Executive Manager, Investigations
Office of the eSafety Commissioner
PO Box Q500, Queen Victoria Building, NSW, 1230
submissions@esafety.gov.au

Saturday 11 September 2021

re: Restricted Access System call for submissions

Dear Executive Manager,

I am writing with the hope to contribute to the drafting of the Restricted Access System declaration.

Yours faithfully,
[REDACTED]

Preface: Online Safety Act 2021

The Online Safety Act 2021 ("the Act") purports to keep Australians safe online, but is a destructive and paternalistic piece of legislation that poses a serious threat to the internet as well as Australians' freedom of expression, access to information, privacy, security and democracy.

> "A key principle underlying the Act is that the rules and protections we enjoy offline should also apply online." *(Draft Online Safety (Basic Online Safety Expectations) Determination 2021 consultation[1])*

This premise of the Act is flawed for multiple reasons.

- It downplays the fact that the offline and online worlds are fundamentally different.
- It seems to imply that online lacks regulation, despite many existing laws that apply online.
- Many rules and protections introduced by the Act go far beyond those that apply offline.

Online spaces can be misused to cause direct harm to people, such as fraud, doxxing and psychological harm; and to coordinate, facilitate or induce offline dangers, such as violence and abuse. However, beyond that, the offline–online safety parity argument falls apart. Physical forces do not transmit through the internet, while information can be easily copied and tends to travel much faster and wider online. Furthermore, people may have online presence but fundamentally exist offline.

The government and service providers are not our parents. Governments have repeatedly abused their powers[2,3,4], and social media providers have demonstrated they are incapable of moderating content[5,6]. Despite that, the Act effectively puts the government (mainly the eSafety Commissioner) and service providers (by compulsion) in parenting roles over all Australians' use of online spaces. The Act also claims extraterritorial jurisdiction to impose its nanny-state measures onto service providers globally.

---

1 https://www.communications.gov.au/have-your-say/draft-online-safety-basic-online-safety-expectations-determination-2021-consultation

2 Ben Smee. "Queensland police officer receives suspended jail sentence for leaking woman's details to violent ex-husband". October 2019. https://www.theguardian.com/australia-news/2019/oct/14/queensland-police-officer-pleads-guilty-leaking-womans-details-violent-ex-husband

3 Bernard Keane. "Data retention scheme is being abused exactly as critics predicted". February 2020. https://www.crikey.com.au/2020/02/25/data-retention-scheme-abuse/

4 Forbidden Stories. "About The Pegasus Project". July 2021. https://forbiddenstories.org/about-the-pegasus-project/

5 Lacey-Jade Christie. "Instagram censored one of these photos but not the other. We must ask why". October 2020. https://www.theguardian.com/technology/2020/oct/20/instagram-censored-one-of-these-photos-but-not-the-other-we-must-ask-why

6 Onlinecensorship.org. "Offline-Online". https://onlinecensorship.org/content/infographics

The Act's approach is not only techno-solutionist[7,8] and extremely paternalistic, it also merely treats the symptoms of social problems but not their root causes. The best approach to harm reduction of online spaces without introducing other harms is debatable, but it is Australian communities and families who must assume primary responsibility to tackle social challenges and promote safe and disciplined use of technology.

Although the Act is capable of improving online safety in some ways, it also would put Australians at risk of other harms. The Act would:

- give the eSafety Commissioner and its delegates sweeping powers void of accountability and transparency, including immunity from liability,
- incentivise or effectively require (depending on level of enforcement) that services proactively remove material using automated tools[9],
- prohibit access to various material, notably "class 1" material, with insufficient exemptions,
- effectively require mechanisms such as facial recognition[10] to gate-keep "class 2" materials,
- systemically undermine end-to-end encrypted communications, and
- require that anonymous account users identify themselves or else face restrictions;

which would apply to almost every digital communication that touches Australia.

These measures risk introducing or worsening material and psychological harms such as rampant censorship, chilling effects, privacy violations, normalisation of surveillance and extreme security risks. The Act risks making the internet a hostile and potentially dangerous place for Australians, especially for marginalised and vulnerable groups of people[11]. Furthermore, the sweeping and unaccountable censorship powers would undermine Australians' freedom of expression and access to information, which are vital for democracy.

It appears that the Act's delegated legislation is being used tactically as a legislative backdoor to introduce unpopular measures the government wants without parliamentary or citizen scrutiny. A notable example is backdoors into end-to-end encryption, which Five Eyes governments are fiercely pushing for[12,13] despite persistent opposition by Australians[14] and civil society worldwide[15]. The Act and its associated bills and explanatory memoranda make no mention of "encryption", but after the Act was solidified into law, "encryption" appears in the exposure draft of the Basic Online Safety Expectations. This raises serious concerns that the Act is drafted and intended to introduce backdoors into end-to-end encryption *by design*.

---

7 Ashali Bhandari. "Feminist Perspectives on Space, Safety and Surveillance: Improving a Woman's Right to the City". March 2021. https://thewire.in/women/feminist-perspectives-on-space-safety-and-surveillance-improving-a-womans-right-to-the-city

8 Digital Rights Watch. "Policy grounded in surveillance won't protect women". June 2021. https://digitalrightswatch.org.au/2021/06/16/policy-grounded-in-surveillance-wont-protect-women/

9 Electronic Frontier Foundation. "Facebook's Most Recent Transparency Report Demonstrates the Pitfalls of Automated Content Moderation". October 2020. https://www.eff.org/deeplinks/2020/10/facebooks-most-recent-transparency-report-demonstrates-pitfalls-automated-content

10 Ariel Bogle. "Porn age filter for Australia recommended by parliamentary committee". March 2020. https://www.abc.net.au/news/science/2020-03-05/age-verification-filter-for-online-porn-recommended-in-australia/12028870

11 Grace O'Brien. "Racial Profiling, Surveillance and Over-Policing: The Over-Incarceration of Young First Nations Males in Australia". Published February 2021. https://www.mdpi.com/2076-0760/10/2/68

12 Department of Justice (US). "International Statement: End-To-End Encryption and Public Safety". October 2020. https://www.justice.gov/opa/pr/international-statement-end-end-encryption-and-public-safety

13 Julie Inman Grant. "End-to-end encryption: a challenging quest for balance". February 2020. https://www.esafety.gov.au/about-us/blog/end-end-encryption-challenging-quest-for-balance

14 Digital Rights Watch. "Research shows Australians deeply concerned by Government's spyware legislation". October 2018. https://digitalrightswatch.org.au/2018/10/25/research-shows-australians-deeply-concerned-by-governments-spyware-legislation/

15 "Security for All". https://securetheinternet.org/

The Act takes a disproportionately punitive approach towards end-users and service providers. Alarmingly, service providers and end-users could be fined 500 penalty units (currently $111,000) and potentially face other punishments in relation to several parts of the Act, and service providers must be able to comply with notices within 24 hours to avoid the civil penalty. Appropriate guidance or warnings, not punishments, should be given to end-users and service providers who put effort into corrective action or compliance with the Act. Punishments should be proportionate and reserved for negligence, refusal to comply and repeat offences.

Although the Act provides a few commendable measures for online safety, I strongly believe the Act must be repealed in its entirety. The government must engage with Australians in a genuine and meaningful dialogue about how it can play a role to reduce online harms, while preserving Australia's human rights, civil liberties, democracy and security, not simply ram the laws it wants through parliament. I make my submission in relation to the Restricted Access System while maintaining my position of opposition to the Act.

Restricted Access System

According to the Act, "class 2" material (broadly speaking, X18+ and R18+ material) must be subject to a Restricted Access System ("RAS") that restricts access by people under the age of 18 ("minors") to such material.

Below, I express my concerns, recommendations and suggestions in relation to the RAS declaration.

Overview of concerns

With the aim of "protecting children from exposure to material that is unsuitable for children", the RAS scheme could risk facilitating and normalising mass surveillance and control of Australians, as well as expose Australians to extreme security risks and wrongly restrict access to information.

The Act appears to incentivise or effectively require an automated mechanism such as biometric recognition or identity document scanning. Using facial recognition technology for the purpose of age verification[16] was suggested by the Department of Home Affairs, in charge of a centralised database "The Capability"[17] designed to catalogue biometric, identity and other information on every Australian.

Compelling the use of an RAS with such characteristics is unacceptable and disproportionate to the purpose of restricting access by minors to material. An RAS that fails to prevent determined minors from accessing age-restricted material but is "good enough" while respecting privacy and security would be better than an RAS that strives for 100% effectiveness at all costs.

Biometric recognition and identity document scanning

Measures such as biometric recognition (such as face, eye, voice, body proportion or fingerprint) and identity document scanning are highly intrusive and pose an extreme security risk to end-users.

Subjecting people to measures such as biometric recognition or identity document scanning is inhumane and should not be the default way of Australian society. Citation of ID solely for age verification purposes by a brick-and-mortar newsagent selling age-restricted material may be acceptable. However, scanning end-users' biometrics and identity documents, especially in an online environment and at scale, would pose serious concerns of intrusive data collection and

16 Chris Duckett. "Home Affairs pushes its face-matching service for porn age verification". October 2019. https://www.zdnet.com/article/home-affairs-pushes-its-face-matching-service-for-porn-age-verification/

17 Simon Lauder. "The Capability: Government's national facial recognition plan raises privacy concerns". December 2015. https://www.abc.net.au/news/2015-12-17/governments-facial-recognition-system-sparks-privacy-concerns/7035980

normalisation of surveillance in Australian society. Such measures would risk making the online world a "show me your papers please" society for Australians, while biometric scanning would also desensitise Australians to encroachment of pervasive biometric surveillance in physical public spaces[18,19].

Collection and storage of biometric or identity information, especially by electronic means, carries risks of such information being abused, hacked or otherwise breached. When (not if) such incident occurs, the damage done to affected end-users would likely be serious, even catastrophic[20,21]. Depending on the nature of such information (name, birthdate, residential address, identity document number, facial photo, etc.) and how much information is compromised, end-users could become victims of identity fraud, deepfakes[22], stalking, violent attacks or other forms of mistreatment. Post-incident, affected end-users would need to obtain new identity documents or even change their name or residence, and since human biometrics cannot be swapped there is no way to remedy leakage of biometric information. The RAS declaration must not cause handling of biometric or identity information.

Privacy risks due to online tracking

Australians should be able to use the internet with privacy and anonymity. Privacy and anonymity allow people to live in society without their activities or persons being targetted by adversarial actors. Privacy allows people to access information, form their own opinions, exercise creativity and communicate without being subject to external scrutiny. Anonymity allows people to express their genuine opinions and political beliefs to the world without fear of retribution; this is why Australian elections use the secret ballot.

Any RAS interaction with The Capability could make the government capable of tracking the activities of Australians who engage with the system. Similarly, an RAS could make RAS operators or service providers capable of tracking Australians' online activities that involve RAS gate-keeping. The RAS declaration must not require nor facilitate surveillance of the online activities of Australians.

Implications affecting access to information and freedom of expression

Although the discussion paper asserts "The purpose of a RAS is not to prevent access to age-restricted content, but to ensure access is limited to people who are 18 years and over.", an RAS could in practice prevent many Australian adults from accessing material. People who have privacy or security concerns, or object to biometric recognition or identity document scanning, would be most affected by access restrictions, unless the RAS declaration properly addresses these concerns.

The Act allows the Commissioner broad discretion to declare unclassified material as class 1 (RC) or class 2 (X18+ and R18+) material. This risks unclassified material being wrongly censored or restricted by "over-classification". Instead, there should be a requirement that if the Commissioner believes that an unclassified material would be classified in a way that makes the material fall under class 1 or class 2, the Commissioner must apply to the Australian Classification Board for classification of the material.

---

18 Perth council facial recognition trial greeted with concern and scepticism. "Elise Thomas". June 2019. https://www.theguardian.com/technology/2019/jun/12/perth-councils-facial-recognition-trial-accused-of-blanket-surveillance

19 Josh Bavas. "Facial recognition quietly switched on at Queensland stadiums, sparking privacy concerns". June 2019. https://www.abc.net.au/news/2019-06-05/facial-recognition-quietly-switched-on-at-queensland-stadiums/11178334

20 Kevin Nguyen. "NSW driver's licence data breach left Sydney health worker 'sickened'". September 2020. https://www.abc.net.au/news/2020-09-02/sydney-man-finds-own-driver-licence-in-nsw-data-breach/12616606

21 Paul Farrell. "The Medicare machine: patient details of 'any Australian' for sale on darknet". July 2017. https://www.theguardian.com/australia-news/2017/jul/04/the-medicare-machine-patient-details-of-any-australian-for-sale-on-darknet

22 ABC Foreign Correspondent. "American Deepfake". June 2021. https://www.abc.net.au/foreign/american-deepfake/13423468

Recommendations

Recommendation 1: Require every RAS to be necessary for and proportionate to the purpose of restricting access by minors to material[23], with proper consideration of privacy and security risks, risks of wrongly denying access to information and limiting freedom of expression, and financial and administrative burden.

Recommendation 2: Prohibit every RAS from using biometric recognition, identity document scanning and similar mechanisms that are inhumane or pose privacy or security risks for end-users.

Recommendation 3: Prohibit every RAS from involving or interacting with The Capability or any other database that catalogues the biometrics, identities, activities or private lives of Australians.

Recommendation 4: Require every RAS to collect no information other than a proof from the end-user that the end-user is not a minor. In particular, prohibit collection of non-essential information and end-user activity information.

Recommendation 5: Require every RAS to use transparent standards, and use free and open-source software instead of proprietary software. Additionally, all software distributed over a network for immediate use (such as on an end-user's device) must be copyleft licensed such that the network-used software's freedoms are protected[24].

Recommendation 6: Require the RAS scheme to be subject to multi-stakeholder and parliamentary oversight.

Recommendation 7: Do not require any RAS to restrict access to unclassified material. Unclassified material that would be classified in a way that makes it class 1 or class 2 should be properly classified before its access is prohibited or restricted.

Examples of mechanisms that may be acceptable

I propose a few examples that (to some degree) limit access by minors to age-restricted material, or make age-restricted material clearly presented as age restricted.

- Affirmative declaration of age by end-users: Prior to account activation or access to age-restricted material, service providers could obtain affirmative declaration that the end-user is not a minor, such as a check box that the end-user must click on.
- Segregation of age-restricted material from other material: Service providers could segregate age-restricted material from other material, such that access to segregated material is strictly opt-in and comes with clear and appropriate content warnings. Segregated material may optionally be further segregated by category (such as sex, gambling, drugs, violence, etc.) so that fine-grained opt-in is possible.
- Parental controls: Service providers could introduce parental controls that allow minors to interact with the service in a limited manner while remaining under parent/guardian control and monitoring.
- Offline visual ID citation: For services that also have offline interaction with end-users, service providers may be able to perform offline visual ID citation for the sole purpose of verifying that an end-user is not a minor and then add that property to the end-user's online account by hand.
- Literacy or numeracy testing: Requiring a correct response to an unguessable reading, mathematics or other problem pitched at an appropriate level could serve as age assurance.

23 Necessary and Proportionate. https://necessaryandproportionate.org/principles/

24 Choose a License. https://choosealicense.com/appendix/#network-use-disclose