence from investigators suggests that victims are confused as to what schemes are available and the correct reporting platform, which highlights a support for more investment or prioritisation in education programs.

Further, the thresholds set for the following complaint schemes are considered inadequate and could be strengthened by factoring in the following components:

- Image-Based Abuse Scheme:
    - Bystanders should be empowered to report material.
    - The scheme should be altered to protect individuals whose images have been digitally altered, including deepfakes and photoshopped images.
    - Content should still be removed even if the depicted person previously consented to the provision of the intimate image on the service but changed their mind later.

- Child Cyberbullying
    - Bystanders should be empowered to report material.
    - Complainants should not be required to report concerns to the service provider before a removal notice can be issued. This opens them up to the potential of retribution and/or amplified abuse.
    - Complaints should be able to be made anonymously.
    - Complaints should be considered for "grooming" behaviour or suspected "grooming".

- Adult Cyber-Abuse
    - Bystanders should be empowered to report material.
    - Complainants should not be required to report concerns to the service provider before a removal notice can be issued. This opens them up to the potential of retribution and/or amplified abuse.
    - Complaints should be able to be made anonymously.

There are also ways that the eSafety Commissioner could manage harmful material in addition to blocking access, including:

- Increased efforts and oversight/monitoring by service providers with mandatory reporting and investigation requirements.
- Education for service providers and penalties for non-compliance.
- Greater and improved liaison with media around the reporting of harmful material.

Additional comment - Online Content Scheme; Illegal and Restricted Content

In relation to Illegal and Restricted Content Complaints, Tasmania's feedback is that these should extend to those made by parties outside of Australia as well.

Class 2 material should not be limited to that which is provided by a service in Australia or hosted in Australia and should be subject to requirements to restrict children's access to age-inappropriate content. Requiring service providers to proactively deliver supportive content for an agreed period after the transgression should be considered for all schemes.

Further, the use of artificial intelligence to automate the identification of unsafe content should be considered for all schemes, rather than relying on the individuals or bystanders to report it. The onus should then be on the provider to explain why their content is safe.

The current powers of the eSafety Commissioner could be extended further to address access to violent pornography through limiting child access to pornography. Consideration should be given to the options available to improve depiction of relational intimacy, safe sex, or the negotiation of consent. Class 2 material should not be limited to that which is provided by a service in Australia or hosted in Australia.

The current powers of the eSafety Commissioner should also be extended to address social media posts that boast about crimes. Feedback suggests the Commissioner does not have sufficient powers to address harmful material that depicts abhorrent violent conduct other than blocking access. There is an opportunity to consider proactive provision of support to individuals who have viewed the content. Notices should also require removal and remedial action.

The existing eSafety education and training activities in this space appears to be already extensive. Wide-spread education and training of peak players, users or interceptors (e.g. parents, school teachers) is vital, and enhanced responsibilities in this area will always help – particularly in being able to reach parts of the community that may not be captured within a focus/target group. The research capability of the Commission is very valuable and the timely release of or access to ongoing research would be helpful to those responsible for policy-setting and legislation development.

Recommendations for what could be done to further promote the safety of Australians online could include further research into the potential impacts of pornography depicting violent, sexist, and racist content, with aggression overwhelmingly directed toward women. More research into the use of artificial intelligence to automate the identification of unsafe content is encouraged.

Penalties and investigation and information gathering powers

Summary current situation: Under the Act penalties generally focus on financial penalties directed at individuals or platforms. The Issues Paper seeks feedback on whether the Act needs stronger investigation, information gathering and enforcement powers.

Tasmanian response: Under the Act, there is no legal obligation upon social media platforms (SMP) to have any user verification methods i.e. they have no idea who or where the person is located allowing unlimited trolling, doxing and abuse until the account is terminated, thus allowing an immediate new account to continue the pattern of behaviour. There is currently no legal obligation placed upon SMP to provide law enforcement with details to enable successful disruption or prosecution of unlawful behaviour online without lengthy legal requests and the like.

Investigators continue to face challenges investigating complaints as they have limited jurisdiction where the facility entity is based i.e. Meta (Facebook, Instagram) based in the United States of America. The existence of a local body whereby legal notices can be served would increase efficiency in investigations. Alternatively, legislative changes could be enacted to enable courts to accept

information from overseas under certain circumstances, or for overseas entities to have an Australian office that can produce evidence on their behalf.

There would be value in consideration of increased monetary penalties to increase motivation to prevent undesired behaviour. There also appears to be a need for stronger take-down powers, and penalties for refusal or ignoring a notice e.g. business disruption sanctions may be a way to obtain compliance. Some of the current penalties do not seem to reflect the seriousness of the harm versus the benefit for the business.

Regular follow-up on the offending behaviour leading to the harm, such as criminal charges, is important to help break the cycle of online crime.

The Act could also have stronger requirements for providers to retain identity information, as well as the ability to share this information with the Commissioner as part of their investigation, and further share that information with jurisdictional law enforcement as appropriate.

International approaches to address online harms

Summary current situation: The Issues Paper identifies online harm as a global problem in which many large providers of digital products and services operate overseas and in markets regulated by multiple governments. The Issues Paper seeks feedback on whether Australia should place additional statutory duties on online services to make online services safer and minimise online harms.

Tasmanian response: For these large providers to be held accountable, additional statutory duties should be imposed upon service providers operating in Australia. Online services should provide and comply with policies and procedures that outline their risk mitigation, investigation and reporting processes. They should also comply with all relevant Acts and policies in the jurisdiction that they operate, such as applying the principles of Tasmania's Child and Youth Safe Organisations Framework set out in the *Child and Youth Safe Organisations Act 2023*.

A mechanism to provide researchers and the eSafety Commission with access to data via mandatory reporting protocols where applicable would be desirable, such as child exploitation material or malicious actions/behaviour reported by users. This would assist to identify trends and areas of concerns so interventions can be put in place before issues become widespread. In addition, researchers/eSafety Commission should be able to access concerns/grievances including investigations and outcomes to allow for thorough investigations to occur.

Regulating the online environment, technology and environmental changes

Summary current situation: The review into the Act acknowledges that Australia is one of many countries regulating online safety in a global regulatory environment that is not confined to national borders. Feedback is sought as to whether Australia should consider a cost recovery mechanism on online service providers for regulating online safety functions.

Tasmanian response: The introduction of cost recovery mechanisms for online service providers is considered likely to be an effective way to assist in regulating online safety functions. Where online services are not complying with their own policies, procedures for risk mitigation and/or safety provisions, then they should be financially liable.

Cost recovery and industry contribution to the establishment and operation of its peak regulatory body occurs in many industries in Australia and it appears to be successful in international settings. Given the expanding burden of protectionist work on the eSafety Commission due to apparent industry apathy or misconduct, it would be reasonable to expect a contribution.

The eSafety Commissioner should have powers to act against content by individuals and groups. Content to be regulated would include content that is being used to denigrate, dehumanise, and maliciously affect individuals or groups including instigating and perpetuating violent and untruthful rhetoric narratives along with inflammatory commentary that is not sourced correctly or misidentified.

It is important that oversight is maintained regarding emerging technologies such as neuroethics, and the potential implications of misuse for those affected. For example, the collation of data for AI and the applications or misinterpretations of information generated from AI, particularly in a criminal context.

In conclusion, thank you for the opportunity to comment on the Review into the *Online Safety Act 2021*. If you have any questions on the feedback provided in this report, please do not hesitate to contact [redacted] who will be able to assist.

Yours sincerely

[redacted]

Donna Adams
COMMISSIONER OF POLICE
SECRETARY