Online Safety Amendment (Social Media Minimum Age) Bill 2024
Response by Cybersecurity Research Group, University of Melbourne

Prof. Benjamin I. P. Rubinstein, A/Prof. Olga Ohrimenko, Dr. Andrew C. Cullen, Dr. Shaanan Cohney,
Dr. Chris Culnane A/Prof. Toby Murray, Dr. Marc Cheong, Dr. Thuan Pham, A/Prof. Xingliang Yuan

Friday November 22, 2024

The proposed amendments seek to address social media use by children and younger teenagers. In this response[1] we focus only on the privacy implications of the proposed legislation.

A lack of specificity in the Bill's implementation language obscures privacy risks

The Bill does not specify how platforms must comply with the minimum age obligation, deferring to a "reasonable steps" test. The implementation details have profound implications to digital privacy and should not be left as an afterthought**.**

The *Exploratory Memorandum* describes a planned Age Assurance Trial that:

> "*...examines available age assurance technologies and methods – such as age verification, age estimation, age inference, parental certification or controls...*"

While this list reflects a range of approaches, we believe these technologies will significantly degrade privacy. These implications will impact all Australians who live and work online, not only those below the minimum age. As researchers working in the field of data privacy and cybersecurity, all methods we know, for verifying age, will magnify privacy vulnerabilities faced by all Australians.

Biometric-based age assurance requires disclosure of hard-to-revoke, highly sensitive information. Such technologies are also unreliable.

The technologies broadly grouped as biometrics-based, where images or other data captured from the user, involve a special kind of personal information. Biometric data cannot be reset like a password if unintentionally released in a data breach and follow an individual for the rest of their life. Most social media platforms in popular use today are based overseas, and any possession of highly sensitive data may be accessible from overseas. False positives and false negatives (erroneously misidentifying individuals or their age) are inherent to biometrics-based approaches[2]. The approaches are susceptible to artificial intelligence evasion by techniques

[1] We note with alarm that respondents were afforded only 1 day. The proposed changes have significant implications to digital human rights. We implore Government to consider fulsome consultation in the Age Assurance Trial.
[2] Reliable benchmark for the state-of-the-art age estimation at time of writing (https://archive.is/ike5I) indicate that the best technology guesses on average only within 3.7 years of someone's true age. This technology performs even worse when evaluated against individuals between the ages of ten and twenty. See https://arxiv.org/pdf/2307.04616v2 (Analysis of Images, Social Networks and Texts 2023) Table 2 and Figure 8

we study, known as adversarial examples[3] , and suffer from collusion and inversion attacks, e.g., recovering biometric data from neural hash[4].

Non-biometric approaches create and centralise severe risks of data breaches

Non-biometric approaches necessarily require implementation of a digital ID or similar. While cryptographic tools can be used to improve the security and privacy of verifying identity and age, it is very difficult to avoid relying on a centralised verification service. A centralised verification service would inherently be able to track and surveil every Australian's online activity. Further, the service would be a lucrative and attractive target for state actors and organised crime. In our expert view, the implementation of this Bill will likely erode the digital privacy of many Australians.

The privacy protections of the Bill are not sufficient to prevent significant data collection

The protections on the use of personal information disclosed through age assurance are welcome, however, any age assurance process is likely to necessitate the collection of equivalent personal information during account registration, which would not be covered by the protections. Notably, an age assurance check is only effective where one validates that the party undergoing assurance checks is the party represented by the online account. *i.e.,* the person undergoing the check is the same person tied to the account. Failure to do so may result in pressure being applied on parents to perform age assurance on their children's accounts. It may also create a market for 16-year-olds to provide age assurance to younger peers. Social media services will likely respond by soliciting identity information during registration–further curtailing the ability of Australians to engage online without divulging their identity. Depending on the technology adopted, particularly where identity is linked to assurance, APP 2 may no longer even apply.

The privacy protections of the Bill would likely not be adequately enforceable or sufficiently deter misuse of collected data

While the Bill places significant penalties on platforms that breach the *Privacy Act 1988,* legislation has not kept pace with either modern online practice, nor how attacks and privacy-enhancing technologies have evolved. The Bill would require that platforms only use age assurance data for the purposes of age assurance (unless consent is given otherwise). We find it difficult to imagine how this would be policed where misuse of personal information is already hard to detect.

As experts within this field, we are concerned that these factors have not been given sufficient consideration. The proposed amendment to the online safety act introduces real and significant risks to all Australians that significantly outweigh the proposed benefits.

[3] A. D. Joseph, B. Nelson, B. I. P. Rubinstein, and J. D. Tygar. *Adversarial Machine Learning*. Cambridge University Press (2019)
[4] J. Madden, M. Bhavsar, L. Dorje, X. Li. "Assessing the Adversarial Security of Perceptual Hashing Algorithms." arXiv preprint arXiv:2406.00918 (2024)