

A POLICY REPORT

BY NEXT GEN TASKFORCE



FILLING IN THE GAPS

DIGITAL DUTY OF CARE

IN AUSTRALIA

FOREWORD



Australia is defining the global landscape on digital governance.

For a middle power, this is no small feat: to chart an approach that balances innovation and safety, and demonstrates that technology policy need not be a choice between competing extremes.

However, sustaining this leadership is measured not only by the laws we pass, but by the talent capable of carrying this legacy forward. Today, Australia has savvy experts and regulators shaping legislative agendas at the top, while enthusiastic young advocates build momentum at the grassroots. Yet, we are missing a middle layer that nurtures the transition from ‘emerging’ to ‘established’.

The Next Gen Taskforce, convened under the Australian Internet Governance Forum, exists to bridge this gap. Through direct engagement in live policy development, the Taskforce gives multidisciplinary, early-career professionals a meaningful pathway into the practice of digital governance, building the leadership pipeline for the field.

Our inaugural report reflects that ambition: optimistic about what is possible, but practical about what it takes to get there.

TABLE OF CONTENT

EXECUTIVE SUMMARY

5

BACKGROUND

6

EXAMINING THE PROPOSED LEGISLATION

14

RECOMMEND -ATIONS

19

SUCCESS & MOVING FORWARD

26

CONCLUSION

30

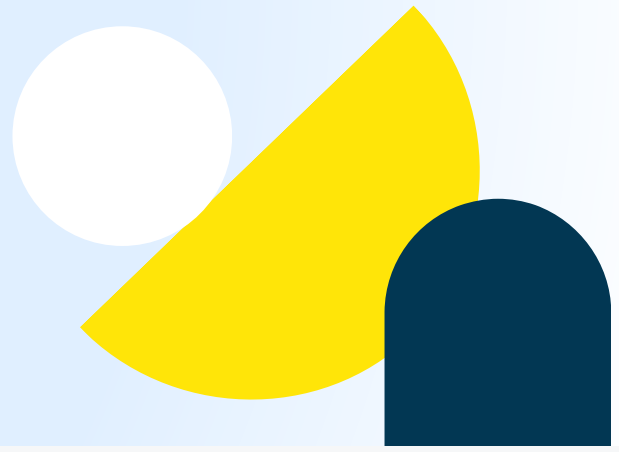
ACKNOWLEDGE -MENT

31

REFERENCES

32

1. EXECUTIVE SUMMARY



The Australian Government has committed to introducing a **Digital Duty of Care** (the "DDOC") to Parliament in late 2026, which represents a material development in the regulation of digital technology. The premise underlying is that responsibility for managing foreseeable risk should rest with those who design and operate digital products, not with the individuals and organisations left to bear its consequences.

This report is intended to assist digital service providers to prepare for that obligation in advance of its commencement.

The Taskforce's position is that digital harm is not an inevitable consequence of technology, but rather from design choices made by service providers. The case studies examined in this report demonstrate that this is a measurable and present concern, while also showcase that safety and innovation are not opposed. Where safety is treated as a design requirement rather than an afterthought, it produces better products and greater trust, which benefit both the public and the providers.

On balance, the Taskforce considers the proposed DDOC to be an appropriate and necessary intervention. Its lean, principles-based structure avoids the resourcing burden and loophole creation; its technology-neutral drafting future-proofs it against a rapidly evolving digital ecosystem; and its proportionate, risk-tiered obligations place the greatest burden where the greatest risk lies. These are strengths the Taskforce urges the Government to preserve as the legislation progresses.

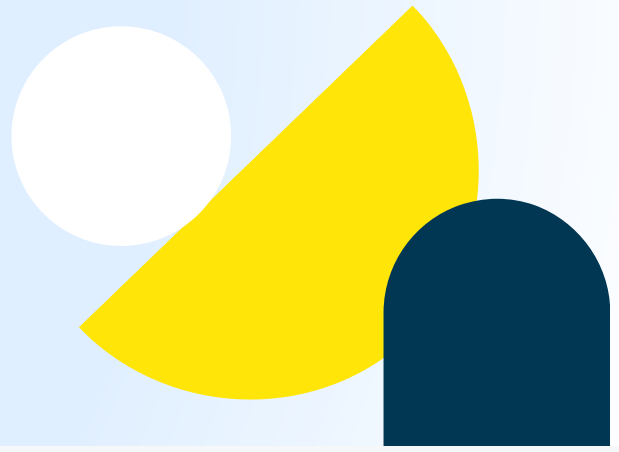
However, the report also identifies limitations in the current drafting of the DDOC. These include the prospect of prolonged legal contestation by well-resourced service providers, the difficulty of distinguishing peripheral design choices from those inherent to a platform's business model, the risk that onerous compliance requirements prompt withdrawal from the Australian market rather than reform, and the limits of Australia's regulatory leverage as a comparatively small market. None of these limitations, in the Taskforce's assessment, undermines the soundness of the legislation. Each requires deliberate mitigation, by Government and industry, if the DDOC is to achieve its intended effect rather than become a compliance exercise diminished by litigation or superficial adherence.



It is on this point that the report makes its central submission. The DDOC has been deliberately drafted to confer considerable flexibility on service providers, which is conditional rather than guaranteed. Where providers apply the duty defensively or performatively, the eSafety Commissioner retains the power to impose mandatory codes that would remove that flexibility and prescribe compliance in more detailed and less forgiving terms. The recommendations set out in this report address risk assessment, design-stage remediation, the treatment of children as likely users, resourcing of Trust and Safety functions, transparency of harm-mitigation outcomes, accessible user controls, market withdrawal, and stakeholder consultation. Rather than framing them as a compliance checklist, these are the terms on which service providers may retain the agency the DDOC currently affords them.

The Taskforce submits that the choice facing service providers is not between compliance and non-compliance. It is between shaping the DDOC's future through substantive engagement now, or having its scope and prescriptiveness determined for them later. Australia has repeatedly demonstrated its capacity to set the direction of global technology policy. Providers that treat the DDOC as an opportunity to lead on safety-by-design will be better placed, in Australia and internationally, than those that treat it as a minimum standard to be met. This report is intended to guide that process.

2. BACKGROUND



2.1 Overview of Digital Technology in Australian Life



Australians spend an average of almost 43 hours online each week across devices

According to Deloitte's Media & Entertainment Consumer Insights Report (2025).

Based on an average full-time work week, this means we spend more time engaging with the digital world than we do at work.

This time encompasses far more than entertainment. Digital services are now integral to how we connect, learn, work, access services and participate in society. For many Australians, there is no meaningful distinction between their “online” and “offline” lives.

This makes online safety fundamentally different from regulating a single product or activity. Digital technology is not homogenous, and therefore creates different risk profiles which can evolve over time.

Nor is digital technology a neutral infrastructure.

As social media age restrictions, pioneered by the Australian Government, continue to gain momentum in other jurisdictions, comparisons with age restrictions on alcohol, tobacco, and other harmful substances have become common. This analogy has some merit, but misses a crucial distinction: while the harmful properties of these substances are inherent to their chemical composition, many digital harms are not inherent to social media itself.

This logic is applicable to other current and emerging digital technologies, too. The design and deployment of digital products, interfaces, and engagement systems determine what people experience, how they behave and the unique risks that they might encounter.

This is the premise behind the Safety by Design principles by the eSafety Commissioner and the emerging Digital Duty of Care. The burden of mitigating risks shouldn't rest solely on individuals or organisations navigating technology, but on the service providers who design, operate, and profit from it.

2.2 Digital Experiences and the Role of Design

The following case studies of young Australians further illustrate the tension outlined in the previous section.



CASE STUDY 1

Flynn was a university student from NSW trying to keep up with his film degree while navigating the isolation of COVID-19 (Hartley & Smith, 2026) .

Unsure of where his life was heading, he spent most of his time doom-scrolling on social media. It was during this period that manosphere content began appearing in his feed, gradually becoming part of his daily media diet. His frustration and insecurity began to escalate into something more aggressive.

“ This is a story about vulnerability, and how it can be exploited by algorithms driven by engagement. ”

eSafety’s latest research (2026) on recommender systems finds that boys and young men can be exposed to manosphere content even when they have not actively sought it out. Seemingly benign content around self-improvement, fitness and masculinity can act as gateways to increasingly extreme material.

A research on the neurobiological impact of prolonged social media use has found that content exploiting insecurity, resentment and the desire for belonging and validation can generate longer engagement, creating incentives for algorithmic systems to amplify it (De et al, 2025).

The scale of this exposure is significant. According to another study by eSafety (2026), almost half of Australian children aged 10–17 have encountered offensive, sexist or hurtful content about girls or women online, while more than one in ten have encountered violent sexual imagery.

The potential consequences are material. A University of Melbourne study of 1,122

young Australians aged 13–17 found that young people who engaged with manosphere content were more likely to hold misogynistic attitudes (Meger, 2026). Among the factors examined, these hostile attitudes towards women were, in turn, the strongest predictor of support for violent extremism.

For Flynn, a turning point came after a video he posted echoing manosphere content went viral, triggering severe public backlash and a mental-health episode. He stepped away from social media, and that distance, alongside professional care and support from his parents and friends, helped him move away from the manosphere.

Flynn was fortunate to have those support systems. But what happens to young men and boys who do not?

This is the gap that Australian founder Hunter Johnson addresses with non-profit The Man Cave. Originally established through in-person workshops, the organisation launched a digital aftercare program on Twitch during COVID-19, where facilitators stream three nights a week and young men are able to ask questions anonymously and engage with one another in a moderated environment.

Based on an Annual Report in 2023, the program has recorded:

51,000+ unique *viewers*
2,500+ boys reusing its *digital check-in* tool
1,000+ *positive* peer-support *comments*

Although the Man Cave does not address the structural incentives driving recommendation algorithms, it demonstrates that social platforms can also be deliberately designed and used to create connection, belonging and healthier models of masculinity.

CASE STUDY 2

Emma is a disability support worker living in a rural Australian town.

With only two psychiatrists available in her community and the mentally demanding nature of her work, she began turning to ChatGPT for support. For Emma, it fills the gaps between Government-subsidised professional appointments without placing the same burden on her friends and family (McLennan, 2025).

“ *This is a story about access, and the risks young Australians are willing to take when help is difficult to reach.* ”

eSafety’s 2026 research surveyed 1,950 Australian children aged 10–17 and found that 78% had used an AI assistant or companion. More specifically:

- 54%** had used one for **personal or social** reasons
- 20%** of those were seeking advice about their **mental health or wellbeing**
- 20%** reported using these tools **daily** or more often
- 32%** had shared personal or potentially **sensitive information** with them

Yet, in a separate investigations, eSafety have identified concrete safety failures among AI companion providers (2026). From providing inaccurate or unregulated advice, to encouraging prolonged interactions that may contribute to dependency, and lacking adequate safeguards around self-harm, sexually explicit content and child sexual exploitation material. The handling of sensitive personal information also raises significant privacy concerns.

But can we preserve what makes AI appealing for mental-health support, such as its immediacy, affordability and flexibility, while embedding it within the safeguards of professional care? Australian innovation is already demonstrating that it is possible.

With two decades of experience as a clinical psychologist, Sally-Anne McCormack founded ANTSA on a model where AI support remains optional, is assigned and

governed by a clinician, and can be reviewed or switched off by them. Its AI tool supports reflection and psychoeducation between sessions, while explicitly not providing independent therapy, diagnosis or crisis care (ANTSA, n.d).

The same principles are being applied at the institutional level. Leading mental health organisation Orygen, in partnership with the University of Melbourne and RMIT, has secured \$5.7 million from Wellcome to develop and evaluate AI-powered youth mental-health care. The two-year project will co-design and test its AI system with young people and clinicians, with a specific focus on personalisation, agency, transparency, safety and collaboration (Orygen, 2026).

Together, these examples demonstrate that when designed with the appropriate safeguards, including clinical expertise, human oversight and user co-design, AI can be a game-changer for healthcare accessibility.

CASE STUDY 3

Amy was walking her dog on Bondi Beach when she was approached by a stranger asking her to go on a date with him (Al-Khouri, 2025).

Little did she know, their interaction was being recorded through Meta smart glasses, with the footage later published to his 1.3 million Instagram followers. Amy became the target of degrading comments about her body and appearance. Her attempts to have the content removed through Instagram's Community Guidelines reporting mechanism were futile.

“

This is a story about consent, and the dilemma when the act of recording is almost invisible.

”

A research found that 17% of Australian smart-glasses owners surveyed in 2024 reported using their devices in prohibited ways, including recording others without consent (Kaviani, 2024).

Smart glasses enable what a new research by University of Sydney described as “ambient capture”. Meaning recording that can happen hands-free and continuously

as the wearer moves through everyday spaces (Stilinovic et al., 2026). This creates a form of peer surveillance in which people can be captured, identified and turned into content without their knowledge or meaningful opportunity to object.

The research also analysed 350 publicly available Instagram videos recorded using smart glasses between 2023 and 2026.

60%

of the interactions analysed were classified as potential **harassment**

43%

exposed women to further **abusive** commentary, including instances where **personal details were revealed**

Yet smart glasses can serve a very different purpose. Through Syntropic Labs, Sydney-based founder Spiro is developing an ambient intelligence home system to help people experiencing cognitive impairment, including dementia and Alzheimer's disease, remain independent for longer. Its forthcoming wearable glasses extend this intelligence beyond the home.

His approach treats autonomy as a core product requirement: raw video remains local wherever possible rather than being uploaded to the cloud; users can access, control and correct information held about them, and retain their data if they cease using the service; and the system is designed to distinguish between circumstances requiring intervention and those in which user agency should be preserved.

Unlike established digital services, consumers are still forming expectations about how smart and wearable devices should operate. For founders like Spiro, one high-profile failure can disproportionately shape public perceptions and create a barrier to adoption. Demonstrable safety therefore matters both for protecting users and for building the trust that emerging technologies need to succeed.

Key Takeaway

The three case studies share a common pattern: design determines whether technology amplifies unique risks or expands connection, access and independence.

They also demonstrate that, alongside Australia's established global leadership in online safety, we have a vibrant innovation ecosystem where safety is embedded into innovation itself. Australia therefore does not have to choose between safety and innovation; framing them as a zero-sum trade-off is a fallacy.

A Digital Duty of Care is an expansion of this logic.

By enshrining obligations for service providers to proactively manage risks and make safety an essential part of digital products and services, the proposed legislation seeks to maximise the benefits while minimising the harms of technology.

2.3 The Australian Online Safety Regulatory Landscape

Australia already has a relatively mature online safety framework, with recent reforms strengthening the role of service providers in proactively preventing and responding to digital harms. Understanding the proposed Digital Duty of Care therefore requires first establishing the scope of the existing framework, how responsibility is allocated and the mechanisms available for enforcement.



Online Safety Act 2021

This is Australia's principal online safety legislation, giving the eSafety Commissioner powers to regulate harmful online content and conduct across a range of services. Its framework includes industry codes and standards covering illegal and age-restricted material, Basic Online Safety Expectations for designated service providers, and specific schemes addressing matters including cyber abuse, image-based abuse and the removal of seriously harmful content.

The Act places obligations primarily on industry, while eSafety monitors compliance and can obtain information, issue notices and directions, accept enforceable undertakings, issue infringement notices, and seek court-ordered injunctions and civil penalties. For breaches of relevant industry codes and standards, penalties can reach \$49.5 million for corporations. The framework also operates alongside other regulators and law enforcement agencies, making clear delineation of responsibilities important where online harms intersect with broader regulatory or criminal matters.



Social Media Minimum Age Restriction Act 2024

The policy adds a targeted access restriction rather than a general safety obligation. Designated age-restricted social media platforms must take reasonable steps to prevent Australians under 16 from creating or maintaining

Its implementation provides an early test of the effectiveness of platform-level age assurance, including the extent to which users can circumvent restrictions through false credentials, alternative accounts or other means. More

accounts; the obligation falls on platforms, not young people or their parents. eSafety determines which services fall within scope and monitors compliance, including through information-gathering notices and assessment of platforms' age-assurance systems. Platforms that fail to take reasonable steps can face court-ordered civil penalties of up to \$54.6 million.

specifically, six months into implementation, eSafety reported that significant numbers of under-16s remained on social media and had issued 23 information-gathering notices to 10 platforms, with investigations moving towards enforcement where evidence supports non-compliance.



Remaining regulatory landscape

Australia has therefore established a substantial regulatory foundation for online safety. Still, the landscape remains largely organised around particular harms, services and regulatory mechanisms.

They do not provide a single, overarching obligation applying consistently across the expanding range of digital products and services through which Australians now interact, including AI, wearable and connected devices, and emerging immersive technologies. Nor do they require service providers to identify and mitigate risks embedded in the design and operation of digital products and services before harm occurs.

The proposed Digital Duty of Care therefore represents an evolution of the existing framework: a broader obligation requiring service providers to proactively identify and mitigate reasonably foreseeable risks arising from the products and services they design and operate.



3. EXAMINING THE PROPOSED LEGISLATION



3.1 What will the Australian Digital Duty of Care include?



Please note that this section, and to a large extent this report, is based mostly on the Issue Paper published by the Department of Infrastructure, Transport, Regional Development, Communications, Sport and the Arts (or 'the Department' for short) in May 2026, and **not the latest Exposed Draft published in September 2026.*

DDOC was developed from recommendations from the **Joint Select Committee on Social Media and Australian Society** and the **Statutory Review of the Online Safety Act**, which included over 2,270 submissions and 72 meetings (including eight roundtables) with stakeholders.

The full list of amendments can be found on the Department's website. These recommendations broadly cover the following areas:

- The adoption of an overarching duty of care, requiring platforms to act with due diligence and employ safety-by-design principles
- Introduction of mandatory risk assessments (routine/annual and when significant changes are made)
- Increased transparency, reporting, record-keeping, and data-sharing requirements
- Stronger requirements around platform-based complaint processes and responsiveness to regulator takedown requests
- Increased maximum civil penalties for non-compliance with the Act
- Provision for the eSafety Commissioner to create mandatory codes further defining what compliance with the Act looks like

3.2 Strengths

The above inclusions together form a comprehensive, well-reasoned legislative approach to the concept of a Digital Duty of Care. The Taskforce observes several particularly promising elements of these recommendations, and anticipates significant positive change will be driven by this legislation.



3.2.1

Lean & principles-based model

Rather than prescribing an exhaustive list of technical requirements, the Duty asks providers to take reasonable, risk-based and proportionate steps to identify and mitigate foreseeable serious harms.

Case Study: UK Online Safety Act

The Taskforce notes the particular strength of the lean approach compared to the UK's recent Online Safety Act regime, which relies substantially on detailed Codes of Practice specifying measures that services can take to comply. These extensive Codes required massive resourcing from the UK Government to develop, and have created safe havens which reduce the overall effectiveness of the legislation.

The Australian Government's approach necessarily places much of the resource burden on service providers to interpret and apply the legislation. The Taskforce considers this an appropriate assignment of responsibility given that service providers derive substantial commercial benefit from their users, including through design choices which may increase user exposure to online harms.



3.2.2

Future-proof structure

The Duty has been designed to be intentionally broad, avoiding language that is overly specific to the current technological environment and instead focusing on universal principles that apply across the tech stack. This ensures that emerging technologies will be captured by the Duty without requiring legislators to continually play catch-up through technology-specific amendments.

The broad, whole-of-ecosystem approach recognises that every online service provider has a role to play. It places responsibility on service providers at all levels of the tech stack, pre-emptively addressing attempts to shift responsibility or claim exclusion from obligations to address online harms.



3.2.3

Proportionate obligations and oversight

The framework is designed to place the greatest burden on services that present the highest risk, seeking to maximise the effectiveness of the legislation without stifling innovation among low-risk and emerging services.

Similarly, graduated enforcement mechanisms allow eSafety to respond to breaches in a proportionate manner, depending on severity and recidivism. This approach ensures minor breaches can be rectified without undue hardship such as large fines or product withdrawal for a first offence.

3.3 Challenges

While the framework provides an overall strong foundation, the Taskforce notes some lingering concerns with both the legislative design and challenges with implementation.



3.3.1

Prolonged legal contestation

A key question for government and eSafety is how the Duty will withstand sustained legal challenge. Big Tech have significant financial and legal resources to contest regulatory decisions, potentially producing years of litigation over the interpretation and application of the Duty.

Taskforce concern

Over time, the framework could become shaped as much by litigation strategy and commercial interests as by its original public-interest purpose. We should consider how to preserve the intent and effectiveness of the Duty as case law develops, and how to avoid the creation of legal loopholes and safe havens.



3.3.2

Challenge to fundamental design philosophies

Many of the harms users face online stem from the fundamental design and business models of certain online services, raising questions about whether these harms can truly be minimised or are inherent to the services themselves.

The DDOC assumes that many harmful design choices are peripheral and can be removed or altered to comply with its requirements. However, it is unclear whether there are pathways for services that would be fundamentally different if required to remove harmful algorithms, features or monetisation models. The Taskforce notes

the difficulty of disentangling harmful design choices from necessary and inherent ones and is concerned that this may result in services whose very design causes harm not being adequately addressed by the Duty.

Taskforce concern

Given how deeply entrenched some risks are to the operating model of these companies, anything short of fundamental reform may not be sufficient to prevent harm to Australians online. However, global online services are unlikely to redesign their approach from the ground up to comply with Australian legislation, claiming undue hardship, launching legal challenges, or withdrawing from the Australian market entirely.



3.3.3 Potential disadvantages to Australian consumers

If implementation is too costly or inconvenient for service providers, they may alternatively opt to withdraw particular features or the entirety of their services from the Australian market. This could place Australia at a disadvantage globally, with restricted access to global services such as AI tools and communication platforms. Additionally, it will not reduce the harm experienced by the many Australians who choose to circumvent geoblocking or similar restrictions - users avoiding the Social Media Minimum Age Restrictions have demonstrated a high degree of willingness to do so.

Taskforce concern

Australia's access to emerging and global standard technology could be reduced through product/service withdrawal, whether initiated by Government restrictions or by tech companies who do not wish to comply with the legislation. Additionally, strict compliance requirements leading to withdrawal may undermine the legislation's effectiveness if the 'unsafe' versions remain available globally and continue to be accessed by Australians



3.3.4 International alignment

Australia is a small market that runs the risk of demanding more with this legislation than its share of the user base justifies in the eyes of service providers, particularly if it differs significantly from other countries. Big Tech companies operate globally and can be difficult for any one jurisdiction to regulate effectively.

While the framing of the DDOC does not appear to be particularly more onerous than international equivalents - and indeed is intended to be less prescriptive than many others - the provision for the eSafety Commissioner to create further codes indicates the legislation could be taken in a more specific direction in future. In addition, many countries do not have an equivalent to the DDOC or similar obligations, limiting the potential for a coordinated international approach.

Taskforce concern

Without appropriate consideration of existing and emerging international approaches, Australia risks diluting its influence over global design norms. Coordinated international pressure is required to strengthen Australia's influence and encourage service providers to adopt safer practices globally.

4. RECOMMENDATIONS



4.1 Service provider agency in applying the legislation

The Digital Duty of Care is designed to put the onus on digital service providers to determine how the duty should be applied to their own services. The recommended amendments specifically avoid creating safe harbours, and encourage service providers to exceed the minimum requirements set out in the Act by taking additional measures to protect their users.

Digital service providers have been awarded considerable flexibility under the duty. The form of their risk assessments, for example, is not prescribed and still under consideration; nor is the definition of a ‘significant change’. If they genuinely and willingly comply with the intention of the Act, they are likely to retain this flexibility, minimising the cost and effort of compliance.

It should be noted, however, that the recommendations would also give the eSafety Commissioner the power to introduce mandatory Codes that tighten the interpretation of the Act; and thus the definition of compliance. Service providers should recognise this as a warning - if they don’t step up and participate fully in creating a safe-by-design environment for their users, the rules may become more prescriptive and reduce their agency in deciding how best to apply them. This could expose them to a greater risk of noncompliance penalties, and may require them to create entirely new processes from the ground up in order to meet a new, stricter definition of compliance. It is therefore in the service providers’ best interest to engage sincerely with the duty of care framework, ensuring their participation is substantive rather than purely performative.

Key Takeaway

If digital service providers don’t take a proactive, holistic approach to their new duty of care responsibilities, they risk facing increased regulatory burden in future.

4.2 Recommendations to digital service providers on the interpretation and application of DDOC

When determining how to meet the requirements under the Digital Duty of Care, we recognise that service providers may seek to apply the interpretation that aligns best with their existing activities and reduces the compliance cost. However, this approach fails to engage meaningfully with the principle underpinning the duty - the need to truly improve online safety for all Australians. In order to prove their status as trusted organisations who are deserving of the agency granted to them under the duty, service providers must uplift their processes and design philosophies entirely, including those processes which aren't specifically named by the Act.

The Taskforce has developed a series of recommendations, outlining key areas which need to be addressed in order to align with the duty of care philosophy. We suggest service providers adopt these recommendations as part of their broader strategy to meet the requirements set out under the Act. These recommendations focus on ensuring not only compliance with the Act, but real positive changes to make the internet safer for end users.



Recommendation 1

Conduct risk assessments early, and in every design process - not as a final checkbox before release

Under the new requirements, entities that carry high risk or reach will be required to complete risk assessments when making significant changes to their service, e.g. the addition of new products or features. We call on service providers to structure their internal processes so that risk assessments are conducted early in the design phase, enabling high-risk design choices to be identified before they become embedded. Identifying risks earlier reduces the complexity and cost of redesigning products or features to mitigate potential harms.

As 'significant change' is not defined under the Act, the Taskforce recommends that service providers apply a broad interpretation of the term. Any change beyond minor bug fixes should be subject to an appropriate level of risk assessment, regardless of how inconsequential it may initially appear. The risk assessment process does not need to be identical for all changes: a major new version should clearly undergo more extensive review than a minor feature addition. However, even small changes can introduce subtle risks that should be appropriately considered before release.



Recommendation 2

Commit to making meaningful design changes to address identified high risks before release

As more risk assessments are conducted, the number of risks identified is likely to increase significantly. Companies must therefore ensure they have processes in place to act on the findings of those assessments. It is not enough to identify and report on a risk; service providers should make a genuine commitment to reducing risks as a priority, even where this requires significant or fundamental redesign of a new service or feature.

Being classified as ‘high risk’ under the DDOC framework significantly increases the reporting burden for service providers. Knowing a product poses significant risks, and choosing to release it anyway, is in itself a risky decision for providers under this new framework. The Taskforce therefore recommends incorporating risk-mitigation redesign into the standard design process for any new feature or product. This should not be reserved for only the highest-risk cases; rather, providers should expect risk assessments to routinely inform design decisions, with appropriate risk-mitigating changes made as a normal part of the development process.



Recommendation 3

Assume all digital services will be accessed by youth, including under-16s, and design accordingly

The DDOC recommends making child safety a ‘primary consideration’ under the framework. The Taskforce seeks to remind providers that having implemented age assurance does not absolve them of this requirement. While Australia has legislated minimum age restrictions for some large online service providers (mainly social media platforms), the Taskforce calls on all service providers to acknowledge that children still make up a portion of their user base and account for their safety when designing their service.

eSafety’s March 2026 Compliance report indicates that 70% of children who were using online platforms prior to the legislation have continued to do so. Online service providers therefore must not rely on the faulty assumption that children are actually restricted from accessing their service. The design and risk assessment process should lend specific consideration to child users, and how risks to children can be minimised.



Recommendation 4

Build or expand a dedicated Trust & Safety function to handle all DDOC-related requirements

While many large companies already have a Trust & Safety team (or equivalent), it is crucial to understand that the demand on these teams will increase significantly once the legislation comes into effect. To ensure that all DDOC-related requirements are met, including those surrounding takedown requests and user-submitted reports, companies should be actively expanding their capabilities in this area now.

Beyond the minimum requirements of the legislation, the Taskforce also calls on service providers to ensure that their Trust & Safety teams are designed thoughtfully and inclusively, and empowered to take action swiftly. Teams should consider the accessibility of their safety mechanisms, across a variety of stakeholders such as culturally and linguistically diverse people, disabled people, youth, and the elderly. Trust & Safety teams should also be utilised to identify recurring issues and drive systemic change. As these teams often sit at the frontline of user harm, providers should draw on their experience to identify recurring harm vectors and inform meaningful improvements across their services.



Recommendation 5

Increase transparency of harm mitigation measures for users

The Digital Duty of Care seeks to uplift not only user safety, but the users' ability to trust in the harm mitigation and abuse reporting mechanisms that service providers have put in place. The Taskforce suggests that digital service providers consider ways in which they can provide a greater deal of transparency into their processes, to reassure users that harm mitigation is something the provider is genuinely concerned with.

The Australian Government's issues paper calls upon services with greatest risk or reach to publish annual transparency reports. While this is only a strict requirement for a small subset of services, it is something all service providers should seek to incorporate into their standard reporting going forward. Those who initially fall outside of the 'greatest risk or reach' categorisation should not take it for granted that they will continue to do so as the legislative approach evolves.

Developing the internal processes needed to publish these reports now will prevent potential compliance issues in future, should a provider find their categorisation has changed.

Achieving genuine public trust in these systems requires greater transparency, an unspoken factor of the DDOC's requirement to maintain an accessible and user-friendly complaints processes.

Service providers should consider how they can improve transparency more broadly, for example by explaining why a user-submitted report was actioned or declined; setting up an accessible internal review process for declined reports; publishing additional aggregate metrics on a quarterly basis to support their annual reports; and releasing information regarding the proactive measures they are taking to reduce harmful content on a systemic level.

Recommendation 6



Increase meaningful, accessible, and transparent user control over individual features of a service, such as algorithmic recommendation systems

6.1 Design features to support user agency and customisation

Users should be given the ability to customise their experience online to the greatest extent practicable, including opting out of features that aren't a minimum functional requirement of the service they are using. Features that carry greater inherent risk, in particular, should be possible to disable or substitute without rendering the service nonfunctional for users. Algorithmic recommendation systems, for example, can be replaced with chronological or tag-based feeds, allowing users to continue engaging with a service without being exposed to the risks associated with those algorithms. Service providers should ensure that opt-outs and/or alternatives are provided for, at minimum, features that were assessed as higher risk.

The ability to functionally opt-out of features that carry risk (or, alternatively, to have to specifically opt-in to the risky features) demonstrates that a service provider has considered the best interests of the user, is doing their due diligence to reduce harm, and is acting in good faith to create a safer online experience as required by the DDOC.

6.2 Increase transparency and accessibility of user controls

Service providers should be measuring the success of their user controls as a harm reduction method based not only on their existence, but also on how many users actually choose to use them. User controls should be easily accessible and clearly communicated to users at the relevant stage of their interaction with the service (e.g. at account creation or when accessing a particular feature). The Taskforce suggests service providers implement a specific settings menu for user controls, which is easily accessible from the base settings menu, clearly and descriptively titled, and includes guidance for users as to what settings are recommended to reduce risk.

The majority of users should not be unaware that these options are available, and users should not need to be tech-savvy in order to locate and select the appropriate settings. In order to align with the Australian Government's vision for a safer online experience, service providers must ensure that their harm-reduction focused user controls are widely understood and utilised.



Recommendation 7

Avoid product withdrawal when modification or redesign is a feasible alternative

The Australian market may be small in the global context, but it is a market abundant with highly engaged, creative, and skilled internet users. Service providers extract a high degree of value from Australian users beyond the simple financial benefits - Australians create content and contribute to communities online in a way that improves the quality of the services they use. While compliance may seem like a higher cost than simply withdrawing from the Australian market, the Taskforce implores providers to consider the unique value their Australian users provide when making this assessment.

Australia has often been at the forefront of technology policy, with measures such as the Social Media Age Restrictions influencing similar approaches in the UK and other jurisdictions. Withdrawing a product from Australia to avoid new regulatory requirements may therefore prove short-sighted if similar legislation is adopted internationally. Service providers should view the Australian approach as an indication of the direction of technology policy and invest in safer product design now, rather than waiting for comparable requirements to emerge in larger markets.



Recommendation 8



Consult widely with diverse stakeholder groups when developing internal risk-assessment frameworks and safety-by-design guidelines

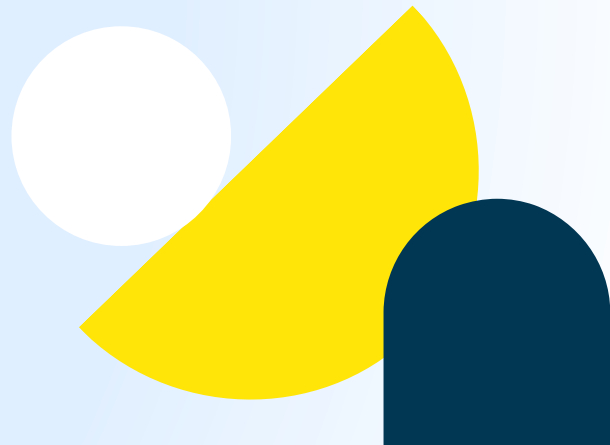
Safety-by-design principles should be informed by the lived experience and expertise of diverse stakeholder groups, including women, LGBTQIA+ people, people with disability, culturally and linguistically diverse communities, and children. Without input from those disproportionately exposed to particular digital harms, providers risk creating frameworks with blind spots that leave certain groups inadequately protected.

To develop effective, evidence-informed frameworks, providers should meaningfully consult with these and other relevant stakeholders throughout the design process.

Consultation should not be performative; stakeholder input should be treated as part of the evidence base informing the design and implementation of safety frameworks.

The Taskforce also encourages providers to consider public consultation alongside targeted stakeholder engagement, allowing broader feedback on proposed frameworks and guidelines. As these frameworks underpin each provider's approach to online safety, getting them right will be critical to both effective harm reduction and demonstrating compliance with the DDOC.

5. SUCCESS & MOVING FORWARD



5.1. Success

Instead of assessing the volume of compliance activity, a successful Digital Duty of Care should produce measurable improvements in how digital products are designed, how risks are managed, and how Australians experience technology.



5.1.1

Safe Products and Services by Design

<i>Risk assessment across the product lifecycle</i>	<i>Proportionate safety interventions</i>
Risk assessment should occur early enough to influence design decisions and continue throughout the product lifecycle. Measure this through the proportion of significant product or feature changes assessed before deployment, the proportion of high-risk issues mitigated before release, and the frequency of previously unidentified risks emerging post-deployment.	Identified risks should translate into substantive design changes that reduce their likelihood or severity. Depending on the risk, safeguards may include warnings, user friction, delays, filtering or functional restrictions. Success should be measured through changes in exposure, frequency and severity of identified harms following intervention.



5.1.2

Greater Autonomy for People Using Technology

<i>Control by design</i>	<i>Effective Recourse</i>
Users should be able to understand and control how they experience technology, including adjusting or	Autonomy also requires meaningful recourse when harm occurs. Providers should report resolution times, the

disabling features where appropriate to their individual circumstances. Success should be measured through the accessibility, comprehension and effective use of these controls, as well as users' ability to access, correct and control personal information.

proportion of complaints resulting in action and whether users receive clear explanations of decisions. The process should be sufficiently accessible that users are not required to navigate complex technical, legal or bureaucratic systems to exercise their rights.



5.1.3

Increased Trust through Observability

<i>Transparency that enables scrutiny</i>	<i>Accountability that has consequences</i>	<i>Co-design and continuous improvement</i>
Trust cannot be established through assurances of safety alone. Providers should publish comparable information on risks identified, safeguards implemented, complaints received and resolved, and the outcomes of interventions. Where decisions affect users, including content removal, providers should provide clear explanations of the relevant rule and basis for the decision.	Observability should enable regulators, researchers and the public to assess whether safety is improving over time. Where providers fail to meet their obligations, consequences should be proportionate to the breach, clearly justified, capable of meaningful impact and imposed promptly following applicable review or appeal processes.	Safety standards should be informed by meaningful participation from people and communities affected by digital harms, especially vulnerable groups that might experience unique harms in comparison to the majority of the population. Consistent reporting and access to evidence should allow those standards to evolve as technologies, risks and user experiences change.

5.2. Moving Forward: Beyond the current DDOC

Although the Taskforce recognises the currently proposed Digital Duty of Care is an important foundation for safer digital technology, legislation alone cannot resolve every challenge identified in this report. Some gaps require action beyond individual service providers or the Australian regulatory framework.

The Taskforce therefore identifies three areas for further development in parallel to the implementation and strengthening of the DDOC.



5.2.1

Strengthen Australia's International Regulatory Leverage

Australia's ability to regulate multinational technology companies is strengthened when acting alongside jurisdictions with aligned objectives.

This is particularly important where providers may otherwise determine that superficial compliance is preferable to substantive changes to product design, particularly where Australia represents only a small proportion of their global user base. Thus, a coordinated international approach could increase the commercial and regulatory incentive for providers to make meaningful changes by creating consistent expectations across markets.

The Taskforce suggests a global coalition which could support:

Regulatory alignment · Information sharing · Coordinated investigations · Common product safety expectations · Joint engagement with multinational technology companies

This should build on existing international regulatory infrastructure, including the UN Internet Governance Forum and the Global Online Safety Regulators Network. Australia is well placed to lead this work, having served as the inaugural convenor of the Global Online Safety Regulators Network, and should seek to develop the network into a stronger mechanism for coordinated regulatory action.



5.2.2

Build a Sustainable Digital Safety Ecosystem

A Digital Duty of Care should not operate solely as a compliance regime. Long-term improvements in digital safety also require investment in the research, capability and community infrastructure needed to identify emerging risks and develop effective

responses.

Australia's experience also demonstrates that safety and innovation need not be competing objectives: when safety is treated as a design requirement, it can produce commercially viable products and services that respects and even expands the wellbeing of people using technology.

The government should therefore invest in an innovation ecosystem that makes safer technology commercially and socially desirable, not simply legally necessary.

The Taskforce proposes consideration of a Digital Safety Fund to amplify:

Independent research · Public education and digital literacy · Community engagement · Responsible technology innovation · Cross-sector collaboration · Youth participation

The objective should be to shift expectations across both sides of the market: consumers should increasingly expect safety as a standard feature of technology, while developers and service providers should see safety-by-design as a source of innovation and competitive advantage.

The Fund could be supported through a combination of government investment and revenue generated through DDOC enforcement. Further consideration would be required as to whether administration should sit with eSafety, an independent grant-making body or a hybrid model.



5.2.3

Give Consumers Greater Visibility of Digital Safety

As digital products become increasingly complex, consumers should not have to rely solely on providers' own claims when assessing the safety of the technologies they use.

The Government should explore an independent, third-party assessment and labelling system that enables consumers to understand the safety risks associated with digital products, services and applications.

The Taskforce suggests a consumer-rating models such as the Energy Rating and Health Star Rating schemes, assessing factors including:

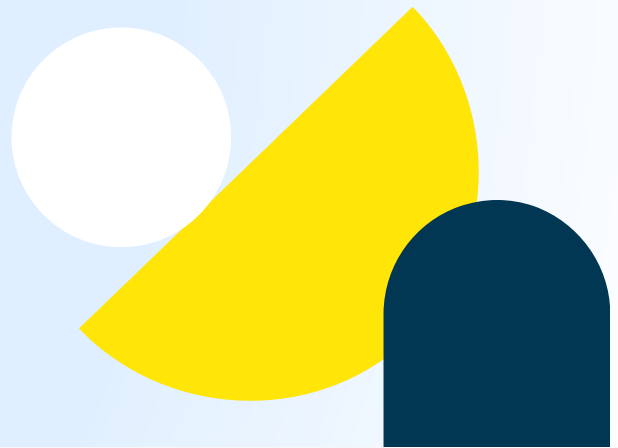
Foreseeable digital harms · Vulnerability of the target user base · Privacy and user autonomy · Effectiveness of safety-by-design measures

Civil society could play a central role in developing and administering such a system,

drawing on its independence from both government and technology providers and its proximity to communities experiencing digital harms. Independent assessment would also help ensure that ratings reflect lived experience and meaningful safety outcomes, rather than simply providers' own representations of compliance.

The framework should be developed through broad multistakeholder consultation to ensure that it is credible, proportionate and capable of adapting to emerging technologies.

6. CONCLUSION



Australia is moving from consultation to legislation, with the Government committed to introducing a Digital Duty of Care to Parliament in late 2026.

The report thus aims to assist service providers in preparing for this transition now, using this period to identify foreseeable risks, test proportionate safeguards and progressively embed safety-by-design across their products and services.

While the immediate obligation will apply to the Australian market, its significance may extend well beyond it. Australia's Social Media Minimum Age reforms have demonstrated its capacity to shape global regulatory expectations. As jurisdictions increasingly move towards proactive harm mitigation, the DDOC could become another test case for how technology providers respond to this shift.

As this report repeatedly demonstrates, safety can itself be an engine for better products, greater trust and wider adoption. The purpose of a Digital Duty of Care is therefore straightforward: to make the prevention of foreseeable harm an integral part of how digital technology is designed and delivered.

7. ACKNOWLEDGEMENTS

With thanks to all members of the NextGen Taskforce for their valued contributions to the development of this report.

Taskforce Members

- Minh Hoang, Chair
- Thomas Lenette, General Secretary
- Elizabeth Pollard, Penholder (Section 3 & 4)
- Chelsea Adams
- Konan Masuda
- Rohan Sachdeva
- Pnina Hagege
- Tritian Young-Glasson
- Alice Dolin
- Priya Virdee-Hero
- Harrison Oates

The Taskforce extends its sincere gratitude to the **auIGF 2026 Multistakeholder Steering Committee (MSSC)** for its continuing guidance and support throughout this project.

Consultations

The Taskforce would also like to extend our thanks to the following stakeholders, who generously shared their time and perspectives through consultation/Q&A sessions:



Australian Government

Department of Infrastructure, Transport,
Regional Development and Communications

TEACH US
* CONSENT

8. REFERENCES

Deloitte. (2025). *Media & entertainment consumer insights 2025*.

<https://www.deloitte.com/au/en/Industries/tmt/research/media-and-entertainment-consumer-insights.html>

Hartley, T., & Smith, E. J. (2026, August 20). The manosphere continues to suck young men in as study reveals troubling findings. *ABC News*.

<https://www.abc.net.au/news/2026-08-20/geek-culture-revealed-as-pipeline-to-manosphere/107050590>

eSafety Commissioner. (2026). *Recommender systems: Position paper*.

<https://www.esafety.gov.au/sites/default/files/2026-05/Recommender-systems-position-paper-May2026.pdf>

De, D., El Jamal, M., Aydemir, E., & Khera, A. (2025). Social media algorithms and teen addiction: Neurophysiological impact and ethical considerations. *Cureus*, 17(1), e77145. <https://doi.org/10.7759/cureus.77145>

eSafety Commissioner. (n.d.). *How common is exposure to content associated with harm among children in Australia?* <https://www.esafety.gov.au/research/the-online-experiences-of-children-in-australia/snapshot-content-associated-with-harm>

Meger, S. (2026). *Misogyny, online radicalisation, and youth extremism in Australia* [Research report]. The University of Melbourne.

<https://outcomes.org.au/resource/misogyny-online-radicalisation-and-youth-extremism-in-australia/>

The Man Cave. (2023). *Annual report 2023*. <https://themancave.life/wp-content/uploads/2024/09/FY23-annual-report.pdf>

McLennan, A. (2025, May 19). Young Australians using AI bots for therapy. *ABC News*.

<https://www.abc.net.au/news/2025-05-19/young-australians-using-ai-bots-for-therapy/105296348>

eSafety Commissioner. (2026). *Talking to machines: Children's experiences with AI assistants and companions*. <https://www.esafety.gov.au/research/talking-to-machines-childrens-experiences-with-ai-assistants-and-companions>

eSafety Commissioner. (2025, February 18). *AI chatbots and companions – risks to children and young people*. <https://www.esafety.gov.au/newsroom/blogs/ai-chatbots->

and-companions-risks-to-children-and-young-people

eSafety Commissioner. (2025). *Findings from transparency notices on AI companion apps: October 2025*. <https://www.esafety.gov.au/industry/basic-online-safety-expectations/ai-services/findings-october-2025>

ANTSA. (n.d.). *Clinical governance*. <https://antsa.ai/governance>

Orygen. (2026, May 21). *Orygen secures \$5.7 million Wellcome funding to advance safe, AI-powered youth mental health care*. <https://orygen.org.au/About/News-And-Events/2026/Orygen-secures-Wellcome-funding-to-advance-safe%2C-A>

Al-Khouri, C. (2025, March 25). US content creator secretly filmed Sydney women with covert sunglasses camera lens. *ABC News*. <https://www.abc.net.au/news/2025-03-25/stateline-us-influencer-secretly-filmed-sydney-women-instagram/104974464>

Kaviani, F., Lyall, B., & Koppel, S. (2024). Exploring social perceptions of everyday smartglass use in Australia. *PLOS ONE*, 19(11), e0313001. <https://doi.org/10.1371/journal.pone.0313001>

Stilinovic, M., Egliston, B., Tan, W., Carter, M., & Gray, J. (2026). *Harm through the “#RizzCam”: Smart glasses, ambient capture and invisible harassment* [Preprint]. OSF Preprints. https://doi.org/10.31235/osf.io/wbj2x_v2

Syntropic Labs. (n.d.). *Syntropic Labs*. <https://www.syntropiclabs.ai/>

eSafety Commissioner. (2026, March 31). *Social media minimum age: Compliance update*. <https://www.esafety.gov.au/sites/default/files/2026-03/SocialMediaMinimumAgeComplianceUpdateMarch2026.pdf>