Microsoft Submission to the Australian Department of Infrastructure, Transport, Regional Development and Communications

Consultation on the exposure draft of the Online Safety Bill

February 2021

Introduction

Microsoft welcomes the opportunity to present a submission on the exposure draft of the *Online Safety Bill* (the Bill), after also having had the opportunity to provide feedback in the earlier consultation on the discussion paper in February 2020. This submission should be read alongside that of the Interactive Games & Entertainment Association (IGEA), which includes a gaming perspective relevant to our Xbox and Minecraft services.

Microsoft's approach to online safety

We recognise that technology companies have a special role to play in helping make the internet safer for everyone. We also recognise that providers should seek to design and operate their services in responsible ways, while anticipating and reducing digital safety risks unique to their services. They should also provide clarity about their terms of use and community guidelines, decision-making processes, and enforcement actions.

Indeed, Microsoft has a long-standing commitment to digital safety, as well as a history of working closely with governments, industry, civil society organisations, and academics to reduce the presence of illegal and harmful online content. Further details can be found in our previous submission on this topic.

While we are proud of our digital safety work, we recognise that voluntary industry efforts are necessary but not always sufficient to address the full range of harms online. We also acknowledge that all stakeholder groups need to do more to address the issues of illegal and harmful content online. Government regulation has an important role to play, and we support the development of principled and carefully calibrated regulatory efforts.

Therefore, we believe providers' obligations should be tailored to the nature and purpose of their services, the relationship between provider and end-users, expectations of users, and the risk profile of the services themselves. For example, regulatory frameworks should not treat online productivity tools (e.g., an online spreadsheet) in the same manner as general-purpose social media services. Providers should adopt and follow safety-enhancing systems and processes that will be most effective for their services.

Providers should also be given incentives to innovate and advance beyond their baseline obligations. For instance, providers that go beyond their responsibilities by actively scanning their services for illegal content or by suspending repeatedly abusive users, for example, should be entitled to an appropriate safe harbour from liability for such actions, so long as they act reasonably and in good faith.

We support a principles-based approach to online safety regulation

We note that the draft Bill largely reflects the Government's proposals outlined in the previous consultation documents. Given that this legislative process is nearing its final stages, we will focus this submission on a smaller number of specific issues.

As a global company, Microsoft has developed a set of harmonised safety principles that inform our thinking on regulatory developments across jurisdictions. We encourage the Australian government to consider these principles, as it considers refinements to the draft Bill.

1. **Operate responsibly.** Digital services play an essential role in promoting digital safety. That means they have an obligation to design and operate their services in responsible ways that anticipate and reduce digital safety risks unique to their services.
2. **Respect the fundamental rights of all people.** The internet is a key enabler of human rights, among them the fundamental right of freedom of expression. It also allows users to access information from a range of sources. Any regulation meant to promote digital safety must also protect these important human rights by preserving the free and open internet.
3. **Maintain an open internet.** Obligations to address digital safety risks should not force digital services to become content gatekeepers. The ability of users to create and share content directly and immediately is what makes the internet so dynamic and enables access to the broadest possible range of information. Although digital services have a responsibility to operate their services safely, making them responsible for what their users say, post, search for, or link to would, for practical purposes, undermine how many services on the internet work and destroy the internet's essence and value.
4. **Draw the line between illegal and harmful content**. Regulation of particular content, including mandatory blocking orders, should be limited to that which government defines as illegal. Elected officials and independent courts—not private companies—should be the decision-makers. They should also be the guarantors of due process where a balancing of rights is required, applying internationally agreed norms and longstanding human rights principles. Regulation to address other digital safety risks associated with harmful, but not illegal, content should focus on systems and processes, including digital services' compliance with their own digital safety commitments to users.
5. **Embrace clarity and transparency**. Any government regulation of content online should clearly define what is regulated and on what services. Ambiguity will chill speech and also force digital services to make subjective decisions on what content to block, what conduct to punish, and under what circumstances. Just as any government regulation should reduce ambiguity, digital services should provide clarity and transparency about their digital safety commitments to users, decision-making processes and enforcement actions.
6. **The internet is global.** Regulation of online content should be harmonised across jurisdictions wherever possible. The global internet benefits everyone. Because of the borderless nature of many digital services, regulatory fragmentation will splinter the internet, deprive users of access to information, and leave digital services less able to protect their users from harm.
7. **Recognize that there's no silver bullet** . . . Digital services should adopt and follow safety-enhancing systems and processes that will be most effective for their services. The law should not, however, require adoption of a specific technology solution nor assume technology exists to solve every problem. There is no technology solution that will keep users absolutely safe.

8. **. . . and there's no one-size-fits-all solution.** Providers' obligations should be tailored to the nature of their services, taking into account the function of the service, relationship between provider and end users, expectations of users, and risk profile of the service itself. In other words, productivity tools should not be treated the same as general purpose social media services.
9. **Incentivize positive action.** Regulation should incentivize digital services to take voluntary steps to protect users from exposure to illegal or harmful content. Furthermore, where digital services act reasonably and in good faith to do more than the law requires, they should not be assumed to acquire "knowledge" of content that then subjects them to liability.
10. **Engage the whole of society**. The fact that criminals and other bad actors weaponize the internet isn't a "technology" problem, or one that the tech sector alone should address. Digital safety requires a whole-of-society approach based on shared responsibilities among services, users, and public authorities.

Specific comments on the exposure draft of the Online Safety Bill

We continue to recommend a differentiated approach to obligations placed on unique services

Our previous submission explained that expanding the cyber-bullying scheme (to services other than social media services) may be neither technically nor practically feasible in some cases. It may, moreover, substantially conflict with global privacy expectations of users. Moreover, it may create conflicts of laws. These comments remain relevant across the various content-related schemes in the draft Bill.

Based on our reading of the Bill, both enterprise and consumer services could be in scope. We have serious concerns about this approach, as the current draft makes no distinction between the purpose, type, or functionality of the services in scope, nor the degree to which a particular service provider has the capability to access the content in order to remove it. For instance, Microsoft has no visibility into the content of emails of any Australian government agency via its use of Outlook. Should Outlook be deemed in scope, Microsoft would be required to take action on any harmful content contained in intra- and inter-agency email accounts.

The purpose and functionality of products deemed in scope within the draft varies significantly, as does the degree to which content is visible, public, shareable, or able to be amplified algorithmically. The risk profile of these products – for example, with respect to cyber-bullying material targeting an Australian child – is not identical. Here are just three examples of how the proposed Bill's lack of distinction between and among services may have consequences that are neither appropriate nor proportionate to the risk. We recommend revisions to the Bill to address these challenges.

First, we are concerned that enterprise services and cloud services provided to public or commercial bodies do not fall under any current exemptions. Infrastructure and enterprise services, including both cloud providers and software-as-a-service, tend to have the least control over third-party content on their systems—by law, by contract, and even by technical capabilities. In most cases, they do not have a direct relationship (contractual or otherwise) with those who upload the content onto the enterprise customer's service. In many cases, providers of these services may not have the technical or legal ability to identify the discrete content at issue within their service, especially when they have neither the right to access customer data, nor the ability to possess unique knowledge of site or service architecture. As a result, only de-platforming an entire service could fulfil a removal request for a single piece of content.

Forcing infrastructure and enterprise service providers to assert such control over customer architecture and end-user content could raise significant privacy, security, and other concerns. To the extent illegal content exists on these services, the obligations to remove or disable access to the content should generally fall as close to where the content is found as possible: First upon the responsible user, if known; then, upon the enterprise customer, third-party provider, or other entity having the most direct relationship with the user uploading the content.

- ➔ We, therefore, recommend that enterprise and cloud services be explicitly excluded from the scope of the Bill.
- ➔ In the alternative, the Bill should be redrafted to recognise that removal notices in relation to enterprise or cloud services are more appropriately directed to the enterprise or other customer, not the underlying service provider.

Second, we note that search engines are not meaningfully distinguished from other services, with the exception of the provisions on link deletion notices. Search engines enable users to find all types of information and provide important public benefits. Search engines differ fundamentally from other services, because they neither host nor share content, nor do they facilitate online social interaction. We, therefore, welcome the fact that a link removal notice can only be provided with respect to class 1 material and, even then, in relatively limited circumstances. Removal is only one moderation option, and link deletion is a relatively blunt instrument.

- ➔ We recommend explicit language that search engines are outside the services subject to the Bill's other provisions.
- ➔ We also recommend in section 124(4)(a) specifying that a link deletion notice cannot be given unless a removal notice has also been given with respect to the same class 1 material.
- ➔ We also recommend incorporating a reasonableness requirement in section 125, clarifying that compliance with a link deletion notice is required to the extent that a person is "reasonably" capable of doing so. This would enable a wider range of providers to comply, without impacting the effectiveness of the requirement.
- ➔ We similarly recommend that "reasonableness" be incorporated into the compliance requirements in sections 67, 80, 91, 111, 116, 121, and 129.

Third, our previous submission urged the Government to limit the obligations to monitor or remove content in the context of private communications, given the heightened privacy concerns that such actions necessarily would imply. We are pleased to see that the Bill does not seek to introduce any proactive monitoring requirements on providers, and that a hosting service provider may only receive a removal notice where the offending material has been communicated to another person. A hosting service provider, however, should still not be a first "port of call."

- ➔ We therefore recommend amending the Bill to clarify that the power under section 110 should not be exercised unless the social media service, relevant electronic service, or designated internet service has first failed to remove the material in line with section 109.

Notwithstanding the above, we remain concerned that no distinction appears to have been drawn between public and private communications. As noted above, the risk of harm and impact to fundamental rights varies between and among services depending on their purpose or functionality.

- ➔ We therefore recommend that the Bill is amended to explicitly remove online storage services and private messaging services from its scope.

We recommend this legislation provide additional clarity on provider obligations to avoid a chilling effect on online speech and expression

The Bill's wide scope also creates challenges for service providers in relation to the basic online safety expectations (BOSE), industry standards, and any other rules that may be determined in accordance with section 151, particularly given that the substance of these obligations are still to be developed. At present, it is not clear how these instruments will be shaped into clear expectations and rules that are proportionate to the size and impact of a service and its relative risk.

There are also few limits on what further obligations may be imposed on providers over time. For instance, section 45 does not limit the expectations that may be imposed through a BOSE determination. The Bill likewise does not limit the matters on which industry codes and standards may be developed. Similarly, it remains unclear what additional matters the Commissioner may take up through the provisions related to "service provider determinations" (section 8)." Thus, at this stage, the Bill appears to postpone the development of a large portion of Australia's online harms regime.

- ➔ We encourage the Department to consider how greater clarity about these obligations might be developed through this Bill (as the primary legislation), rather than in subsequent standards, determinations, or other instruments.

Legislation should seek to avoid any "one-size-fits-all" solutions; as such, industry should be enabled to develop codes in an appropriately differentiated fashion. Regarding the development of industry codes, we welcome the fact that: (a) the Bill allows for co-regulatory measures, and (b) the applicability of various industry codes or standards will depend on "which section of the online industry is involved" (section 138(2)).

For instance, we welcome the effort to create distinct categories within the BOSE framework for social media services, relevant electronic services, and designated internet services. However, even within those categories, there remains a wide range of services with different functionality, purposes, and users.

The BOSE, industry standards, and any other rules should be developed in a proportionate and appropriate fashion that accounts for the diversity of online services and their different risk profiles. We recommend a focus on systems and processes, appropriate to different service types. Each provider must take into account the nuances of their own services, user expectations, and other factors to avoid undue negative effects, including those related to fundamental rights.

While the current list of matters that may be dealt with by industry codes and standards (at section 138) does focus on procedures, the processes in the examples may differ significantly across services and may not universally applicable. For instance, it makes sense to distinguish general-purpose user-generated content platforms from services used primarily for personal storage of private files.

This distinction is particularly relevant with regard to consumer cloud storage services. Cloud storage services are often used to store highly personal data. While these services may enable the sharing of content to a small circle of colleagues or business associates, this functionality stands in stark contrast to services where broad public access to user-generated content is the default. Imposing the same content moderation obligations on consumer cloud storage services as on a publicly accessible social media service would therefore be unnecessary and disproportionate, with potential negative consequences on fundamental rights such as the right to privacy.

- ➔ We, therefore, strongly recommend building a clear expectation into the Bill that the BOSE, industry codes and standards, and any other rules will be proportionate and differentiate between service offerings. These subsequent instruments should focus on robust systems, policies, and processes (including providers' compliance with their stated digital safety commitments, as specified in their terms of use and community guidelines), not on specific product design mandates. Policymakers may wish to consider an approach similar to that proposed for the European Digital Services Act which emphasizes process and transparency over mandated outcomes.
- ➔ We also recommend that the timeframes and elements of the content-specific regimes account for the varying risk profiles of different providers. A smaller platform that seldom receives removal notices or similar requests may not be as well-positioned to respond as swiftly as larger companies with established or centralised processes.

Further specific comments on the BOSE:

- ➔ We also recommend that the Bill be amended to clarify that "reasonable steps", in relation to the core expectations in section 46(1)(a), depend on the nature of service and its intended purpose and audience.
- ➔ We also recommend that section 46(1)(b) be removed from the Bill. Given the breadth of services that may be within the scope of these provisions, and the frequency with which changes could feasibly be made to settings on those services, the requirement to consult the eSafety Commissioner seems likely to become highly burdensome for both the Commissioner and for providers.
- ➔ Given the breadth of services that could be in scope even for a single company, we suggest that section 47 (on consultation) include an obligation to take reasonable steps to make relevant companies aware of a draft determination.

In general, the periodic and non-periodic requirements to report on compliance with the BOSE risk becoming unduly burdensome, especially for small providers. At this stage, it is also unclear which range of services may be required to report, as well as which topics will be included. As recommended above, further developing these proposals now (rather than after the passage of the primary legislation), would give providers more certainty about measures they will need to put in place.

- ➔ We recommend amending section 49(5), which outlines the matters the Commissioner must consider when deciding whether to give a periodic reporting notice. We suggest adding whether the information being requested is already publicly available and the size of the provider (including presence in Australia). We suggest similar mandatory considerations also apply to decisions to issue non-periodic reporting determinations (section 56).
- ➔ Given the potential burden, we also recommend that compliance with reporting requirements is to the extent that a provider is "reasonably" capable of doing so (sections 50, 53, and 60).

Further specific comments on industry standards and codes:

- ➔ To recognise the requirement for public consultation, the need to work across industry, and the potential multiplicity of industry codes that may be developed at one time, we recommend extending the time by which a code must be registered to at least 12 months from 6 months (section 137(2)).
- ➔ Section 140 provides the Commissioner with the sole authority to determine whether they are satisfied with an industry code. This leaves some uncertainty for industry when preparing these codes and does not necessarily recognise that the code will have already undergone cross-industry and public consultation. The Commissioner must also be consulted in the

course of a code's development. As a result, we recommend that this section be amended to provide more clarity about how the Commissioner will assess a code before registering it.

- ➔ Given the consultative process for developing an industry code and the requirement to register industry codes, we view the Commissioner's power to determine an industry standard as a reserve power that is unlikely to be used. If this is the policy intent, we recommend reconsideration of whether vesting such authority with the Commissioner undermines the purpose and spirit of having an industry code in the first place.
- ➔ Not all providers will be members of an industry association, so to enable consultation when the Commissioner has requested a code, we also recommend extending the period for which a consultation must run (section 141(2)).

We suggest further oversight mechanisms are necessary, given the significant expansion of powers vested in a single unelected official and the potential impact on fundamental human rights

A notable feature of the Bill is the degree to which it extends the eSafety Commissioner's powers. Our previous submission noted the importance of recognising the potential human rights' impacts of decisions taken under online safety legislation and recommended that decisions should be subject to oversight by the Australian Parliament.

We believe the regime could benefit from further checks and balances, including improving transparency of government actions, addressing specific risks from conflicts of law, and elevating consideration of the potential impact on fundamental human rights. Decisions on online content can be difficult, given the importance of intent and context, and reasonable people may reach differing views. As it stands, the Bill places many of these decisions in the hands of one unelected Australian government official. We continue to recommend that the Department consider where additional Parliamentary scrutiny may be appropriate, as well as the following changes:

- ➔ Transparency by competent authorities in relation to their rules and procedures and on decision-making is critical so that providers can understand how supervisory and enforcement processes operate. We, therefore, suggest that the Bill also incorporate specific reporting requirements for the Commissioner. This might include, for example, reporting on the number and type of notices that have been issued, formal warnings issued, information sought by the Commissioner from providers, and on the processes followed to develop new or amended determinations. Further transparency about the Commissioner's decision-making and how fundamental rights have been balanced through those processes would also be useful.
- ➔ To provide an alternative perspective on BOSE, industry standards, or rules developed by the Commissioner in accordance with Division 8 of the Bill, we continue to recommend that the Commissioner be required to complete and publish a human rights' impact assessment on each instrument. This would transparently enable the identification of any potential adverse human rights impacts, and should be based on a consultative, multi-stakeholder process.
- ➔ In addition to the Commissioner's existing powers to request information, the new information-gathering powers in Part 13 seem to lack reasonableness or proportionality requirements. We recommend including some limitations on the information that may be requested from a provider (i.e., specifying the types of information that can be requested), narrowing the circumstances in which information may be requested (i.e., introducing a threshold test higher than relevance to the operation of the Bill), and making it clear that data will not be requested from providers in situations where a conflict of law could arise. This presents particular challenges if enterprise and cloud services remain in scope. As we have

previously noted, requests for such information should be directed to the enterprise or other customer, not the provider.

We would also like to raise a concern about the "name-and-shame" approach taken in parts of the legislation, and particularly how it appears to apply to providers regardless of good faith efforts to prevent or remove harmful content. Providers cannot prevent every instance of potentially harmful content. We suggest considering how these provisions could be reframed to support cooperation and reasonable, good faith efforts to prevent and respond to harm when it occurs. Requiring the Commissioner to first consult with providers can help incentivise those efforts and reduce the risk of over-removal of users' content.

- ➔ We recommend reshaping the provisions of service provider notifications in the cyber-bullying, non-consensual sharing of intimate images, and adult cyber-abuse schemes [sections 73(2), 85(2) and 93(2)] to incentivize good faith efforts to address harm. These provisions currently appear to give the Commissioner the ability to issue a public statement about a service even if the offending material has been removed—whether voluntarily, by the end user, or in compliance with a removal notice. This would not reflect accurately industry efforts to comply with the Australian legislative regime. We recommend that this approach is taken only where providers have failed to reasonably comply with a removal notice.
- ➔ Equally, if the provisions for service provider notifications remain (section 48), we recommend that the Bill specify that the eSafety Commissioner consult with a provider before preparing a statement. This would notify a provider of the Commissioner's view that it has contravened the BOSE, giving the provider an opportunity to explain and/or cure. This applies equally to the notifications in relation to reporting (sections 55 and 62).

Conclusion

Microsoft again thanks the Department for the opportunity to provide comment on the exposure draft of the Bill. We look forward to continuing to engage on this Bill, including through the Parliamentary process. We are also available to discuss this submission.