Friday November 12, 2021

[redacted]
Department of Infrastructure, Transport and Regional Development and Communications

[redacted]
Online Safety, Media and Platforms Division

By email: OnlineSafety@infrastructure.gov.au

[redacted]

The Digital Industry Group Inc. (DIGI) welcomes the opportunity to engage with the Australian Government on the draft Online Safety (Basic Online Safety Expectations) Determination 2021 (BOSE), and we thank you for your consideration of this submission.

By way of background, DIGI is a non-profit industry association that advocates for the interests of the digital industry in Australia. DIGI's founding members are Apple, eBay, Google, Linktree, Meta, Twitter and Yahoo, and its associate members are Redbubble, Change.org and Gofundme.

DIGI's vision is a thriving Australian digitally-enabled economy that fosters innovation, a growing selection of digital products and services, and where online safety and privacy are protected. DIGI shares the Government's strong commitment to online safety; not only is it a part of our organisational vision, we have invested in the development of *The Australian Code of Practice on Disinformation and Misinformation* which we developed and oversee. DIGI is also working with a diverse range of industry participants to develop industry-wide mandatory codes under the he *Online Safety Act 2021* (the Act), in partnership with the Office of the eSafety Commissioner (the Office).

Furthermore, our members have and continue to make major longstanding investments in the safety of their users and the community. At the end of this submission, we outline a high level overview of our relevant members' work in this area. As leading technology companies that have experienced both the benefits and immense challenges of user-generated content, they have long codified their own basic online safety standards through policies that usually exceed the thresholds of applicable laws.

In line with this commitment to online safety from DIGI and its members, we broadly support the Government's objective of establishing a set of basic safety standards that go beyond the takedown schemes of the Act. We agree that those takedown schemes should exist only as a safety net, and that online safety must occur at the platform level through safety by design.

DIGI's founding members are leading technology companies, and they are often the companies that are popularly associated with such reforms; however we note that the scope of the BOSE is far wider and extends to every website in Australia and every messaging service, including text messages. It is important to note that the scope is not limited to technology companies – but rather every business with a website accessible in Australia – nor to companies of a certain size, so therefore includes small businesses. The investments in online safety that DIGI's founding members are able to make are arguably

not a standard that can be achieved across such a breadth of organisations, and this needs to be reflected as the BOSE is finalised.

In that context, the aim of our submission is to strengthen the clarity and operationalisation of the BOSE across diverse industry participants by recommending more targeted expectations. That is also because several provisions in the BOSE will pose implementation challenges, user privacy and security concerns for both large and small technology companies alike. We are also concerned about the lack of oversight of the BOSE, should disagreements arise from Internet users or industry participants about its implementation.

Again, we wish to underscore our strong commitment to online safety, as well as privacy and cyber security. We believe there are ways to address our concerns with the BOSE without compromising the Government's commitment to online safety. To that end, we would like to work constructively with you and your office and the Department in resolving these issues before the BOSE comes into force.

We thank you for your consideration of the matters raised in this submission. Should you have any questions, please do not hesitate to contact us with any questions.

Best regards,

[redacted]

Sunita Bose
Managing Director, DIGI
[redacted]

[redacted]

Dr. Jennifer Duxbury
Director of Policy, Regulatory Affairs & Research, DIGI
[redacted]

Table of contents

The BOSE's scope is not "basic" for all websites, messaging and digital services in Australia

We understand the BOSE articulates "the Government's minimum safety expectations of online service providers, establishing a benchmark for providers to take proactive steps to protect the community from abusive conduct and harmful content online."[1] It is intended that reporting by service providers and public statements by the Commissioner about service providers' non-compliance with the BOSE will "provide much needed transparency about the level of harm occurring on services used by Australians and help to drive improvements in online safety practices by industry."[2]

Digital services covered under the BOSE include "social media services," "designated internet services," and "relevant electronic services." These are very broad and overlapping categories that are not confined to large technology companies, but encompass any service that allows users to share user generated content, including blogs and community online forums; every website or app that is accessible to Australian users including those used in non-technology sectors and by small business; all email and online messaging and gaming services, including text messages.

DIGI agrees that a range of digitally-enabled businesses should be required to embed reasonable baseline safety standards into their operations, and to be transparent about those standards. **Clearly illegal content, especially materials relating to child sexual exploitation, abuse and terrorism, must be eradicated as quickly as possible.**

However, the general scope of the BOSE as drafted in Section 6 cannot reasonably be described as a "basic" standard for the broad range of organisations it covers. In this section of the submission we discuss our four key concerns about the implications of that scope for digital services and Australian end-users. In summary, these concerns are as follows:

a. The BOSE is not sufficiently clearly targeted at protecting users' safety and extends into areas beyond the scope of the legislation itself.
b. The general safety expectation outlined in Section 6 of the BOSE in combination with expectation 11 could be interpreted as expecting service providers to engage in extensive surveillance of Australians online behaviour (including private communications), and to rapidly remove lawful content content from websites, messaging and digital services;
c. The approach to harmful online materials and activities in the BOSE requires service providers to override existing standards of harm in the general law, creating different safety standards for online and offline materials and activities by Australian end-users;
d. Operationalising the broad scope of the BOSE will pose implementation challenges for all websites, messaging and digital services in Australia.

At the end of this section we outline our recommendations about how the Government might consider addressing these concerns.

[1] Department of Infrastructure, Transport, Regional Development and Communications, *Draft Online Safety (Basic Online Safety Expectations) Determination 2021 Consultation Paper,* July 2021,2.
[2] Ibid 4.

BOSE scope needs to be more clearly targeted at protecting users' safety

Policy intent of BOSE

While DIGI supports the introduction of basic online safety expectations, **the breadth of the expectations as set out in the current drafting of the BOSE do not achieve the policy intent of creating a minimum safety benchmark** for the range of online businesses that are within the scope of the directive. The general expectation outlined in Section 6 of the BOSE sets out the baseline safety standard for all websites, messaging, and digital services with an Australian user base. Section 6 (1) states a core expectation that regulated services will take reasonable steps to ensure that end-users are able to use the service in a "safe" manner. Section 6(2) adds an additional requirement that service providers take "reasonable steps" to proactively minimise unlawful or potentially unlawful materials activities on their services and also to proactively minimise potentially harmful and lawful materials and activities. **There is no definition of "safety' or "harm" in the BOSE or the Act to assist service providers interpret what this expectation means.**

Broad scope of BOSE

Our analysis of Section 6 of the BOSE is that in its current form, the expectations require all service providers within scope to **anticipate**, and respond to, a vast and uncertain realm of material or activity that is unlawful or harmful or **may be potentially unlawful or harmful.** The broad scope of the BOSE is reinforced by the consultation paper that states: "The core expectations are principles-based and intended to be read in their broadest sense,"[3] and "the Expectations can be applied to the range of material or activity on a service that is or may be unlawful or harmful."[4] While Sub-section 6(1) of the BOSE is directed at the protection of end-users, there is no such limitation in Sub-section 6(2). In the Frequently Asked Questions (FAQ) about the BOSE, the Department further explains the scope of the BOSE is intended to capture emerging forms of harmful material and behaviours that service providers "are best placed to identify."[5]

If the BOSE is to be interpreted as the Department has stated, the requirement on service providers to detect and deal with harmful and potentially harmful online materials and activities, encompasses many categories of lawful speech, contrary to users' reasonable expectations that their use of online services will not be restricted, unless that use is unlawful or contrary to the service provider's terms of use. Because there is no clear standard for harm in Section 6, it can be interpreted to capture any material that is the subject of a user complaint, even if that material is lawful.

Broadly interpreted, Section 6 is also not limited to harm, and potential harm, to individuals' mental or physical well being; rather it covers the gamut of unlawful or harmful activity online, including that which is already regulated under existing laws such as consumer protection, competition, copyright and defamation laws. Interpreted in its broadest sense, sub-section 6(2) not only encompasses harms, and potential harms, to end users that are due to online activity but could, for example, include harms and potential harms to businesses, political processes such as elections, the economy or institutions.

Attempts by the Government to comprehensively regulate all possible forms of online harm in online safety legislation have met with public opposition in the past, and amendments were made to the Online Safety Act 2021 to address these. For example, as a Senator recently explained in a Senate Estimates

[3] Ibid , 2.
[4] Ibid, 4.
[5] Department of Infrastructure, Transport, Regional Development and Communications, *Frequently Asked Questions: Basic Online Safety Expectations*, October 2021, 2

hearing, amendments to the Online Safety Act were made during the parliamentary process to ensure the standard of harm for adult cyberbullying therein aligned with the serious harm standard in the *Criminal Code Act 1995*.[6] Section 5 of the Act provides that "serious harm" means serious physical harm or serious harm to a person's mental health, whether temporary or permanent. Serious harm to a person's mental health is further defined in Section 5 to include:

(a) serious psychological harm; and
(b) serious distress;
but does not include mere ordinary emotional reactions such as those of only distress, grief, fear or anger.

In its current form, the Act is focused on addressing unlawful materials and activity that cause serious harms such as adult cyberbullying, non-consensual sharing of intimate images, exposure to abhorrent violent materials, and Class 1 (Refused Classification) materials. The Act also addresses specific harms to children from cyberbullying, and exposure to age inappropriate materials (Class 2 materials). These kinds of materials and activities are also largely unlawful under existing laws. The Act does not envisage that the BOSE determination will revise the safety standards required by the Act[7]. We are concerned that, as drafted, the wide scope of the BOSE effectively overrides the harm standards of the Act.

We appreciate service providers' obligations to address harmful and potentially harmful behaviour under the BOSE is limited to taking "reasonable steps". The Department has helpfully clarified that this does not mean service providers must take all reasonable steps, a clarification which we suggest could usefully be included in additions to the BOSE[8]. However, as explained by the FAQ, the "reasonable steps" limitation only impacts the kinds of *measures* that must be taken by service providers, and does not limit the broad scope of *materials and activities* that are subject to the BOSE.

Service providers are incentivised to remove lawful user content and conduct surveillance

Cumulatively, the expectations in the BOSE that service providers take "proactive" steps to keep users safe create strong incentives on service providers to favour the fast removal of material (including lawful material), without allowing time to adjudicate whether there is a reasonable basis for doing so.

As drafted, the BOSE encourages service providers to exercise an abundance of caution when assessing if materials and activities are within the expectations. When in doubt about whether materials and activities are within scope of the expectations, many service providers will likely find it easier to remove users' content and/or online accounts under the BOSE. Indeed, Section 6(3) **encourages service providers to "remove (as applicable) material or activity on the service that may be unlawful or harmful"**. This poses a risk for end users' ordinary everyday interactions online with implications for political communication. There is a serious risk that controversial news content and content from activists, including online petitions and protests which serve an important democratic function in holding governments and the private sector to account, will be removed. Arguably, under the BOSE, service providers should remove almost any politically controversial subject matter that prompts or is likely to prompt a response of "distress, grief, fear or anger", because the BOSE has a lower threshold of harm than the threshold of "serious harm" in the Online Safety Act (quoted above).

[6] *Environment and Communications Committee*, Tuesday, 26 October 2021, Senate, 40-41.
[7] See Section 46 of the Act which sets out requirements for the BOSE.
[8] Ibid 2.

A question to this effect was recently raised in Senate Estimates.[9] In response, the Department stated that Section 233 provides a safeguard of political communication[10]; However, we would argue that the strength of that safeguard is questionable as Section 233 merely provides that the Act "does not apply to the extent (if any) that it would infringe any constitutional doctrine of implied freedom of political communication". This provision only acknowledges the existence of the constitutional doctrine (which in any event cannot be undermined by legislation) and does not mitigate the strong incentives in the Act and the BOSE for service providers to take down lawful content.

The Department's broad interpretation of Section 6(2) requires service providers to intercept potential unlawful or harmful activity on their services. As the Consultation paper explains, the BOSE includes an "expectation that service providers do more to assess and anticipate risks of harm facilitated by their services and take proactive and preventative action or 'reasonable steps' to mitigate those risks."[11] **How is every website or messaging service in Australia going to be able to anticipate potentially harmful behaviour? How can service providers reasonably meet this requirement unless they implement mass surveillance and analysis of users' online activities?** We welcome efforts at clarity in relation to these questions in the FAQ, but ask that this be reflected in the legal instrument itself to provide business with legal certainty, rather than any supplementary documents such as explanatory memorandums that have no legal standing.

The BOSE creates disparity between regulation of offline and online businesses

The broad scope of the BOSE arguably creates different harm standards for the same material based on whether the material is disseminated online and offline. This has significant implications for end-users who may be denied access to materials online that they can lawfully access offline, for example via traditional broadcast media. It is often said by the Australian Government that the rules for the online world should mirror that of the offline, and while we appreciate and accept the broad philosophy behind that argument, by requiring the online industry to take a pre-emptive approach to curtail materials or activity that are "potentially harmful or unlawful" creates a significant disparity between the regulation of offline and online businesses in the Australian market. **The assumption here is that materials online that are considered to be safe for the Australian public to view offline are rendered unsafe by the mere fact they are distributed online.**

For example, defamation law regulates publication of material which harms a person's reputation. As drafted, the BOSE can be interpreted to require online service providers to protect users from being exposed to materials, such as satirical comedy, that may cause hurt to an individual's feelings but fall short of the standard for defamation under the Uniform Defamation Law. In the offline world, of print publications, and broadcast media, the same type of material would be freely accessible unless it is established to have harmed an individual's reputation, to the degree required under applicable defamation law. We believe the BOSE safety standard should ensure that equivalent offline and online materials and activities are subject to the same safety standard.

Again we note that the limitation in Section 6 that service providers need only take "reasonable steps" to deal with harmful or potentially harmful material does not rectify this problem, since it only affects the measures that must be put in place and not the requirement that service providers address any type of materials that may be harmful.

---

[9] *Environment and Communications Committee*, Tuesday, 26 October 2021, Senate, 68.
[10] Ibid.
[11] *Draft Online Safety (Basic Online Safety Expectations) Determination 2021 Consultation Paper,* July 2021,5.

Operationalising the BOSE's broad scope on all websites, messaging and digital services in Australia

A threshold question for all websites, messaging and digital services who are subject to the BOSE is how they should interpret and comply with such a wide and subjective scope of online materials and activities in practice. In our view, this "catch all" approach is likely to undermine the effectiveness of the BOSE, since it is unclear how all websites, messaging and digital services can realistically fulfill their obligations to comply and report about their compliance under Section 49 of the Act.

An additional problem for businesses is assessing if their services fall within the scope of the BOSE. The Act outlines eight industry sections, three of which are covered by the BOSE, yet some services may not obviously identify within the section as defined. For example, it is not clear whether "a designated internet service of any kind" would also include an "app distribution service". The overlap between the industry sections under the Act, reflected in the BOSE, makes it difficult for businesses to assess how certain aspects of the law apply to them, where the law differentiates between different industry sections. It would be helpful if the Office could provide published guidance on this point.

Furthermore, the BOSE does not clearly articulate how businesses with different risk profiles will be treated for the purpose of assessing their compliance with the expectations. At present there are no exemptions from the BOSE for low risk services such as business-to-business services. Nor are there any public interest exemptions, from the BOSE, for example for online news media[12].

There also needs to be recognition that different sized businesses have different capabilities to meet the BOSE, which is recognised by the eSafety Commissioner's Safety by Design initiative. For smaller businesses with limited resources, it will be extraordinarily difficult to set up systems and processes to evaluate the harm or potential for harm of material and activities on their services. Even large technology companies with advanced automated systems and large enforcement teams will struggle to operationalise this open-ended nature of this expectation, because there is no consensus as to the meaning of "harm" with which they can train their systems and teams. We recommend that the Department test the practicality and impact of each of the proposals in the BOSE for a range of businesses in these different categories, and make adjustments accordingly.

Recommendation: the BOSE be amended to better target policy intent of protecting users' safety

Based on our analysis of the current drafting of the BOSE we recommend that amendments be made to Section 6 so that it is more clearly targeted at protecting users' safety.

We suggest that the scope of Section 6(2) be amended as follows:

> The provider of the service will take reasonable steps to implement systems and processes that mitigate the risk of harm to end users caused by materials and activities on the service. For the purpose of section 6(2), materials and activities are those materials and activities on a service that may be subject to a removal notice by the Commissioner under the Act.

[12] Note that there is a limited exemption in Section 104 of the Act for certain types of content being included in news materials, but news organisations would still be required to ensure that their reporting is not harmful, or potentially harmful or unlawful.

This would capture the range of materials and activities that are set out in Section 46, that the Government considers are particularly harmful to Australian users and fall within the scope of the expectations:

a. Cyberbullying material targeted at an Australian child;
b. Cyber-abuse material targeting an Australian adult;
c. Image-based abuse material;
d. Class 1 and Class 2 content; and
e. Material depicting, promoting, inciting or instructing in abhorrent violent conduct.[13]

This approach also enables service providers to take an approach to compliance that is based on the risk of harm to end-users on their services, enabling the expectations to be more effectively operationalised across the diverse range of websites, messaging and digital services that will be regulated. A more targeted focus on material and activities that are harmful under the Act ensures an appropriate balance between the need to provide a baseline safety standard and users' freedom of speech. This approach also ensures an appropriate level of parity in the safety standard for equivalent materials and activities in offline and online environments.

Recommendations in this section

1. We ask that the Government consider refining the scope of the BOSE as follows:
    a. Amend section 6(2) to provide "The provider of the service will take reasonable steps to implement systems and processes that mitigate the risk of harm to end users caused by the materials and activities on the service . For the purpose of section 6(2), materials and activities are those materials and activities on a service that may be subject to a removal notice by the Commissioner under the Act". This aligns with the safety standard embodied in the Act and ensures that laws that do not concern the safety of end users sit outside the scope of the Act.
    b. The BOSE should make clear that the "reasonable steps" requirement does not require service providers to take all reasonable steps, as set out in the FAQ.
    c. There are clear exemptions for businesses that are low risk such as business to business services or where there is a strong public interest in protecting freedom of expression, such as news media.
    d. Prior to implementation, the Department should test the practicality and impact of each of the proposals in the BOSE for a range of businesses with different risk profiles and of different sizes, and make adjustments as needed.

All service providers must seek Commissioner's guidance on how they will meet the Government's expectations

All websites, messaging and digital service providers subject to the BOSE are required to seek guidance from the Commissioner (Section 7(1) of BOSE) and report on their compliance (Section 49 of the Act). There is an additional requirement that service providers will have regard to any relevant guidance

[13] *Draft Online Safety (Basic Online Safety Expectations) Determination 2021 Consultation Paper,* July 2021,5.

material made available by the Commissioner (Section 7(2)).The Act also provides for the Commissioner to publish statements about the performance of service providers in meeting the Government's expectations, giving the Office power to publicly "name and shame" businesses that do not comply with the Offices' guidance (Section 48 of the Act).

The requirement in Section 7(1) means **every website or app that is accessible to Australian users must proactively seek the Commissioner's input on how they will meet these expectations.** We question the practicality of this requirement; it is impossible to even quantify the number of websites, messaging and digital services in Australia, let alone the time it would take for each of those services to consult individually with the eSafety Commissioner. This is not an efficient use of Government resources.

The broad powers of the Commissioner to seek compliance reports from businesses under the Section 49 Act and the BOSE in effect enable the Commissioner to conduct spot audits to assess if service providers are following the Office's guidance. While the Commissioner can publish adverse statements if service providers do not follow their guidance, service providers have no right to reply if they disagree. The "closed door" guidance the Commissioner provides individual service providers under the BOSE also has the potential to have unintended and hidden anti-competitive effects, especially if the guidance were to impose an unequal compliance burden on competitor companies.

In the FAQ, the Department has stated that the eSafety Commissioner will consult with industry concerning the guidance that her Office will provide concerning the interpretation and operationalisation of the BOSE[14]. We welcome such consultation, noting that it is critical that it is broad-based and includes large, medium sized and small service providers with diverse business models, and companies outside of the technology industry; this consultation should occur before the BOSE is finalised. However, such consultation does not overcome the issues with the lack of transparency in implementing that guidance under Section 7(1). In order to ensure the guidance provided by the Commissioner is transparent and consistent, it is best provided via published guidelines that are visible to all service providers. We therefore suggest that the requirement that service providers individually seek the Commissioner's guidance in Section 7(1) be redrafted so that service providers may, in their discretion, seek the Commissioner's guidance but any such guidance will be published to the extent reasonably practical. Section 7(2) should be amended so that service providers take into account published guidance from the Office.

Recommendations in this section

2. We ask that the Government amend the BOSE as follows:
    a. The requirement for service providers to individually seek the Commissioner's guidance on compliance is removed (noting that, as drafted, this requirement applies to every business in Australia that has a website, and every digital service used by Australians);
    b. Before the BOSE is finalised, we ask for additional consultation between the broad range of service providers regulated by the BOSE and the Commissioner concerning the guidance that will be provided by the Commissioner about compliance; and
    c. That the requirement on service providers to seek guidance from the Commissioner is replaced by a right for service providers to seek guidance provided that this is published to the extent reasonably practical. Section 7(2) should be amended so that service providers take into account published guidance from the Office.

[14] *Frequently Asked Questions: Basic Online Safety Expectations*, October 2021,2.

Need for eSafety Commissioner powers to be subject to reasonable oversight and review

One of the most concerning features of the overall regime established under the Act is the lack of symmetry between the extensive requirements imposed on businesses to be transparent and accountable to the eSafety Commissioner regarding their compliance with the rules, and the very limited transparency and accountability of the Commissioner in enforcing that compliance. Service providers and the public currently must largely rely on questioning of the Office in Senate Estimates to obtain information about how the Office is approaching the exercise of powers under the Act. Through this mechanism, we know that the Office has never exercised some of the powers granted six years ago under the Enhancing Online Safety Act 2015. For example, the eSafety Commissioner recently confirmed in Senate Estimates that, to date, her Office has not exercised powers to issue a remedial direction requiring service providers to implement a restricted access system to ensure minors do not access Class 2 materials[15]. Nor has the eSafety Commissioner utilised powers to request users to takedown materials but has only used those powers against service providers.[16]

While the Act has of course passed, and Section 183 imposes some minimum reporting obligations on how often the Commissioner is utilising powers under the Act, industry participants do not have clarity about how their compliance with the Act is assessed by the Commissioner. This places the burden of interpreting the Commissioner's expectations of compliance with the Act entirely on service providers. There is, for example, no requirement on the Office to publish guidance about how they approach the classification of online materials under the online content scheme in part 9 of the Act. This guidance is critical in setting expectations for service providers and end-users under the Act and about how the Commissioner exercises powers to take down Class 1 and 2 online materials, including materials the Office considers unsuitable for minors under the age of 18. It is also critical to setting clear expectations for service providers and end-users under the forthcoming codes under the Act, and the BOSE, to the extent these require service providers to address Class 1 and Class 2 materials.

Under the Act and the BOSE, the eSafety Commissioner is charged with making an assessment as to whether online materials or activities may be harmful or unlawful[17].The broad range of industry participants in scope cannot rely on Senate Estimates or supplementary documents, such as explanatory memorandums or FAQs, to ascertain how the BOSE will be administered; their rights and responsibilities must be reflected in the laws themselves. Information about how the BOSE is administered is essential for ensuring public confidence in the BOSE, and for assessing the need for and scope of new powers being granted to the eSafety Commissioner. As noted above, service providers must seek guidance from the Commissioner on how they should comply with BOSE under Section 7(1), but the Commissioner is not required to provide that guidance publicly, nor is there any recognition that the powers to provide guidance or demand reports should only be used in a proportionate and targeted way. While we are supportive of incentives being placed on service providers to encourage them to keep pace with emerging challenges in the usage of technology, the combination of the guidance powers in Section 7(1) of the BOSE with the powers in Section 48 of the Act to "name and shame" service providers, creates an unacceptably broad level of discretion in the Commissioner, without adequate restrictions or oversight.

[15] *Environment and Communications Committee*, Tuesday, 26 October 2021, Senate 47
[16] Ibid 39
[17] *Environment and Communications Committee*, Tuesday, 26 October 2021, Senate,71.

There is no provision in the Act for service providers to contest the assessment of the Commissioner under Sub-section 7(1) and (2) of the BOSE regarding the kind of material or activity on online services that may be unlawful or harmful. The limited opportunity for service providers to challenge the Commissioner's decisions creates an additional incentive for service providers to take an over-cautious approach to removal of content. While we understand that the Commissioner will provide an internal review mechanism for certain decisions, we are unaware of whether this has been established or how it will operate. There is a need to increase industry and public confidence that the eSafety Commissioner's powers will be exercised in a transparent manner and subject to reasonable oversight and review. While the current Commissioner may reasonably and sensibly exercise her powers today, clarity in the law will create certainty in the event of changes in personnel at the Office in the future and mitigate the possibility that these powers could be used improperly.

To specifically address the issues raised about the accountability and transparency of the eSafety Commissioner's powers under the BOSE, we suggest that the Commissioner be required to request and (with consent) publish service providers' responses to statements about their performance in meeting the Government's expectations under Section 48 of the Act. The details of the Office's internal review process must also be finalised and communicated to industry participants before the BOSE is finalised. To strengthen the overall accountability of the eSafety Commissioner, we suggest that the Minister for Communications exercise the power in Section 188 of the Act to direct the Commissioner to develop and publish a detailed regulatory action policy. At a minimum, this policy should include details of the internal review process that can be utilised by service providers and guidelines used by the Office to assist them in decision-making under the Act, such as the scope and limits of online materials that are classed as Class 1 or Class 2 in Part 9 of the Act. This will help ensure that future changes to the powers of the Commissioner under the Act draw upon evidence of past practice.

Recommendations in this section

3. We ask the Government to:
    a. Amend the BOSE to require the Commissioner to publish service providers' responses to statements about their performance in meeting the expectations of the BOSE.
    b. Ensure that the Commissioner's internal review mechanism is established and details of how it will operate are made available to service providers prior to the implementation of the BOSE.
    c. Issue a direction by the Commissioner from the Minister under Section 188 of the Act, requiring the Commissioner to develop and publish a detailed regulatory action policy that includes details of the internal review process and guidelines used by the Office to assist them in decision-making under Act.

Need for a flexible approach to compliance for different sized online businesses

The BOSE does not differentiate between the variety of services that are regulated across the websites, messaging and digital services in scope. Many of these services will have varying resources and capabilities to meet the expectations in the BOSE as drafted. Those with more limited resources may well find the cost of providing services in Australia is not viable, particularly if they are caught simply because they offer a global product accessible in Australia. Taking a one-size-fits-all approach to regulation also

risks leaving small and medium sized businesses at a competitive disadvantage to larger platforms, compromising the Government's goal of being a leading digital economy by 2030.

The FAQ states that service providers can take a flexible approach to compliance and can "choose to undertake different steps that work best" for them[18]. If this is the intent of the law, then this clarification should be incorporated in the drafting of the BOSE so that it is clear that smaller businesses are not necessarily required to invest in technology at the same level as larger platforms, nor are they required to have the extent of consultation required with the Office to comply with the expectations. Again, the diverse range of industry players require certainty in the law itself, not by way of supplementary FAQ or materials that are subject to change outside of a parliamentary process.

Recommendations in this section

4. We ask the Government to amend the BOSE to be consistent with the BOSE's FAQ so that it is clear that service providers of different sizes and resources can take a flexible approach to compliance and can "choose to undertake different steps that work best" for them.

Expectations may undermine end-users' confidence about their online privacy and security in private environments

User expectations for private messaging

While the BOSE should inspire Australians' confidence about their safety online, we believe that this well intentioned initiative has not given sufficient regard to their reasonable expectations of privacy and security. We are especially concerned about the implications of the BOSE for private messaging. As drafted, Section 6 of the BOSE encourages service providers to take proactive steps to prevent material or activity that is or may be unlawful or harmful. In addition to the broad scope of Section 6, Section 11 of the BOSE requires that service providers take reasonable steps to minimise certain materials such as cyber-bullying material targeted at an Australian child; cyber-abuse material targeted at an Australian adult; non-consensual intimate images of a person, and class 1 (Refused Classification) material. When applied to private messaging, the requirements in Section 6 and 11 could be interpreted as setting an expectation that service providers surveil and scrutinise private conversations between Australian adults that may be "potentially harmful", as well as those that may convey the more serious kinds of harmful materials in Section 11 .

We believe that the BOSE needs to be modified to make clear that the reasonable steps standard in Section 6 and 11 will accommodate differences in private communications from those that are more public. The FAQ contains a statement that platforms will not be required to monitor users' private communication, but can simply ascertain whether their activity is harmful or unlawful by analysing data and trends. However, in order for service providers to collect data that is sufficient for them to assess whether users' private communications are lawful or harmful, service providers may need to monitor user's private communications. Private messaging works very differently to publicly accessible services like websites or social media. Users of messaging services have greater expectations about the privacy

[18] Department of Infrastructure, Transport, Regional Development and Communications, *Frequently Asked Questions: Basic Online Safety Expectations*, October 2021

of their personal communications and will likely be seriously concerned about the imposition of regulations that require businesses to monitor or control the content of their messages.

The Government currently has significant powers of surveillance to detect unlawful activities online. For example, the Assistance and Access Act 2018 (Cth), International Production Orders Act 2021 (Cth) and the Surveillance Legislation Amendment (Identify and Disrupt) Act 2021(Cth) provide law enforcement agencies and intelligence organisations with far-reaching powers to access any network, system, device, or user accounts covertly and, where required, with the assistance of the service provider. To put it simply, in our view, the scope of the existing pieces of law do not indicate that there is any need for further intrusive monitoring of users' private communications beyond those given to enforcement agencies. We are concerned that any further powers, particularly directed at the surveillance of private messaging, will not strike an appropriate balance between safety and preserving the secure user environment that a thriving democratic society depends upon.

User expectations for encryption

The BOSE contains a specific expectation in Section 8 that service providers that use encryption with their services will implement processes to detect and address material or activity on the service that is or may be unlawful or harmful.[19] This implies that providers of encrypted services have special obligations to go further in scrutinising users' online interactions than providers of services that are more public.

Encryption is a vital part of modern electronic communications as it allows two or more parties to securely and confidentially engage with each other in many forms of communication and online activities. The ability to encrypt (and subsequently decrypt) communications underpins almost every online activity, from speaking on a mobile phone, messaging friends, accessing Government services to online banking, shopping and web browsing. It is fair to say that most of the common online activities that so many Australians engage with numerous times each day would not exist in their current form, or not at all, if not for the security that encryption affords. Encryption, and end-to-end encryption in particular, is critical for enabling civil rights activists, human rights defenders, and dissidents the ability to communicate securely. In the words of the UN Special Rapporteur for Human Rights, "Encryption and anonymity provide individuals and groups with a zone of privacy online to hold opinions and exercise freedom of expression without arbitrary and unlawful interference or attacks.'[20]

DIGI recognises that encryption in the wrong hands can be mis-used by a small minority of highly motivated individuals to conceal illicit and criminal activities, and may necessitate some changes in the approach of industry, Government and law enforcement bodies to keeping Australians safe. However, it is important to ensure that these changes (and regulatory instruments like the BOSE) do not weaken the overall security or privacy or disrupt the normal activities undertaken by the majority of law-abiding users, which participate in and contribute to creating a "safe online environment. While many of DIGI's founding members are investing in innovative approaches to these challenges, these innovations may be beyond the resources of the broad scope of websites, messaging and digital services in scope under the BOSE.

In light of these challenges, rather than incentivising service providers to engineer vulnerabilities into their products and services in ways that weaken users' privacy and security, the expectations should instead incentivise service providers to work closely with law enforcement and regulators to support specific investigations. Since online services are inherently global in accessibility, and encryption is also widely

[19] Section 8

[20] David Kaye, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression,* UN Human Rights Council, June 2015

used around the world, incentives to engineer vulnerabilities into service design will likely just shift bad actors onto other platforms. There are many opportunities for Government and industry to work together to reduce challenges confronting law enforcement and regulators regarding obtaining digital evidence without resorting to creating systemic vulnerabilities. The Center for Strategic and International Studies' report on *Low-Hanging Fruit Evidence-Based Solutions to the Digital Evidence Challenge*[21], is an effective and relevant resource for recommendations to improve the ability of law enforcement to obtain digital evidence.

Recommendations in this section

5. We ask the Government to amend the BOSE so that it is clear that the expectations in Section 6 and 11 of the BOSE do not require businesses to monitor or control the content of users' private messages on email, chat or messaging service..
6. We suggest that the expectation in Section 8 be replaced by an expectation that service providers will work cooperatively with relevant government agencies to improve the capability of law enforcement and regulators to obtain digital evidence of unlawful activity on anonymous accounts in response to warrants and other notices, under applicable laws.

Service providers are incentivised to remove anonymous accounts

The BOSE contains an expectation that service providers will take reasonable steps to prevent anonymous accounts from being used for unlawful or harmful materials or activities (Section 9 (1)). The steps service providers can take to meet this expectation include having processes that require verification of identity or ownership of accounts (Section 9(1) (b)).

The expectation regarding anonymous accounts assumes that anonymity is frequently correlated with unlawful or harmful behaviors, and increases safety risks to Australians online. Yet, end-users may have valid reasons for anonymity. For example, as the eSafety Office acknowledges, a valid reason for anonymity and identity shielding is to protect users from unwanted contact. eSafety encourages children only to use their given name, a nickname or an avatar online instead of a full real name which makes it more difficult for sexual predators and scammers to interact with them[22]. As acknowledged by the UN Special Rapporteur for Human Rights, the ability for users to remain anonymous online can also be an important means for keeping them safe and promoting human rights[23]. For example, anonymity enables activists to expose repression, corruption and hate, and allows stigmatised and abused communities to find safety and support when revealing their real-world identity could expose them to harm.

There are difficulties in developing an appropriate balance between users' privacy and the need to prevent anonymous accounts from being used for unlawful or harmful materials or activities. The FAQ for the BOSE states that service providers will be expected to take "reasonable steps" to detect, prevent and remove the ability of suspended or banned users to exploit anonymity or identity shielding to re-register

[21] William A. Carter and Jennifer C. Daskal, *Low-Hanging Fruit Evidence-Based Solutions to the Digital Evidence Challenge: A Report of the CSIS Technology Policy Program* (July 2018) available at https://www.csis.org/analysis/low-hanging-fruit-evidence-based-solutions-digital-evidence-challenge
[22] https://www.esafety.gov.au/about-us/tech-trends-and-challenges/anonymity
[23] Ibid

under a different or fake/imposter account[24]. It further states that users will not be required to divulge their real identities to service providers. The clarification in the FAQ that the BOSE does not require users to disclose their identity is not helpful, as service providers (rather than users) are the subject of regulation, and Section 9(2) explicitly suggests that service providers can require users to verify their identity to meet this expectation.There is potential for this expectation to conflict with the universally accepted privacy best practice of data minimisation that forms part of the Australian Privacy Principles under the Privacy Act 1988 (cth). Data minimisation is also a key principle of the Consumer Data Right. An additional concern is that this "one size fits all" requirement does not take into account the way privacy laws currently operate in order to protect anonymous users on certain services. For example, where third party apps are developed for a particular service, the operator of that service may not hold information about the accounts of app users, in order to comply with data minimisation requirements under privacy laws[25].

The BOSE, and other reform programs, must question the appropriateness of requiring online services to verify users' identity as a means to overcome the challenges associated with anonymity. Rather than imposing a requirement that is specific to anonymous accounts, we suggest that a better approach is to focus the expectations on service providers cooperating with law enforcement and regulators to obtain digital evidence of unlawful activity on anonymous accounts in response to warrants and other notices, under applicable laws.

Recommendations in this section

7. The expectations in the BOSE concerning anonymous accounts should be replaced by an expectation that service providers will work cooperatively with relevant government agencies to improve the capability of law enforcement and regulators to obtain digital evidence of unlawful activity on anonymous accounts in response to warrants and other notices, under applicable laws.

The expectation that service providers cooperate must not encourage further sharing of harmful materials

The expectation in Section 10(1) of the BOSE that service providers will consult and cooperate with other service providers to promote safe use is acceptable in principle, however industry needs more clarity about what such cooperation entails in practice. The implication in Section 10(2) is that websites, messaging and digital service providers must share data with each other. This was further reinforced by characterisation of this provision during Senate Estimates where it was stated that:

> *"The Basic Online Safety Expectations, as currently drafted and out for consultation, say that we expect platforms to consult and cooperate between each other to promote safe use. So that might be if you had a pile-on from a collection of organisations all directed at that trans youth organisation, for example, we would expect platforms to share information about hashtags that are trending, work together to disable those hashtags, disable tags* ***so that those groups*** *aren't targeted as directly and aren't copping as much of that abuse online."*

[24] Department of Infrastructure, Transport, Regional Development and Communications, *Frequently Asked Questions: Basic Online Safety Expectations*, October 2021

[25] See *Australian Privacy Principles:* AAP 3 and APP 11.

Using the example above, it is unclear how services would do this without sharing the personal information of the trans groups to other services which will be needed to determine if the same group has an account on other products and services. Furthermore, given that the scope of BOSE is a limitless set of websites, messaging digital services in Australia, to comply with the law as drafted would require the widespread sharing of this personal information, or making it available in some way for all services to access it. This may not be consistent with the privacy policies of the services in scope, nor in line with Australian users' expectations of privacy. It may also be prohibited by Australian and international laws. We believe this section poses major privacy concerns, and needs to be rethought.

Section 10(2) could further be interpreted to require service providers to preserve a large amount of data about users' behaviors on their site (for example, child sexual abuse materal, intimate images or cyberbullying) and share that data with other providers, without any mechanism for oversighting how this will "promote" safe use of services. It is worth noting that the eSafety Commissioner's recent position paper[26] outlining expectations for the codes under the Online Safety Act (Codes) contains an overlapping expectation that service providers will cooperate with each other in relation to harms caused by Class 1 and Class 2 materials. Care needs to be taken to ensure that this expectation does not encourage websites, messaging and digital services to share, and share widely, every occurence of unlawful and harmful material that is targeted at an end-user, together with information about the end user's identity.

We also note that the expectation around cooperation must be respectful of platforms' ability to develop their own policies, especially on sensitive issues like political speech and hate speech, where the law does not set clear standards. Different platforms have different policies on these types of issues and what might violate one, might not violate others.

DIGI notes that the eSafety Commissioner has successfully engaged with service providers on several important initiatives that promote best practices concerning the safe use of services such as the Safety by Design initiative.[27] Several of DIGI's members are also involved in international initiatives to combat serious online threats including the Global Internet Forum to Counter Terrorism, which have evolved in conjunction with the Christchurch Call to Action, an initiative that governments, technology platforms, and civil society organisations committed to after the devastating March 2019 Christchurch terrorist attack. They are also active members of several coalitions, such as the Technology Coalition, the ICT Coalition, the WeProtect Global Alliance, and INHOPE and the Fair Play Alliance, that bring companies and NGOs together to develop solutions that disrupt the exchange of child sexual abuse materials online and prevent the sexual exploitation of children. These types of initiatives provide a targeted and appropriate forum for relevant service providers to share best practices and support the government's policy goals. Rather than requiring broad-based, unsupervised, data sharing amongst service providers, we suggest that the Government's policy focus should be on enabling the eSafety Commissioner to improve the range of existing forums that enable service providers to cooperate in promoting online safety, in a targeted, open and transparent manner. For example, consideration could be given to establishing a roundtable forum for industry to examine the possibility of cooperation on emerging areas of concern such as volumetric attacks. This would, we suggest, provide an appropriate venue for addressing the policy concern that the BOSE make provision for service providers to identify and respond to emerging harms.

Recommendations in this section

[26] Office of the eSafety Commissioner, *Development of industry codes under the Online Safety Act: Position Paper*, available at https://www.esafety.gov.au/about-us/consultation-cooperation/industry-codes-position-paper
[27] https://www.esafety.gov.au/about-us/safety-by-design

8. The expectations in the BOSE concerning cooperation amongst service providers should be focused on improving the existing venues for cooperation that have been established to support the work of the Commissioner in progressing the government's online safety goals. For example, consideration could be given to establishing a roundtable forum for industry to examine the possibility of cooperation on emerging areas of concern such as volumetric attacks.

Volume of complaints is not an accurate measure of the safety of services

The BOSE empowers the eSafety Commissioner to obtain information from service providers about the number of complaints during a specified period (not shorter than 6 months) about breaches of the service's terms of use (section 16). Service providers must comply with the request within 30 days after the notice of request is given (Section 20). This expectation assumes that complaints data about the number of complaints made is a reliable indicator of safety or of compliance with the BOSE. This is highly questionable since complaints can be made for non-safety related reasons, including to harass and bully other users via mass reporting. For example, user flagging tools are often gamed by users on content that they simply dislike. As such, not all user complaints are vaid and upheld. As drafted, Section 16 covers any type of complaint about terms of use, not necessarily those that are safety related, because of the broad range of service providers regulated by the BOSE. The terms of use, for websites for example, will contain different kinds of requirements, depending on the purpose of the site including requirements relating to advertising, the sale of products and services, intellectual property protection, use of data and systems. We believe that this requirement should be removed on this basis, noting that the eSafety Commissioner already has extensive powers to compel service providers to report on their compliance with the BOSE under Part 3, Division 3 of the Act (Sections 49 to 62).

Recommendations in this section

9. We ask the Government remove Section 16, since it does not provide a reliable indicator of end-users safety and adequate reporting on compliance with the BOSE is set out in Part 3, Division 2 of the Act

Overlap and inconsistencies between BOSE and other regulations

Expectations concerning minors' access to Class 2 materials

The BOSE requires service providers to ensure that technological or other measures are in effect to prevent access by children to Class 2 material provided on the service in Section 12. Class 2 material is online content that would be classified as unsuitable for minors under the Online Content Scheme in the Act. At present, there are four separate Government initiatives under the Act concerned with issues relating to the protection of children from online harms. Three of these, namely the Codes, the BOSE, and the RAS are being developed under the Act. The fourth, The Age Verification (AV) roadmap, deals broadly with pornography (which includes Class 2 materials) follows on from a parliamentary inquiry into age

verification for online wagering and online pornography[28]. A fifth initiative concerning online age verification is part of the government's privacy reform program through the Online Privacy Bill. The draft RAS declaration and Privacy Bill have only recently been released on October 25 2021, and the consultation on the AV roadmap is ongoing.

According to the FAQ, the BOSE differs from new industry codes under the Act and the updated Restricted Access System (RAS) Declaration, which focus on material covered by the Online Content Scheme in the Act. Yet, Section 12 of the BOSE covers the same subject matter of the RAS declaration and Codes, both of which are concerned with ensuring minors do not access Class 2 materials. The eSafety Commissioner's Position paper on the Codes, released on September 29 2021, asks that industry to defer finalising a Code dealing with minors' access to Class 2 materials until December 2022. The Government's expectations about the age-gating of content must be consistent, practical and achievable for the diverse range of service providers subject to the BOSE. It is currently unclear the extent to which the Government requires diverse service providers to change the design and engineer changes to their services under these initiatives. Any major changes to a service provider's technology and practices will take time and will need to be informed by a consistent and uniform policy position about the Government's requirements. We therefore suggest that expectation 12 should not be included in the BOSE, since the concerns about minors' access to Class 2 materials are being dealt with by the RAS declaration, the Codes and the Age Verification roadmap.

Recommendations in this section

10. The expectation in Section 12 of the BOSE that technological measures be in place to restrict minors' access to Class 2 Materials should be removed, as the concern about the access of minors to Class 2 materials is being dealt with many other regulatory initiatives, including the Restricted Access System Declaration, the Codes and the AV Roadmap.

Need to reconcile BOSE with Codes/Standards under the Act

As foreshadowed above, there are a number of different regulatory instruments contemplated by the Act, including industry codes to address Class 1 and 2 materials, standards which can be imposed by the eSafety Commissioner, and the BOSE. The eSafety Commissioner's position paper on the Codes[29] contains requirements for them that directly overlap with several of the requirements of the Code such as Section 11 b) (minimisation of Class1 materials) and Section 12. There is also potential for future standards to overlap with any of the expectations in the BOSE. The BOSE consultation paper explains the differences between the industry Codes (currently under development), the standards, and the BOSE. The BOSE Determination is intended to provide flexibility for service providers to meet the Government's basic online safety expectations. In contrast, the purpose of industry Codes is to set out "binding self-regulatory procedures directed at ensuring class 1 and class 2 material is limited to services accessible to Australian end-users". As drafted, it is unclear as to whether compliance by service providers with the Codes (or standards introduced by the Office should the Codes not meet their requirements) will meet the BOSE requirements. We suggest that this be addressed by a simple amendment that makes clear that where a service provider meets the relevant standards in Codes or Standards, they meet the standards of the BOSE.

[28] *Protecting the age of innocence Report of the inquiry into age verification for online wagering and online pornography* (House of Representatives Standing Committee on Social Policy and Legal Affairs, Febuary 2020)

[29] Office of the eSafety Commissioner, *Development of industry codes under the Online Safety Act: Position Paper*, available at https://www.esafety.gov.au/about-us/consultation-cooperation/industry-codes-position-paper.

Recommendations in this section

11. We ask the Government to include a section in the BOSE which provides that where a service provider meets the relevant standards in Codes or Standards they meet the standards of the BOSE.

Need for a whole of Government approach to digital regulation

We note that several of the concerns DIGI raises in this submission about the uncertain scope and impact of the BOSE are compounded by the array of additional online safety initiatives progressing in parallel, via separate consultation processes conducted by the Office of the eSafety Commissioner. Many of the concerns raised in this submission are equally applicable to the development of the AV, RAS and the Codes. There is an urgent need for the suite of initiatives to be closely examined together, and streamlined to ensure that the diverse range of industry participants that must comply with the BOSE – which includes all websites, digital and messaging services accessible to Australian users – can feasibly comply with the overall regime the Government aims to establish.

As currently drafted, the BOSE contains several provisions that overlap with, or are inconsistent with the Act, the Age Verification Roadmap (AV), the Restricted Access System Declaration (RAS), and the development of industry-wide codes in relation to Class 1 and Class 2 materials (the Codes). For example, under the Act and the BOSE, service providers may be required to remove all types of Class 1 material. However, the Commissioner's position as stated in their position paper on the Codes[30] is that an identified subclass of Class 1, termed "Class 1b (fetish practices)" can be treated as Class 2 materials, and therefore do not need to be removed by service providers under the Codes.

Not only is there a need for a clear and consistent policy approach to protecting minors across these separate initiatives,we also think that any form of age assurance is challenging from a privacy perspective, since it will require additional data collection from users. We suggest that the eSafety Commissioner's Roadmap on age verification should be developed in consultation with industry and consultation with the Office of the Australian Information Commissioner. There are considerable challenges to implementing age verification on messaging platforms where service providers may encrypt services to protect users' privacy and security and not have full visibility into the content being shared; therefore, in developing this guidance, we ask that consideration be given to exempting messaging services from this expectation.

We also note the potential for the BOSE to clash with other policy initiatives by the Government in the area of defamation[31], cybersecurity[32] and privacy[33]. There is a risk of direct conflict between the Government's expectations that online businesses will protect users' privacy and the introduction of new online safety rules that require online businesses to reduce the level of user privacy on their services or collect vastly increased volumes of personal information. The inconsistent approach of some of these reform proposals to the subject-matter of regulation is also a concern. For example, the definition of social media

[30] Office of the eSafety Commissioner, *Development of industry codes under the Online Safety Act: Position Paper*, available at https://www.esafety.gov.au/about-us/consultation-cooperation/industry-codes-position-paper.
[31] See *Discussion paper-Attorney General's Review of Model Defamation provisions-Stage 2* (NSW Government, 2021).
[32] See *Strengthening Australia's cyber security regulations and incentives* (Commonwealth, 2021)
[33] See Privacy Legislation Amendment (Enhancing Online Privacy and Other Measures) Bill 2021 (Online Privacy Bill).

services in the Act is different to the definition proposed in the Privacy Legislation Amendment (Enhancing Online Privacy and Other Measures) Bill 2021. These examples highlight the need for a coordinated, whole-of-Government approach to digital regulation that is proportionate, and scalable.

Recommendations in this section

12. We ask the Government to develop a coordinated, whole-of-Government approach to digital regulation. Specifically, the BOSE needs to be closely examined for consistency with the RAS, AV, and eSafety's guidance on the Codes, the privacy reform program of the Attorney General's Department, and the cyber security reforms being progressed by the Department of Home Affairs.

Relevant DIGI member work

DIGI's members have and continue to make major longstanding investments in the safety of their users and the community. In this section, we outline a non-exhaustive, high level overview of our relevant members' work in this area. As leading technology companies that have experienced both the benefits and immense challenges of user-generated content, these companies have long codified their own basic online safety standards through policies that usually exceed the thresholds of applicable laws. In line with this commitment to online safety, they broadly support the Government's objective of establishing a set of basic safety standards that go beyond the takedown schemes of the Online Safety Act.

Policies

Every DIGI member has policies outlining restricted content and user behaviour on their platforms, which are regularly updated to ensure they reflect emerging patterns of abuse. While policies vary on a product basis, at a high level, DIGI member services remove and restrict:

- Hate speech that attacks or maligns a group of people based on their protected class status;
- Content that promotes or glorifies violence;
- Bullying, harassment or abuse that directly threatens another person;
- Promotion of self-injury or suicide;
- Non-consensual sharing of intimate images;
- Child sexual exploitation;
- Various other illegal content in the markets where they operate.
- Strict content policies in relation to pornographic content.

Moderation of user-generated content

The industry has also heavily invested in reporting tools and content moderation teams for user-generated content to ensure policy-violating content is surfaced and promptly actioned. DIGI's founding members maintain extensive review teams that operate to swiftly take appropriate action with user and community reports of policy-violating content. They also have expedited processes and protocols in place for urgent reports from law enforcement bodies, and for other content that requires rapid response.

Technology

The industry has and continues to invest in technology to detect and prevent the dissemination of policy-violating content. This includes, but is not limited to:

- Image hashing, such as PhotoDNA to report and identify child sexual exploitation material.
- Machine learning algorithms that identify potentially problematic content before many people have consumed it and trigger a human review, as appropriate and proportionate to the risk to end-users and the service they provide.
- A hash database that is shared amongst members of the Global Internet Forum to Counter Terrorism (and eligible members of the Hash Sharing Consortium) listing all known examples of terrorist content. At present this database includes almost 100,000 distinct pieces of content. Companies also coordinated within hours of the Christchurch terrorist attacks adding more than 1000 visually-distinct videos related to the attack to the collective hash sharing database. Crucially these were shared with smaller businesses that can benefit from this type of technology and information exchange.

Private & public sector collaboration

In addition to the sharing of online safety technology noted above, several DIGI founding members are pioneering a range of collaborative efforts across the industry, and with governments and with civil society, to address a wide range of issues related to online safety.

As one example, the aforementioned hash database is one example of industry collaboration that is occurring through the GIFCT, which aims to prevent terrorists from exploiting digital platforms. The GIFCT is an NGO designed to prevent terrorists from exploiting digital platforms. Founded by Facebook, Microsoft, Twitter, and YouTube in 2017, the goals of the GIFCT are threefold: (i) building shared technology to prevent and disrupt the spread of terrorist content online, (ii) conducting and funding research by international experts, and (iii) sharing information and best practices with businesses of all sizes to assist them in managing this content on their platforms. Since 2017, GIFCT's membership has expanded beyond the founding companies, and it has become an independent organisation led by Nicholas Rasmussen.

As one of several of its workstreams, the GIFCT has developed The Content Incident Protocol (CIP) to respond to emerging and active terrorist events, and assess any potential online content produced and disseminated by those involved in the planning or conducting of the attack. When GIFCT declares the CIP is in force, all hashes of an attacker's content are shared in the GIFCT among its members, and a stream of communication is established between them. The first CIP was activated on October 9, 2019, following the shooting in Halle, Germany.

Similar collaborations exist in other online safety issues such as child protection. As noted, relevant DIGI members are active in several coalitions, such as the Technology Coalition, the ICT Coalition, the WeProtect Global Alliance, and INHOPE and the Fair Play Alliance, that bring companies and NGOs together to develop solutions that disrupt the exchange of child sexual abuse materials online and prevent the sexual exploitation of children. In addition, they also proactively deliver social programs in the community with expert academic and civil society partners for a holistic approach to complex online safety challenges.

The depth and breadth of DIGI members' online safety work demonstrates our commitment to online safety, which we share with the Australian Government. We believe there are ways to address DIGI's concerns with the BOSE without compromising the Government's commitment to online safety, and we look forward to our continued work in this endeavour toward our shared goals.