

**Do you think that Social Media Companies owe their users, participants, partners, and content creators a duty of care?**

Yes. We believe social media companies owe a duty of care to their users because their products have both positive and negative impacts on users' mental health.

**Do you support raising the age for social media access?**

While we recognise and share the concerns of parents, caregivers and policy-makers about social media's harmful effects, our submission offers alternative approaches to age restrictions because we believe these will be more effective and impactful in addressing the mental health-related harms associated with social media use by young people.

We argue in favour of mandating safety-by-design approaches to social media product development so that user wellbeing is prioritised over user engagement. As well, there is a need to equip young people and their parents, care-givers and other support networks with the skills and knowledge they need to create safe social media environments.

**Can you explain to us what you mean by limiting infinite scrolling – how would this look practically?**

Infinite scrolling has been identified as a problematic feature of some social media platforms by mental health researchers, clinicians and by young people.

We think this is an area where young people, digital technology and online safety experts, the eSafety Commissioner and platforms (under compulsion if required) should co-design an effective, workable model to redesign a range of social media functionality that prioritises engagement over wellbeing.

**What other kinds of restrictions or regulations do you think Government should mandate for Australian children under 18 or under 16 years of age?**

We believe social media companies, and the design of their products, are an appropriate target for regulation. For example, governments could explore opportunities to mandate that platforms prioritise verified mental health content, make algorithms transparent and give users the ability to reset their algorithm.

We support the eSafety Commissioner's safety by design principles and think they should be applied to social media platforms, including to limit sticky or addictive features that prioritise users' engagement over their wellbeing.

While there is significant scope for new safety features, it is also clear that many young people (and parents) are unaware of existing features or are unsure how to access them. These features must be made easily accessible and, for users under 16, be defaulted on.

### **What kind of data do you think would be useful to researchers with respect to mental health?**

The following information would be useful for our researchers to understand how social media impacts youth mental health more fully:

- **Time and activity type**
  - How much time users spend on specific platforms, segmented by demographic information and activity type:
    - Scrolling
    - Posting
    - Communicating with others
  - Behaviours from the individual that are indicative of addiction e.g., going back and forth from the app or platform, frequency of checking the app or platform
  - Behaviours that are indicative of distress – ie searching self harm, or suicide related content
- **Usage frequency:**
  - How often users spend in individual sessions
  - Any patterns in usage (e.g., late-night usage).
- **Content types:**
  - Data on the type of content that users interact with, and the valence (negative, positive, neutral) of the content along with the time taken, action and engagement type regarding that content
- **Changes in patterns of use:**
  - Information about shifts in behaviours on platforms (e.g., marked increase in passive scrolling, or dramatic changes in post frequency), that could be indicative of a change in emotion.
- **Mental health content:**
  - Types of mental health content accessed
  - What time of day the mental health content was accessed
- **Moderation mechanisms:**
  - Explanation of moderation processes for platforms as they currently are and the threshold for identifying/removing harmful content
  - How many views does the harmful material get (on average), before it is removed

- **Continuous assessment**

- Continual assessment of the platforms as they may change their features, UX/UI, and algorithms, which means the content of the data access and what is necessary to keep the public safe may change as well.

**What key information should these reports include to ensure they are both informative and actionable for policymakers, researchers, and users?**

The following in-depth content should be made available to policymakers and researchers on request:

- **Content prioritisation:**

- How algorithms make content decisions, and what personal information from users informs this process.
- The different algorithms used to determine what is shown, and the algorithm used to identify content that is harmful, and the criteria used to define harmful content.
- Demographic trends outlining what popular content is currently prioritised to different groups and the rationale for why these patterns exist.
- Explanation about how ethical issues (e.g., mental health, behaviour change) are considered in the design and implementation of algorithms, and the policies (if any) in place to safeguard users.

- **Customisable features such as content preferences and screentime limits:**

- Percentage of users in different demographics who currently use each customisable feature including:
  - Content preferences like see more/ less of this
- Percentage of users who comply with these features (e.g., how many users exit the app after reaching their screentime limit or receiving bedtime reminders).
- Data to indicate if these features actually function effectively:
  - If a user states they are not interested in seeing specific content, to what extent is this content actually restricted for how long?
  - How often is the function to delete data used by individuals?

- **Advertisements:**

- Demographics: data outlining how advertisements are prioritised to specific demographic groups.
  - What are the processes for reporting advertising funding and what is done in response to that
  - Rates of engagement and uptake rates for different advertisements segmented by demographic, geographic information

- Regulations: provide current rules and guidelines for what can be advertised and to who
  - How successful are current regulations on this (if they exist) and how many times has action been taken when inappropriate advertising has been flagged?
- Revenue: provide advertising revenue, segmented by demographic groups.
- **Existing safety audits**
  - Some tech companies already have workshops with experts to help determine their policies, responses, detection methods for distress and suicide, however, this isn't mandated, and they don't have to use the results. We would like information on existing safety audits that have taken place from different social media companies and the results/ impact of these

**Have you worked with the social media/tech companies to develop tools to direct people to resources/help/support? Such as pop ups directing to help or support? Similar to Butterfly work with Meta to develop pop ups when people search eating disorder content they will get redirected to support. If so, could you provide examples of this and any data measuring success or not.**

ReachOut and Black Dog Institute work with social media and technology companies to connect people to information and support within platforms, and to pathways to further support outside of platforms.

Examples include:

- ReachOut has partnered with Snap on on-platform activations that promoted mental health and wellbeing. These include Snap's Here For You in-app portal, which provided support by surfacing resources to users searching topics related to friendship, stress and self-esteem. This portal was a significant success, with more than 132,000 users directly accessing ReachOut resources on these topics. A key – and welcome – feature of this portal was that it was structured to take users out of the Snapchat platform to engage with support resources on ReachOut's platform, providing easier pathways to further topics and support options.
- ReachOut partnered with Meta to amplify social and emotional wellbeing content supporting Aboriginal and Torres Strait Islander young people in the lead up to and aftermath of the Voice to Parliament referendum. During the peak campaign period (mid-July to mid-Oct) this support contributed to our overall reach of 3.5 million across social media platforms, and more than 50,000 direct on-site service engagements.

- The Black Dog Insitutute developed The Lived Experience Hope Exchange suicide prevention initiative as a part of the Centre of Research Excellence in Suicide Prevention (CRESP) in partnership with tech platforms including Google Ads, Reddit and Microsoft Bing. The project utilises series of advertising campaigns that are distributed via social and digital channels to identify individuals at risk of suicide by triggering ads when users search for suicide-specific keywords. The website offers emergency help links, community stories, and calming resources. The impact of the Hope Exchange has been significant. From the funding provided to us by Reddit, in 14 days, our ads were shown in communities where distress and suicide is often discussed, resulting in over 49,000 clicks on an ad explicitly targeting individuals experiencing suicidality or acute distress. Our estimates suggest that up to a third of people who visited the BDI website pages had not previously sought help. These results indicate that this project offers an engaging, far-reaching and cost-efficient way to reach significant numbers of people at risk of suicide