

Commissioner Meeting Brief

Meeting with Google (YouTube)

Brief#: CMRB_25/0046

To	The eSafety Commissioner
Cleared by	s 22, Executive Manager, ICE Branch
From	s 22 Manager, Industry Supervision
Meeting Details/Subject	Meeting with YouTube
Date of meeting	Thursday 25 September 2025
Time of meeting	1.45PM – 2.45PM
Location	1600 Amphitheatre Pkwy, Mountain View, CA 94043
Subject	YouTube's implementation of the social media minimum age restrictions
Purpose	<p>To:</p> <ol style="list-style-type: none">1. Discuss eSafety's preliminary assessment on whether YouTube is an age-restricted social media platform2. Discuss the released SMAR regulatory guidance3. Seek an update on how YouTube plans to meet its obligations, if it is an age restricted social media platform
External Attendees	
Internal Attendees	<p>Julie Inman Grant</p> <p>s 22</p> <p>s 47E(c), s 47F</p>

Social Media Minimum Age

Key points

- As part of our consultation to develop regulator guidance, eSafety began seeking meetings with relevant online services, including YouTube, from July (we first wrote to Google on 21 July seeking a meeting as part of our consultation). Google declined.
- Of the 6 services named by the Prime Minister and Minister, YouTube is the only one who has not engaged with eSafety and participated in consultation to inform the regulatory guidance.
 - All platforms saw the final Rules at the same time when they were released on 30 July, following the public release of eSafety's advice to the Minister on 23 June.
- On 15 August eSafety sent Google a [letter](#) setting out eSafety's expectations for 10 December, noting that eSafety would be undertaking assessments to give a preliminary assessment of whether YouTube is an age restricted platform with obligations.
- A response was received from YouTube on 29 August in which they stated that it does not consider itself to be in scope of the SMMA and maintains that it is a video-sharing service distinguishable from traditional social media.
 - On 4 September you met with YouTube and Google Regulatory Affairs virtually to discuss SMAR. YouTube reiterated its view that it is not an age restricted social media platform with obligations and did not provide any new updates. You invited them to provide any information to inform our assessment.
- On 16 September eSafety released the Reasonable Steps Guidance outlining a principle based and technology neutral approach.
- On 19 September **YouTube wrote to you stating that they do not consider YouTube to be an age restricted social media platform**, based on the specific characteristics of how it is used by Australians. They state YouTube's primary purpose is as a video-sharing library and streaming platform, and any social media like features are not a significant purpose of the service, but ancillary to the primary purpose of video-sharing and streaming.
- The letter does not provide any data or evidence as to their statement of usage by Australians.

Key messages

Assessment

- Offer to discuss YouTube's initial assessment that it's not an age-restricted social media platform and walk them through the assessment framework released on 4 September, which is designed to help services determine whether they meet the criteria.

- eSafety is undertaking a preliminary assessment of YouTube. YouTube will have procedural fairness in this process.

Regulatory guidance

- Offer to provide an update on the regulatory guidance, which is based on our consultations (summaries now available on the eSafety website) as well as the final report of the AATT (released on 1 Sept to all stakeholders simultaneously); all companies will receive the regulatory guidance at the same time.
 - Reasonable steps consist of systems, technologies, people, processes, policies and communications which support compliance with the obligation. The guidelines discuss each of those elements, with a particular focus on age assurance.
 - eSafety considers it is reasonable for platforms to take a layered approach across the user journey and to implement a range of measures to meet the SMMA obligation.
- Whether a provider has taken reasonable steps will include an assessment of the holistic impact of all steps taken by a provider, across the service. Measures will not be evaluated in isolation. This is about systems and processes, not individual accounts.
- As set out in our 15 August letter, the initial focus for 10 December should be detecting and deactivating existing accounts held by under 16s – with proper advance warning and supportive messaging, as well as the ability to appeal if a mistake has been made – and taking reasonable steps to prevent those underage users from immediately creating new accounts (including relying on signals already in use to prevent suspended account holders from coming back on the service).
- Self-declaration by itself will not be sufficient. eSafety expects YouTube to monitor the number of new accounts being created from 10 December. If there is a significant increase, from 10 December, YouTube should consider if further steps are needed to comply with the legislation.

US age assurance

- eSafety welcomes understanding:
 - whether YouTube is considering applying similar approaches in Australia
 - what third-party tool is being used for selfie-based age estimation and how it ensures privacy and accuracy

BOSE

Key points

- On 23 July 2025, you met with s 47F

s 47E(d)

-
-

Expectations. Ahead of the report 2 deadline in September, Google and eSafety agreed on alternate metrics. eSafety is currently assessing Google's response to report 2. Both parties agreed to continue discussions, including at upcoming consultations and international forums.

Key concerns from report 1

- BOSE report 1 identified several safety gaps in Google's services:
 - Google does not use CSEA livestreaming detection tools across all parts of its services.
 - Google does not deploy CSEA detection tools across all key surfaces of their services
 - Many Google services do not use language analysis tools to detect sexual extortion.
 - Google confirmed it had made no improvements to these safety gaps in its response to report 2.

Preliminary improvements in report 2

- Compared with report 1, Google's report 2 included some positive developments
 - Google is using more hash matching tools to detect known CSEA on Gemini training datasets.
 - Google commenced blocking URLs linking to known CSEA material on YouTube.
 - Google Introduced in-service user reporting options for CSEA in Gmail and harmful content in Messages.
 - Google started ingesting signals from Lantern and joined the Take It Down initiative to help reduce the spread of CSEA material.

Key messages

- Acknowledge the positive tone and willingness to re-engage shown by Google in the July 30 meeting and subsequent engagements regarding BOSE. However, note that beyond BOSE, this has not been the case.
- Acknowledge Google's notable improvements between periodic reporting notices

Phase 1 App Store Code

Key points

- One of the Taskforce's priorities are high reach, large services that can play gatekeeper roles, such as app stores.
- On 22 August 2025, eSafety notified Google that a formal warning had been issued to Ome.TV for breaching the RES Standard. The Ome.TV app remains available on the Play Store.

- s 47G

s 47E(d)

- s 47G

Key messages

- eSafety expects Google to demonstrate sustained improvement in transparency and compliance across their obligations under the App Code.

Phase 2 Codes

Key points

- On 9 September, eSafety registered 6 Industry Drafted Codes in relation to age-restricted material (the Phase 2 Codes). These Codes will come into effect in 6 months, from 9 March 2026. eSafety previously registered 3 of the Codes on 27 June 2025.
- The Codes focus on age-restricted material as defined under the National Classification Scheme (as opposed to SMAR which focusses on account creation). The priority material

covered by the Codes is online pornography, self-harm material, high-impact violence material and simulated gambling.

- Google is a member of DIGI and was directly involved in drafting the Codes.

Key messages

- eSafety seeks early engagement with Google on how it will implement the Search Engine Services Code, registered on 27 June 2025.
- Age assurance measures must be in place by June 2026 for account holders, to filter out online pornography and high-impact violence from search results.
- For non-logged-in users, default safety measures must apply, including:
 - Blurring of explicit images in search results.
 - Crisis support material surfaced in response to certain search terms.
- The Code also requires consideration of technical accuracy, robustness, reliability, fairness, and privacy impact when implementing age assurance.
- Google may also need to address safety guardrails or age assurance for high-impact generative AI features, such as AI companion chatbots capable of generating sexually explicit or self-harm-related material.

Risks and sensitivities

- s 47E(d)

-

Background

September 4 meeting

- On 4 September you met with s 47F [REDACTED]
[REDACTED]
[REDACTED] virtually to discuss SMAR.
 - eSafety confirmed it will proceed with its assessment using the published self - assessment tool and encouraged YouTube to share any information that may assist.
 - YouTube shared that its current age estimation trial in the United States applies protections across YouTube, Search and Ads. Users can correct age inferences using a government ID or a selfie.
 - YouTube confirmed that a third-party tool is being used to support age inference correction but did not specify which tool.
 - eSafety reiterated that platforms are expected to begin deploying existing age assurance tools in Australia by 10 December and take reasonable steps to prevent circumvention, such as VPNs.

U.S Age Assurance Trial

- In July 2025, YouTube announced that it is rolling out a machine learning-based age estimation system to a small group of U.S. users where the goal is to differentiate between teens and adults to apply age-appropriate protections, regardless of the user's declared birthdate.
- YouTube has used similar technology in other markets and this builds on previous initiatives like YouTube Kids and supervised accounts for teens.
- The system uses AI to infer age based on search behaviour, viewing history, and account longevity
- YouTube will apply protections such as disabling personalised advertising, activating digital wellbeing tools, and adding safeguards to recommendations (including limiting repetitive views of certain content) if a user is identified as a teen
- Users incorrectly flagged as under 18 can verify their age using a credit card or government ID.
- Only users inferred or verified as 18+ can access age-restricted content.
- YouTube will monitor user experience and collaborate with creators to ensure ecosystem-wide benefits.

Google Credential Manager API on Android and Google Wallet

- Google is integrating age verification into its Credential Manager API and Google Wallet, allowing users to store and share digital credentials with apps and websites.

- In early 2025, Google launched ID Pass in the UK and US, enabling users to verify age and identity using a passport and facial video. The ID Pass is encrypted in Google Wallet but is not a government-issued ID.
- To set up the ID pass, Google uses a valid passport and a brief video of the user's face to verify it's the user. Google uses the passport information to confirm that the passport is valid and accurate, reviews the video to confirm the user is a real person, and compares the video to the passport photo to confirm the user is the owner of the passport.
- Google uses Zero-Knowledge Proof (ZKP) technology to allow users to verify age without revealing identity. ZKP is being integrated across Google Wallet and other products, with partners like Bumble using it for age assurance.
- In July 2025, Google announced a partnership with Sparkasse (Germany) to support EU age verification, though progress is unclear. Google plans to open source ZKP for broader adoption.

Behind the Screen

- According to eSafety's Behind the Screen report (Feb 2025):
 - YouTube was the most-used platform by children under 13, with 68% of surveyed 8–12-year-olds reporting use. Many accessed it without an account or via Family Link.
 - Google shared that it uses machine learning to estimate age based on account activity. Suspected underage users are given 14 days to verify or add supervision before account deletion. Verification options include government ID, credit card, or selfie-based age estimation.
 - Google was the only major platform that did not provide key safety metrics, such as the number of underage reports, enforcement actions, or response times—flagged by eSafety as a concern.
 - Reporting underage users on YouTube requires navigating a legal issues webform, including citation of laws, legal name, and detailed explanation.
 - Google reported 15,500 account bans between Jan–Jul 2024 but did not specify whether these were triggered by proactive tools or user reports.