Submission to the Department of Infrastructure, Transport, Regional Development and Communication on the proposed Online Safety Bill 2020

14 February 2021

Samantha Floreani

I welcome the opportunity to submit comments to the Department of Infrastructure, Transport, Regional Development and Communications concerning the Online Safety Bill 2020. While I commend the overarching objectives of the bill in creating safe and accountable online spaces, and in particularly protecting children from harm caused by predatory online behaviour, I have some significant concerns that some of the powers included in this Bill will undermine the digital rights of Australians, and exacerbate harm for marginalised and vulnerable groups, which I have outlined below. I have also included some recommendations at the end of this submission.

1. **The Bill perpetuates regressive ideas about sexual impropriety**

Section 8 is underpinned by the idea that determining “offensiveness” of content can be judged by “standards of morality, decency and propriety generally accepted by reasonable adults.” This approach, which also underpins Australia’s classification system, by its very nature privileges and upholds the status quo. This approach inherently centres heteronormative and cisgendered experiences.

The new categories of Class 1 and Class 2 material are extremely broad, and rely on the outdated and antiquated classification system of the National Classification Code. Under Australia’s classification system, many sexual practices of queer and disabled communities are disproportionatley deemed “Refused Classification” and thus perpetuate ableist and heteronormative ideas by restricting the kind of pornography that is and is not available to Australians.[1] Applying this system of classification to online spaces is excessive, and likely to quash the diversity in freedom of expression on the internet.

Sex and sexuality are also important parts of humanity and expression. The broad categories that this Bill relies on are underpinned by a moral standpoint of what is and is not “harmful” or “offensive,” rather than an evidence-based approach. The implications of this Bill is that sex, sex workers, and sexualised content is inherently harmful. It places consensual sexual content in the same realm as that of terrorism and abhorrent violence, which is completely misguided and regressive.

---

[1] See, ‘If not a fist, then what about a stump? Ableism and heteronormativity within Australia’s porn regulations,’ Ryan Thorneycroft, *Porn Studies,* 7:2, 152-167 2020. Available at: https://www.tandfonline.com/doi/abs/10.1080/23268743.2020.1713872

2. The Basic Online Safety Expectations incentivise technology companies toward over-compliance, censorship, and harmful technological "solutions"

Section 46 requires the expectations to specify that service should minimise Class 1 material, and take reasonable steps to prevent children from accessing Class 2 material.

Due to the massive scale of online content, these expectations are likely to push companies toward implementing automated processes for determining content that may be covered by these broad categories. There is plenty of evidence that demonstrates that algorithmic systems are not effective at telling the difference between content that is and is not harmful.[2] Instead, they disproportionately target content created by Black[3], Indigenous[4], fat[5], disabled[6], and LGBTQ+ people[7]. This is largely because these communities are still wrongfully identified as deviations from the "norm," which will only be exacerbated by the regressive notions of morality and propriety which are elevated in s 8.

Further, we have already seen the results of SESTA/FOSTA in the US that these kinds of expectations incentivise technology companies to over-comply to avoid penalty, resulting in blanket removal of any sexual content, rather than the more technically challenging task of identifying genuinely harmful content. This kind of censorship obviously harms freedom of expression online, and censorship *always* harms marginalised and vulnerable groups the most. It is essential that this Bill doesn't cause excessive harm on its path to reducing harm for children online.

The requirement to take 'reasonable steps' to prevent children from accessing Class 2 content also creates concerns about the potential technological approaches that may follow. We have already seen suggestions to use facial recognition technology for age verification to access porn sites in 2020.[8] Applying a technological "solution" such as this to a social issue such as porn, is both ineffective (people find work-arounds for even the most sophisticated schemes), and also create significant privacy and data protection risks for children and adults alike. The Bill should not incentivise companies to implement technological "solutions" that will create more harm.

---

[2] For example, see this Electronic Frontiers Foundation report in response to Facebook's transparency report:
https://www eff org/deeplinks/2020/10/facebooks-most-recent-transparency-report-demonstrates-pitfalls-automated-content

[3] The algorithms that detect hate speech online are biased against Black people:
https://www.vox.com/recode/2019/8/15/20806384/social-media-hate-speech-bias-black-african-american-facebook-twitter

[4] Facebook has repeatedly banned Ingigenous Activists: https://onlinecensorship.org/content/infographics

[5] Instagram photo censorship
https://www.theguardian.com/technology/2020/oct/20/instagram-censored-one-of-these-photos-but-not-the-other-we-must-ask-why

[6] Facebook algorithm rejects ads, discriminating against disabled people:
https://www.nytimes.com/2021/02/11/style/disabled-fashion-facebook-discrimination.html

[7] This community based research shows that LGBTQ+ communities are most impacted by algorithmic bias on social media https://saltyworld.net/algorithmicbiasreport-2/

[8] Porn age filter for Australia:
https://www.abc.net.au/news/science/2020-03-05/age-verification-filter-for-online-porn-recommended-in-australia/12028870

3. The Online Content Scheme will harm sex workers online

It is clear that the Online Content Scheme will significantly impact the livelihood and safety of sex workers across Australia. As a result of the coronavirus pandemic, many sex workers, sex educators, and sex-positive activists have migrated their work to online spaces to comply with coronavirus restrictions. This Bill is likely to force many sex workers offline, which can lead them tou unsafe working environemnts, in turn causing more harm.

Further, there is very little transparency for those who wish to comply with the requirements, which makes it extremely difficult for content creators to understand what they can and cannot publish online. This, combined with a lack of a clear and effective appeals process for people to challenge take-down notices issued by the eSafety Commissioner is likely to result in disproportionate and unfair targeting of sex workers, with minimal pathways for redress or appeal.

I would like to extend my support to the submissions made by Assembly Four and Scarlet Alliance, with regard to the specific harms that this Bill will cause to the sex worker communities in Austraila.

4. The Bill threatens to undermine encryption

Part 13 and 14 of the Bill provides the Commissioner with Information Gathering Powers and Investigative Powers which allow them to gather information about the identity of an end user of a service, as well as “any documents” that may be considered relevant.

Given that ‘relevant electronic service’ includes email, instant messaging, SMS and chat, without mentioning end-to-end encrypted messaging services, it is possible that the Commissioner’s information gathering and investigative powers would extend to encrypted services.

The current eSafety Commissioner has already publicly argued that end-to-end encryption creates places for abusers to hide, saying that it makes “investigations into child sexual abuse more difficult.”[9] Encryption is an essential tool for digital security that protects everyone from surveillance by malicious actors and cybercrime, protects people’s privacy, including victim-survivors of domestic violence, confidentiality of journalists, safety of activists, lawyers and reporters. It is also essential for Australia’s cybersecurity strategy. Any move to undermine encryption weakens the security of all Australians.

It is essential that compliance with this Bill does not create ways to compel providers to restrict or weaken encryption. The Bill needs additional clarification to clearly indicate that it will not undermine encryption.

I would like to extend my support to the submissions made by Digital Rights Watch and Electronic Frontiers Australia.

---

[9] https://www.esafety.gov.au/about-us/blog/end-end-encryption-challenging-quest-for-balance

Recommendations

- **Remove Part 9 containing the Online Content Scheme from the Bill.** The issue of content moderation online should be handled separately to the other goals that this Bill seeks to achieve with regard to mitigating specific and directed harms against adults and children. Content moderation has already been shown to be a complex issue that requires a more nuanced approach than what is provided for in Part 9 of this Bill. The current way that Part 9 is drafted is so broadly that any adult content could be captured under the scheme. The attempt to control availability of consensually produced sexual content on the internet in this way is not an issue of safety, but of morality.
- **Include a sunset clause.** Given the proposed level of discretion given to the eSafety Commissioner, there needs to be an opportunity to review whether these powers are working well, and decide if the legislation should be renewed or revisited.
- **Establish a multi-stakeholder oversight board for activity covered by the Bill.** There is an international consensus that content moderation and take-downs require robust oversight and accountability to prevent abuse of power. The Board should be made up of the groups most impacted by the proposed laws, including sex workers and activist, and meet regularly, or at least annually, to closely examine how decisions are being made by the Commissioner's office.
- **Require transparency reporting on complaints and take-downs.** There should be quarterly, or at least annual, reporting of across all the powers prescribed to the Commissioner by the Bill. This includes the categories of content take-downs, complaints received (vs actioned and escalated), and blocking notices issued, including the reasoning. This will allow for public and Parliamentary scrutiny over the ultimate scope and impact of the Bill.
- **Articulate a meaningful and timely appeals process.** Individuals must retain their rights under the Bill which should include the ability to challenge removal notices in a timely manner, without having to seek an external judicial process to bring accountability to the Commissioner. Especially in cases where removal may directly impact income and livelihood, affected individuals should be able to seek remedy from the eSafety office if the removal is unjustified or arbitrary, including monetary damages as appropriate.
- **Include an explicit assurance that ISPs and/or digital platforms will not be expected to weaken or undermine encryption in any way to comply with any parts of this Bill.**