15 October 2021

Snap Inc. response to Draft Online Safety (Basic Online Safety Expectations) Determination 2021 consultation

Thank you for the opportunity to respond to this consultation on the Draft Online Safety (Basic Online Safety Expectations) Determination. We support the aims of the Australian Government's strategy of protecting people from harmful or illegal content and interactions online, and we welcome the proactive and collaborative approach that the Government and the eSafety Commissioner (eSC) have taken to the development and implementation of the Online Safety Act. Snap has been a strong supporter of the eSC since its inception; our longstanding focus on safety by design is closely aligned with the eSC's own priorities, and indeed we worked closely with the eSC on the development of their safety by design tools and resources, which launched earlier this year.

Before we turn to our perspective on the draft Basic Online Safety Expectations (BOSE), it may be helpful to provide a brief introduction to Snap, our approach to safety, and our views on the most effective model for digital regulation.

Snapchat and our approach to safety

Snap Inc. is a camera and technology company that, as well as designing wearable video technology and augmented reality software, owns and operates the visual messaging application, Snapchat. While Snap is still a significantly smaller company than the established tech giants that have dominated online media for the past decade, we are growing, with 293 million people globally now using Snapchat every day.

Snapchat has intentionally been designed very differently to traditional social media. At a high level, we use two principles that guide our design process: safety by design, which is about ensuring the safety of our community, and privacy by design, which focuses on data minimisation and protecting user data. Product counsel and privacy counsel are fully involved in the development cycle of any new product and feature at Snap, from outset to release. This up-front focus on safety and privacy by design is reflected in the build of Snapchat. Unlike traditional social media, Snapchat does not offer an open news feed where unvetted publishers or individuals have an opportunity to broadcast hate or misinformation, and we don't offer public comments that may amplify harmful behavior. Snapchat is at heart a visual messaging application, designed for private communications (either 1:1 or in small groups), with the aim of encouraging users to interact with their real friends, not strangers. Snapchatters' friends lists are only visible to themselves, by default you cannot receive a message from someone who you haven't accepted as a friend, and you can never share your location with someone who isn't your friend. Public areas of Snapchat - our Discover page for news and entertainment, and our Spotlight tab for the community's best Snaps - are curated and pre-moderated, ensuring harmful content is not surfaced to large groups of people. Most content on Snapchat is also designed to delete by default - this means that default settings are that messages and Snaps are deleted

from our servers once they've been opened, while Stories are deleted after 24 hours. This limits how widely content can be shared.

Snap's perspective on effective online safety regulation

We support the case for effective online safety regulation, based around broad principles that companies of all sizes are able to follow and implement proportionately, as relevant to their service and risk profile. Regulation in this area is most effective when it focuses on the principles or outcomes companies should deliver, setting out "what" objectives are to be achieved, without being too prescriptive on "how" companies should achieve these. There is incredible variety in the size, resources and service models of different online platforms. A principles-based approach accommodates this variety and allows for innovative, effective approaches to be developed, while focusing on what is most important: the safety of users.

Where regulation goes wrong is where it becomes overly prescriptive and complex, focusing too much on process rather than outcomes, and assuming that there is a uniform "one size fits all" approach which will work for all online services. Ultimately the companies who are best served by overly prescriptive, complex regulation are the largest firms, with the largest compliance teams, who can easily deal with the bureaucracy involved, which smaller companies (and in particular start-ups and scale-ups) would really struggle to comply with. The Australian Competition & Consumer Commission (ACCC)'s 2019 Digital Platforms Inquiry highlighted structural problems with Australian digital markets, with certain companies identified as having dominant positions in the market, with adverse effects for consumers and businesses. Prescriptive regulation risks exacerbating these imbalances by disproportionately harming smaller challenger companies and strengthening the advantages of those largest players.

Snap comments on the draft BOSE

Overall, we welcome the Government's approach as set out in the draft Determination and accompanying consultation paper. Efforts to establish a common online safety benchmark for all online safety providers are welcome, and should help to drive up the standard across the industry (particularly for smaller companies). Critically, the Government is clear that the BOSE are not intended to "prescribe how these expectations will be met. Indeed, they have been crafted in a way that allows flexibility in the method of achieving these expectations." This approach is in keeping with the principles-based approach to regulation which we have outlined above. The Government's recognition that "service providers are best placed" to identify harmful content on their services, and to choose the best way to address these in the most responsive way, is significant.

While we have a few substantive comments on the draft Determination, set out below, it is worth emphasising that **we consider this a well-considered, practical and effective approach which will have a positive impact on improving online safety in Australia.**

Expectations regarding safe use

The core expectations regarding taking reasonable steps to ensure that users are safe are clear and reasonable, and the supporting examples of reasonable steps are helpful without being constrictive. Likewise, some level of consultation with the eSC about the steps a platform is taking to protect users, and having regard to relevant eSC guidance, is sensible and

appropriate, as is the expectation of a level of joint work and cooperation with other platforms on safety issues (indeed many platforms already do this via existing fora such as the Technology Coalition). Continuing to take reasonable steps to develop and implement processes to detect and address illegal or harmful content, as well as developing further improvements and innovations, if opting to implement end-to-end encryption on elements of an online service, is aligned with Snap's own approach and commitments in this area.

With regard to the additional proposed expectation around anonymous accounts, we agree with the basic expectation that services should take reasonable steps to prevent accounts being used to carry out harmful or illegal activity. Further examples of the kinds of steps that platforms can take include limiting the ability of non-verified users to publicly broadcast content, or - if there are repeated examples of harmful or illegal activity - blocking the device associated with a deleted account from being able to create a new one. Regarding the second example of a "reasonable step" outlined by the Government, though, we should be clear that the debate around requiring ID to use online platforms is a complex and multifaceted one, with several unresolved policy and data privacy challenges. The UK Information Commissioner's Office (ICO), for example, has highlighted the data security risks of the mass collection and retention of identification such as passports or drivers licenses. This should not be considered the only way to achieve such an expectation.

Expectations regarding certain material and activity

We agree with the core expectations that providers should take reasonable steps to minimise provision of certain illegal or harmful material as specified in the Determination, as well as ensuring that measures are in place to prevent access by children to Class 2 material under the National Classification Scheme. **We would recommend adding a reference to the safe build and design of a service under the "reasonable steps" outlined by the Government for the expectation around Class 2 material.** Limiting the ability for harmful and illegal content to be surfaced to users through product design and controls is a critical tool in meeting this objective, and one that is aligned with the priorities and advice of the eSC.

Regarding the proposed "reasonable step" around age assurance, it is worth noting that international regulators have highlighted the complexities and challenges around age assurance and verification, including the aforementioned privacy risks around collecting identification, as well as the unproven nature of new facial recognition technology. While levels of age assurance - a broad term that refers to the spectrum of methods that can be used to inform about a user's age online - may be helpful in meeting this expectation, this should not be considered the only way to do so.

Expectations regarding reports and complaints

We agree with the intent behind these expectations: creating clear community guidelines and terms of service, putting tools in place for users to report violations of these or safety concerns, and then acting upon such reports quickly, is vitally important to keeping people safe online.

Snapchat provides clear and easy-to-use in-app reporting tools across the app so users can quickly report any content or interaction they find concerning; our global, 24/7 Trust & Safety team reviews any reports for violations of our Community Guidelines and takes appropriate action.

This meets the objectives of enabling users to report, and make complaints about, specified material and violations of our Community Guidelines, within one streamlined and unified process. Good practice in this area recommends as few reporting categories as possible, making them as clear and simple as possible, and as few steps or menus to pass through as possible. Reducing friction in the reporting process has been found to greatly increase the number of reports completed. Our process generally requires only two or three clicks for a user to submit a report and enable our team to promptly review.

Making certain information accessible

We agree that providers should make their key terms of use and policies available and accessible to users. Having clear terms and conditions in place, and enforcing these consistently and transparently is an important part of keeping users safe. Snap's Community Guidelines and Terms of Service are publicly available online, easily accessible via the Snapchat app itself and are routinely surfaced to the user in a variety of ways, including on occasion within the content itself. All Snapchatters are made aware of these at account registration. They are also written in straightforward language and as free of legal jargon as possible. We should note, however, that as these documents form part of a binding legal agreement between company and user, they cannot be so simplified as to render them unfit for purpose.

Snap's online Privacy Centre and Safety Centre, also both accessible within the app, are designed to be accessible and visually appealing as well as informative. Instead of just setting out lengthy privacy or safety policies, these provide a range of pieces of bite-size information which are easy to read and digest, written in friendly and informal language, and using icons to tell the story.

Expectations regarding record keeping

Transparency about the amount of harmful and illegal activity identified or reported on online services, and platforms' response to these, is an important part of ensuring accountability. Snap currently goes significantly further than most major platforms in terms of our approach to transparency reporting, by providing country-specific breakdowns of our response to different kinds of harmful activity in our Transparency Report, including a dedicated page for Australia. Our reports focus on key, comparable quantitative information about our response to different types of harmful content. Getting to this stage has been an iterative process: building the tools, systems and teams to enable users to report harmful content, to ensure we are acting on it quickly, and then to set out our response took several years. We will continue to iterate and improve on our approach and include more and more detail with each report. We consider these reports the most effective way of providing valuable, comparable information to the eSC about our efforts to keep Snapchatters safe.

Regarding the proposed additional expectation that services should keep records of reports for five years - our Transparency Reports for previous periods will continue to be publicly available and accessible into the future, to give an overall picture of the prevalence of harmful and illegal content on Snapchat (including specifically for Australia) and our response to this. However, we do not typically retain information on individual reports for lengthy time periods. Snap is committed to data minimisation - this means our products and features are designed to collect

and store as little user data as possible. We apply short retention terms to data and content, which prevents abuse of data from occurring. This approach is aligned with the priorities and objectives of leading privacy regulators around the world. Storing sensitive information about reports of harmful content for such a lengthy time period runs counter to privacy best practice.[1]

We would recommend that while the Government should expect platforms to provide and share transparency information showing their overall efforts to keep users safe, they should not be required to retain sensitive data or information for lengthy time periods if this runs counter to internationally accepted best practice in protecting the privacy of users, their content and data.

Expectations around dealing with the Commissioner

We agree that platforms should be responsive in providing requested information from the eSC. As above, Snap is committed to transparency and already provides detailed information about the prevalence of harmful and illegal content related to Australia in a dedicated page on our Transparency Report. In most instances this will be the most useful source of information in response to the eSC's questions, although of course we are happy to provide more detail on our efforts and approach as required.

We consider that information about reports from users related to harmful content, and platforms' response to these, is the most appropriate metric for the eSC to measure online safety efforts, and we hope the Government takes this into consideration when finalising its Determination.

We would be happy to provide the eSC with a designated contact person for the purposes of the Act.

Conclusion

Thanks again for the opportunity to respond to this consultation. We look forward to further engagement with the Australian Government and the eSC on the implementation of the Online Safety Act.

If you would like to discuss any aspect of this response, please don't hesitate to get in touch.

[1] It is worth noting that Snap has well-established processes for supporting valid law enforcement investigations internationally, including in Australia. Our 24/7 dedicated law enforcement operations team regularly responds to lawful requests from Australian law enforcement to support their investigations, and honours valid law enforcement preservation requests to enable preservation of specific data beyond normal retention periods, when appropriate.