

11 January 2022

## **Snap Inc. Submission to the Inquiry into Social Media and Online Safety**

### **Introduction**

Thank you for the opportunity to present a submission to the Select Committee's inquiry into social media and online safety.

We welcome and support the Committee's efforts to investigate these issues. Computing and technology have become increasingly integrated into our daily lives, and this trend will only continue to accelerate. It is right that policymakers should carefully analyse the decisions that technology companies make as they design the services which we use, and the impact that this has on our safety and security.

At Snap, our goal is to ensure that people can engage in a creative, safe and positive way online. Since the company's founding a decade ago, we have carefully considered the architecture of our Snapchat app, the design of our products and features, our content and conduct policies and their enforcement. We hope that our insights in this space - set out in Part 1 of this submission - will be of use to the Committee as you conduct your inquiry.

We also welcome the Committee's analysis of the various legislative and regulatory initiatives being developed by the Australian Government. We believe that regulation is necessary, and we support the case for online safety regulation that improves the safety of users while also ensuring that the technology sector - and particularly new and innovative challenger companies - can continue to flourish. Our thoughts and suggestions for the development of online policy and regulation in Australia can be found in Part 2 of this submission.

### **Part 1: 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 306 million people globally now using Snapchat every day (over 5 million of those in Australia).

Snapchat has intentionally been designed very differently to traditional social media. At a high level, we use two principles to help guide our design process: **safety by design**, which is about prioritising 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 product and feature development lifecycle, from conception 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, misinformation, or violent content, nor do

we permit public comments that may amplify harmful behavior. Snapchat is at heart a visual messaging application, designed for private communications (either 1:1 or in limited-size groups), with the aim of encouraging users to interact creatively with their real friends, not strangers. 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 numbers of people. Most content on Snapchat is also designed to delete by default: this means that default settings are such that messages and Snaps are deleted from our servers once they've been opened, while Stories are deleted after 24 hours. This further limits how widely content can be shared.

The approach that we have taken - focusing on encouraging communication between close friends rather than strangers, limiting virality and the ability to broadcast - contrasts markedly with the approach taken by traditional social media platforms. Broadly, such platforms aim to operate as digital "town squares," encourage the public broadcast of user-generated content, and rely heavily on artificial intelligence and automated moderation to identify harmful activity reactively.

Credible, external organisations can attest to the overall success of Snap's approach:

Snap is a signatory to the European Commission's [Code of Conduct on countering illegal hate speech online](#), which undertakes an annual collection of reports from NGOs specialising in the reporting of hate speech online. In the Commission's last two annual reports evaluating compliance of the Code, the 39 participating NGOs submitted zero reports of illegal hate speech on Snapchat.

The Conseil supérieur de l'audiovisuel (CSA), the French audiovisual regulator, produces an [annual review of reports of misinformation in France](#) across a range of participating online platforms. In 2020, there were only four reports of misinformation in France on Snapchat; for two other platforms this figure was in the millions.

### **1.1. Our approach to content moderation and enforcement**

Our approach to safety on Snapchat is much broader than purely reactive content moderation. A key part of this is a focus on **prevention**. As set out above, protections and features which prevent people from broadcasting content publicly without it first being moderated, or which prevent strangers from contacting people they don't know, are embedded into the app's design.

Where harmful content and activity takes place, we have effective systems and processes to act quickly. We provide easy-to-use in-app reporting tools so users can notify us of potential safety issues, and our global Trust & Safety team works 24/7 to review user reports and take appropriate action.

Anyone who signs up to use Snapchat agrees to abide by our [Community Guidelines](#). These are publicly available online, and are very simple and clear about what activity is prohibited on Snapchat. When reviewing user reports, our Trust & Safety team will make an assessment of whether the content or user reported has violated these Guidelines, and then take action as appropriate. Actions can include warning the account, deleting the content in question,

terminating the account and deleting the user's data, and/or reporting to law enforcement, depending on the severity of the violation.

We use technological tools to detect abuse and support the team. For example, we use PhotoDNA and CSAI Match technology to identify known images and videos of child sexual exploitation, and we report any instances to the National Centre for Missing and Exploited Children (NCMEC) in the U.S., which coordinates as appropriate with international law enforcement, including the Australian Federal Police. That said, the majority of our content moderation operations involve reviews by a dedicated team of expert analysts.

**Transparency:** Snap publishes bi-annual transparency reports detailing our response to illegal and harmful content on Snapchat. Our [latest report](#) shows that the median turnaround time - the time taken to action a user report - for **every** key category of harmful content was less than 30 minutes, and in many instances even less than that.

Currently, Snapchat is the only major platform to provide country-specific breakdowns of content reported and enforced, including a dedicated page for [Australia](#).

We are committed to continuing to make these reports more comprehensive and informative to the many stakeholders who care deeply about online safety and transparency.

Our approach to content moderation and enforcement has been influenced by long-standing regulatory frameworks that govern broadcast and telecommunications. For instance, when you talk to your friends on the phone, you have a high expectation of privacy, whereas if you are a public broadcaster with the potential to influence the minds and opinions of many, you are subject to different standards and regulatory requirements.

On Snapchat, while we curate and moderate public areas of the app (Discover and Spotlight), we believe that when it comes to our community's private conversations, they have a justifiable expectation that we are not monitoring their every communication, and that is not something we do. There is the limited and targeted exception of the technology mentioned above, which detects known images and videos of illegal child sexual exploitation and abuse imagery. Snapchatters can also quickly and easily report Snaps (visual messages) in private conversations, or report an individual user, which will flag for our team to review.

## **1.2. Protecting young people on Snapchat**

Given Snapchat's popularity among teenagers, we have dedicated significant time and resources to help ensure that younger people have a safe experience when using the app. Through our safety and privacy by design frameworks, we have made a range of design choices to help keep teenagers safe:

- We intentionally make it harder to find others on Snapchat compared to other platforms. For instance, Snapchatters' friends' lists are only visible to themselves; it's never possible to view another user's friends' list.

## Snap Inc.

- Snapchat does not facilitate messaging from strangers, and by default you cannot receive a message from anyone who you have not accepted as a friend on the app. Friend recommendations are expressly designed to prevent sensitive information including friends lists and geographical information from being shared.
- Location sharing on Snap Map is off by default, and there is no option on Snapchat to share your location with anyone other than your friends, or a designated sub-set of your friends. We named the default “off” location-sharing setting “**Ghost Mode**”, with a clear accompanying ghost icon, to ensure that it would be understood by younger individuals, to help them make informed choices about whether to use Snap Map, whether to share their location and, if so, with whom.
- Through our simple, intuitive and easy-to-use in-app reporting tools, we give Snapchatters the ability to quickly and easily report any content they find concerning. We recognise that young people may be reluctant to report harmful content online, and are continually looking for ways to encourage reporting. In 2021, we launched “**Safety Snapshot**”, a dedicated channel on our Discover page which provides advice for users on keeping their accounts secure in a creative and visually accessible way, designed to appeal to young people; one episode focused on debunking common myths about reporting.
- We make no effort - and have no plans - to market Snapchat to children, and individuals under the age of 13 are not permitted to create Snapchat accounts. When registering for an account, individuals are required to provide their date of birth, and the registration process fails if a user inputs an age under 13. We have also implemented a new safeguard that prevents Snapchat users between 13-17 with existing accounts from changing their birthday to an age of 18 or above. Specifically, if a minor attempts to change their birth year to an age over 18, we will prevent the change.
- We are working to develop tools that will offer parents insight into who their teens are connecting with on Snapchat, while respecting the teen’s privacy. The tools are intended as a conversation-starter among parents/guardians and teens.

**Research:** Snap regularly conducts research to better understand our audience. Importantly, this research has informed steps we have taken in the app to better protect our community.

In 2019, we conducted a programme of research looking at how teens and young adults (people aged 13-24) approach mental health and wellbeing across a range of areas. Our research showed that young people are incredibly attuned to, and deeply affected by issues around mental health and wellbeing. It also showed that friends are considered the first line of defence when dealing with these challenges. Rather than going to parents or therapists, young people often turn to their friends first, and having someone to share with was identified as an important first step in addressing these problems.

As a platform designed for communication with close friends, we felt we had a unique opportunity to help address some of these issues, and this insight helped inform our “**Here For You**” initiative, launched in 2020, which we continue to develop and improve.

“Here For You” surfaces in-app support to Snapchat users who may be experiencing mental health or emotional issues - or who are curious about those issues and may want to discuss them with a friend. We have worked with *R U OK?* and *Project Rokit* in Australia to design content tailored for younger people, which links to expert resources and advice, and we are looking forward to building on this offering in 2022.

## **Part 2: Online safety policy and regulation in Australia**

Snap works closely with governments and regulators around the world on the development of laws designed to protect people online. We welcome regulation, but believe it is critically important that it is crafted in a way that ensures new and innovative approaches to safety and privacy can be developed and implemented, and that smaller challenger companies can continue to grow and flourish.

We support the case for coherent, effective online safety regulation, ideally focused in one comprehensive and clear regulatory framework. This is the model being pursued in the European Union through the Digital Services Act, and in the UK through the draft Online Safety Bill.

Online regulation is most effective when it is based on broad principles that companies of all sizes are able to follow and implement proportionately, as relevant to their service and risk profile. Such regulation focuses on the principles or outcomes companies should deliver, setting out “what” objectives are to be achieved, without being too prescriptive as to “how” companies should achieve them. 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.

Regulation goes wrong when it becomes overly prescriptive and complex, focusing too much on process and outputs rather than impact and outcomes. Another shortfall comes in assuming a uniform, “one-size-fits-all” approach exists that 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 that can easily deal with the bureaucracy involved, while smaller companies (and in particular start-ups and scale-ups) would really struggle to comply. 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.

### **2.1. A crowded regulatory environment**

There are a wide range of pieces of new, proposed and existing legislation and regulation in Australia that cover online safety, yet set different and often overlapping requirements on online platforms. A non-exhaustive list includes the **Abhorrent Violent Material Act**, which sets obligations on removing content and reporting to law enforcement; the **Online Safety Act**, the implementation of which includes requirements both under the Government-drafted Basic Online Safety Expectations (BOSE), as well as under a range of codes of practice developed by

industry associations in concert with the eSafety Commissioner; the draft **Privacy Bill**, which includes problematic obligations around age verification and parental consent; and the draft **Social Media (Anti-Trolling) Bill**, which proposes new obligations on collecting, storing and providing personal user information to other users to support defamation proceedings.

The combined effect of this crowded landscape is to introduce confusion, particularly for smaller companies who cannot rely on large compliance, policy and legal teams to help them make sense of the competing obligations, codes and guidance under each piece of legislation. There is a real danger that smaller companies will fall by the wayside; perversely, this will only serve to entrench the advantages enjoyed by the largest companies who have inspired much of this legislation.

There is clearly a need for online regulation in Australia, but we do not consider that the current, highly complex and crowded landscape is the best way to achieve a safer, healthier and more civil online experience for Australians. Instead, we would recommend that the Government looks for opportunities to simplify and rationalise this landscape, around one central and comprehensive regulatory framework (similar to the models being developed in the EU and UK). The **Online Safety Act** could be a good vehicle for this. We consider that the Act, and the Government's position as set out in the BOSE, represent a well-considered, practical and effective approach which will have a positive impact on improving online safety in Australia. Critically, the BOSE are explicitly principles-based and allow for the flexibility that is vital for effective online regulation.

**Snap recommendation 1:** The Government should look for opportunities to simplify the crowded and complex online safety regulatory environment in Australia as a means of supporting innovation and competition. The Online Safety Act could be broadened to establish one central, principles-based regulatory framework for online platforms, echoing the models being developed in the EU and UK.

## 2.2. Ending anonymity online?

Several of the measures which the Government is currently proposing to introduce seem designed to end the right to anonymity online, and would, if approved, essentially mandate tech companies to collect and store people's IDs as a requirement for using online platforms:

- The Privacy Bill contains a requirement for platforms to “verify” the age of users. Age verification is commonly understood to mean determining a person's age with a high level of certainty by checking against verifiable records of data (e.g., through the collection of IDs).
- The Bill also includes an obligation for platforms to “obtain parental or guardian express consent” for users under the age of 16. This would require the collection and retention of IDs en masse for Australians. For such a model to be workable, services would need to verify the identity of both users under the age of 16 as well as their parents or guardians.

- The Social Media (Anti-Trolling) Bill proposes a “complaints scheme” where platforms would be expected to pass on the personal information of users to complainants who feel they have been defamed online.

Allowing people to engage with each other and to express themselves freely has been fundamental to the development of the internet. Sometimes, staying anonymous online is the only way that people can do so, including if they are living in politically repressive countries, are in abusive relationships, or are whistleblowers seeking to expose corruption. Legislating to end that freedom in Australia could have profound consequences, not just for internet users in Australia, but also in terms of the message sent to more autocratic countries around the world. This is before we even touch on the risks of discrimination to members of minority groups who may be less likely to have a Government ID, or the data security risks of mandating technology companies to collect and store people’s IDs and personal information en masse.

There are alternatives that are less privacy-invasive and ethically concerning, but which would still accomplish the Government’s objectives of ensuring Australians are safer online:

- **Age assurance, not age verification:** Instead of asking platforms to “verify” the age of users, the Government should require platforms to develop age assurance techniques. Age assurance is a more comprehensive term that describes a range of methods and approaches to provide an adequate level of assurance that children are unable to access adult, harmful or inappropriate content; and to estimate or establish the age of a user so that a service can be tailored to the needs and protections appropriate to their age. This age assurance model is increasingly being proposed in regulatory and legislative initiatives around the world. Given the issues set out above with age verification, the most privacy-intrusive form of age assurance, most international Governments and regulators have not recommended age verification as a requirement for online services, except for services that are explicitly directed at adults (such as pornography or gambling sites).
- **Parental tools, not parental consent:** Many online platforms and services have either developed, or are in the process of developing, tools that give parents and guardians more visibility about what their children are up to online, and opportunities to spark dialogue to help spot and mitigate risks. Instead of enforcing the mass collection and retention of people’s IDs to develop a “consent” mechanism, the Government should focus on encouraging the development of controls that empower parents to partner with their children in navigating the digital world.
- **Focusing on the systems and processes that help keep users safe:** The best protections and counter-measures against trolling online are not seeking to end anonymity to support defamation proceedings. In any case, the ability to sue someone for defamation is something that is only available to a tiny proportion of internet users. Instead, the Government should focus on ensuring that online platforms have effective systems and processes in place to prevent and protect against abuse, to respond quickly and effectively to harmful content or activity when it does occur, and to account for their actions to an independent regulatory body (in Australia, the eSafety Commissioner). Again, the Online Safety Act is the best vehicle for ensuring this.

**Snap recommendation 2:** The Government should abandon proposals that seek to end the freedom to anonymity online, and instead focus on requiring platforms to take systemic measures which will help ensure the safety and security of Australians online: the implementation of age assurance, parental controls, and systems and processes to help keep users safe.

## **Conclusion**

Thank you again for the opportunity to present a submission to the Select Committee's inquiry. These are critically important issues, and we hope that our insights into Snap's approach to safety, and our thoughts on the development of online safety policy and regulation in Australia, will be of use to the Committee as you develop your findings.

At Snap, we are always striving for new ways to keep our community safe, and we have more work left to do. Online safety is a shared responsibility, spanning a host of sectors and actors. We believe that regulation is necessary - and we are committed to helping governments design effective and lasting online regulation - but regulation alone won't get the job done. Technology companies must take responsibility and actively protect the communities they serve. If they don't, the Government must hold them accountable.

We are committed to doing our part in concert with safety partners including our Safety Advisory Board, technology industry peers, international governments and regulators, and civil society. From technology-focused and awareness-raising initiatives, to research and best practice sharing, we work closely with a wide range of organisations dedicated to keeping people safe online. We also know that there are many complex problems and technical challenges across our industry, and we remain committed to working with partners and policymakers, including the Government and Parliament in Australia, to identify robust industry-wide solutions.