Online Safety Act 2021 Review Submission 20 June 2024

We appreciate this opportunity to make a submission to the Online Safety Act 2021. Our submission will follow the following format:

1. **Summary Statements**
2. **Comments relating to the Summary Statements are made with reference to "Part 7 – Summary of consultation questions included in this paper".**

1. Summary Statements

a. There is a mismatch between Tech Company values and community values. Tech Companies provide services and amenities for the community and should not act or consider they have a licence to act in ways that contradict and disregard community values, despite what they proport to manage as per their terms and conditions.

b. The underlying issue is the Tech Companies business model, that creates a conflict of interest whereby these Companies do not effectively moderate their platforms. The underlying business model is the problem, namely:
   i. Via the illusion of free services, the aim is to have total market dominance through Behaviour Modification of any person using these platforms.
   ii. Behaviour Modification enables monetising online posts and use of apps. (Ref: https://journals.sagepub.com/doi/full/10.1177/2053951718820549) (Ref: Yanis Varoufakis 'Techno Feudalism')
   iii. Behaviour Modification enables monetising online activity, which in turn gives owners of Tech Companies an "Extractive Power" to command those lacking digital control (market dominance), and to generate surplus value for their owner's benefit. This power facilitates Tech Companies dictating terms of use to businesses that access their services, as well as potentially constraining the extent to which Governments perceive they can regulate the Tech Companies without risking community & business users negative perceptions of the impacts on them and their data. (Ref: Yanis Varoufakis 'Techno Feudalism')

c. Prevention is as important as penalties and sanctions for Tech Companies. Any financial penalties are ideally severe enough and linked to a Company's Global Gross Turnover to encourage compliant governance.

2. **Comments:** with reference to "Part 7 – Summary of consultation questions included in this paper" the following numbered comments are made adjacent the relevant questions, and listed below:

Part 2 – Australia's regulatory approach to online services, systems and processes

1. Are the current objects of the Act to improve and promote online safety for Australians sufficient or should they be expanded? Ref Comments 1 and 2
2. Does the Act capture and define the right sections of the online industry?

3. Does the Act regulate things (such as tools or services) that do not need to be regulated, or fail to regulate things that should be regulated? Ref Comment 1, 2, 3, 6, 7 and 8
4. Should the Act have strengthened and enforceable Basic Online Safety Expectations?
5. Should the Act provide greater flexibility around industry codes, including who can draft codes and the harms that can be addressed? How can the codes drafting process be improved?
6. To what extent should online safety be managed through a service providers' terms of use? Ref Comment 3, 6 and 7
7. Should regulatory obligations depend on a service providers' risk or reach?

Part 3 – Protecting those who have experienced or encountered online harms

8. Are the thresholds that are set for each complaints scheme appropriate?
9. Are the complaints schemes accessible, easy to understand and effective for complainants?
10. Does more need to be done to make sure vulnerable Australians at the highest risk of abuse have access to corrective action through the Act?
11. Does the Commissioner have the right powers to address access to violent pornography?
12. What role should the Act play in helping to restrict children's access to age inappropriate content (including through the application of age assurance)?
13. Does the Commissioner have sufficient powers to address social media posts that boast about crimes or is something more needed?
14. Should the Act empower 'bystanders', or members of the general public who may not be directly affected by illegal or seriously harmful material, to report this material to the Commissioner?
15. Does the Commissioner have sufficient powers to address harmful material that depicts abhorrent violent conduct? Other than blocking access, what measures could eSafety take to reduce access to this material? Ref Comment 1
16. What more could be done to promote the safety of Australians online, including through research, educational resources and awareness raising? Ref Comment 1

Part 4 – Penalties, and investigation and information gathering powers

17. Does the Act need stronger investigation, information gathering and enforcement powers?
18. Are Australia's penalties adequate and if not, what forms should they take?
19. What more could be done to enforce action against service providers who do not comply, especially those based overseas? Ref Comment 1
20. Should the Commissioner have powers to impose sanctions such as business disruption sanctions?

Part 5 – International approaches to address online harms

21. Should the Act incorporate any of the international approaches identified above? If so, what should this look like?
22. Should Australia place additional statutory duties on online services to make online services safer and minimise online harms?
23. Is the current level of transparency around decision-making by industry and the Commissioner appropriate? If not, what improvements are needed?
24. Should there be a mechanism in place to provide researchers and eSafety with access to data? Are there other things they should be allowed access to?
25. To what extent do industry's current dispute resolution processes support Australians to have a safe online experience? Is an alternative dispute resolution mechanism such as an Ombuds scheme required? If so, how should the roles of the Ombuds and Commissioner interact?
26. Are additional safeguards needed to ensure the Act upholds fundamental human rights and supporting principles? Ref Comment 4

Part 6 – Regulating the online environment, technology and environmental changes

27. Should the Commissioner have powers to act against content targeting groups as well as individuals? What type of content would be regulated and how would this interact with the adult cyber-abuse and cyberbullying schemes? Ref Comment 5
28. What considerations are important in balancing innovation, privacy, security, and safety?
29. Should the Act address risks raised by specific technologies or remain technology neutral? How would the introduction of a statutory duty of care or Safety by Design obligations change your response? Ref Comment 1, 5, and 6
30. To what extent is the Act achieving its object of improving and promoting online safety for Australians?
31. What features of the Act are working well, or should be expanded?
32. Does Australia have the appropriate governance structures in place to administer Australia's online safety laws?
33. Should Australia consider introducing a cost recovery mechanism on online service providers for regulating online safety functions? If so, what could this look like? Ref Comment 1

Comment 1

We note the role of the eSafety Commissioner as the independent regulator for online safety, and we support the Commissioner's Mission and eStrategy 2022-2025.

We believe an industry Self-Regulated Code of Practice is not adequate.

Online platforms are known to have the ability to monitor their practices but choose when to do so, with examples of failing to police their own terms and conditions.

Ref: https://www.tandfonline.com/doi/full/10.1080/10383441.2022.2138140

A mandatory and regulated Code of Practice is needed for Companies to operate with a Social Licence, and a Duty of Care and Safety Design obligations are essential elements.

Such a Code of Practice needs to apply to Companies no matter what technology they use.

Because of the evolving and developing nature of technology, an evolving and developing Code of Practice is required. Companies should be mandated to contribute funding to a Research Foundation for the purpose of research/publishing about both the good and the harm of online language and behaviour.

In addition, Companies need to demonstrate use of current data/research from such a Research Foundation for management of online language/behaviour. Any institution could have differing models, e.g., it could focus on research only, but more importantly be involved in monitoring and managing Company practices.

Examples of Research Groups

The Social Media research Institute https://www.smri.world/about-us/history

The Swinburne University Social Innovation research Institute https://www.swinburne.edu.au/research/institutes/social-innovation/

Harvard University Berkman Klein Centre

https://cyber.harvard.edu/story/2021-07/berkman-klein-center-to-launch-three-year-institute-for-rebooting-social-media

Harvard University Berkman Klein Centre Institute for Rebooting Social Media

https://cyber.harvard.edu/programs/institute-rebooting-social-media

Critically, these practices outlined above need to be regulated by an Independent Auditor, independent of Government, but answerable to Parliament, e.g., an Online E & AI Safety Institute.

Companies that comply and are cooperative will help develop better practices for all involved.

A regulatory Body is required to manage Companies that do not comply or obfuscate the process outlined above. Such regulation needs to involve severe penalties including removing the Social Licence, e.g., deregistration. The regulatory Body could be established with a Social Accountability Act which includes a diverse group of community stakeholders that grade corporations according to an index of social worthiness. Consistent low levels of social worthiness would trigger a public inquiry.

Ref Yanis Varoufakis "Techno Feudalism" page 197.

(We note the Online Safety Research Program in the Dept of Infrastructure ended 30-06-2023.)

Comment 2

Online harm and mental health symptoms have been well documented.

Ref: UNSW Sydney – Prof Michael Slater School of Social Sciences.
https://www.researchgate.net/profile/Michael-Salter-2

Ref: https://www.abc.net.au/news/2024-05-30/short-statured-community-call-out-online-abuse-facebook-groups/103896180

Ref: https://www.abc.net.au/news/2022-10-04/mental-health-teens-screens-research/101495990

Harmful language online is the symptom.

The underlying cause is the business model of the online platforms whose aim is to monetise the production of behaviour modification.

Ref: https://hbr.org/2022/01/the-psychology-of-your-scrolling-addiction

Ref: Yanis Varoufakis 'Techno Feudalism' page 176.

> To quote 'Techno Feudalism' page 178-179: 'If fascism taught us anything, it is our susceptibility to demonising stereotypes and the ugly attraction of emotions like righteousness, fear, envy and loathing that they arouse in us. ……. the internet brings the feared and loathed other closer, right in your face. And because online violence seems bloodless and anodyne, we are more likely to respond to this other online with taunting, inhuman language and bile. Bigotry is technofeudalism's emotional compensation for the frustrations and anxieties we experience in relation to identity and focus. Comment moderators and hate-speech regulation can't stop this because it is intrinsic to cloud capital, whose algorithms optimise for cloud rents, which flow more copiously from hatred and discontent.'

Ref: https://www.abc.net.au/news/2024-05-17/eu-says-instagram-and-facebook-are-too-addictive-in-probe/103859680

Ref: https://www.science.org/doi/10.1126/sciadv.add8080

Another example of Companies conflict of interest regarding monitoring their sites adequately is the "Rabbit Hole Effect" resulting from the algorithms Companies design, such that users follow algorithmic recommendations to videos more extreme than the video they were watching.

Ref: https://cyber.fsi.stanford.edu/news/study-finds-extremist-youtube-content-mainly-viewed-those-seeking-it-out

Comment 3

Further to the fundamental issue of the online platform business models causing harmful online language, is the illusion of their free services. These 'free' services could be replaced with a micropayment platform, e.g., Netflix's subscription model combined with the British National Health Services' principle of universal provision.

Ref: https://help.netflix.com/en/node/24926

Ref: https://commonslibrary.parliament.uk/the-most-civilised-thing-in-the-world-the-political-foundations-of-the-nhs/

Comment 4

Human rights are not mutually exclusive of each other, and by pursuing one right (freedom of speech and privacy) does not give the right to violate other human rights (freedom from discrimination, torture, cruelty, injury to the body or mind). Each human right is not a stand alone right irrespective of the other rights, and one cannot have one right without consideration to the other rights.

Therefore, owners of online platforms cannot use a human right in isolation to protect their own position. However, it is recognised owners of online platforms do not want to rescind their control of the internet, rather they want to continue to control the internet and social media, via the means outlined above, and thereby continue to make vast sums of money by the distribution of hate speech which in turn promotes hate speech.

The growth of hate speech online is illustrated by the UN Ref
https://news.un.org/en/story/2023/01/1132597

Comment 5

AI needs to be open, accountable and transparent. When humans make errors they are accountable. When AI makes an error, then how is the AI accountable? It is unacceptable that AI designers are willing to unleash AI on the community when they have advised the community that they do not know how it works. The community needs more rigor from these Companies before experimenting on the community with AI.

Ref: International Scientific Report on the Safety of Advanced AI

From the UK:
Department for Science, Innovation and Technology and AI Safety Institute

Published
17 May 2024

'The interim report highlights several key takeaways, including:

- General-purpose AI can be used to advance the public interest, leading to enhanced wellbeing, prosperity, and scientific discoveries.
- According to many metrics, the capabilities of general-purpose AI are advancing rapidly. Whether there has been significant progress on fundamental challenges such as causal reasoning is debated among researchers.
- Experts disagree on the expected pace of future progress of general-purpose AI capabilities, variously supporting the possibility of slow, rapid, or extremely rapid progress.
- There is limited understanding of the capabilities and inner workings of general-purpose AI systems. Improving our understanding should be a priority.
- Like all powerful technologies, current and future general-purpose AI can be used to cause harm. For example, malicious actors can use AI for large-scale disinformation and influence operations, fraud, and scams.
- Malfunctioning general-purpose AI can also cause harm, for instance through biassed decisions with respect to protected characteristics like race, gender, culture, age, and disability.
- Future advances in general-purpose AI could pose systemic risks, including labour market disruption, and economic power inequalities. Experts have different views on the risk of humanity losing control over AI in a way that could result in catastrophic outcomes.
- Several technical methods (including benchmarking, red-teaming and auditing training data) can help to mitigate risks, though all current methods have limitations, and improvements are required.
- The future of AI is uncertain, with a wide range of scenarios appearing possible. The decisions of societies and governments will significantly impact its future.'

Ref: We note the AI Soul Summit and AI Global Forum's press statements.

AI Governance needs to address the following aspects:

1. AI generated images, text and documents must be clearly labelled that they are AI generated and by whom.
2. AI designers and people who deploy AI are accountable for its use.
3. Ethics and Integrity aspect incorporated at every level, stage in the development and deployment of AI.
4. We support the Federal Government's proposal to criminalize AI fake generated pornography.
5. The same online regulation as any online platform.
6. Strict laws to prevent AI designers using personal data, copyrighted and intellectual property without express permission.

Non-compliance with these principles by companies and individuals needs a well-resourced regulatory Body to issue removal, modification of content, penalties and conduct investigations and prosecutions. Such a regulatory Body could be part of the Audit process above in Comment 1.

AI can be used as a beneficial tool, however currently we are at the mercy of agents promoting vilification, fake, untrue and discriminatory comment. The result of this chaos is a danger to social coherence and greatly undermines trust in the political process and democracy.

Ref: https://builtin.com/artificial-intelligence/risks-of-artificial-intelligence

Comment 6

Are Online Platforms and Social Media Companies 'publishers' or 'platforms'?

Online Platforms and Social Media Companies claim they are not publishers and argue privacy and free speech should take precedence at the expense of other human rights, as well as abrogating responsibility and liability for defamatory and hateful online posts.

However, these companies become distributors of user content as well as distributors by the republication of user online comments and images. As a result, they need to be subject to the same laws, expectations, requirements as any publisher/distributor in any media. The same laws need to apply as a community standard and not be dependent on which technology is used.

Ref: https://www.eff.org/deeplinks/2020/12/publisher-or-platform-it-doesnt-matter

> To quote: 'Historically, there is some legal distinction between "publishers" and more passive "distributors" of others' speech, and "distributors" is perhaps what those who yearn for "neutral platforms" are referring to. But "distributors" was just a subcategory of "publishers" and both bore liability.
>
> So, what is the legal difference between "publishers" and "distributors"?
>
> One is always a "publisher" of their own words, the stuff they write and say themselves. That is completely uncontroversial. The controversy and confusion arise around *republication* liability, the idea that you are legally a "publisher" of all statements of others that you republish even if you accurately quote the original speaker and attribute the statement to them. So, if you accurately and directly quote someone in an article you have written, and the quoted statements defame someone, you can be liable for defamation for republishing those statements. This applies to any content in your publication that you did not write yourself, like letters to the editor, advertisements, outside editorial, wire service stories, etc. Legally, you are responsible for all of these statements as if they were your own creations.'

Comment 7

Tech Companies that decide to change their terms and conditions need to expressly ask users for new permission without affecting their other services. Users must be given the option to opt-out without affecting services.

Ref: https://www.abc.net.au/news/2024-06-10/instagram-facebook-train-meta-ai-tools-no-opt-out-australia/103958308

Comment 8

Each person has the right to ownership of their own digital data.

Reclaiming ownership of our individual online identity is an important goal. This could be achieved with the establishment of a Digital Bill of Rights (or other title?) that guarantees:

a. Tech Companies that need our data need to pay for it.
b. The right to choose which of our data to sell and to whom.
c. The right to access data to adjust or delate our record. For example, data is deleted following online purchases. This could assist with the reduction of cybercrime hacking historical and current data.
d. Platforms can only use location based on IP address when using the platform rather than precise GPS tracking.

Such a Digital Bill of Rights could be the basis of objectives for a Regulatory Body to authenticate a person's identity with their digital identity, similar in nature to a driver's licence or passport. Such an online ID could be used/needed when posting online or using generative AI to authenticate the identity of the person responsible for the post. Unfortunately, no system is without criminal activity, so digital ID would be at risk of theft. However, as with a stolen credit card, or other identity document, there could be a system for notification, cancellation and re-establishment.

Ref: Yanis Varoufakis 'Techno Feudalism' page 198.

Ref: https://journals.sagepub.com/doi/full/10.1177/2053951718820549