

**Australian Information
Industry Association
Response to the Online Safety Amendment
(Digital Duty of Care) Bill 2026 - Exposure
Draft**

About the AIIA

The Australian Information Industry Association (AIIA) is the nation's peak body for those in the digital ecosystem, leading strategic policy and advocacy to shape a thriving digital sector. Through strong engagement with government, industry, and the broader community, the AIIA ensures the voice of its members informs decision-making on technology, innovation, and digital capability.

Membership provides direct access to influential networks, premium events, and opportunities to collaborate on initiatives with the sector's best and brightest to drive industry growth, improve productivity, and secure Australia's place as a global technology leader. AIIA members access real collaboration, real connections, and real outcomes.

Support for prevention with clear industry responsibilities

The Australian Information Industry Association (AIIA) supports the Government's objective of preventing online harm and strengthening accountability for the design and operation of digital services. Australians should be able to use technology with confidence, and children require protections that reflect their age and vulnerability. Clear, effective regulation also supports responsible investment and adoption across Australia's digital economy.

The proposed duty offers an opportunity to embed safety throughout product development and operation. eSafety's Safety by Design framework already provides a foundation through provider responsibility, user empowerment and accountability. AIIA supports building on that approach through enforceable, proportionate obligations and practical guidance.

Our recommended amendments would connect responsibility to control, clarify the harm and design provisions, and establish a workable transition. This matters across the technology supply chain. A company running an online platform can control what content it recommends to users. Businesses that install equipment or provide internet connections usually cannot control those recommendations or see what users are doing. The legislation should recognise those differences expressly.

Furthermore, the AIIA has undertaken important work in this space, bringing industry, government and civil society together to advance the conversation on specific digital harms. The AIIA supports a proactive duty that translates the urgency of protecting people into measures that work. Our investigation of specific harms has illustrated both the need for intervention and the difficulty of designing effective safeguards across everyday technology. The government needs industry's technical expertise alongside independent research, law enforcement and protection systems. Collaboration should strengthen accountability and accelerate effective action.

1. Match obligations to control and the service provided

Proposed ss 25A (scope and responsibility), 25H (the “reasonably practicable” test) and 26 (the digital duty of care) should connect obligations to the functions performed, the provider’s material contribution to risk and its practical capacity to prevent or mitigate harm. Responsibility should extend to foreseeable misuse that the provider can reasonably address. Where responsibilities overlap, providers should cooperate, and contractual arrangements should not displace statutory duties.

Section 26 should be framed as a systems-based obligation to establish, implement and maintain effective safeguards and processes that, so far as reasonably practicable, provide a safe online environment. Providers should be accountable for the adequacy of measures within their control or reasonable influence, rather than treated as guarantors against every act of a user or malicious third party. This is consistent with the approach industry understood was under consideration through the earlier Digital Duty of Care Issues Paper.

Clarify the broad equipment limb in section 25A(1)(h), including its application to manufacturers, resellers, installers and maintenance providers. The legislation should clearly explain which technology suppliers are covered and what is expected of each. Businesses that manufacture, sell, install or maintain equipment have different levels of control over its use. Practical examples would help ensure each business takes appropriate safety measures within its control, while identifying where responsibility rests with another provider.

Clarify that day-to-day control under s 25A(2)(b) requires substantive authority over the operation, configuration or safety settings of the regulated service. It should not arise merely because an employee or contractor performs routine technical administration under another person’s direction.

The Act should provide a targeted statutory exclusion or reduced obligations for closed enterprise services, assessed by their actual functions and risks. Relevant considerations should include whether access is confined to authenticated business or government users, whether the service is offered to the public or enables public dissemination, the provider’s lawful visibility of user content, and its control over configurations that materially contribute to relevant harm. B2B status alone should not determine scope. Other demonstrably low-risk services should have access to transparent, conditional relief under s 25A(3), while a high-risk service should not be exempt merely because its provider is small.

2. Preserve proportionality throughout the duty

AIIA supports a risk-based, systems-focused duty and the “reasonably practicable” standard in s 25H. The same standard should govern measures required under s 26(3). Section 26A should require providers to identify reasonably foreseeable risks of harm within the scope of the digital duty, rather than all risks of any kind. The depth, frequency and documentation of assessments should be proportionate to the service’s risk, reach

and capacity to influence harm. Annual full assessments should be confined to prescribed higher-risk services; lower-risk or standardised services should be reviewed following a material change or at proportionate, risk-based intervals. Providers should be able to reuse materially equivalent assessments completed under existing codes, standards or other legal regimes.

Compliance should be assessed by whether the provider took reasonably practicable measures to identify, prevent and respond to foreseeable harm. The occurrence of harm is relevant evidence, but does not, without more, establish a breach. Where multiple persons are responsible for the same service under section 25A(2), the reasonably practicable standard should be applied separately to each person by reference to that person's functions, knowledge, technical capability, contractual authority and practical influence over the relevant risk.

The guidance published on 26E should explain acceptable evidence, residual risk and the treatment of emerging threats. The reasonable reliance on relevant standards and tested controls as evidence of compliance should be recognised, subject to their suitability and effective implementation.

3. Make harm categories clear and context sensitive

AIIA supports strong protection against child sexual exploitation, grooming, sexual violence and other serious harms. Sections 25C–25E should distinguish harmful encouragement or facilitation from legitimate reporting, education, prevention, research and survivor testimony. The current application may capture legitimate material without sufficient attention to context. These contextual considerations should not operate as categorical exemptions, but should inform whether material encourages, facilitates or causes the harm addressed by the relevant provision.

Provide statutory contextual safeguards which preserve access to legitimate sexual health information, eating-disorder recovery resources and reporting about violence, while maintaining prohibitions on abusive material. Clarify the meanings of harm and serious harm, including cumulative exposure and effects on vulnerable groups, so businesses can apply consistent thresholds.

Build on section 27B's requirement to obtain the Commissioner's advice. Before expanding harm categories, require published evidence, consultation with affected groups and industry, a rights and implementation assessment, and reasonable notice. Extend these safeguards to service determinations under section 25B(2) and tools under section 26(4). Any urgent exception should be time limited and subject to prompt review.

4. Target harmful design and support meaningful user choice

Sections 25F and 25G should focus on harmful configurations, deployment patterns, default settings and optimisation objectives, rather than deeming neutral technical features harmful in themselves. Restrictions should be evidence-based and proportionate to the risk presented. Providers should be permitted to demonstrate that an alternative

design achieves an equivalent or stronger safety outcome. Enforceable restrictions remain appropriate where a configuration presents an unacceptable risk to children.

The European Commission's minors guidelines offer a useful comparator: they address persuasive design and age-appropriate controls within a proportionate framework. Australian rules should require demonstrable protection while allowing providers to show that an alternative design achieves the required outcome. The comparator informs implementation. It does not establish that any particular Australian restriction will be effective.

AIIA supports meaningful choice over personalised feeds where there is a material connection between the covered feature, the content it delivers and harms within the scope of the duty. Instruments under s 26(4) should therefore be limited to higher-risk services or features using evidence-based criteria addressing risk, reach, design and user vulnerability; they should not automatically capture routine product or entertainment recommendations. For covered services, the instruments should require clear options, persistence of the user's selection and protection against interfaces that undermine that choice. Personalised discovery should remain distinguishable from safety filtering, spam prevention and delivery of content a user has chosen to follow.

5. Integrate assessments into product development

Under Section 26, businesses should be able to update the relevant parts of an existing risk assessment when a material change affects the service's risk profile. Routine updates should receive checks proportionate to their risks, with decisions and reasons recorded. Urgent security or safety fixes should follow a faster, documented process, with a prompt review afterwards.

The Government should provide assessment templates suited to service type and risk. Allow reuse of evidence from equivalent overseas assessments and existing security, privacy and AI governance processes, supplemented where needed to address Australian users' circumstances. Certification alone should not establish that all online safety risks are managed. Any determination under s 26A(2)(F)(iii) should be developed through consultation and should recognise established risk, security and AI governance frameworks. Implementation of a relevant standard should constitute evidence capable of supporting compliance where the standard is suitable, effectively implemented and supplemented by a service-specific assessment; it should not operate as an automatic safe harbour.¹

The legislation should clarify that retaining assessments for six years does not require retention of unnecessary personal data or harmful source material. Providers should

¹ Standards Australia, *AS ISO 31000:2018, Risk management - Guidelines*; *AS/NZS ISO/IEC 27001:2023, Information security, cybersecurity and privacy protection - Information security management systems - Requirements*; *AS ISO/IEC 42001:2023, Information technology - Artificial intelligence - Management system*; and *AS ISO/IEC 23894:2023, Information technology - Artificial intelligence - Guidance on risk management*.

generally retain their reasoning, control evidence and review history, using minimised or de-identified supporting information where feasible.

6. Give AI providers a practical allocation of responsibilities

Clarify which AI services are covered by s 25A(1)(j), including whether sharing AI-generated material must be an inherent function of the service and how the provision overlaps with other service categories. Responsibilities should be allocated among model developers, service providers, deployers, system integrators and organisations configuring AI within their own environments according to each party's control over model development, guardrails, system configuration, deployment context, user access and monitoring. Guidance should include examples covering consumer chatbots, AI companions, workplace assistants, embedded AI services and tools operating on users' devices.

Guidance should explain how AI providers can test for predictable misuse, protect against sexual exploitation and dangerous advice, monitor harmful exchanges and involve human support where appropriate. Assessments should consider how risks develop over a conversation and after significant changes to the AI, not just individual responses. These proposed measures should reflect each service's risks and capabilities.

NIST's Generative AI Profile provides a cross-sectoral resource for incorporating trustworthiness into AI design, development, use and evaluation.² AIIA recommends accepting relevant evidence produced through such frameworks, mapped to Australian online safety requirements. This would help businesses invest in substantive safeguards and avoid reproducing the same evidence in incompatible formats.

7. Protect privacy and security alongside online safety

Retain section 25H's privacy consideration and section 26(7)'s exception for lawful private communications solely between consenting adults. Clarify that the duty itself does not authorise unlawful interception or require providers to weaken encryption or create systemic security vulnerabilities. The exception should not be implemented through indiscriminate inspection of private communications to establish whether it applies.

Develop joint eSafety and OAIC guidance on age assurance and safety data. Require data minimisation, purpose limitation, secure handling, limited retention and accessible challenge mechanisms for incorrect age decisions. Permit privacy-preserving approaches where they provide adequate assurance for the relevant risk. Assess whether a proposed safety intervention creates additional exposure to identity theft, surveillance or exclusion.

8. Enable independent research with enforceable safeguards

AIIA supports the objective of enabling legitimate independent research, but does not support the proposed data-access framework in its present open-ended form. Essential

² Autio et al, 'Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST Trustworthy and Responsible AI', National Institute of Standards and Technology, Gaithersburg.

limits should be fixed in the Act, including eligible researchers, approved purposes, relevant categories of data, necessity and proportionality, privacy, security and confidentiality protections, grounds for varying or refusing access, reasonable cost recovery and independent review. Subordinate data-access rules should address technical and administrative implementation and should not be capable of materially expanding the scheme's substantive scope without further parliamentary scrutiny.

Access should use the least intrusive method reasonably capable of meeting the approved research purpose. This should include controlled research environments, de-identified or equivalent datasets and restrictions on exporting raw data where necessary to protect users, security, legal privilege or commercial confidentiality.

Provide a prompt mechanism to resolve disputes about scope, feasibility, security and reasonable costs. Permit an equivalent dataset or access method where it meets the research objective with less risk. A provider should also be permitted to suspend or refuse a particular access method where it can demonstrate that access would create a material cybersecurity risk, contravene another legal obligation, or expose information belonging to customers or third parties without adequate protection. Any refusal should be supported by reasons and remain subject to the dispute resolution process. Commercial confidentiality should receive effective protection without becoming a general veto over scrutiny.

For fictitious research identities under Division 3, ss 205G–205L, and particularly the operative authorisations and safeguards in ss 205H and 205J–205L, the Act should define the scope of authorised conduct, necessity, approved research purposes, ethics oversight, prohibited conduct and incident-reporting requirements. Any authority capable of displacing other laws, agreements or provider policies should extend no further than is necessary for the approved research. Providers should retain ordinary defensive measures against conduct that cannot reasonably be verified as authorised research.

9. Make rapid intervention precise and reviewable

AIIA supports action against abusive synthetic intimate imagery and faster help for people facing harm. For section 86B, define fake nude material clearly and publish how predominant purpose or use will be established. Distinguish abusive services from legitimate general-purpose tools that are misused. Notices should identify the app, website, links and required action precisely, with reasons, geographic scope and a process for correcting errors. Where practicable and consistent with urgent victim protection, the operator of the affected app or website should receive notice and reasons and should have an avenue to provide relevant information or seek review. Lack of prior notice should not prevent urgent action where delay would materially increase harm.

Retain the capability qualification, power to allow longer compliance periods and Tribunal review. For reposted material, require sufficient identification and confirmation that the altered material still meets the applicable harm test in context. Do not infer an

indiscriminate monitoring obligation from the ability to issue a further notice. Publish authenticated notice formats and urgent escalation channels.

10. Provide a coherent transition and accountable enforcement

New Schedule 1 obligations requiring providers to establish compliance systems should commence no earlier than 12 months after Royal Assent or, if later, 12 months after the final instruments, exemptions and guidance necessary for compliance are published. Urgent victim-protection and enforcement provisions that do not depend on those new systems may commence earlier. Where instruments are subsequently changed or delayed, affected providers should receive a corresponding adjustment period.

Publish a transition map for the repeal of Basic Online Safety Expectations and online content scheme codes and standards. Recognise effective existing controls and evidence, explain what must change, and avoid gaps in protection. Establish a joint regulatory work programme covering online safety, privacy, cybersecurity and relevant AI requirements, with consistent requests for evidence where lawful.

Resource an SME assistance service and practical assessment tools. Consult providers, child protection organisations, researchers and affected communities on guidance. Publish an enforcement approach that distinguishes serious or persistent failures from remediable documentation defects. Preserve decisive action for urgent harm and deliberate non-compliance, alongside proportionate remediation for businesses acting in good faith.

Retain review rights for remedial directions and transparency notices. Require consultation before publishing reports to identify protected information, without allowing providers to suppress legitimate findings. Build on section 192G by expressly protecting security-sensitive details that could enable abuse.

Evaluate the regime using exposure to serious harm, repeat victimisation, effectiveness of complaints and the quality of mitigation. Include erroneous restrictions, privacy effects and compliance costs. Removal counts alone cannot demonstrate that Australians are safer. Publish a baseline and require an independent review after three years of substantive operation, with recommendations tabled in Parliament.