Submission to the Department of Home Affairs' Consultation on Horizon 2 of the 2023-2030 Australian Cyber Security Strategy

29 August 2025

Introduction

Global Shield Australia welcomes the opportunity to provide a submission to the Department of Home Affairs' Consultation on Developing Horizon 2 of the 2023-2030 Australian Cyber Security Strategy.

Global Shield Australia is an independent, non-profit policy advocacy organisation dedicated to reducing global catastrophic risk. We take an all-hazards approach to preparedness, supporting governments to enact and effectively implement policies that prevent and prepare for all forms of risk. This includes a particular focus on threats from emerging technologies, securing critical infrastructure, and considering how cyber security and critical infrastructure policy can be part of a system-wide solution to reducing risk.

This submission focuses on Shield 4 of the Australian Cyber Security Strategy (Protected Critical Infrastructure), particularly the 'last resort' powers found in the *Security of Critical Infrastructure Act* (2018) (**SOCI Act**). It also highlights the key trends and strategic factors that Horizon 2 of the 2023-2030 Australian Cyber Security Strategy (**Horizon 2**) will need to engage with, and provides recommendations for Shield 3 (World-class Threat Sharing and Blocking), given their interaction with themes considered under Shield 4.

Contents

The Outlook for Horizon 2

Discussion paper question

1. What trends or technology developments will shape the outlook over the next few years and what other strategic factors should Government be exploring for cyber security under Horizon 2?

The Horizon 2 time period will be a critical time of change in Australia's cyber security landscape. Rapid technological development combined with significant geopolitical uncertainty creates a challenging strategic environment that Australia and our cyber industry needs to be prepared to engage with.

As the Government develops actions and initiatives for Horizon 2, it will need to particularly ensure these are responsive to a cyber security environment that is being transformed by advances in artificial intelligence (**AI**). As the Cyber and Infrastructure Security Centre (**CISC**) has recognised "*AI technology will almost certainly make elements of cyber intrusion operations more effective and efficient, which may lead to an increase in the frequency and intensity of cyber threats*".[1]

Beyond just exacerbating cyber risk, AI will also create new risk vectors, test—and potentially break—existing governance arrangements, and re-shape the relationship that government, business and citizens have with information technology. The most impactful AI **trends and developments** relevant to the cyber security environment under Horizon 2 include:

1. **Rapid and increasing AI capability gains:** recent frontier AI models have surfaced new capabilities such as to assist the identification and exploitation of zero-day vulnerabilities, generate malware, and enhance spear-phishing and targeted campaigns.[2] These capabilities are only going to deepen as the frontier develops, democratising access to advanced cyber attack tools and making it easier for novice rogue actors to exploit unknown vulnerabilities. This includes through AI-enabled social-engineering and undermining of public information sources, reducing trust in the authenticity and reliability of online services.
2. **Agentic AI behaviour:** AI systems that can plan and act with minimal or no human supervision are moving into production and consumer facing use-cases. Without proper supervision, these systems could exploit known misconfigurations, discover new vulnerabilities, and amplify the speed and scale of cyber attacks.
3. **Systemic risk from a concentration of AI foundation model developers:** the most capable foundation and frontier AI models are being developed by a small number of overseas firms. These models have the potential to be deployed in a variety of contexts and settings creating systemic risk from common failure modes and leaving Australia exposed to foreign regulatory choices or supply chain disruptions. The July 2024 CrowdStrike incident, where a single faulty update disabled around 8.5 million Windows devices globally, demonstrates how global monocultures can translate into cross-sector outages.[3]
4. **Integration of AI into critical sectors:** the integration of AI into energy, healthcare, transport, finance, and government systems presents enormous potential benefits. AI tools can

[1] CISC, 'Factsheet for Critical Infrastructure: Artificial Intelligence in Critical Infrastructure', June 2025, https://www.cisc.gov.au/resources-subsite/Documents/artificial-intelligence-factsheet.pdf.
[2] See, e.g., Saskia Laura Schroer et al, 'SoK: On the Offensive Potential of AI', 3rd IEEE Conference on Secure and Trustworthy Machine Learning, 24 Jan 2025, https://arxiv.org/pdf/2412.18442.
[3] Joe Tidy, 'CrowdStrike IT outage affected 8.5 million Windows devices, Microsoft says', BBC News, 21 July 2024, https://www.bbc.com/news/articles/cpe3zgznwjno.

strengthen defence through faster detection and response to incidents. But they can also widen attack surfaces and introduce novel failure modes and vulnerabilities (such as model hallucinations, poisoned data sources, and prompt injections).

5. **Safety and trustworthiness assurance falling behind capability development:** at present safety and security evaluation of AI models is uneven and underinvested in. This presents a trust-gap that could hinder uptake or result in uneven or inappropriate deployment practices.
6. **Risk to authenticity and trust from deepfakes and synthetic media:** highly realistic images, audio, and video from generative AI systems threatens to erode trust in digital channels and media sources.

There are also associated **strategic factors** that must inform the Government's approach to Horizon 2. These will require proactive and forward-leaning Government action to preserve Australia's strategic agency, limit dependence on foreign firms and global supply chains, and manage risks from geopolitical competition and hostile actors:

1. **Understanding and protecting AI and cyber security supply chains:** AI models and their artefacts (weights, training data sets, evaluation suites, prompts, and fine-tuning) need to be treated as the complex supply chain systems they have become. This will require increasingly robust supply chain management systems and frameworks to ensure transparency, accountability, and trust in AI.[4]
2. **Reducing exposure to foreign AI supply chains and regulators:** as a likely importer of advanced AI models and systems, without deliberate action Australia may default to being a rule-taker on AI regulations and rely on foreign regulators to secure AI deployments domestically. This also presents an opportunity for Australia to build domestic third-party assurance capacity that can evaluate safety, robustness, misuse-resistance, and overall controllability of AI systems.
3. **Keeping pace with international developments around AI alignment and safeguards:** The science of making AI systems provably aligned to human interests is advancing, including methods to ensure AI is 'provably safe'.[5] Australia should track and, where appropriate, fund this line of work to reduce our exposure to systemic risk.
4. **Rising risk of AI-enabled catastrophic crises:** in addition to safeguarding existing systems, the Government needs to engage in a rigorous and comprehensive risk assessment and planning for catastrophic AI-enabled incidents.
5. **Geopolitical uncertainty and shifting regional norms:** Australia's ambition to be the partner-of-choice in the Indo-Pacific[6] will be tested as traditional alliances and international norms shift. Demonstrating leadership within the region and in defence of shared interests will be key to maintaining Australia's role. This will also support our overall resilience by mitigating risk from interconnectedness and dependencies (such as in relation to critical inputs, submarine cables, software and hardware supply chains, and data centre access).

[4] See, e.g., work such as the Trustable AI Bill of Materials (TAIBOM), 2025, https://taibom.org/.
[5] See, e.g., David Dalrymple, 'Safeguarded AI: Constructing Guaranteed Safety', Advanced Research and Invention Agency, 2024, https://www.aria.org.uk/media/3nhijno4/aria-safeguarded-ai-programme-thesis-v1.pdf.
[6] See, e.g., DFAT, 'Australia in the World - 2025 Snapshot', February 2025, https://www.dfat.gov.au/publications/international-relations/australia-world-2025-snapshot.

Shield 4: Protected Critical Infrastructure

Discussion paper question

33. How effective do you consider the SOCI Act is at protecting Australia's critical infrastructure from cyber attacks? Are the current obligations proportionate, well-understood, and enforceable?

Global Shield Australia considers that the current obligations in the SOCI Act in regard to so-called 'last resort' powers could be better understood across government, industry and more broadly (**Box 1**).

The recent reforms to the SOCI Act provide the Government with important authorities to prepare for and respond to a national crisis or catastrophic incident, including in the cyber domain. The last resort powers, in particular, are powerful and necessary tools for enhancing the preparedness and resilience of Australia's critical infrastructure. However, more can be done to increase understanding and certainty in relation to when, why, and how the government might decide to use these powers.

Box 1. Key SOCI Act Powers

Part 3A of the SOCI Act contains a set of significant government capabilities designed for '**last resort**' use. These can be used in response to a serious incident that has, or is likely to, impact critical infrastructure, and has a material risk of seriously prejudicing Australia's social or economic stability, defence or national security. Once triggered, Part 3A allows for a range of responses including information gathering, the giving of mandatory directions, and intervention requests to the Australian Signals Directorate to intervene in relation to serious cyber security incidents (such requests also require agreement from the Prime Minister and Defence Minister).

Separate to the Part 3A powers, under **Part 3 of the SOCI Act**, the Minister can also issue a pre-emptive direction to certain critical infrastructure entities to do or refrain from doing specific acts where the Minister is satisfied that there is a risk that an act or omission would be prejudicial to security. **Part 2A of the SOCI Act** also empowers the Government to order remedies of 'serious deficiencies' in critical infrastructure risk management programs.

To fully realise the value and effectiveness of the SOCI Act powers, Global Shield Australia recommends that the Government:

1. **Strengthens planning and guidance on the use of the last resort and associated powers**
   In 2024-25, the CISC's Regulatory Posture aimed to balance education and awareness raising activities with compliance activities.[7] This focus on uplift has helped industry to understand and implement their obligations, and embed sound security and risk management practices into the critical infrastructure sector. As CISC moves into 2026 and beyond, and given the expansion of the SOCI Act in 2024, a full examination of the range of enforcement powers is warranted, with a particular focus on the last resort powers, to understand what further planning and guidance is needed in regard to their use.

[7] See CISC, 'SOCI Compliance Regulatory Posture 2024 and beyond', CISC News, 6 March 2024, https://www.cisc.gov.au/news-media/archive/article?itemId=1176.

Such an examination should look to ensure these powers are well-understood within government and also within industry, and that there are clear guidelines and policies around when they could be triggered and how they could be used (particularly in a catastrophic crisis).[8]

Importantly, this guidance should *not* avoid considering and specifying scenarios of when and how these powers may be used. While the last resort powers are and should always be truly last resort in nature, frank communication to industry of the government's ability and willingness to use them to protect Australians is essential. This will help to set expectations regarding the risk environment, enhance sector-wide preparedness, and minimise surprises when it does become necessary to make use of these powers.

2. **Recognises and plans for reactive and proactive applications of SOCI Act powers**
   As part of the recommended examination of the last resort and associated powers, attention should be paid to both their *reactive* and *proactive* application. The proactive use of the SOCI Act powers is likely to be particularly complex and politically sensitive. Action early in or prior to an impending crisis can be critical to avoid high-consequence impacts, particularly for fast-moving threats such as those associated with AI failures or misuse. Clear criteria, thresholds, and decision-making protocols should be developed for such proactive use, including in relation to engagement with the Australian Security Intelligence Organisation for security assessment purposes. While it is correct that the primary strategy should be to foster strong risk management practices within critical infrastructure operators, the Government must also plan for the contingency that these are not sufficient.

3. **Undertake communication to industry and stakeholders regarding the last resort powers**
   While certain operational details must remain classified, releasing non-sensitive guidance on the scope, intent, and general decision criteria for last resort powers to industry would help asset owners and operators understand how these powers may affect them, and align their risk management programs accordingly. Avenues such as the Trusted Information Sharing Network provide a secure method for communicating necessary information to stakeholders without compromising security interests.

   This transparency serves a dual purpose: it not only prepares industry for potential government action and also incentivises them to enhance their own security posture to prevent such intervention from becoming necessary. The outcome would be more efficient and effective regulatory action for government, industry, and the safety of the Australian public.

4. **Conduct exercises and scenario testing**
   To ensure operational readiness, the Government should conduct regular stress-tests of the SOCI Act powers, including joint exercises with the critical infrastructure sector. This should include testing of the practical activation of last resort powers under both reactive and proactive scenarios. Regular tests of these powers are key to ensuring they remain fit-for-purpose, that operational gaps can be identified, and that agencies and industry understand how and when these powers could be used to safeguard Australia.

[8] Noting CISC's existing guidance, such as CISC, 'Government Assistance' (4 April 2025), https://www.cisc.gov.au/how-we-support-industry/government-assistance; CISC, 'Incident Response: Government Assistance Measures' (April 2025), https://www.cisc.gov.au/resources-subsite/Documents/cyber-incident-response-government-assistance-measures.pdf.

Discussion paper question

36. What support would assist critical infrastructure owners and operators to mature their cyber and operational resilience practices? What role should government play in enabling uplift, including through tools, guidance or incentives?

The SOCI Act requires critical infrastructure owners and operators to maintain robust Critical Infrastructure Risk Management Programs (**CIRMPs**). These play a central role in ensuring that Australia's critical infrastructure is ready and able to respond to a range of threats and hazards. With the expansion of the SOCI Act to cover all-hazards, the Government can play a more active role in ensuring that these programs are more than just compliance checklists, and that they reflect the scale, complexity, and catastrophic potential of modern threats.

To that end, Global Shield Australia recommends that:

1. **The CISC integrate catastrophic risk into its Annual Risk Review**
   The CISC's Critical Infrastructure Annual Risk Review is a key assessment of all-hazards threats relevant to critical infrastructure. However, there is significant scope for future editions of the Annual Risk Review to better address potential catastrophic and worst-case scenarios. This will help ensure that owners and operators pay sufficient attention to how catastrophic incidents could impact their assets. In the absence of a national risk assessment, the Annual Risk Review should explicitly integrate significant and credible catastrophic threats (such as large-scale cyberattacks on operational technology, AI-enabled systemic failures, and concurrent hazards from natural and geopolitical shocks).

2. **Guidance is issued on the nature and relevance of catastrophic and AI risk to critical infrastructure security**
   The CISC has undertaken excellent engagement with industry to help it better understand and comply with their SOCI Act obligations. CISC's factsheets, seminars and information sharing mechanisms provide key guidance that is helping uplift the security of Australia's critical infrastructure. As part of this work, the CISC should issue guidance on catastrophic threats, with worked examples of plausible scenarios and sector-specific control expectations. As AI is increasingly deployed into critical infrastructure, there will also be a need for clear guidance around identifying and managing risk presented by the deployment and development of AI within CIRMPs. This should also include translating high-level threats into practical, sector-specific scenarios and response action plans.

3. **Support is provided to establish and grow the domestic AI assurance industry to certify AI deployments in critical infrastructure**
   Sovereign capability to assess and certify safe and secure AI will be critical for ensuring Australia is not dependent on foreign third-party providers or regulators to guarantee the safety of our critical infrastructure. As such, developing and supporting a domestic AI assurance industry is key. This will provide options for critical infrastructure owners and operators to assess their AI deployments and not leave them reliant on overseas offerings or the primary providers of AI systems. Support for a domestic AI assurance industry could

include: support for certifications through regulatory and procurement processes, co-funding assurance pilots and industry-academic partnerships, and recognition of third-party certification in compliance assessments.

4. **The Government undertake or support mapping of cyber and AI supply chains**
   Gaps in critical infrastructure resilience can often stem from hidden vulnerabilities or dependencies in upstream supply chains. This can be a particular problem for cyber and AI systems, which are often made up of complex systems of suppliers or involve proprietary models where there is little visibility of the provenance of their components or operation. The Government can support a better understanding of supply chain exposures by providing a consistent methodology for mapping cyber supply chains, identifying critical dependencies, and assessing how failure of or vulnerabilities in third-party providers could impact operations.

Shield 3: World-class Threat Sharing and Blocking

Discussion paper question

30. Are the roles and responsibilities of government and industry clear for cyber security in a conflict or crisis scenario? What activities, such as cyber exercises, could Government undertake to make you feel better prepared to respond in a cyber conflict or crisis?

The roles and responsibilities of government and industry for cybersecurity in a conflict or crisis scenario are not sufficiently clear, particularly concerning catastrophic or systemic cyber and AI incidents that could threaten national security. While plans like AUSCYBERPLAN exist,[9] their integration with broader, all-hazards national crisis arrangements could be improved, as could their promotion across and outside of government.

Industry requires clarity on the specific triggers for government intervention, the practical support it can expect to receive during a severe national cyber crisis, and what it may be obliged to do to assist the response—this is especially true in regards to the SOCI Act powers. To ensure preparedness, stress-testing and exercises are also required, including with all levels of government, industry, and other stakeholders.

To enhance preparedness and clarify these roles and responsibilities, Global Shield Australia recommends the following:

1. **Undertaking a national cyber risk assessment that explicitly includes catastrophic cyber risk**
   A comprehensive cyber risk assessment is essential to understand Australia's existing vulnerabilities, the state of our strategic environment, and the key threats on the near-term horizon. Such an assessment must cover the full range of threats and hazards, including catastrophic scenarios. The findings of this assessment can then inform better response planning, as well identifying where capabilities need to be built and roles need to be clarified. This includes the AUSCYBERPLAN and AUSCATPLAN.[10] The outcomes of the risk assessment should include unclassified summaries that can be made available to private industry and the public, because the utility of these plans depends in part on the cooperation of actors uninvolved in their classified development.

2. **Reviewing, mapping, and clarifying existing powers and roles and how these would be used in a crisis**
   A formal review of all legal powers and operational roles relevant to conflict or crisis scenarios is needed to ensure they remain appropriate and are well understood across sectors. The Commonwealth, States and Territories, and local governments have a range of legal powers that could be activated or used during a truly catastrophic crisis. Understanding how these powers relate to each other, how each could be best used in different crisis scenarios, and where gaps are or clarity is needed is essential. The outcome of this review should be a clear and public articulation of all levels of governments' legal powers, response capacities, triggers for their use, and reforms needed to enhance and clarify their use. This is particularly

[9] National Office of Cyber Security, Australian Cyber Response Plan (AUSCYBERPLAN), June 2025, https://www.homeaffairs.gov.au/cyber-security-subsite/files/australian-cyber-response-plan.pdf.
[10] See Department of the Prime Minister and Cabinet, Australian Government Crisis Management Framework (AGCMF), Coordination for Extreme to Catastrophic Crises (Tier 4), https://www.pmc.gov.au/resources/australian-government-crisis-management-framework-agcmf/crisis-coordination/coordination-extreme-catastrophic-crises-tier-4.

important to ensure that escalation points between the various crisis management plans are clear.

3. **Strengthening communication and engagement with industry and community stakeholders on crisis planning, particularly in relation to catastrophic crises**
   To ensure industry and the broader community feels prepared to respond to a crisis, active and regular communication is key. Communication prior to a crisis is essential so that Australians understand what powers the government has, when these will be used, and how they will ensure the safety and security of the community. This is also needed to ensure businesses are ready to cooperate with authorities in the use of these powers and understand their role in delivering safe and effective outcomes in a crisis.

4. **Conducting regular, multi-stakeholder stress tests and exercises**
   To ensure plans for conflict or crises are practical and effective, regular stress testing is essential. This testing must include the most severe, but plausible, crisis scenarios. We recommend a program of regular, joint exercises involving government agencies, critical infrastructure operators, and other key stakeholders. These exercises must go beyond conventional incident response and simulate catastrophic and systemic failure scenarios, such as those involving AI-driven attacks or where operational technology is compromised at scale. These tests are essential for identifying vulnerabilities in our national crisis arrangements and ensuring the SOCI Act powers are operationally ready to be used if ever required.

Conclusion

Horizon 2 is an opportunity for Australia to prepare for and navigate significant changes in the cyber security landscape. Proactive, forward-leaning action by government is needed to maintain Australia's strategic agency and strengthen national resilience. This means building a clear understanding of Australia's risk profile, vulnerabilities, and opportunities; preparing clear planning and guidance around when and how the Government will act, including clarifying powers and duties where necessary; and continuing engagement with industry and other stakeholders to communicate and test our collective ability to respond to a range of scenarios, including catastrophic crises.

Global Shield Australia appreciates this opportunity to contribute to the Department of Home Affairs work. We would also welcome any further engagement to discuss these matters in more detail.

About Global Shield Australia

Global Shield Australia is the Australian office of an independent, international policy advocacy organization dedicated to reducing global catastrophic risk. For more information on this submission or our work, please contact [redacted]

[redacted]