Google Australia Pty Ltd
Level 5, 48 Pirrama Road
Pyrmont, NSW 2009
Australia

google.com

26 February 2021

Committee Secretary
Senate Standing Committees on Environment and Communications
PO Box 6100
Parliament House
Canberra ACT 2600

BY EMAIL ec.sen@aph.gov.au
ONLINE SAFETY BILL INQUIRY

Dear Committee Secretary,

I am writing with respect to the Committee's inquiry into the recently tabled Online Safety Bill. This Bill was introduced into the House of Representatives a mere ten (10) days after the public consultation period on the Exposure Draft of the Bill closed.

Google made submissions to both of the public consultations relating to the Online Safety Act. Our most recent submission on the Exposure Draft of the Bill makes several constructive suggestions for amendments (designed to enhance the law) and also poses some questions for further deliberation.

Noting the Committee's abbreviated timetable for this inquiry, I confirm Google's interest in this legislation and attach our most recent submission in response to the Exposure Draft of the legislation for your consideration as you undertake this inquiry.

Yours faithfully,

Samantha Yorke
Government Affairs and Public Policy

Google Australia Pty Ltd
Level 5, 48 Pirrama Road
Pyrmont, NSW 2009
Australia

google.com

Sunday 14 February 2021

Online Safety Branch, Content Division

Department of Infrastructure, Transport, Regional Development and Communications

GPO Box 594

BY EMAIL: OnlineSafety@infrastructure.gov.au

Google and YouTube welcome the opportunity to provide feedback on the exposure draft of the Online Safety Bill (the "Bill"). We acknowledge that this exposure draft has been released on the back of an earlier consultation on a discussion paper previewing many of the additional powers and expansion in scope of the Office of the eSafety Commissioner.

We believe the Internet has had, on balance, an immensely positive impact on society. Our mission is to organise the world's information and make it universally accessible and useful. We build tools that are a force for creativity, learning, access to information, and much more. They have enabled economic growth, boosted skills and opportunity, and fostered a thriving society.

We recognise, however, that the Internet is also at times exploited by bad actors. We take the safety of our users very seriously, and we are committed to ensuring that illegal and harmful content that appears on our platforms is dealt with as quickly as possible.

Google is supportive of regulation, where it is carefully crafted and appropriately tailored. And indeed we haven't waited for regulation to address problematic content online. We have made significant investment in technology and human resources, and we have engaged with policymakers in Australia and around the world on the appropriate oversight for content sharing platforms, such as social media and video sharing sites.

Background

We welcomed and participated in the earlier consultation on the Online Safety Act discussion paper, culminating in constructive suggestions on how an Online Safety Act could move us towards a truly effective framework to foster online safety for all Australians and reflect emerging best practice in the regulation of online content. In particular, we offered the following recommendations:

- Government should acknowledge that there is a shared responsibility to foster online safety between industry, government, parents / carers, NGOs and civil society.

- The focus of Basic Online Safety Expectations (BOSE) should be on practical best efforts and overall processes, while avoiding being overly prescriptive.

- Any preemptive and preventative action recommended under the BOSE should be coupled with a 'Good Samaritan' framework that incentivises companies to take these proactive measures without risking liability for occasional missteps in that process.

- Transparency reporting requirements should be flexible, and, if there are to be any sanctions attached to them, they should focus on systemic failures.

- Any expansion to the scope of services subject to both the cyber bullying and cyber abuse schemes should be carefully limited and tailored, recognising relevant differences between services. Rules that make sense for social networks, for instance, do not necessarily make sense for other types of platforms or services.

- If the cyber abuse scheme were to be extended to adults, it is crucial that the definition of relevant content be tied to the Criminal Code.

- Regarding removal turnaround times, we strongly suggest that a more workable standard would be one that instructed online platforms to remove content "with all due speed," "without undue delay," or "expeditiously" and without a fixed 24 hour turnaround. We also call attention to the numerous comments made by the eSafety Commissioner that businesses typically do respond expeditiously to requests to remove content.

- The proposed accreditation scheme for safety tools does not provide clear utility. It would entail considerable resources to set up and administer, and would be very slow.

- On the subject of blocking terrorist and extreme violent material online, appropriate legislative instruments already exist to address these issues efficiently, and, to the extent any new instruments are introduced, it is essential that they be narrowly tailored to address only those 'worst of the worst' platforms and services that willfully and systematically fail to respond to valid legal removal requests regarding specific items of identified content.

- For ancillary services, any additional powers should specifically focus on notice-and-takedown of specific illegal material.

- In the context of governance, any increase in the powers and responsibilities of the Office of the eSafety Commissioner should be accompanied by a formal framework of multi-stakeholder oversight into the policy direction and decisions being made by the Office.

We were pleased to see some of these suggestions reflected in the drafting of the Bill and acknowledge that others have not been incorporated.

General Comments on the Bill

Impact on existing laws

Parts 5 and 6 of the Bill address the same subject matter as the existing Enhancing Online Safety Act 2015 and the Enhancing Online Safety (Non Consensual Sharing of Intimate Images) Act 2018 respectively. Similarly, Part 9 appears to amend Schedule 5 of the Broadcasting Services Act 1995. However, there is no reference within the Bill to how Parts 5, 6 and 9 interact with, amend or indeed replace the existing legislation. We would welcome further clarity on the Government's intentions with respect to existing legislative instruments and their future once the Bill is passed into law.

Scope of services

As noted above, any expansion to the scope of services subject to both the cyber bullying and cyber abuse schemes should be carefully limited and tailored, recognising relevant differences between services. Rules that make sense for social networks, for instance, do not necessarily make sense for other types of platforms or services.

However, in its current construction, the schemes would appear to apply to a wide range of other sorts of services, such as messaging services, email, application stores, and business-to-business services that serve as providers for other hosting services. The scope of the Bill should be carefully considered and narrowed.

As a case in point, the scope as defined in the draft Bill also would also seem to cover cloud based infrastructure platforms that third party businesses use to provide services to their clients. In its simplest form, cloud infrastructure and platform services are a type of technology where customer or business data and files are stored on servers at a data centre owned by a cloud service provider. This provides an alternative to using up the limited storage space on a personal computer or a company's servers. For business to business services, customers of cloud service providers - and not cloud service providers themselves - have ownership and control over the content they put on the cloud. The cloud provider typically does not have visibility into its customers' content to meet the privacy, security, and regulatory demands of its customers (and of their end customers), and to comply with existing laws and regulations governing cloud based services. Even if something was flagged by an external observer, it is often impossible for a cloud provider to remove individual pieces of content. Therefore, a request from the eSafety Commissioner to remove one single piece of content could result in a cloud infrastructure and platform service provider being mandated to remove a customer's entire website and we suggest that this is not a desirable outcome.

Therefore, compliance with certain obligations contained within the Bill will be challenging if not impossible for Google's Cloud business due to technical limitations on how Google can and should moderate business client content. For example, Sections 66(1)(f), 79(1)(g), 90(1)(h), and 110(1)(f) require hosting services to remove a single piece of content upon receiving a notice from the eSafety Commissioner.

Similar challenges would exist within, for instance, app distribution platforms like Google Play. There, too, the app platform operator does not have the ability to remove individual pieces of content from within an app.

Our submission in response to the Online Safety Act discussion paper suggested that the scope of the Bill be limited in scope to content sharing services, like social media and video sharing services, which have the principal purpose of helping people to store and share content with the public or other broad audiences, over which the platform provider does not have editorial responsibility.

Along with clarifying and narrowly tailoring the services in scope, we also suggest providing for clear exclusions for the avoidance of doubt, including for app distribution services, Cloud based infrastructure services, and communications services (e.g. messaging, chat or other applications where users expect a greater degree of privacy).

In addition, there is a caveat contained in Section 13 that states that, in the context of social media services, social interactions do not include business interactions. We would like to see a similar caveat be included in Section 17 on hosting services. Google's consistent approach to this topic is that governments should contact the customer in the first instance seeking access to the data and if the Government progresses this Bill with Cloud based infrastructure services within scope, we request that a provision be included that requires the eSafety Commissioner's Office to issue a notice to the specific client of a hosting service.

Turnaround Time

We are committed to tackling illegal content. We estimate that we spent over $1 billion in 2020 on content moderation systems and processes and we continue to invest aggressively in this area[1]. It's a complex task, and–just as in offline contexts—it's not a problem that can be solved by one silver bullet solution. Rather, it's a problem that must be managed in combination with other efforts, and we are constantly refining our practices. As a result, Google achieves generally expeditious removal, particularly of harmful content.

In our submission in response to the Online Safety Act discussion paper, we expressed a desire to better understand why there is a perceived need to reduce the turnaround time that exists under the Enhancing Online Safety Act 2015 from 48 hours to 24 hours, particularly when the eSafety Commissioner has made repeated references to the fact that most platforms remove content upon receiving a request from her Office very promptly.

[1] https://blog.google/outreach-initiatives/civics/our-work-2020-us-election/

Some take-down requests can be complex and necessarily take time to assess thoroughly. A complainant may not initially provide sufficient information; there may be questions as to the complainant's authority to make the complaint; consideration of whether there is a possible exception created by the material being shown for educational or documentary purposes; or simply the difficulty of assessing whether material has crossed the line of impropriety in the often-nuanced cases that we face nowadays, among other issues; each of which can take time to resolve and can only be accommodated by a flexible requirement. Specifying an exact turn-around-time, regardless of complexity of case, provides an incentive for companies to over-remove, thereby silencing political speech and user expression. In addition, quick and prescriptive turn around times and unexpected spikes in volume place a significant pressure on content reviewers / moderators (who are already looking at difficult content) to make quick decisions about content that in some cases are incredibly nuanced and complex. Indeed, focusing on the speed with which content is removed as a measurement of success may not actually reflect the public policy objective of minimising widespread exposure to a piece of inappropriate or harmful content.

Of course, we have observed regulatory frameworks in other countries adopt a 24 hour turnaround time. Germany's NetzDG law for instance, requires social media platforms only, after receiving notice, to exercise a local take down of "obviously illegal" content (e.g. a video or a comment) within 24 hours after notification. Services have 7 days to remove content that is not "obviously illegal" and even longer if the content is referred to an accredited self-regulatory body for review. However, the NetzDG law demonstrates that the quality of takedown requests can vary wildly. As our Transparency Report notes, 76.62% of reported items of content were neither removed nor blocked because the content did not actually violate YouTube's Community Guidelines or the criminal statutes referred to in NetzDG.[2] Spending time evaluating such a high volume of spurious complaints takes reviewers away from reviewing content that does violate YouTube's Community Guidelines or local law.

Relevantly, the French Constitutional Council's decision[3] to overturn key provisions of that country's online hate speech law ("Loi Avia") as unconstitutional pivoted on the short turnaround time for content removals. By imposing a 24 hour turnaround time and thereby putting the onus for analysing content solely on tech platforms without the involvement of a judge, within a very short time frame, and with the threat of hefty penalties, the Court concluded the law "infringe[s] upon the exercise of freedom of expression and communication in a way that is not necessary, suitable, and proportionate".

We argued, in our response to the Online Safety Act discussion paper, that a more workable standard would be one that instructed online platforms to remove content "with all due speed," "without undue delay," or "expeditiously" upon receipt of a clear and specific notice. Such a standard would allow platforms to provide the necessary human oversight, seek guidance, and consult legal doctrine before making a considered decision to remove content.

[2] https://transparencyreport.google.com/netzdg/youtube
[3] https://www.conseil-constitutionnel.fr/decision/2020/2020801DC.htm

If the Government is determined to codify an explicit numeric turn around time, we strongly suggest making a distinction between clearly illegal content on social media services (consistent with international precedents) - that could be subject to a 24 hour turnaround time - and simply illegal or legal but harmful content on designated Internet services, electronic communications services and hosting services subject to a longer turnaround time (assuming best efforts to act as quickly as possible). Such a differentiation would be consistent with the NetzDG law in Germany and would offer more flexibility in considering complex removal requests that may take longer to resolve.[4]

'Good Samaritan' provision

As noted in our submission in response to the Online Safety Act discussion paper, content regulation best practice suggests that preemptive and preventative action taken by tech platforms and services should be coupled with a 'Good Samaritan' framework. To the extent companies that take action to proactively detect and remove content may incur liability for their failure to catch and take action on specific items of illegal content, the risk of liability creates a perverse incentive for companies to either refrain from taking reasonable preventive action, or to over-remove legitimate content in the course of moderating. 'Good Samaritan' protections address this concern by giving protection for platforms to seek out and remove illegal or harmful content, without risking the loss of liability for occasional failures in that process. Any new law should ensure businesses can continue to invest in responsible proactive detection methods, without incurring an increased risk of legal liability in so doing and we suggest that such a provision be included within the Bill.

Governance

We appreciate the importance of appropriate oversight measures for how illegal and harmful content is addressed online. In our submission in response to the Online Safety Act discussion paper, we suggested the establishment of a multi-stakeholder body to oversee the Office and its decisions, especially to ensure the proper balance and respect of rights such as freedom of expression and opinion. We appreciate that the Office undertakes significant consultation with experts and stakeholders today, however we suggest that this consultation be formalised, mainstreamed and ongoing.

Irrespective of the Government's preferred approach, we believe that there are a number of central principles which should be considered:

- True regulatory independence: we believe it is very important that any oversight body in this area is truly independent. It is important that, in circumstances where the Government is proposing to issue instructions over the content of new codes or standards, adequate protections are in place to ensure that this independence is not threatened. While an oversight body's remit and powers should be clearly defined, it should be required to consult on the best ways of issuing guidance and codes of

[4] https://twitter.com/daphnehk/status/1354125120831418368

practice in order to ensure they are technology driven, platform agnostic, operationally sustainable and create a clear path to compliance for the platforms involved.

- Consultation with companies, experts, and other stakeholders: An oversight body would ensure that experts are consulted, and that any code or decision / determination is subject to an economic or human rights impact analysis. This would ensure that the requirements of the codes or decisions are technically viable, based on evidence of actual levels of harm, and economically and legally feasible.
- The Government could establish a multi-stakeholder forum involving representatives from companies and other relevant stakeholders, which could provide direct expertise from the field to make the regulator's decisions more effective and up-to-date with the existing social, legal and technological environment. The newly formed eSafety Advisory Council (formerly the Online Safety Consultative Working Group) could serve this purpose.
- Establish a formal industry board: To ensure industry is properly consulted, we propose the establishment of a Forum with representatives from industry, including companies of many different sizes. The Forum would provide input to appropriate codes of practice, and help set best practice for industry. To give the public confidence in the robustness and independence of this process, the minutes from meetings of the Forum could be published publicly. We understand that the Commissioner informally engages with many different industry organisations; perhaps this engagement can be formalised as a broader industry stakeholder forum?
- Reasonable expectation of ability to comply: Companies covered by the scope of the new framework also need a reasonable expectation that they can comply with any proposed regulation (for example, by avoiding mandating the use of technological solutions that would be inappropriate for some services or harms - as we explain earlier in this response).

Specific feedback

1. Our submission in response to the discussion paper in 2020 suggested that the definition of serious cyber abuse material reference the existing definitions contained within the Criminal Code. In the absence of a definition that complements the Criminal Code, and to reflect the intention to capture serious cyber abuse material, the definition contained Section 7(1) should additively build on each limb. We suggest the inclusion of the word "and" after the end of 7(1)(a)(iii), 7(1)(b) and 7(1)(c). In addition, we suggest that it would be helpful in the explanatory memorandum to provide further guidance on specific types of behaviour that are intended to be captured by the Bill, such as doxing, specific threats to commit serious harm, malicious attacks or ridicule, unwanted sexualisation, content repeatedly shared with the primary intention of harassment. This guidance should make it clear that the new law is not intended to regulate, for example, potentially defamatory content, in relation to which comprehensive and nuanced state-based regulation already exists (and which is itself in the midst of a reform process).
2. The reliance on standards generally held by reasonable adults in Section 8 to determine what is offensive appears nebulous and highly subjective.

3. We are confused by the distinction between Sections 10 and 11. All content that is provided (Section 10) or posted (Section 11) on a social media service, electronic service or designated internet service will be provided / posted by an end user, so why is there a distinction being made between these two sections? Could they be collapsed into one?
4. The definition of intimate image in Section 15 remains rather broad and, we respectfully suggest, out of step with community expectations. For example, a body part that is dissociated from any identifying features is considered an intimate image under Section 15(6) and a fully clothed person photographed without attire of religious or cultural significance is considered an intimate image under Section 15(4).
5. After working with the eSafety Commissioner on the non-consensual sharing of intimate images over the past two years, there is a category of image that the Bill and Commissioner should consider providing greater clarity about. Sections 16(c), 77(1)(d), 77(1)(f), 79(1)(d), 79(1)(f), 85(1)(d) and 85(1)(e) all refer to absence of consent for an image to be published. We have been asked in the past to consider instances where an individual has previously consented to intimate images being taken and published, often as part of a professional photo shoot where the individual has presumably been compensated, and has subsequently changed their mind. There should be further consideration of whether such instances should qualify for removal under the Bill (or indeed the existing Enhancing Online Safety (Non-Consensual Intimate Images) given the lack of information available about the commercial relationship between the individual and whomever took the photo, the terms under which they were compensated etc. We note that Section 86 includes a list of exempt posts and suggest that consideration is given to whether this category of image might be included there.
6. As mentioned above in the general comments section, the obligations on hosting services in Section 17 appear to capture business uses of enterprise cloud services. There are technical reasons why Google cannot remove single pieces of content from client accounts using the Google Cloud Platform. We would like to see an interpretive statement, much like that contained in Section 13, that explicitly exempts business use of hosting services from the operation of the Bill.
7. Section 42(2) describes how own motion investigations by the Commissioner will be conducted and as currently drafted there is no mention of any procedural fairness principles such as a right of appeal or transparency. We suggest that such principles be explicitly included so as to ensure the administrative integrity of any such investigations.
8. Section 49(4) describes the minimum intervening gap between requests by the Commissioner for reports. We would appreciate this being extended to 90 days and an obligation introduced on the Commissioner to determine whether the information being requested has already been made publicly available. Google and YouTube publish regular transparency reports and Google participates in the Lumen project, an independent research project conducted by the Berkman Klein Centre for Internet and Society at Harvard University studying cease and desist letters concerning online content. Preparing custom bespoke reports will likely require significant effort from a number of different teams across Google and YouTube.

9. Section 52(2) describes how the Commissioner can make determinations for periodic reporting through legislative instruments. It is not clear whether the Commissioner has law making powers under legislation. It would also be prudent to explicitly state that the Commissioner should have requested a report under Section 49(2) at first instance and if a report is not forthcoming only then can an application for a determination be made.
10. It appears that there are no penalty provisions for non compliance with end user notices (Section 71). If this is correct, it would be good to understand why this is the case and have a public debate about whether this is an appropriate public policy objective for this legislation. We consider that penalties for individuals who are actually responsible for publishing abusive content serve an important deterrent purpose and motivate changes in behaviours.
11. Sections 73(2), 85(2) and 93(2) identify the circumstances under which the Commissioner can publicly highlight organisations that have failed to comply with two or more service provider notifications. Two instances of non-compliance is a low threshold for a large business that routinely receives notifications from the Office; could you please explain why two was considered the appropriate number?
12. Sections 77(2), 79(2), 88(2), 109(2), 110(2), 124(2) and 128(2) relate to how content is identified in removal notices. Google and YouTube require the exact URL and in some cases a screenshot in order to facilitate a removal and therefore we request the removal of the words "So long as is reasonably practicable…" from all of these sections.
13. Can systems that are not designated under Section 108(1) still be used to prevent access to material that is unsuitable for children? For example, YouTube uses declared age to determine what content is accessible and does not present such content to users who are not signed in to YouTube at all (as in these cases we do not know the age of the user).
14. Section 120(1)(g) describes how hosting services can restrict access to class 2 material. We would like to better understand what types of access restrictions are envisaged here. Hosting services may not have access to information about individual users (including their age). Would access to content by invitation only be sufficient to meet expectations?
15. Divisions 7 and 8 describe how industry codes, standards and determinations can be determined. Given the diverse nature of the services within scope of the Bill and the dynamic and evolving nature of these services, we think it appropriate that, in first instance, industry be given the opportunity to develop codes of practice *before* the Commissioner sets a standard or makes a determination about industry rules. The Commissioner will have a clear role in overseeing any industry codes (and indeed can request that codes be developed on specific issues) and in ensuring that industries are accountable to commitments made within codes. It is reasonable to expect that the Commissioner would only need to intervene and set a standard or make a determination if a specific industry has failed to develop a code to the Commissioner's satisfaction or has demonstrated a lack of regard to obligations contained within a code through non-compliance.
16. Division 9 outlines new powers for the Commissioner to request that the Federal Court ban a service from operating within Australia if they have committed two or more

contraventions in a 12 month period. This again sets an extremely low threshold for preventing a service from operating within Australia, could you please explain how this threshold was arrived at? We respectfully suggest that a 'systemic' failure to comply is a more proportionate threshold.

17. Section 124(4)(b)(i) explains the circumstances under which a link deletion notice can be issued to a search engine operator and Section 128(4)(b)(i) similarly explains the circumstances under which an app deletion notice can be issued to an app distribution service. In both instances, could words be added that require the Commissioner to produce evidence of the one or more removal notices that have been previously issued?

Thank you once again for the opportunity to comment on this legislation.

Yours sincerely,

Samantha Yorke
Government Affairs and Public Policy