SQUIRE PATTON BOGGS

Phase 2 Online Safety Codes

New Compliance Obligations Ahead

Australia – November 2025

At a Glance

What's Changing?

From December 2025, the eSafety Commissioner's Phase 2 Online Safety Codes introduce new obligations for online services to manage access to lawful but restricted content – including pornography, simulated gambling, graphic violence and restricted AI-generated material.

Who's Affected?

The codes apply across eight sectors of the online stack, including social media, messaging apps, websites or applications, app stores, hosting providers, search engines, device manufacturers, internet service providers (ISPs) and generative AI services.

What's Required?

Providers must risk assess their services and, where required, implement reasonable age assurance and access control measures to prevent children from accessing restricted content. These may include age verification, biometric age estimation and parental consent. In addition, some providers must continually seek to improve their online detection tools for such content, as well as implementing appropriate safety and reporting tools to reduce the likelihood and number of children viewing such content.

In this article, we break down what these changes mean for businesses by providing practical steps, legal considerations and sector-specific guidance.

The eSafety Commissioner's Phase 2 Online Safety Codes will begin to take effect from December 2025, imposing new and far-reaching obligations on online and digital service providers. While Phase 1 dealt squarely with unlawful material such as child sexual exploitation and terrorist content, Phase 2 addresses "lawful but restricted" material (i.e. pornography, simulated gambling, graphic violence, self-harm material and, increasingly, restricted content generated by AI chatbots) and captures a broader range of entities in its net.

Who Do the Codes Apply To?

The Phase 2 codes apply across nine distinct sectors of the online ecosystem, capturing services at every layer of the technology stack. These sectors include:

Code	Description	Effective Date
Internet Carriage Services (ICS)	ISPs that provide services to end users.	December 2025
Hosting Services (HOS)	Services that host stored material provided by either the service itself or another party.	December 2025
Internet Search Engine Services (SES)	Platforms that collect, index, organise and return search results, including AI-generated results.	December 2025
Social Media Services (SMS)	Platforms enabling social interaction, content sharing and discovery.	March 2026
Messaging Features Within Social Media (MMS)	Instant messaging features within social media services that allow for private communication.	March 2026
App Distribution Services (ADS)	Platforms that allow users to download apps via a carriage service.	March 2026
Equipment Providers (EQP)	Manufacturers, suppliers and installers of internet-enabled devices and operating systems.	March 2026

Code	Description	Effective Date
Designated Internet Services (DIS)	A broad category including websites, applications and file storage services that do not fall within another category (e.g. social media or relevant electronic services).	March 2026
Relevant Electronic Services (RES)	A broad category including, but not limited to, communications relevant electronic services (e.g. instant messaging services and some email services), gaming services and dating services.	March 2026

What Do the Codes Require?

At the core of Phase 2 is a general requirement to risk assess potential access to harmful material and determine what steps are "reasonable" to prevent access by children, depending on the nature of your service. Unlike the clear prohibitions in Phase 1, the Phase 2 codes shift responsibility to businesses to assess the risks of Class 1C and Class 2 material within their services and justify the measures they adopt. As a reminder, Class 1C and Class 2 material includes content that is age inappropriate for children (like pornography, high-impact violence and material relating to self-harm) but not illegal.

The codes acknowledge that "reasonable" steps, in the context of age assurance and access control, will depend on the type of service, the functionality offered and the risks posed. It will also depend on the risk tier of the relevant service. The age assurance measures envisaged under the codes were clarified as part of the Age Assurance Technology Trial report (which we reported on in September) and include:

- Age verification – encompassing ID or payment card checks
- Age estimation – using biometric cues such as voice or facial analysis
- Age inference – drawing on account data or user behaviour
- Parental control and consent mechanisms
- Successive validation, which combines two or more of the methods above

In addition, the codes contain other obligations related to promoting and protecting online safety. These obligations vary depending on the applicable code but, where end users are involved, they can include:

- Appropriately drafted terms and conditions prohibiting age-inappropriate content
- Effective reporting mechanisms for breaches of terms and conditions
- Tools, features and settings that limit children's receipt of unsolicited material (including Class 1C and Class 2 material), such as blocking direct messages (DMs) and automatically blurring age-inappropriate content

The codes do not prescribe a single method for meeting any of these obligations. Instead, they expect providers to consider their service, assess the level of risk and adopt proportionate controls.

What Do the Codes Mean for Providers?

With staggered commencement dates approaching quickly, businesses may want to start considering how the codes will affect them and their operations by taking the following steps:

1. Conduct a High-level Risk Assessment

All providers should assume they are in scope unless clearly excluded. The first step is to assess the risk of the service by reference to the relevant code, including by reference to features like:

- The purpose of the service and its functionality
- The likelihood of children accessing, or being exposed to, restricted content
- The types of content hosted, surfaced or distributed
- Existing safeguards or moderation tools

Risk levels are generally set at three tiers: Tiers 1, 2 and 3.

Unlike the *Privacy Act 1988* (Cth) (Privacy Act) threshold (AU$3 million in annual revenue), the Online Safety Code obligations apply to providers within the nine previously mentioned categories regardless of enterprise size. Given the breadth of the DIS and RES codes in particular, if a provider operates any sort of communication service, file sharing service or website or app (even with limited upload and download functionality), it is important to conduct and record the results of your risk assessment.

2. Evaluate Age Assurance Options

Providers should consider age assurance mechanisms that are appropriate to the risk identified.

These may include:

- Age inference using existing user data
- Parental consent workflows
- Biometric estimation or ID verification for high-risk content

For platforms without the data richness of social media, age inference alone may not be sufficient. Additional steps, such as onboarding changes or content gating, may be required. Alternatively, services may wish to consider other forms of age assurance, such as those that rely on off-site behaviour. For example, the UK's Ofcom guidance on highly effective age assurance promotes "email-based age estimation" – which analyses other online services where a user's email address has been used to help assess user age, on the basis that certain services (e.g. mortgage brokers) are more likely to be accessed by adults than children.

3. The Codes Capture More Than Just Content Distributors

Responsibility under the codes does not rest solely with content platforms – even infrastructure-level services may need to implement safeguards or support age assurance mechanisms. Equally, each layer of the technology stack has different risk exposure. For example, the HOS code specifically recognises that the role of hosting services is "different from other service providers who may have a direct relationship with end-users".

It is critical for services to understand the measures that they are required to implement and – especially for services that regularly interact with other providers – what the codes mean for their relationship with those providers. For example, the ADS code requires an app distribution service provider to "take appropriate action" where a third-party app provider has breached its agreement with the ADS provider by failing to implement appropriate age assurance measures and access control measures in relation to high-impact or gambling apps. Apple and Google's recent decision to remove OmeTV from their Australian app stores after contact from the eSafety Commissioner may help to show the Commissioner's expectations in this regard.

4. Document and Justify Your Approach

Even low-risk Tier 3 providers must maintain records of their risk assessment and rationale. While formal mitigation may not be required for these entities, the expectation is clear that, at a minimum, risk assessment is mandatory and documentation is key to ensure your business is covered.

5. Context Is Key

In many ways, Class 1C and Class 2 materials pose more challenges than materials which are outright illegal. Many of the user-facing codes require certain online services to employ enough trust and safety personnel to oversee the safety of a service. Those trust and safety personnel will need to ensure that their approach to moderating content is appropriate to the audience that they are serving and the particular risk posed to children.

Additional Considerations

Privacy Law Intersections

While the codes require age assurance and data-driven risk assessments, providers must ensure that these measures do not breach other Australian laws, particularly the Privacy Act. This is especially relevant for smaller providers that may not meet the AU$3 million threshold for mandatory compliance under the Privacy Act but still need to handle personal information responsibly when implementing age assurance.

The potential conflict between online safety and privacy is well-understood. We recommend that providers generally craft their age assurance methods to be appropriate to the risk posed by the service and limit the data collected to that risk. For example, a gaming service with limited communications functionality that is targeted to children under 16 may not need to apply age verification.

Conclusion

These codes are comprehensive and contain obligations which will affect a number of digital service providers, including those who may not have considered online safety as part of their compliance requirements. For assistance in conducting or actioning risk assessments or anything else, please get in touch.

Authors

Tanvi Mehta Krensel
Partner, Sydney
T +61 2 8248 7810
E tanvi.mehtakrensel@squirepb.com

Eman Mourad
Paralegal, Sydney
T +61 2 8248 7831
E eman.mourad@squirepb.com