2 August 2018

Ms Lynelle Briggs AO
Australian Public Service Commissioner
co/ Director, Online Content and eSafety Section
Department of Communications and the Arts
GPO Box 2154
Canberra ACT 2601

By email: onlinesafety@communications.gov.au

Dear Ms Briggs AO,

The Digital Industry Group Inc (DIGI) is the industry body representing the digital industry in Australia. Our members include Facebook, Google, Oath and Twitter who collectively provide various digital services to Australians ranging from Internet search engines to digital communications platforms.

As more of our daily lives are played out online, one of our core missions as an industry is to ensure the internet is a safe and respectful place, and we undertake a multipronged approach to ensure people are having positive experiences and meaningful connections when using our services.

We appreciate the opportunity to make a submission into the statutory review of the *Enhancing Online Safety Act 2015*. If you have any questions or require additional information, please let me know.

Kind regards,

Nicole Buskiewicz
Managing Director
DIGI

DIGI Submission to the review of the *Enhancing Online Safety Act 2015*

Executive Summary

The safety and wellbeing of people who use our services is the industry's top priority. Across the board, we undertake a multipronged approach to ensure people feel safe online using a combination of policies, tools, education, and outreach, and have been doing so for over a decade before the establishment of the Office of the eSafety Commissioner ('eSafety Office') in 2015. Since that time there have been many developments in the way in which digital platforms and electronic service providers enhance the safety of their users, and significant investment continues to be made by both industry and community groups in this space. However, we are concerned that certain activities of the eSafety Office duplicate these efforts, and suggest the role of the Office is refocused towards education and awareness raising amongst parents, and behavioural change.

The evolution of online safety: 2015-2018

Online safety is a dynamic space, and DIGI members have long invested in designing products that are safety enhancing. Even in the three years since the eSafety Office was established, there have been significant developments in the way illegal content is managed across digital platforms. In particular, the industry has become increasingly sophisticated at using technology to increase users' safety online in addition to the existing reporting and blocking tools that have existed for many years. While the sheer volume of content[1] and the need to establish context makes it difficult to proactively identify every piece of content that is in violation of a member's policies the industry is constantly striving to improve. We are seeing a number of technology based-online safety trends emerge, including:

- **Image hashing**, which works by taking a fingerprint or 'hash' of the image that is then used to prevent the image from appearing in other places on the internet. For example, PhotoDNA is now used across the industry to report and identify child sexual exploitation material. Facebook is also piloting image hashing technology to automatically prevent content that's been flagged as image based abuse from being uploaded.
- **Machine learning algorithms**, which proactively identifies potentially problematic content before many people have viewed it and triggers a human review. Automation,

[1] 400 hours of new content is uploaded onto YouTube every minute, 500 million Tweets a day on Twitter (Twitter blog (2014). Retrieved 31 July 2018 from <https://blog.twitter.com/official/en_us/a/2014/the-2014-yearontwitter.html>.)

image matching, and other tools have been particularly successful in proactively identifying terror related, suicidal, and image-based abuse content, and surfacing it for review and removal.
- **Targeted education** and awareness raising for online safety, which is conducted by targeting a particular audience to receive that safety message using the platform, like Facebook's Parents Portal.

While the industry continues to invest in and operate best practice notice and takedown schemes, it's important to note that the emergence of technology-based responses means companies are now relying less and less on reporting to address potentially offending content and resolve complaints. It also underscores the importance of legislative safe harbours, which are crucial if the industry is to continue to develop innovative technology-based solutions to online safety. For example, during the eSafety Office consultation on the *Non-consensual Sharing of Intimate Images Bill 2017*, DIGI strongly advocated for a carve-out from intermediary liability for those providers who have policies that prohibit this kind of content and have a prompt and effective removal processes in place. We were disappointed to see this recommendation was not adopted and remain of the view that such protections are essential if the Government seeks to encourage responsible digital platforms and other service providers to continue to invest in such solutions.

Complaint handling and the cyberbullying notice scheme

As mentioned above, DIGI members have traditionally operated well established (and legislated in some parts of the world) 'notice and take down' processes to manage content that violates each platform's terms of service or community guidelines. By way of background, this is a process that allows any of the millions of people who use our services to flag content that may violate our policies. DIGI members maintain extensive review teams that operate around the clock to swiftly take appropriate action with reports. All reports are reviewed and actioned by real people, who undergo extensive training when they join and are regularly trained and tested beyond this initial training so they can correctly action a report. Members triage complaints dealing with the most serious cases first (those that relate to real world consequences like suicide and child abuse), and consider a number of enforcement options (for example, content can be removed or age gated, features can be limited, and accounts can disabled).

DIGI members widely promote the existence of these processes, including to parents, schools, governments, and users and reporting problematic content directly to the platform operator continues to be the fastest and most efficient way of getting content reviewed and ultimately removed. The creation of a new reporting process within the eSafety Office for cyberbullying content not only duplicates the content removal processes already in place on digital platforms, but also risks confusing users about where they should report. We note that the number of reports being received by the eSafety Office are very low - 305 complaints received in the 2016-2017 financial year - which demonstrates that the digital platforms are adequately

managing cyberbullying content through their own reporting tools. By way of comparison, the Office of the Australian Information Commissioner received 2,494 complaints in the 2016 - 2017 financial year.

All social media platforms or relevant electronic service providers operating in Australia fall within the purview of the Office of the eSafety Commissioner and the cyberbullying civil notice scheme. Notably, no civil penalties (formal notices) have been levied under this scheme in the three years since it started. Oftentimes the platform was already in the process of removing the content when the complaint came in from the eSafety Office. In terms of informal approaches relating to cyberbullying from the eSafety Office, each DIGI member company has received a very low number of reports over the last three years. We also note that the eSafety Office hasn't issued any end-user notices relating to cyberbullying. The discussion paper appears to suggest that the end user notice scheme is a fall back remedy when a social media service or relevant electronic service provider does not remove content. Rather, we consider end user notices to play a critical role in deterring abusive online behaviour and changing the way people are treating each other. The eSafety Office should consider these notices as a significant deterrence that is independent of action taken by a social media service or relevant electronic service provider.

We also observe that this is not the only context in which the eSafety Office is duplicating existing support services. For example, we understand that efforts by the eSafety Office to support services that support people in domestic violence situations have duplicated some of the work already being undertaken by organisations like WESNET, a national women's peak advocacy body which works on behalf of women and children who are experiencing or have experienced domestic or family violence.[2] Additionally, in-school education programs (through the Virtual Classrooms) duplicates work being done by many non-governmental organisations (NGOs) that specialise in child protection.

Role of the eSafety Office in 2018 and beyond

Since 2015, the remit of the eSafety Commissioner has been significantly expanded, not only in terms of who is covered (from children to all Australians), but also in terms of wide ranging enforcement powers. For example, the *Non-consensual Sharing of Intimate Images Bill 2017* (currently before Parliament) proposes giving the eSafety Commissioner additional enforcement powers, such as seeking court orders, issuing search warrants, and information gathering. We believe such powers are best left to law enforcement agencies, which have experience with the legal processes involved in determining guilt and intent, and to whom we already disclose metadata following legal requests for information for crimes that involve a DIGI member service.

We urge caution in treating the eSafety Office as a first responder to victims of bullying or family violence similar to a law enforcement agency. The eSafety Office was established with the twin

[2] WESNET (2018). Retrieved on 31 July 2018 from < https://wesnet.org.au/>.

purposes of creating a national leadership role (in the Commissioner) to promote online safety for children and to operate the complaint handling mechanism. Since its inception, the remit of the eSafety Office has been expanded significantly leading to a stretching of resources and a plethora of disparate initiatives that are outside of the scope of the enacting legislation.

In acknowledgement of the fact that many of the eSafety Office's activities are duplicating efforts being made by other organisations, DIGI respectfully submits that the role of the eSafety Office is reorientated to focus more on (1) education and awareness raising amongst parents, and (2) behavioural change.

1. On the first issue of education and awareness raising, each social media service or electronic service provider has a number of safety protections and tools that are made available to users, yet many claim to be ignorant of these protections. Many child protection organisations are working with children in schools to teach them about respectful relationships and encouraging responsible digital habits; perhaps the eSafety Office could support these efforts rather than compete with them. We consider that the eSafety Office could make a significant impact on children's online safety by working with industry and NGOs to collaboratively educate children and parents about the range of tools currently available to them to assist in creating an optimal online environment.
2. We acknowledge that legal and criminal frameworks are only one part of the solution when it comes to reducing instances of cyberbullying, image-based abuse, and other kinds of bullying and harassment online. However, we must remember that these are behavioural problems, and in order to effect change, we must focus on interventions that will encourage perpetrators to question their instinctive reactions to challenging situations, and prompt them to take a different course of action. This ambitious goal of behavioural change appears to be perfectly suited to a wide reaching Government campaign. For example, with respect to safe and responsible online behaviour, particularly in relation to young people, there should be an increased focus on in-school education for students and parents/guardians and awareness raising to support the prevention of negative behaviours online.

Overall, we believe there is a broader educational role for the eSafety Office especially when it comes to promoting existing complaint-handling mechanisms and online safety initiatives by industry and civil society. Online safety is a joint effort between government, industry and the community, and each sector undertakes many valuable initiatives in this space. We would encourage greater awareness-raising for these initiatives rather than duplicating the significant investment already made by other groups.

Legislative changes

In addition to the comments above, DIGI members would like to propose the following legislative changes to the *Enhancing Online Safety Act 2015* ("the Act"):

1. Remove the tiering of social media services. The current distinction between tiers 1 and 2 has no practical difference. If a social media service is compliant with the basic safety requirements set out in the Act, they should be treated by the eSafety Office as a responsible actor.
2. Remove the need for social media services or relevant electronic service providers to formally apply to be placed in tier 1. If a social media service or relevant electronic service provider meets the basic safety requirements set out in the Act, the Commissioner should have the power to deem the service a responsible actor. This would be a much simpler way of identifying distinguishing responsible companies in the Australian market.