Consultation on the Online Safety Bill 2021

Global Partners Digital submission

About Global Partners Digital

Global Partners Digital is a social purpose company dedicated to fostering a digital environment underpinned by human rights.

Introduction

GPD welcomes the opportunity to make a submission to the Senate Standing Committees on Environment and Communications on the Online Safety Bill 2021 (the Bill). We recognise the legitimate desire of the Australian government to tackle unlawful and harmful content online, and we believe that the majority of the proposals put forward in the Bill are reasonable and sensible. Based on our analysis, however, we believe that particular aspects of the Bill, if taken forward in their current form, may pose risks to individuals' right to freedom of expression and privacy online and could be inconsistent with Australia's international human rights obligations.

In this response, we set out our concerns and make a series of recommendations on how the Bill could be amended to address and mitigate these risks. We believe these recommendations, if accepted and reflected in the final legislation, would greatly help safeguard the rights to freedom of expression and privacy online.

Framework for analysis of the Online Safety Bill

Our analysis of the Bill is based on international human rights law, specifically the International Covenant on Civil and Political Rights (ICCPR) which was ratified by Australia in 1980. Article 19 of the ICCPR guarantees the right to freedom of expression, including the right to receive and impart information and ideas of all kinds regardless of frontiers. Article 17 of the ICCPR guarantees the right to privacy and provides that "no one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence". Restrictions on the right to freedom of expression or privacy are only permissible when they can be justified. In order to be justified, restrictions must meet a three-part test, namely that the restriction: (1) is provided by law; (2) pursues a legitimate aim; and (3) is necessary and proportionate, which requires that the restriction be the least restrictive means required to achieve the purported aim.

It is important to remember that Australia's obligation to ensure that these rights are not unjustifiably restricted exists both in relation to restrictions which stem from the actions of the state itself as well as those caused by third parties, such as private companies. As such, it makes no difference from the perspective of the individual affected whether any restrictions are imposed and enforced directly by the state (e.g. through creating criminal offences which are enforced by the police and the courts) or through third parties, particularly when the third party is acting in order to comply with legal obligations.

Human rights analysis of the Online Safety Bill

Overarching Elements

We welcome the inclusion of Section 233(1) of the Bill, which states "This Act does not apply to the extent (if any) that it would infringe any constitutional doctrine of implied freedom of political communication". This reference is critical considering the potential risks to freedom of expression posed by the proposed legislation, but it fails to ensure protection for freedom of expression consistent with that right as protected under international law. Australia's constitutional protections, including the implied freedom of political communication, are far more limited than the right to freedom of expression as protected by Article 19 of the ICCPR. The text of Section 233(1) could be strengthened with additional reference to Australia's international human rights obligations.

Recommendation 1: We recommend that Section 233(1) be amended to include explicit reference to Australia's obligations under Article 19 of the International Covenant on Civil and Political Rights. For example, Section 233(1) could be amended to read: "This Act does not apply to the extent (if any) that it would infringe any constitutional doctrine of implied freedom of political communication or the right to freedom of expression under Article 19 of the International Covenant on Civil and Political Rights."

We welcome the fact that Section 24(1) of the Bill contains a standalone reference to the Convention on the Rights of the Child and provides that the Commissioner must, as appropriate, have regard to the Convention in the performance of functions under the Act and in relation to Australian children. The Convention on the Rights of the Child sets out a broad range of rights to which all children are entitled. Section 24 would therefore require the Commissioner to have regard to the full array of rights provided for in the Convention, including a child's right to freedom of expression (Article 13) and the right to privacy (Article 16). Nonetheless, it would be beneficial to include a more clear-cut recognition of the full range of rights encompassed under the Convention.

Recommendation 2: We recommend that Section 24(1) be modified to include explicit reference to the full range of rights provided for under the Convention on the Rights of the Child. For example, Section 24(1) could be amended to read: "The Commissioner must have regard to all civil, political, economic, social and cultural rights enumerated in the Convention on the Rights of the Child in the performance of functions: (a) conferred by or under this Act; and (b) in relation to Australian children".

We also believe that there should be a similar duty on the Commissioner to consider the ICCPR. This is because there are a number of functions and duties of the Commissioner that will engage the rights protected by the ICCPR. We suggest that the Bill contain a standalone reference to the ICCPR, which requires the Commissioner to have regard to it in the performance of functions under the Act.

Recommendation 3: We recommend that the Bill contain a standalone reference to the International Covenant on Civil and Political Rights, which requires the Commissioner to have regard to the Covenant, including the rights to freedom of expression and privacy, in the performance of functions under the Act. For example, Section 24 could be amended to include a new subsection: "(2) The Commissioner must, as appropriate, have regard to the International Covenant on Civil and Political Rights in the performance of functions: (a) conferred by or under this Act; and (b) in relation to Australians."

Basic Online Safety Expectations (BOSE)

We are pleased that Section 47 of the Bill would place obligations on the Minister to consult with the public and have due regard to comments. However, we believe that it would be preferable if the Minister were required to consider specific factors when making a Section 45 determination, including that they make determinations as narrowly as possible. This would help ensure that companies are only required to satisfy the Basic Online Safety Expectations where there is clear evidence or risk of harm. The scope of companies required to adhere to these expectations should be proportionate to the type of service and the amount of harm that takes place on a respective service. Blanket decisions on whole classes of services, as described in Section 45, are unlikely to constitute a narrowly tailored and proportionate response.

Recommendation 4: We recommend that the Bill require the Minister to make BOSE determinations on which companies are within scope as narrowly as possible. The scope of companies required to adhere to these expectations should be proportionate to the type of service and the amount of harm that takes place on a respective service.

Furthermore, we believe there is insufficient detail on particular elements of the core expectations outlined in Section 46 that may pose potential risks to individuals' freedom of expression and privacy without additional clarification. Section 46 indicates that service providers must take "reasonable steps" to ensure that end-users are able to use a service in a safe manner, and provide for clear and identifiable mechanisms for users to report and make complaints about various forms of illegal and harmful material. There is an expectation that, in determining what are such "reasonable steps", that the provider will consult with the Commissioner. However, there is no indication that the Commissioner would be required to provide guidance on what actions or processes satisfy core expectations and we recommend this be added within Section 46. This would be beneficial for services to be able to consult with the Commissioner to ensure that their decisions do not impermissibly restrict users' right to freedom of expression.

Recommendation 5: Companies that are required to adhere to the BOSE and take "reasonable steps" set out in the core expectations should be provided an opportunity to request further guidance from the Commissioner where they believe that upholding the core expectations might undermine their ability to safeguard freedom of expression or privacy.

Services must also take "reasonable steps" to minimise the extent to which cyber-bullying material, cyber-abuse material, non-consensual intimate images, class 1 material, and abhorrent violent material are available on their services. These core expectations could encourage proactive monitoring of content and the unintentional removal of permissible content. Given the scale of content which is generated and shared online, companies will increasingly turn to automated processes, including AI, to meet their obligations. The risk here is that automated processes will detect and remove content that is not actually unlawful or harmful in a particular context. Automated processes have had some success in relation to content moderation with types of images, including the ability to identify copies of images that have already identified by humans as constituting child sexual abuse and exploitation. However, automated processing has been less effective when identifying speech or less specific forms of unlawful or harmful content, such as cyber-bullying material or cyber-abuse material.

Recommendation 6: Section 46 should clearly indicate that taking "reasonable steps" to minimise illegal and harmful forms of content does not require a service to use automated

processes to proactively monitor and remove content. If automated decisionmaking is undertaken to meet core expectations, this should be accompanied by requirements to ensure the use of open source tools, transparency around standards, and appropriate appeals mechanisms.

We are particularly concerned that the core expectations may pose risks to encryption and individuals' right to privacy. While the Bill does not reference encrypted services, any requirement in the BOSE to filter or monitor material which applied to encrypted and other private channels would almost certainly amount to an unjustifiable restriction on individuals' right to communicate privately. This is because such services would need to remove or weaken privacy-enhancing technologies, such as encryption, in order to be able to filter or monitor material content which is generated or shared using them. We therefore suggest that the Bill explicitly note that companies are not required to cease, restrict or in any way weaken their use of encryption or other privacy-enhancing technologies to satisfy core expectations.

Recommendation 7: Section 46 should clearly indicate that companies' "reasonable steps" to satisfy core expectations do not include the filtering or monitoring of material, if doing so would require a service to restrict or in any way weaken their use of encryption or other privacy-enhancing technologies.

Section 46 also indicates that designated services will have to take "reasonable steps" to ensure that technological or other measures are in effect to prevent access by children to class 2 material. The lack of clarity around "reasonable steps" poses a potential risk to individuals right to privacy. There are particular technologies that could be used to satisfy this expectation, such as facial recognition technology, but these types of technological solutions involve the processing of large amounts of data, often personal data, when employed for identification, profiling or age verification. We recommend that service providers are not required to employ any form of technology that may pose risks to individuals' right to privacy in order to satisfy the "reasonable steps" element of this core expectation.

Recommendation 8: Section 46 should clearly indicate that taking "reasonable steps" to ensure that technological or other measures are in effect to prevent access by children to class 2 material does not require service providers to use technologies that pose risks to freedom of expression or privacy, such as facial recognition technologies. If these technological measures are to be pursued they should be accompanied by sufficient safeguards, including comprehensive data protection measures being taken by those who collect or process any personal data, and oversight by a competent authority or regulatory body.

The Bill would empower the Commissioner to establish reporting requirements under the BOSE, which would mandate services report on their compliance with one or more of the specified expectations. It would, however, be beneficial if these reports also contained relevant information on how freedom of expression and privacy were protected by particular services. This type of transparency requirement has been proposed in the UK Online Harms White Paper, and the recently released full government response provides that "[c]ertain companies will also need to produce transparency reports, which are likely to include information about their measures to uphold freedom of expression and privacy". These reports will include "information about the measures and safeguards in place to uphold and protect fundamental rights, ensuring decisions to remove content, blocking and/or delete accounts are well

founded".[1] This approach would be considered as best practice and should be emulated in the Bill.

Recommendation 9: We suggest that, as proposed in the UK Online Harms White Paper, reporting requirements include information on how companies are respecting freedom of expression and privacy on their services.

Takedown Schemes

We are concerned over the lack of adequate appeals mechanisms for removal and blocking notices made by the Commissioner for private companies and end-users. International human rights law requires that any person whose rights or freedoms are violated has the right to an effective remedy (Article 2(3) of the ICCPR). We appreciate that the Bill does not propose making private entities decide whether a particular piece of content is lawful or unlawful, and that Section 220 would provide private companies and end-users the ability to challenge decisions in the Administrative Appeals Tribunal. However, while the decision making of a public body can provide a far greater level of transparency and accountability, additional opportunities to challenge take-down notices or other types of decisions should be provided for within the framework. These additional appeals mechanisms, specifically those between the Commissioner and particular companies or end-users, would be beneficial since civil proceedings and other forms of redress are often cumbersome, time-intensive and expensive. Meaningful opportunities to challenge decisions should be readily available and accessible to the public before resorting to the courts.

Recommendation 10: The proposed takedown and blocking schemes should enable all end-users and private companies the opportunity to challenge decisions made by the Commissioner before resorting to the court system. The Commissioner should have the resources available to provide an effective remedy, which should include the ability for content to be reinstated.

We are concerned that private companies will be required to take action in a reduced time frame upon receiving a removal notice from the Commissioner. We understand the need to quickly respond to take-down notices for image-based abuse, cyber abuse, cyber-bullying and seriously harmful content, but believe that the shorter 24 hour period may not be practical for certain companies, particularly smaller companies or newer types of relevant electronic services, such as online gaming services, which are not included in the scope of the existing framework. Many of these smaller companies and newer services will not have the capacity or dedicated structures to respond in such a short time-frame.

Recommendation 11: Smaller companies and newer services should be provided with a more flexible time frame when they are unable to comply with the 24 hour take-down requirement. They should also be able to seek assistance from the Commissioner if they are unable to develop the necessary internal structures to be able to respond to notices.

[1] Online Harms White Paper: Full government response to the consultation (Dec 2020), available at: https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response

Information Gathering Powers

We are concerned that the information gathering powers provided for in Part 13 of the Bill may pose a potential risk to individuals' right to privacy without additional clarification on the scope of these powers. Section 194 empowers the Commissioner to obtain the identity or contact details of an end-user from a person when they are the provider of a social media service, relevant electronic service, or designated internet service. We understand that the Commissioner must have sufficient information gathering powers to effectively carry out its functions, but are nonetheless concerned that the threshold for issuing a written notice to a particular provider is relatively low. The Commissioner only needs to "believe on reasonable grounds" that the information is relevant to the operation of the Act. This threshold is inconsistent with the language used on page 12 of the 'Online Safety Bill - Reading Guide', which instead notes that this power will only be used to obtain contact details or the identity of an end-user "if necessary". We suggest that this higher threshold be incorporated into the Bill.

Recommendation 12: Section 194 of the Bill should be modified and replaced with the language provided for in the Reading Guide. Specifically, the threshold required for issuing a written notice should be whether the Commissioner determines such information to be "necessary to the operation of this Act", as opposed to when the Commissioner simply believes on reasonable grounds that the information is relevant.

Moreover, Section 195 requires that the provider of a social media service, relevant electronic service, or designated internet service comply with written notices "to the extent that the person is capable of doing so" or they could face a substantial fine of 100 penalty units. It is unclear what is meant by "to the extent that the person is capable of doing so" and whether services which use end-to-end encryption would fall within the scope of this exception.

Recommendation 13: Section 195 of the Bill should clearly indicate that a person does not need to comply with a written notice under Section 194 to the extent that it would require the provider to decrypt encrypted communications, or to cease, restrict or in any way weaken their use of encryption.