



# The National AI Plan & Children's rights

## *Advancing children's rights through AI policy & guidelines*

Published March 2026  
Rys Farthing



## Summary

In December 2025, the Department of Industry, Science and Resources released the *National AI Plan for Australia* ('the Plan'). The Plan presents significant opportunities to advance children's rights in the digital world, where it is effectively implemented.

Like all technologies underpinning the digital world, AI systems and platforms have not been designed with children's rights in mind. Significantly, the governance of AI so far has not seen children's safety, privacy nor best interests enshrined. The consequences of both of these failures in digital platform governance perhaps provide reason to pause for thought. Failures to integrate children's rights into the initial design of popular platforms, compounded by hesitations and delays in creating effective practices to govern platforms, resulted in the proliferation of harms to the extent that the Australian Government eventually opted to remove children under 16 from platforms. This will not be an option for AI. Governments around the world — including Australia — are now at a pivotal moment when it comes to AI governance, and making sure the National AI plan works for children is critical.

This paper reflects a discussion with 16 experts in children's rights, consumer law, online safety and AI safety science in late February 2026. The discussion focussed on four commitments made in Action 7 of the Plan and their impacts on children's rights, including:

- **Consumer protections for AI-enabled goods and services.** Introducing consumer safety standards and liability for AI goods *and* services could enhance the impact of the ACL on AI products as children use them. Product safety obligations, including for AI goods and services, on digital platforms that act as market places for these products (as intermediaries) would provide stronger protection for children. Likewise, a Digital Ombuds and other avenues for complaint and redress that has a remit over digital platforms, including for the supply of AI goods and services could help advance children's rights.
- **Reducing online harms through reforms, codes and standards within the *Online Safety Act*.** Revising the Act to include a singular, overarching Digital Duty of Care that replaces the current BOSE would go some way towards ensuring meaningful accountability under the Act. The Codes under the Act would need to be replaced, and the industry classifications revised and broadened.
- **Advancing the science of AI safety.** A broad AI Safety research agenda is needed that directly considers children's rights, including their rights to participation, access and safety. This agenda needs to address issues affecting children directly, and children's contexts, such as procurement of safer AI for schools. Co-design and children's participation need to be a central part of this, as does civil society collaboration. A researcher access scheme would ensure that critical scientific data is available for transparency and independent analysis.
- **The establishment of an AI Safety Institute.** The Institute presents an opportunity to embed children's rights into its foundational work programme. Its mandate to generate technical research into AI risks and harms, and to engage with regulators, should explicitly include children's safety, privacy and participation as core priorities from the outset.

These measures would enhance accountability and transparency of AI services and products, and help secure children's rights.

## Contents

Introduction .....	1
Opportunities to advance safety within the AI Plan .....	8
Consumer law, AI & Children’s rights .....	8
Online safety law, AI & Children’s rights.....	11
AI Safety research & Children’s rights .....	15
Conclusions & Recommendations .....	18

### About the organisations and academics

Digital Rights Watch is a not-for-profit that exists to ensure fairness, freedoms and fundamental rights for all people who engage in the digital world.

The Alannah & Madeline Foundation is a national not-for-profit organisation dedicated to keeping children and young people free from violence and trauma wherever they live, learn and play.

The Australian Child Rights Taskforce is a civil society coalition of over a hundred organisations and individuals committed to the protection, promotion and fulfilment of the rights of all children and young people in Australia.

Professor Julian Sefton-Green is a professor at REDI (Research for Educational Impact), a strategic research centre in the Faculty of Arts and Humanities at Deakin University. He is also a fellow of the Australian Research Council’s Centre of Excellence for the Digital Child, which supports research, advocacy and public understanding about digital childhood for young children aged 0 to 8.

Professor Judith Bessant is a member of The Social Equity Research Centre based in the School of Global Urban and Social Studies, RMIT University, Melbourne.

Professor Rys Farthing is a Professorial Research Fellow at the News and Media Research Centre at the University of Canberra, and author of this paper. The Centre advances public understanding of the changing news and media landscape, and advocates for a media system that builds trust, inclusivity and diversity.

### Acknowledgements

We would like to thank the Minderoo Foundation and the Allanah & Madeline Foundation for their support in the development of this report.

## Introduction

### The National AI Plan and its origins

In December 2025, the Department of Industry, Science and Resources released the *National AI Plan for Australia*.<sup>1</sup> The Plan announced the Government's approach to infrastructure, innovation, skills and regulation around the emerging AI economy in Australia.

The Plan follows years of debate across the political sector and civil society about the potential for, and the need for regulation for safer AI, including, for example:

- The 2019 publication of *Australia's AI Ethics principles*.<sup>2</sup> Broadly aiming to align with the Sustainable Development Goals, these principles were developed after consultation with industry, civil society and the public. They highlight, among other things, requirements for providers to be responsible for the outcomes of the AI systems that they design, develop, deploy and operate. One of the eight principles was specifically rights-centric, noting that 'AI systems should respect human rights, diversity, and the autonomy of individuals'. This would, presumably, include children's rights.
- The August 2024 release of the *Voluntary AI Safety Standards*, and supporting guidelines for implementation, that were designed to help organisations develop and deploy AI systems in Australia safely and reliably.<sup>3</sup> A consideration of children's rights is implied in the development of risk assessments recommended by the Standard. The guiding questions prompt organisations to think about the context in which they are using AI and to think carefully about whether or not their AI system interacts with or affects 'people who have extra forms of legal protection (such as children)'.
- In June 2023, the Government opened consultation on 'Safe and Responsible AI in Australia' to identify gaps in governance and regulation, to meet the emerging scale of risks associated with the rapid uptake of generative AI. An interim response to this was released in early 2024,<sup>4</sup> including the establishment of an AI experts group, leading to a proposals paper that outlined potential mandatory guardrails for AI use in high risk settings.<sup>5</sup> The potential capacity of AI products and services to create Child Sexual Abuse Material was a key risk identified and addressed in the paper.

There were strong indicators that risk-based, safety-focussed regulations were emerging.

Internationally, there is growing momentum around AI governance frameworks that explicitly address children's rights. The *EU AI Act (2024)* classifies AI systems intended to be used in education, or that interact with children in consumer-facing products, as high-risk, triggering heightened conformity assessment obligations. The United Kingdom's ongoing work on AI

---

<sup>1</sup> Department of Industry, Science and Resources 2025 *National AI Plan for Australia* <https://www.industry.gov.au/publications/national-ai-plan>

<sup>2</sup> Department of Industry, Science and Resources 2019 *Australia's AI Ethics Principles* <https://www.industry.gov.au/publications/australias-ai-ethics-principles>

<sup>3</sup> Department of Industry, Science and Resources 2024 *Voluntary AI Safety Standards* <https://www.industry.gov.au/publications/voluntary-ai-safety-standard>, noting that these have been updated and replaced in Oct 2025

<sup>4</sup> Department of Industry, Science and Resources 2024 *Safe and responsible AI in Australia consultation Australian Government's interim response* [https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public\\_assets/safe-and-responsible-ai-in-australia-governments-interim-response.pdf](https://storage.googleapis.com/converlens-au-industry/industry/p/prj2452c8e24d7a400c72429/public_assets/safe-and-responsible-ai-in-australia-governments-interim-response.pdf)

<sup>5</sup> Department of Industry, Science and Resources 2024 *Safe and responsible AI in Australia Proposals paper for introducing mandatory guardrails for AI in high-risk settings* [https://storage.googleapis.com/converlens-au-industry/industry/p/prj2f6f02ebfe6a8190c7bdc/page/proposals\\_paper\\_for\\_introducing\\_mandatory\\_guardrails\\_for\\_ai\\_in\\_high\\_risk\\_settings.pdf](https://storage.googleapis.com/converlens-au-industry/industry/p/prj2f6f02ebfe6a8190c7bdc/page/proposals_paper_for_introducing_mandatory_guardrails_for_ai_in_high_risk_settings.pdf)

regulation similarly foregrounds the need for age-appropriate design principles to be applied to AI products. Australia has an opportunity to draw on these frameworks in the development and implementation of its own Plan.

Likewise, there were strong indicators that policy around AI adoption around economic growth was emergent. In 2021, the Government released *Australia's AI Action Plan*, which outlined a productivity focussed vision for AI adoption across the economy.<sup>6</sup> Following this, the Productivity Commission outlined the capacities of AI to drive economic growth in their 2025 Interim Report *Harnessing data and digital technology*,<sup>7</sup> and the ACCC noted the need for effective competition in the AI sector to drive innovation and growth for Australian businesses.<sup>8</sup>

The release of the National AI Plan indicated the start of a coordinated response from the Government, outlining how they intend to meet expectations of safety-focussed guardrails, as Australia drives for wide-scale adoption of AI systems to push for growth.

### The National AI Plan and children's rights

The plan does not explicitly outline how AI affects children's rights, or indeed how it may advance them. However, there are a number of propositions within the plan that may, with the right implementation, work to advance children's rights to protections:

- **Advancing the science of AI safety:** The Plan notes that *'safety research underpins the reliability and trustworthiness of AI systems. The government is engaging domestically and internationally to build expertise and understanding of the capabilities and risks of advanced AI systems, to inform when and how to respond.'* Research on how the capabilities and risks of AI systems affect children can help ensure that both innovation and regulation advance children's rights.
- **Consumer protections for AI-enabled goods and services:** The Plan notes that the recent *'Review of AI and the Australian Consumer Law'*<sup>9</sup> found that *Australians enjoy the same strong consumer protections for AI products and services as they do for traditional goods and services, including safety protections. The Government will consult with states and territories on minor opportunities to clarify existing rules that the review identified and progress the changes when appropriate.'* Children are consumers of AI products and services, and indeed in principle enjoy additional protections as consumers.<sup>10</sup> The effective application of Australian Consumer Law could enhance the protection of children using AI enabled services and products.
- **Reducing online harms through reforms, codes and standards:** The Plan notes that *'the government addresses AI-related risks through enforceable industry codes under the Online Safety Act 2021 and by criminalising non-consensual deepfake material. Further restrictions on 'nudify' apps and reforms to tackle algorithmic bias are also being considered.'* While industry codes developed under the *Online Safety Act* are typically

---

<sup>6</sup> Department of Industry, Science and Resources 2021 *Australia's AI Action Plan*  
[https://wp.oecd.ai/app/uploads/2021/12/Australia\\_AI\\_Action\\_Plan\\_2021.pdf](https://wp.oecd.ai/app/uploads/2021/12/Australia_AI_Action_Plan_2021.pdf)

<sup>7</sup> Productivity Commission 2025 *Harnessing data and digital technology* <https://www.pc.gov.au/inquiries-and-research/data-digital/>

<sup>8</sup> ACCC 2025 *Digital platform services inquiry 2020-25*

<https://www.accc.gov.au/inquiries-and-consultations/finalised-inquiries/digital-platform-services-inquiry-2020-25>

<sup>9</sup> The Treasury 2025 *Final report – Review of AI and the Australian Consumer Law*  
<https://treasury.gov.au/publication/p2025-702329>

<sup>10</sup> See for example, Reset.Tech 2023 *Can the consent model improve the digital world, especially for young people?*  
<https://apo.org.au/sites/default/files/resource-files/2023-05/apo-nid322895.pdf>

drafted by industry, and are notoriously lacking when it comes to advancing children's rights,<sup>11</sup> there may still be an opportunity to use the development of these codes, and the potential for eSafety drafted Industry Standards, to improve protections.

- **The establishment of an AI Safety Institute**, which is intended to strengthen the Government's ability to 'respond to AI-related risks and harms, and keep Australians safe'. This includes the ability to generate technical research about the risks and harms of systems and models, and to engage with regulators to share relevant findings. This would include, presumably, risks and harms to children.

In addition to these broad commitments, the Plan creates an opportunity for participation in policy development, and could help to shape their rights to meaningful access in the digital world.

In short, the implementation of the Plan has the capacity to significantly shape the way children experience the digital world, and to advance the full suite of children's rights. Understanding how the Plan and children's rights interact, and the opportunities and risks within the policy, is therefore critical.

## AI & children's rights

The digital world is increasingly a space where young people's rights can be promoted, or harmed. The Committee on the Rights of the Child, in their General Comment around children's rights and the digital world, notes that:<sup>12</sup>

*The rights of every child must be respected, protected and fulfilled in the digital environment. Innovations in digital technologies affect children's lives and their rights in ways that are wide-ranging and interdependent, even where children do not themselves access the Internet. Meaningful access to digital technologies can support children to realize the full range of their civil, political, cultural, economic and social rights.*

This includes AI, and the General Comment notes '*The digital environment is constantly evolving and expanding, encompassing information and communications technologies, including digital networks, content, services and applications, connected devices and environments, virtual and augmented reality, **artificial intelligence**, robotics, automated systems, algorithms and data analytics, biometrics and implant technology*' (emphasis added).<sup>13</sup>

A recent joint statement from the UN Committee on the Rights of the Child and the International Telecommunications Union, and others, outlines some of the challenges and opportunities the rise of AI poses for children's rights.<sup>14</sup> They note that emerging technologies, such as AI, are:

---

<sup>11</sup> See for example, Reset.Tech 2024 *Does co-regulation function in children's best interests?*  
<https://apo.org.au/sites/default/files/resource-files/2024-10/apo-nid328873.pdf>

<sup>12</sup> Committee on the Rights of the Child 2021 *General Comment 25: Children's Rights in Relation to the Digital Environment*  
<https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

<sup>13</sup> Committee on the Rights of the Child 2021 *General Comment 25: Children's Rights in Relation to the Digital Environment*  
<https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation>

<sup>14</sup> International Telecommunications Union 2025 Joint Statement on Artificial Intelligence and the Rights of the Child  
[https://www.itu.int/dms\\_pub/itu-d/opb/str/D-STR-CYB\\_JOINT-2025-PDF-E.pdf](https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-CYB_JOINT-2025-PDF-E.pdf)

*...fundamentally changing the world and affecting present and future generations of children. They have created unprecedented opportunities for children and for the realization of their rights as enshrined in the (Convention on the Rights of the Child). At the same time, AI can pose profound challenges to the realization of the rights of the child. Risks extend both to direct interactions between children and AI systems and to the ways in which AI systems impact children indirectly.*

Figure one outlines a selection of case studies that present risks and opportunities from AI for children, to highlight this duality.

The Joint Statement outlines areas of responsibility to ensure that AI advances children's rights despite not being designed with them in mind. Firstly, it recognises that the Australian Government — like all state parties to the Convention on the Rights of the Child — has obligations to take legislative, administrative and policy measures to ensure that AI is governed in ways that promote children's rights. This includes identifying key responsibilities within Government, key budget allocations and key areas for monitoring to ensure that rights are protected. The National AI Plan bears some of these hallmarks. However the plan, as it is published now, fails to specifically acknowledge children's rights or outline how children's best interests will be achieved. The implementation of the plan then, will be key to evaluating how Australia has lived up to these commitments.

Secondly, the Joint Statement notes that businesses and providers have obligations to realise children's rights when it comes to AI, and that State parties need to ensure that businesses meet these obligations.

*States should protect against child rights violations in the context of AI within their territory and/or jurisdiction by third parties, including business enterprises. States should also set out clearly the expectation that all business enterprises domiciled in their territory and/or jurisdiction respect children's rights, throughout their operations with respect to the development and deployment of AI.*

The Joint Statement clearly stipulates that action enhancing children's safety, privacy, participation and best interests are necessary, from both States and businesses, to advance children's rights.

A recurring theme in discussions of children and AI is the tendency to frame children primarily as at-risk, 'risky', or risks to be managed.<sup>15</sup> This is either as vulnerable subjects of harm, or as vectors of problematic behaviour. A rights-based approach demands a fundamentally different framing. Children are rights-holders, and the question is not whether their behaviour can be managed, but whether the systems they inhabit uphold their rights. This distinction has direct implications for regulatory design. It moves accountability from individual children and families to the platforms, providers and systems that shape their digital experiences.

The development and deployment of AI faces two familiar challenges. Firstly, like all technologies underpinning the digital world, AI systems, tools and platforms have not been designed with children's rights in mind, and secondly, the governance of AI so far has not seen children's safety, privacy nor best interests enshrined. Both of these challenges were also features of early debates about digital platform governance, and the failures of platform governance perhaps provide an important point of reflection. Failures to integrate children's rights into the initial design of popular social media platforms, combined with hesitations and delays around strong regulation and enforcement, saw harms proliferate to the extent that

---

<sup>15</sup> See for example, Adam Jordan (2017). Risky Children: Rethinking the Discourse of Delinquency and Risk. *Journal of Thought*, 51(1–2), 31–46. <https://www.jstor.org/stable/90010894>

the Australia Government opted to wholesale remove children from platforms. This will not be an option for AI, which unlike social media platforms is horizontally integrated across health, education, employment, justice and civic systems. The window for embedding children's rights into AI governance is narrow. Learning from the social media experience means acting on the architecture of these systems now, rather than retrospectively managing the harms they generate. Governments around the world — including Australia — are now at a pivotal moment when it comes to AI governance, and making sure the National AI Plan works for children is critical.

### **About this paper**

This paper reflects a discussion held at a workshop with 16 experts in children's rights, consumer law, online safety and AI safety science in February 2026. The workshop encouraged participants to think through three specific proposals in the Plan — advancing the science of AI safety, consumer protections for AI-enabled goods and services, and reducing online harms through reforms, codes and standards — framed in terms of how these actions could affect children's rights. The discussion has implications for the potential workplan of the AI Safety Institute, and how it might best address children's rights in its functioning.

This workshop was held under the Chatham House Rule, and this paper reflects the key areas of discussion and points of agreement among participants.

## Known risks and opportunities of AI for children:

### Three examples of opportunities

- Benefits to education: AI offers significant opportunities for developing personalised learning resources. For example, AI has been introduced into NSW classrooms already. In 2025, the Department for Education NSW rolled out NSWeduChat — a generative AI tool owned and designed by the NSW Department of Education — to NSW schools and students from years 5 to 12. NSWeduChat is designed ‘to respond to students with guidance and by asking open-ended questions that encourage them to explore and share their thoughts and reasoning’.<sup>16</sup> Many children around the world have reported using Generative AI tools for similar educational purposes.<sup>17</sup>
- The right to play: Children are using AI to create new ways to play in the modern world. British evidence found that 40% of children aged 8-12 were using AI to play.<sup>18</sup> In experimental settings, this has been found to be particularly beneficial for neurodiverse children.<sup>19</sup>
- Future skills development: Early AI literacy is a potential future skill for the next generation. AI is estimated to create 200,000 paid jobs by 2030,<sup>20</sup> and by 2024, 1.8% of job ads already state that the job requires AI skills.<sup>21</sup> AI literacy, and familiarity with AI systems and agents, could be a critical employability skill for young people entering the workforce.

### Three examples of risks

- Encouraging harmful behaviour: Chatbots are known to create safety and security risks for children. For example, in 2024, a 14-year-old took his own life after encouragement from a chatbot companion. His family brought legal action and settled their claim with CharacterAI.<sup>22</sup> This case is not isolated, and the risks of Chatbots for those experiencing mental health issues is well documented.<sup>23</sup> The risks from companion AI chatbots must also be understood in the context of a documented loneliness epidemic among children and young people. Where social infrastructure is lacking, commercial AI companion services may position themselves as a solution. This dynamic targets and creates particular vulnerability.

<sup>16</sup> NSW Government Education 2025 *Student use of NSWeduChat* <https://education.nsw.gov.au/schooling/parents-and-carers/artificial-intelligence/student-use-of-nsweduchat>

<sup>17</sup> UNICEF 2025 *Children's perspectives on their best interests and AI* <https://www.unicef.org/innocenti/stories/childrens-perspectives-their-best-interests-and-ai>

<sup>18</sup> The Alan Turing Institute 2025 *Understanding the Impacts of Generative AI Use on Children* [https://www.turing.ac.uk/sites/default/files/2025-05/combined\\_briefing\\_-\\_understanding\\_the\\_impacts\\_of\\_generative\\_ai\\_use\\_on\\_children.pdf](https://www.turing.ac.uk/sites/default/files/2025-05/combined_briefing_-_understanding_the_impacts_of_generative_ai_use_on_children.pdf)

<sup>19</sup> Ali Fahad Aldakahil 2024 'Investigating the impact of an AI-based play activities intervention on the quality of life of school-aged children with ADHD' *Research in Developmental Disabilities* <https://doi.org/10.1016/j.ridd.2024.104858>

<sup>20</sup> Tech Council of Australia 2024 *AI to create 200,000 jobs in Australia by 2030* <https://techcouncil.com.au/newsroom/ai-to-create-200000-jobs-in-australia-by-2030/>

<sup>21</sup> PWC 2025 *How AI is impacting Australia's jobs and workers* <https://www.pwc.com.au/services/artificial-intelligence/ai-jobs-barometer.html>

<sup>22</sup> Agence France-Presse 2025 'Google and AI startup to settle lawsuits alleging chatbots led to teen suicide' *The Guardian* <https://www.theguardian.com/technology/2026/jan/08/google-character-ai-settlement-teen-suicide>

<sup>23</sup> Chris Stokel-Walker 2025 'AI driven psychosis and suicide are on the rise, but what happens if we turn the chatbots off?' *BMJ* doi: <https://doi.org/10.1136/bmj.r2239>

- Using and creating harmful content: Generative AI products have been known to both include CSAM in their training models,<sup>24</sup> and often to create synthetic CSAM when prompted.<sup>25</sup> While accurate figures may be difficult to develop, the rise of synthetic CSAM appears to have been exponential:
  - The Internet Watch Foundation found a 26,362% rise of photo-realistic AI videos of child sexual abuse in 2025<sup>26</sup>
  - In 2024, the National Centre for Missing and Exploited Children reported an alarming 1,325% increase in reports of CSAM<sup>27</sup>
  - The Office of the eSafety Commissioner noted a 218% increase in reports of AI-generated CSEA material from 2023 to 2024.<sup>28</sup>
  
- Privacy violations: AI systems train on children’s data, and also process children’s data as they function. Often the ‘training process’ violates children’s data rights. For example, Australian children’s data has been included in data training sets without consent.<sup>29</sup> Moreover, AI services often process children’s data without appropriate consent — for example, many retailers use AI enabled facial recognition software to scan the faces of consumers entering shops, which would include children.<sup>30</sup> The capacity for data breaches in data heavy AI systems also raises concerns, especially with regards to EdTech products.<sup>31</sup>

Figure 1: Some examples of known risks and opportunities of AI use for children.

<sup>24</sup> See for example, David Thiel & Jeffrey Hancock *Identifying and Eliminating CSAM in Generative ML Training Data and Models* <https://purl.stanford.edu/kh752sm9123>. However the full scale of the issue is relatively unknown. The training data sets for commercial AI models are proprietary, and not open for public inspection. This research was conducted on LAION-5B, one of the few large-scale, open datasets.

<sup>25</sup> Chris Vallance 2026 ‘Elon Musk’s Grok AI appears to have made child sexual imagery, says charity’ *BBC* <https://www.bbc.com/news/articles/cvg1mzlyxeo>

<sup>26</sup> Internet Watch Foundation 2026 *AI becoming ‘child sexual abuse machine’ adding to ‘dangerous’ record levels of online abuse, IWF warns* <https://www.iwf.org.uk/news-media/news/ai-becoming-child-sexual-abuse-machine-adding-to-dangerous-record-levels-of-online-abuse-iwf-warns/>

<sup>27</sup> Emma Vaughan 2025 *2024 in Numbers* <https://www.missingkids.org/blog/2025/nmec-releases-new-data-2024-in-numbers>

<sup>28</sup> Office of the eSafety Commissioner 2025 *Generative AI & Child Safety* <https://www.esafety.gov.au/newsroom/blogs/generative-ai-and-child-safety-a-convergence-of-innovation-and-exploitation>

<sup>29</sup> Human Rights Watch 2024 *Australia: Children’s Personal Photos Misused to Power AI Tools* <https://www.hrw.org/news/2024/07/03/australia-childrens-personal-photos-misused-power-ai-tools>

<sup>30</sup> See for example, Choice 2022 *Kmart, Bunnings and The Good Guys using facial recognition technology in stores* <https://www.choice.com.au/data-protection-and-privacy/data-collection-and-use/how-your-data-is-used/articles/kmart-bunnings-and-the-good-guys-using-facial-recognition-technology-in-store>

<sup>31</sup> See for example, an American case study of a data breach. Danai Nhando 2026 ‘Unmasking EdTech’s Surveillance Infrastructure in the Age of AI’ *Tech Policy Press* <https://www.techpolicy.press/unmasking-edtechs-surveillance-infrastructure-in-the-age-of-ai/>

## Opportunities to advance safety within the AI Plan

### Consumer law, AI & children's rights

#### Recommendations

To unleash the capacity of the ACL to ensure children's rights are respected by AI services and products would require:

- A revision of the ACL to introduce consumer safety standards and liability for AI goods *and* services
- The establishment of a Digital Ombuds service, or at the least, a Digital Ombuds for children. This would need to be adequately resourced and staffed
- A review of Child Safety Standards,<sup>32</sup> and how they may apply to AI service and products operating in Australia
- Guidance from the ACCC about what unconscionable conduct looks like on AI services and products for children, which could also be informed by Child Safety Standards
- Complementary regulation that sits outside the ACL, that would advance its functioning, is needed. This includes:
  - Regulation that includes requirements for transparency, such as:
    - User data about the number of Australian users accessing a service or product
    - AI service or provider data about risks and harms
    - A researcher access schemeThese would most likely be located outside of the ACL in other regulations, such as through reforms of the *Online Safety Act*
  - Privacy reforms — including a strong Children's Online Privacy Code and tranche II of the *Privacy Act* reforms — that addresses and limits data sharing, and includes a prohibition on targeting children (including the whole pipeline of data collection)
- Outside of the ACCC, the network of regulators associated with AI service and products — the DP Reg group — needs to be enabled to fully implement and enforce the regulations already in place
- Within the regulatory landscape, we need to adopt a regulatory approach that moves away from 'directed at children' requirements to 'likely to be accessed by' requirements. This will require guidance from regulators about what the likely to be accessed standard looks like in Australia
- At a civil society level, we need to ensure that we are not contributing to the mystification of AI, by talking about AI products and services, and AI companies and providers rather than talking about "AI". This will help to reframe the discussion away from the idea that technology is a 'black box' that is separate from the humans who created it, and towards the business models that underpin it

---

<sup>32</sup> For example, the National Office for Child Safety 2018 *National Principles for Child Safe Organisations* <https://www.childsafety.gov.au/resources/national-principles-child-safe-organisations>

## Discussion

The ACL has the capacity to improve children’s rights when they are users of AI products and services, but there are many challenges to implementing it effectively. The ACL provides protections to users of a product or service, offering them the right of a consumer guarantee. This includes children where they are users of products. This applies even if they were not the original purchasers of the product.

Where children are users of a product, they may be entitled to a more 'rigorous' consumer guarantee. This is particularly relevant given the common industry claim that digital products and services were unaware children were using their services, or that their products were not directed at children. This is not necessarily a valid argument under the ACL as products and services are meant to consider unexpected use, which would presumably include unexpected use by children. It does highlight however, the value of having data about children’s use of specific AI products and services. A move from a focus on services ‘directed at children’ to ‘likely to be accessed’ by children would also help.<sup>33</sup>

The consumer safety guarantee outlines that products need to be safe, and services need to be rendered with reasonable care and skill, but have limited requirements for safety. AI is typically classed as a service, but there were discussions around whether a more appropriate classification would see them formally represented as a product, which could create product safety rules for them. There is a precedent for this. The ACL defines “software” as a product not a service. In principle, meeting the *Voluntary AI Safety Standards*<sup>34</sup> should demonstrate reasonable care and skill, but they potentially lack nuance.

Accountability to the *Voluntary AI Safety Standards* also presents a challenge, as does accountability to consumer law in general. Regulators seem to have a preference to issue guidelines and voluntary codes and standards, which can be easily overlooked by AI providers and platforms without clear consequences. However, there appears to be a welcome appetite from the ACCC to expand their scope and remit over digital platforms, emerging from the *Digital Platforms Inquiry*, followed by an announcement that digital services and products were going to be a priority focus for the regulator across 2026/27.<sup>35</sup>

Regardless, currently consumers — including children, parents and guardians — are often left without a suitable means to seek recourse when something goes wrong. The ongoing discussions around a potential Digital Ombuds highlights one avenue that can allow consumers to seek some redress.

Transparency and the need for evidence was also noted as potential barriers to seeking recourse with regulators such as the ACCC. Currently, it can be difficult to prove harm was caused by certain products or services because the technology underpinning the services and products they provide are proprietary. This makes determining any causal relationship between the product or service and harm difficult. Two possible solutions to this were discussed:

- Stronger investigative powers for regulators, or potentially a more muscular approach to using the powers that they have, to surface evidence and data around harms
- Researcher access regimes, to allow public interest research to identify the connections between products-services and harms.

---

<sup>33</sup> See for example, Reset.Tech 2025 *The likely to be accessed test and the Children’s Online Privacy Code* <https://apo.org.au/node/331748>

<sup>34</sup> Department of Industry, Science and Resources 2024 *Voluntary AI Safety Standards* <https://www.industry.gov.au/publications/voluntary-ai-safety-standard>, noting that these have been updated and replaced in Oct 2025

<sup>35</sup> ACCC 2026 *ACCC’s compliance and enforcement priorities update 2026-27 address* <https://www.accc.gov.au/about-us/news/speeches/acccs-compliance-and-enforcement-priorities-update-2026-27-address>

The role of Child Safety Standards<sup>36</sup> were also discussed, as each State has their own set of standards, and they are organisationally focussed (i.e. provider focussed). However, these are largely focussed on “children’s services” and implemented sector by sector. The contents of these Standards may however, provide insights for how AI providers *ought* to realise children’s right to safety, and could provide clarity for understanding what unconscionable conduct looks like.

There are specific use-cases of AI which raise child rights issues, such as AI used in education settings, or healthcare settings. Specifically:

- On top of all of the challenges that any AI product or service poses, in these cases there are specific governance issues that need particular consideration. While children have the right to participate in decision-making processes that affect them, under article 12 of the *Convention on the Rights of the Child*, it is unclear how children and young people are involved in the governance of these AI products and services (or indeed, if they are even aware of their deployment).<sup>37</sup>
- AI chatbots are being sold into schools as replacement counsellors, but the health outcomes of these products are often questionable. It is also unclear if they are licensed the same way as health care products should be.
- The regulation of EdTech products overall raises questions about children’s right to privacy, and to the ‘commercialisation’ of the learning experience. AI stands to exacerbate these challenges.

Often, products deployed in child-centred spaces are ‘general use’ products because specialised products for children have not been developed. That said, we note that child-focussed products might not be desirable in a market where the business model is inherently exploitative, and that potentially service providers themselves can also be cautious of the additional regulatory oversight they attract.

Privacy law also has a role to play in advancing children’s rights as AI use expands. AI training models both include children’s data, which raises consent and privacy issues, and also create data about children, by for example enhancing data sets. There is an opportunity to place guardrails on both of these processes through the Children’s Online Privacy Code and the much anticipated tranche II of the *Privacy Act* reforms.

Lastly, how we describe and define ‘AI’ matters. Firstly, as described above there are consequences from representing AI as a product or a service, and secondly because language and discourse frame policy and potential solutions. Understanding that ‘every product is a process’ and every process embeds inherent values helps to disarm the myth that ‘AI’ is an incomprehensible black box that is independent or separate from humans and beyond our capacity to manage. This description of AI highlights the role of business models, and human decision making, in defining and shaping the AI products and services that end up being used by children.

---

<sup>36</sup> For example, the National Office for Child Safety 2018 *National Principles for Child Safe Organisations*  
<https://www.childsafety.gov.au/resources/national-principles-child-safe-organisations>

<sup>37</sup> Three levels of ‘AI’ products and services were discussed: AI services where it in principle, it is potentially clear to children that AI is used, such as Companion services and chatbots; existing services that use embedded AI technology, such as search services where children may or may not be aware that AI is enabling their functionality, and; services and products where AI use is entirely obscured such as use in health care provision.

## Online safety law, AI & children's rights

### Recommendations

- The Online Safety Act (OSA) needs to be revised, to include:
  - The introduction of a Digital Duty of Care (DDOC). The DDOC must supersede the BOSE and its codes, or at the very least, it must be made clear that the DDOC is not limited to the BOSE
  - The DDOC must also be singular and overarching to ensure it covers all systems and processes that platforms deploy, and to cover all risks — both those we currently know the OSA cannot cover, and potential future unimaginable risks
    - Advertising systems and monetisation systems, as well as content recommender systems, must be covered by the DDOC
  - The current industry classification scheme needs to be replaced with a broader focus on all information services or at least expanded to include AI services and products beyond chatbots. This is to avoid repeating the limitations of the current industry classification scheme that struggles to capture the horizontal and vertical nature of AI
  - Requirements for heightened transparency. This could include:
    - Proscribed annual transparency reports, including data such as the number of Australian child users and information about risk
    - A researcher access scheme
    - Increased investigative powers for regulators
  - A licensing scheme or onboarding requirements that allow additional requirements to be placed on platforms and providers, and increases accountability. Requirements for children's participation should be part of any licensing scheme.
- Supportive regulation that sits outside the OSA, but would be necessary to advance its functioning, such as:
  - Reforms to the *Privacy Act*, which need to include; prohibiting targeting to children; the right to request data removal for children, and; direct marketing provisions (especially when it comes to ads for nudifying apps for example). In addition, consideration needs to be given to how the *Privacy Act* could address the use of children's data in AI training sets overall
  - Exploring the capacity of Commonwealth and State based criminal laws regarding 'sexting', when it comes to spicy Companion AI chatbots
  - Addressing algorithmic bias, which the National AI plan references, may need a focus on anti-discrimination laws and processes that currently sit outside the OSA.
- Legislative and regulatory reforms could be supported by:
  - Undertaking or commissioning an analysis of what a holistic child-rights approach to AI products and services would look like, including participation and access rights as well as protection issues
  - The development and use of industry standards — in the International Organisation for Standards (ISO) model of Standards — as would certification schemes against these standards.
- At a civil society level, we need to develop a vision of agency for children throughout the policy development cycle, supported by critical literacy skills to support and enable children to participate.

## Discussion

The *Online Safety Act* (OSA) has limited capacity to improve children's rights on AI services and products. The OSA's core focus is on removing the 'worst of the worst' illegal content, and protecting children from accidental exposure to pornography, but many of the risks AI poses to children do not fit neatly within these categories.

We noted that, for example, one of the key requirements that AI chatbot style services may face is requirements for age gating and the deployment of age assurance technology. Regardless of the legitimate questions of efficacy around age assurance technologies and their impact,<sup>38</sup> this approach places a lot of policy eggs in the age gating basket, with few other requirements. Two issues arise from this; firstly, it is absence of protections for children over any minimum age requirements (i.e. 16- and 17-year-olds have no protections), and; secondly, if age assurance and age gating does single-handedly deliver the panacea of benefits policy-makers have intended, children will remain exposed to risks.

Other limitations to the OSA that may limit the ability of online safety law to improve protection on AI products and services were also noted, especially with regards to:

- The systems and processes platforms deploy, that sit beyond a content focus. For example, advertising approval systems and monetisation systems were noted as absent from the OSA
- The industry classification scheme, which fails to reflect both the vertical and horizontal nature of AI services and products
- The harms that the OSA focuses on are relatively narrow, and concern content harms (especially from content of a sexual nature) and sexual exploitation and abuse. While these are critical issues, they do not reflect the totality of risks children face from AI systems and services.

For example, the use of children's data — and in particular their images — in Generative AI products was discussed. Children are not able to consent to the use of their images for training AI, and beyond this, often have no knowledge or 'say' about whether their images are shared online (where parents or guardians post them). This means children's images often appear in data sets without their consent, assent or knowledge.<sup>39</sup> These images are often used to create new likenesses, or alter to depict children in deepfake contexts. This is a risky capacity, that children had 'no say' in the development of.

One potential way to 'overhaul' the focus of the OSA, and to ensure its coverage beyond content and beyond the current industry classification scheme would be the implementation of an effective Digital Duty of Care (DDOC). A DDOC has the capacity to be defined in terms of users' rights, and this could include children's rights where they are users. However, to achieve impact, A DDOC would need to either supersede the current Basic Online Safety Expectations (BOSE) and their codes, or at the very least outline that the duties are not limited to the BOSE and that compliance with the BOSE alone does not equate to fulfilling obligations under the DDOC. A strong risk assessment framework would need to be implemented to ensure the DDOC model creates proactive, systemic change.

Further the DDOC would need to extend beyond the Online Contents Scheme to focus on all of the systems digital platforms, and AI systems and products deploy, in order to truly be effective. In effect then, the OSA would need to comprise of a new DDOC, with new Industry

---

<sup>38</sup> See for example, page 52 of *Age Check 2025 Age Assurance Technology Trial*  
[https://www.infrastructure.gov.au/sites/default/files/documents/aatt\\_part\\_d\\_digital.pdf](https://www.infrastructure.gov.au/sites/default/files/documents/aatt_part_d_digital.pdf)

<sup>39</sup> Human Rights Watch 2024 *Australia: Children's Personal Photos Misused to Power AI Tools*  
<https://www.hrw.org/news/2024/07/03/australia-childrens-personal-photos-misused-power-ai-tools>

Standards covering all information services, and a (potentially refreshed) Online Contents Scheme.

There are some parallels and precedents for this in the Scams Prevention Framework, which places broad requirements on platforms to detect and disrupt scams across all the processes a system deploys (noting that the focus is on scam content here, and the OSA would need to move beyond content).

There was discussion about the narrowness of a Statutory Digital Duty of Care, which limits focus to the digital world, and places liability on platforms and providers rather than the Government, when compared to the broader concept of a Duty of Care. This was particularly apparent when it came to the Government's strong desire to challenge the notion that it owed a duty of care to young people with regards to climate change. However, there was a desire not to see the perfect as the enemy of the good.

The OSA review process also provides the opportunity to include requirements for licensing and 'onshoring' that could help drive up accountability processes, as well as increased requirements for transparency and researcher access that could help with gathering evidence of risks and harms.

Outside of the OSA, support for children's safety could emerge from:

- Improved privacy protections for children. For example, a stronger enforcement of the rules around direct marketing or data sharing could protect children from some harmful uses of AI services and products used in advertising delivery. These protections could be strengthened through either the Children's Online Privacy Code and the much anticipated tranche II of the *Privacy Act* reforms.
- More effective enforcement of, or revision to, Commonwealth and State Criminal Codes, specifically when it comes to 'spicy' Companion AI Chatbots. There was a lack of clarity about if spicy discussions between Chatbots and children under 18 may invoke criminal codes around 'sexting'.
- Amendments to anti-discrimination laws to effectively address algorithmic bias (which is noted in passing in the National AI Plan). This also requires an explicit recognition that algorithmically mediated discrimination can compound existing disadvantages for children, especially for First Nations communities, children with disability, children in out-of-home care, and children in low-income households. Any review of anti-discrimination law in the context of AI should explicitly consider these intersections.

Beyond the legal system, the development and use of 'traditional' industry standards in the International Organisation for Standards (ISO) sense could be useful, as would certification schemes against these standards.

Lastly, the need for a rights enhancing approach to developing and implementing online safety regulations was stressed, with the need for broad based participation from children and young people. This is in keeping with Article 12 of the *UN Convention on the Rights of the Child*. We noted the development of the Youth Reference Group at the Office of the eSafety Commissioner as a positive step, but discussed the need for meaningful children's participation to include and extend beyond this. Firstly, as a formal mechanism the Terms of Reference for the group and its members can have a limiting effect on public commentary, and secondly, there is a need for participation right across the 'policy stack' of AI governance debates — from school and university focussed policies to health care providers and so on. The lack of youth input in the early development of the *Social Media Minimum Age Act* was noted as a low-point for youth engagement in policy development. Young people's exclusion from these policy conversations is also notable given the documented role social media plays in the formation of political identity among adolescents. AI systems that shape what

young people see and engage with online have significant implications for civic development. This is a dimension of children's rights that falls outside the current framing of online safety law.

## AI Safety research & children's rights

### Recommendations

- The research agenda emerging under the National AI Plan needs to foster research into:
  - A child-rights framework for understanding the risks and capabilities of AI products and services on children
  - Procurement of safe AI within schools, child-care and health-care settings
  - Human security as a safety issue
  - Developing model training sets that are appropriate for children and built for children, both in Australia and globally
  - A future-focussed sandbox that allows researchers to test Generative AI guardrails, and other 'unimaginable' risks
- Participation and co-design needs to be built into the research agenda. Civil society is skilled and able to support this, beyond developing a potential 'youth reference group' for the AI Safety Institute.
- Support for civil society to meaningfully undertake AI safety research outside of the AI Safety Institute, including:
  - A researcher access scheme and requirements for publicly available data
  - Different or enhanced funding models for AI safety research as it affects children's rights.
- At a civil society level, we need to adopt a language of 'risks and capabilities' from AI services and products, to ensure that the extent of integration of technology in young people's lives is recognised.
- Meaningful participation requires structural independence from government. Youth panels convened under the auspices of regulatory bodies are valuable but limited in their capacity for public advocacy. Civil society should consider developing an independent child and youth advisory body on AI governance that can engage across the full policy stack. This ranges from procurement decisions in schools through to national regulatory frameworks. It is essential that this is adequately resourced to support genuine co-design.

### Discussion

There is a current gap in knowledge about the risks of AI products and services for children, and how we should understand them from a child rights perspective. Specifically:

- Understanding risks also requires understanding capabilities of AI products and services for children, and this capability focus needs to be identified as an equal knowledge gap. A focus on 'risks' alone can be both individualising,<sup>40</sup> and present only one side of the story from a child rights perspective
- The current 5C's framework was largely developed for digital platforms, and there were questions about if it is appropriate as a foundation for understanding

---

<sup>40</sup> Noting that in this individualised model, if children are risky or at-risk, perhaps platforms, services and providers could be considered as perpetrators

- A broad range of risks and capabilities are under investigated. For example, risks arising from the attention-economy and human-safety from the cognitive perspective are rarely addressed when it comes to children.

To be effective, AI Safety research would need to address this knowledge gap, and develop a guiding framework to identify the issues and the opportunities from AI products and services from a child rights perspective. These should inform the work of the AI Safety Institute from the outset.

Outside of the need for research around risks and capabilities for individual children, there is also a need for safety research that focuses on the systems that children inhabit, such as education systems, child-care systems and health systems. For example, safety-focussed research could have substantial impact by exploring educational procurement models and how schools can safely evaluate and purchase AI services and products.

The AI Safety Institute should be given specific investigative powers to test the capacity of generative AI systems to produce child sexual abuse material. Without the ability to test guardrails, it is impossible to evaluate whether platforms are meeting their obligations. The Oxford Internet Institute and similar bodies provide a model for what this research capacity could look like. However, it requires a legal framework to support it.

Consideration needs to be given to both the current state of ‘safety and risks’ when it comes to research into AI products and services, but also the future of safety. Many issues are currently emergent or unimaginable, and any research agenda needs to hold space and flexibility to be able to address these. As an example, the inability of researchers around the world to test the safeguards in generative AI when it comes to the production of synthetic CSAM were discussed. These risks were unimaginable 5 years ago, emergent 4 years ago, but have recently come to a head. Without a future-focussed sandbox in place, researchers have largely been left unable to test Generative AI guardrails.

Consideration of safety needs to extend across the technology life-cycle. This includes exploring risks and capabilities emerging from; the model training stage, to ensure children’s rights are respected as foundational models are developed; safety mechanism development and deployment, to explore for example how post-inference, input and output filtering function; monitoring systems to ensure adequate thought is given to how these models operate in practice and how they are used, and; collaboration between regulators, industry and civil society.

Whatever the focus of the research, children’s right to participate needs to be respected. Youth-led research, as well as youth consultative research, has significant capacities to generate new insights and understandings, and needs to be considered as part of the broader research agenda. Children are experts in their own lives, and often share critical insights only when their experiences can be centred in the research agenda.

There is a need for AI safety research that would sit outside the proposed AI Safety Institute. For example, civil society can develop and deliver impactful research, as can academia. Fostering this requires a deliberative approach to collaboration, capacity building and resourcing. For example:

- Transparency requirements around publicly documented risks assessments and data about risks and harms, and a researcher access scheme, could provide critical evidence to researchers
- Collaboration and capacity building around research for evaluation could help to develop a model for understanding how AI products and services advance or threaten children’s rights

- Existing research funding models could also include priorities for AI safety research as it affects children's rights.

Alongside research into safety, a comprehensive approach would require fostering enhanced critical AI literacy among children. Civil society may be well placed to develop and deliver such literacy schemes, with resourcing, in comparison to an AI Safety Institute. There is a strong role for research documenting evidence of 'what works' in digital literacy. A focus on digital literacy however, needs to reflect the learnings from the past, and ensure that 'digital literacy' is not deployed as an intervention that avoids the need for strong regulatory action. Digital literacy skills can function only as the 'last line of defence' after regulatory reforms have been implemented and safety-by-design techniques adopted.

## Conclusions & Recommendations

The discussion here offers a 'blueprint' for the way the National AI Plan could be developed and implemented that would advance children's rights. While detailed recommendations are provided in each section, the discussion above identifies four key priorities:

- **Consumer protections for AI-enabled goods and services.** Introducing consumer safety standards and liability for AI goods *and* services could enhance the impact of the ACL on AI products as children use them. Product safety obligations, including for AI goods and services, on digital platforms that act as market places for these products (as intermediaries) would provide stronger protection for children. Likewise, a Digital Ombuds and other avenues for complaint and redress that has remit over digital platforms, including for the supply of AI goods and services could help advance children's rights.
- **Reducing online harms through reforms, codes and standards within the *Online Safety Act*.** Revising the Act to include a singular, overarching statutory Digital Duty of Care that replaces the current BOSE could go some way to ensuring meaningful accountability under the Act. The Codes under the Act would need to be replaced, and the industry classifications revised and broadened.
- **Advancing the science of AI safety.** A broad AI Safety research agenda is necessary that incorporates children's rights, including participation, access and safety rights. This must address issues affecting children directly, and children's contexts, such procurement of safe AI in schools and universities. Co-design and participation of children and young people need to be a central part of this, as does civil society collaboration. A researcher access scheme would ensure that critical scientific data can be surfaced.
- **The establishment of an AI Safety Institute.** The Institute presents an opportunity to embed children's rights into its foundational work programme. Its mandate to generate technical research into AI risks and harms, and to engage with regulators, should explicitly include children's safety, privacy and participation as core priorities from the outset.

Finally, there was an agreement that at a civil society level, the discussions held at this workshop should be the start of ongoing deliberations. More organisations, and more focussed discussions — that are inclusive of children — are necessary.

