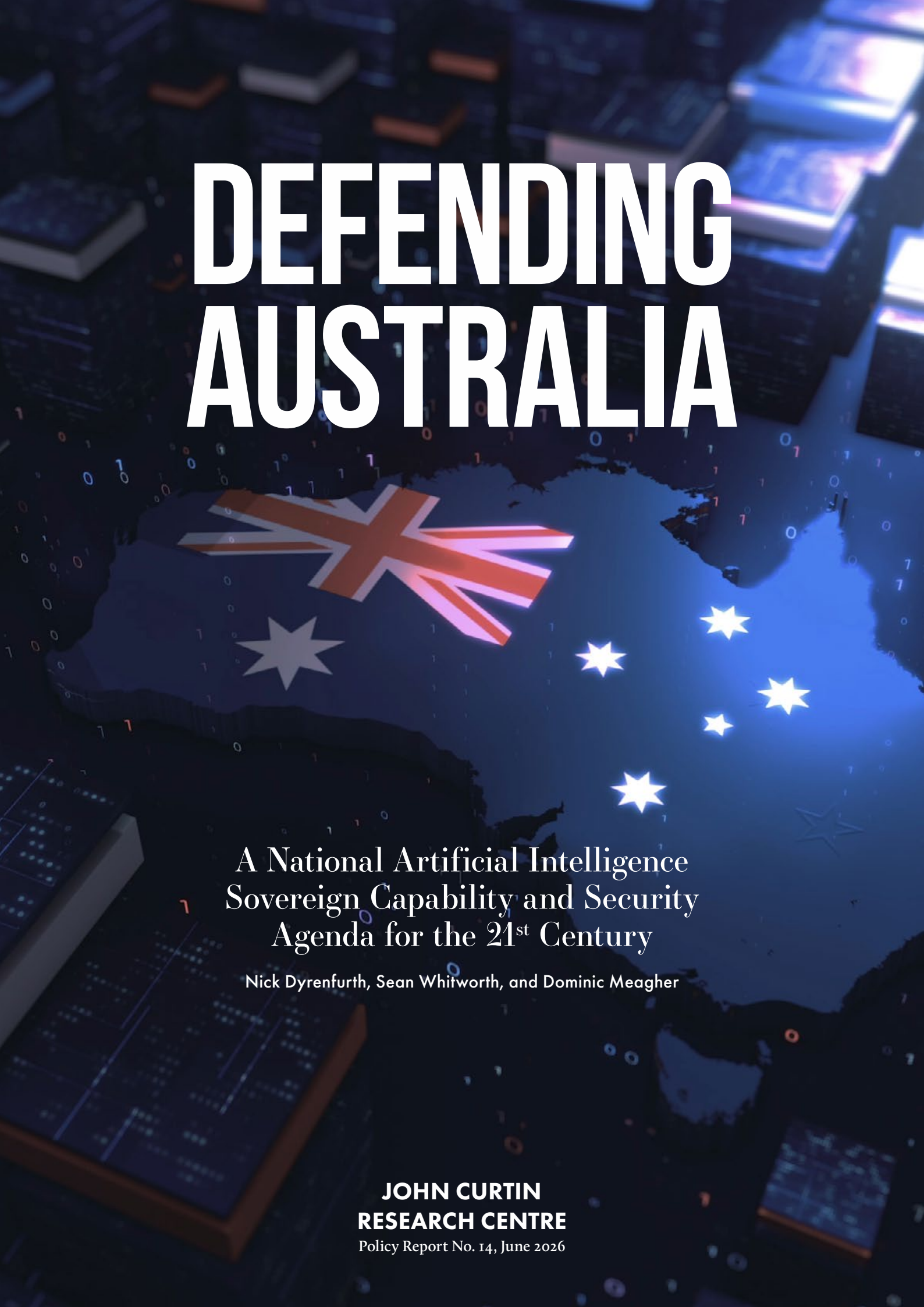

DEFENDING AUSTRALIA

A National Artificial Intelligence Sovereign Capability and Security Agenda for the 21st Century

Nick Dyrenfurth, Sean Whitworth, and Dominic Meagher

JOHN CURTIN RESEARCH CENTRE
Policy Report No. 14, June 2026

About the authors

Dr Nick Dyrenfurth is Executive Director of the John Curtin Research Centre and a columnist for The Australian newspaper. Nick has written and edited thirteen books, spanning politics and history. Nick authored the JCRC Policy Reports, Make Australia Fair Again (2017), Super Ideas: Securing Australia's Retirement Income System (2018), Rental Nation: A Plan for Secure Housing in Australia (2020), Private Gain, Public Pain: Taking Stock of Privatisation, Outsourcing and Contracting Public Sector Work (2021) and coauthored (with Adam Slonim) Artificial Intelligence and the Future of Work (2019), as well as (with Dominic Meagher) the JCRC's July 2025 discussion paper, Innovation Nation: Boosting Productivity and the Fair Go in the Age of Artificial Intelligence, and (With Sean Whitworth and Dominic Meagher) For All of Us: Making Artificial Intelligence Work for Working People (2026).

Sean Whitworth is a researcher and policy contributor to the John Curtin Research Centre whose work focuses on the political economy of technology, industrial policy, and Australia's economic sovereignty. Sean co-authored (with Nick Dyrenfurth and Dominic Meagher) the JCRC Policy Report, For All of Us: Making Artificial Intelligence Work for Working People (2026) and contributed to the JCRC policy report Safer, Stronger, Sovereign: Preparing Australia for Climate Disasters.

Dr Dominic Meagher is an economist and public policy specialist at the Australian National University, where he is a Research Fellow at the Centre for Climate & Energy Policy in the Resources, Environment & Development Department at the Crawford School of Public Policy. He is also a non-resident fellow at the JCRC, where he served as the centre's Deputy Director and Chief Economist from 2021 to 2025.

Nick, Sean and Dominic wish to thank the following who contributed to researching this report or read drafts: Lachlan Kerwood-McCall, Gerard Dwyer, and Jamileh Hargreaves.

Defending Australia: A National Artificial Intelligence Sovereign Capability and Security Agenda for the 21st Century, John Curtin Research Centre Report no. 14, June 2026.

ISBN: 978-1-7635462-2-6

Foreword

Gerard Dwyer, National Secretary,

Shop Distributive and Allied Employees Association

Artificial intelligence is reshaping the world of work at a pace not seen in generations. It is changing how jobs are organised, how decisions are made, and where power sits in our economy. As this report shows, AI is no longer a technological upgrade — it is becoming the operating system of modern capitalism, embedded in the systems that determine employment, wages, productivity and control.

Defending Australia: A National Artificial Intelligence Sovereign Capability and Security Agenda for the 21st Century highlights a central proposition: Australia is increasingly dependent on foreign-governed AI systems that underpin defence, critical infrastructure and the democratic information environment. Access to key technologies and data is now shaped by decisions made outside Australia, creating risks that traditional policy or procurement cannot mitigate.

This is not just a question of national security or economic policy — it is a question of whether working people have a meaningful say in the technologies that shape their livelihoods.

Left unchecked, the current trajectory of AI deployment risks repeating patterns workers have seen before: offshoring capability, eroding secure jobs and concentrating economic power in a small number of global corporations. Across sectors such as banking and telecommunications, AI has already been accompanied by outsourcing, deskilling and the hollowing out of domestic expertise. When the knowledge needed to run core systems sits offshore, accountability becomes weaker and Australian workers lose influence over how those systems operate and affect them.

Sovereignty in the AI era is not an abstract concept — it is about control over the systems that shape economic life and about ensuring that those systems operate in the interests of the Australian community. That cannot be achieved without placing Australian workers, and the organisations that represent them, at the centre of the solution. Government, business and unions must work together to shape how AI serves the Australian community.

A key strength of this report is its recognition that sovereign capability is built not only through infrastructure but through people. Compute, cloud platforms and data systems are only sovereign when Australian workers have the skills, authority and institutional backing to govern them. The question is not whether Australians can use AI, but whether we can understand it, challenge it, adapt it and control it in the interests of the nation and working people.

That is why the report places such emphasis on industrial policy and workforce development. A sovereign AI future must create secure, high-quality jobs, rebuild domestic capability and ensure the benefits of technological change are broadly shared. It must reject models of AI deployment that replace labour, strip out expertise and leave workers exposed to opaque, unaccountable systems.

The report also rightly highlights the role of trade unions as essential institutions of democratic and economic resilience. Workers are often the first to see how AI systems operate in practice — where they fail, where they create risk and where they undermine fairness. The ability of working people to organise, bargain

collectively and assert their rights is fundamental to this agenda.

The proposed National AI Sovereignty Compact reflects this reality. By bringing government, industry and unions together on equal footing, it recognises that effective governance requires worker knowledge, enforceable rights and genuine participation in decisions about technology deployment. Without this, sovereignty risks becoming a technical concept rather than a lived reality in workplaces.

Managing the economic transition is critical. AI will transform industries and displace jobs if left unchecked. A responsible approach requires enforceable standards:

- Transparency and explanation for automated decisions affecting employment
- Consultation before major technological change
- Guaranteed retraining and redeployment pathways
- Industrial frameworks ensuring workers share in productivity gains

These are the foundations of a sustainable and legitimate transition.

Australia faces a strategic choice. We can remain passive consumers of foreign-controlled technologies, or we can take a different path — investing in sovereign capability, rebuilding domestic expertise, and ensuring that the next wave of technological change strengthens, rather than weakens, the position of working Australians.

This report sets out a credible pathway. Sovereignty in the AI era will not be delivered by markets alone; it requires coordinated action across government, industry and the labour movement, backed by sustained investment and clear policy direction.

The central message is simple: a sovereign AI future is not about where systems are hosted or who owns infrastructure, but about who controls AI in the economy, in workplaces and in society. Ensuring working people have a real voice in that future is the precondition for getting it right. Getting it right will mean economic structural independence, enhanced national capability and an AI future that serves the Australian people—not the other way around.

Contents

Executive Summary

Artificial intelligence (AI) has become a defining instrument of state power, the technological foundation of national security, critical infrastructure, economic coordination, and democratic life. Sovereignty in the AI era is therefore determined not only by territorial control or military capability, but by whether a nation can design, host, govern, and secure the AI systems that underpin its essential functions.

Australia is structurally exposed to this transformation. Critical frontier AI capabilities – from foundation models and compute infrastructure to cloud platforms and data pipelines – are concentrated in a small number of foreign firms ultimately governed by foreign political and regulatory systems. The April 2026 Anthropic–Trump dispute demonstrated this with unusual clarity: when Anthropic refused to grant unfettered access to its frontier models, the US Government designated the company a supply chain risk and ordered all federal agencies to cease using its systems. The episode revealed a structural truth: access to frontier AI can be revoked, restricted, or repriced by foreign political decision, with no recourse for dependent states.

For Australia, the implications span every domain of national life. Defence intelligence, autonomous systems, logistics, and cyber defence increasingly rely on foreign governed AI. Critical infrastructure, such as energy, water, transport, food supply chains, telecommunications, finance, emergency services and health, depend on cloud platforms subject to foreign legal jurisdictions. The democratic information environment is mediated by foreign platforms whose algorithms shape public discourse and institutional trust. In each of these domains, operational continuity and institutional integrity are dependent on systems Australia does not own, cannot fully audit, and may not be able to secure.

This report argues that Australia must treat AI as a sovereign capability challenge grounded in industrial strategy, not a technology policy issue. The question is not whether AI is beneficial or risky; it is whether Australia can build the industrial foundations required to operate AI systems under Australian authority and law.

The analysis is organised around three pillars:

- **Pillar 1: National Security and Defence Sovereignty** – AI is now a core component of defence capability. Intelligence, surveillance and reconnaissance (ISR), autonomous systems, logistics, cyber defence, and decision support depend on models and compute governed offshore. Sovereign hosting, model assurance, and domestic integration capacity are essential for operational continuity in a crisis.
- **Pillar 2: Critical Infrastructure Resilience** – Essential services increasingly rely on AI systems embedded in foreign cloud platforms vulnerable to foreign legal claims, opaque supply chains, and foreign political decisions. Infrastructure resilience requires sovereign cloud and compute, sovereign data governance, and domestic systems integration capacity.
- **Pillar 3: Democratic Sovereignty and Social Cohesion** – AI mediates public discourse, political communication, and institutional legitimacy. Synthetic media, personalised manipulation, and opaque platform algorithms create vulnerabilities that foreign governed systems are neither incentivised nor required to address.

Across all three pillars the conclusion is the same, sovereign AI capability is built on sovereign industrial capability. Compute, cloud, data governance, model assurance, systems integration,

and workforce pipelines are the industrial foundations of national resilience.

To address these challenges, the report proposes ten recommendations spanning a National AI Capability and Security Strategy; a National AI Sovereignty Compact; mandatory data sovereignty requirements, sovereign cloud and compute, and a national AI assurance framework; strategic procurement reform; a national workforce pipeline; platform governance, public sector accountability, and an AI risk register; AUKUS AI capability integration; an international governance engagement strategy; a sovereign AI research and development (R&D) program; and an AI economic transition and labour compact.

A coordination architecture assigns clear responsibilities to the Department of the Prime Minister and Cabinet (PM&C), the Australian Signals Directorate (ASD), the Department of Finance (Finance), the Department of Industry, Science and Resources (DISR), the Department of Defence (Defence), the Department of Home Affairs (Home Affairs), the Attorney-General's Department (AGD), the Department of Foreign Affairs and Trade (DFAT), the Department of Education (Education), the Department of Employment and Workplace Relations (DEWR), the Australian Communications and Media Authority (ACMA), and the Australian Electoral Commission (AEC), ensuring that sovereign AI capability is developed, governed, and deployed in a way that is strategically aligned, operationally secure, and democratically accountable.

Australia stands at a defining moment in its strategic autonomy. The decisions made in the coming years will determine whether we remain a downstream consumer of foreign AI systems or build the sovereign capability required to defend our interests, strengthen our institutions, and secure our democratic future.

Introduction

AI, Sovereignty, and the New Strategic Order

Artificial intelligence has become a defining instrument of state power. No longer a discrete technology or sector, AI now functions as the technological underpinning of national security, economic coordination, critical infrastructure, and democratic life. In this environment, sovereignty is increasingly determined by control over the industrial systems that produce, host, and govern AI, not simply by who uses it.

For Australia, this transformation is unfolding amid intensifying geopolitical competition. The United States and China are openly instrumentalising AI as tools of national security strategy and industrial ambition. Their decisions about access, governance, and permissible use of frontier AI reflect their own political priorities. Dependent states have no influence over these decisions, yet their security architectures and economies are increasingly built on the systems they govern.

The clearest demonstration of this came in April 2026, when the US Department of Defense sought unrestricted access to Anthropic's frontier models for applications including mass domestic surveillance and autonomous weapons development. Anthropic refused. The U.S. Government designated the company a supply chain risk and ordered all federal agencies to cease using its systems immediately. Firms that complied, including OpenAI and Google, retained access to federal integration pathways. The one that resisted was excluded overnight. The episode revealed a structural truth: access to frontier AI is ultimately contingent on the political decisions of foreign governments, and companies that resist those decisions can be removed from the ecosystem without warning.

This report describes Australia's current position as a cloud hostage condition: critical systems are physically hosted, legally governed, and operationally controlled outside Australian jurisdiction, meaning that even services which appear domestically integrated remain subject to foreign political authority. This is not a failure of procurement or policy; it is the predictable consequence of building sovereign functions on foreign governed infrastructure. Cloud platforms, frontier AI models, data architectures, and the integration layers connecting them are overwhelmingly concentrated in foreign firms whose legal obligations run to their home governments, not to Australia.

This report argues that Australia must treat AI as a sovereign capability challenge grounded in industrial strategy. The question is not whether AI is beneficial or risky; it is whether Australia can build and maintain the industrial foundations – compute, cloud, data governance, model assurance, systems integration, and workforce capability – required to operate AI systems under Australian authority and law. Sovereign capability in AI is not about autarky or isolation. It is about ensuring that the systems underpinning national security, essential services, and democratic governance are built on a domestic industrial base capable of sustaining them.

The analysis that follows is organised around three domains where AI now shapes Australia's sovereignty: national security and defence, critical infrastructure, and democratic integrity. Each presents a distinct mechanism of exposure, and a corresponding set of industrial capability actions government can take to build sovereign capacity and reduce strategic dependency.

Australia stands at a strategic inflection point. The decisions made in the next five years will determine whether we remain a downstream consumer of foreign AI systems or build the sovereign

industrial capability required to defend our interests, strengthen our institutions, and secure our democratic future.

Artificial intelligence is rapidly redefining national security and defence capability. No longer limited to discrete platforms or support functions, it now operates as a foundational layer across military systems, intelligence architectures, cyber operations, and critical infrastructure. AI enables largescale data processing, faster decision-making and the integration of complex operational environments, reshaping the character of conflict and the ways states organise and project power.

This evolution is occurring within a broader shift in the strategic environment. Contemporary competition increasingly unfolds below the threshold of conventional conflict, through cyber operations, economic coercion, technological dominance, and information manipulation, all of which AI accelerates. It enables more precise cyber intrusions, strengthens intelligence analysis, supports autonomous systems, and facilitates sophisticated influence operations. As a result, the boundary between civilian and military domains is eroding, with critical infrastructure and commercial data ecosystems now central to national defence.

In this context, the ability to develop, deploy and secure AI systems is becoming a decisive source of strategic advantage. States that control advanced AI capabilities and the data, infrastructure, and talent behind them, are better positioned to defend their systems, deter adversaries and shape emerging technological norms. Those reliant on externally controlled systems face constraints on autonomy and heightened exposure to disruption, coercion, and compromise.

For Australia, this presents a dual challenge. AI can act as a force multiplier, extending operational reach, improving logistics, strengthening intelligence, and enhancing responses to both conventional and nonconventional threats. Yet Australia's dependence on foreign owned AI systems, cloud infrastructure and data architectures introduces systemic vulnerabilities across defence and critical infrastructure, raising fundamental questions about control, resilience, and the ability to operate independently in a contested environment.

Pillar One: AI, National Security and Defence Sovereignty

1.1 The Mechanism of Sovereignty Risk

The April 2026 Anthropic-Trump episode described in the Introduction confirmed a structural reality that extends well beyond a single company or a single directive. Access to frontier AI is ultimately governed by the political decisions of the jurisdictions in which those systems are developed and hosted.[1] Commercial agreements, technical safeguards and corporate policies offer no protection when national security powers are invoked. Foreign hosted AI systems may be subject to extraterritorial legislation like the US Clarifying Lawful Overseas Use of Data Act (CLOUD Act) or China's National Security Law, enabling overseas authorities to compel access to data stored or processed through foreign owned infrastructure.[2] For Australia whose defence, infrastructure and public sector systems increasingly rely on foreign governed AI, access can be revoked, restricted or repriced without notice or recourse. Not because of technical failure but because of political decisions made elsewhere.

The mechanism of exposure operates through three channels:

- **Political authority** – Foreign governments can compel AI companies to provide access, modify systems, or restrict availability under national security legislation. These powers apply extraterritorially and can override commercial agreements or technical safeguards.

- **Commercial concentration** – Frontier AI capability is concentrated in a small number of firms who's strategic and financial priorities may not align with Australia's interests. Their decisions about access, pricing, and product direction can reshape entire capability systems.

- **Infrastructure control** – Cloud platforms, compute infrastructure, and model hosting environments are governed by foreign legal frameworks, even when physically located in Australia. This creates unavoidable exposure for defence, intelligence, and critical infrastructure data.

Together, these channels create a form of strategic dependency qualitatively different from previous eras of defence procurement. When the systems that underpin defence capability are governed by external political authorities, Australia's freedom to operate is constrained by decisions made elsewhere.

1.2 AI as the Strategic Engine of Defence Capability

AI is transforming defence from an asset centric model to an integrated system-of-systems architecture in which operational advantage depends on the ability to sense, analyse, decide and act faster than an adversary. Modern military power is increasingly defined not by individual assets but by the systems that connect them. Modern warfare unfolds in a fast moving, information saturated battlespace, with threats emerging simultaneously across maritime, air, land, cyber and space domains. In this environment, the decisive advantage lies in the ability to integrate information, detect emerging threats and coordinate responses across multiple systems in real time. AI enables this integration, fusing data from diverse sources, predicting the behaviour of adversaries, and supporting commanders with actionable decision options at machine speed.

The implications are visible across the core functions of Australian defence capability. Intelligence, surveillance and reconnaissance (ISR), which underpins Australia's ability to monitor

activity across the Indo-Pacific, increasingly depends on machine learning systems capable of processing data at scales beyond human capacity. AI systems can analyse satellite imagery to detect changes in infrastructure or military deployments, identify irregular maritime activity, process signals intelligence and integrate disparate sensor inputs into a coherent operational picture.[3] Australia's geography and maritime dependence make these capabilities particularly important for maintaining situational awareness and supporting deterrence.

Autonomous and semiautonomous systems are similarly becoming central to defence operations, particularly through uncrewed aerial, surface and underwater systems that extend operational reach, reduce risk to personnel and enable persistent surveillance in contested environments. These systems rely on AI for navigation, target recognition, threat detection, mission execution and coordination with other platforms.[4] Likewise, logistics increasingly depend on AI enabled optimisation systems capable of forecasting maintenance requirements, identifying supply bottlenecks, allocating fuel and munitions, and modelling disruptions to readiness across dispersed theatres of operation.[5] In a country defined by strategic distance and long supply chains, such capabilities are vital to maintaining operational resilience, and they depend on continuous monitoring, anomaly detection and rapid isolation of disruptions before they cascade across networks.

Cyber defence further illustrates the extent to which AI is reshaping military capability. Actors are already employing AI to automate vulnerability discovery, generate sophisticated phishing campaigns and scale intrusion attempts. Defensive systems must therefore operate at machine speed to detect anomalous behaviour, identify malicious code and support automated containment and remediation. These risks are amplified by increasingly sophisticated statebased cyber and intelligence operations. Russian and Chinese cyber espionage networks have repeatedly penetrated government systems, critical infrastructure and private organisations for intelligence collection and strategic positioning.[6] North Korea and Iran employ cyber enabled theft, disruption and asymmetric tactics to offset conventional capability gaps.[7] AI accelerates this by analysing large datasets, generating targeted intrusion campaigns and enabling rapid exploitation across interconnected systems. As Australia's digital infrastructure is deeply integrated with foreign technologies, opaque supply chains and limited visibility over system vulnerabilities compound these exposures.

AI is also transforming command and control. Decision support systems now integrate information across domains, model potential courses of action and compress decision timelines by providing commanders with realtime situational awareness and operational recommendations.[8] In this environment, cyber resilience is no longer a supporting function but a core determinant of operational viability.

Across these domains, capability is increasingly determined by the quality of data, model performance, compute availability, network resilience and the ability to continuously update systems. Traditional defence acquisition models, built around long asset life cycles and periodic upgrades, are therefore being disrupted by a technological environment in which capability depends on continuous software iteration and model adaptation. This also means that human oversight, legal compliance and ethical accountability must be embedded into system design, particularly where AI informs time sensitive or lethal decisions.[9]

This transformation introduces a new category of sovereignty risk. When critical defence functions depend on foreign governed AI, Australia becomes dependent on foreign update cycles, foreign security practices and foreign political decisions. ISR, autonomous navigation, logistics optimisation, cyber defence and command support tools all require models that are reliable, auditable and continuously available under crisis conditions,

yet reliance on externally governed systems means access could be delayed, degraded or restricted at precisely the moment it is most essential. Australia may also be unable to verify whether models have been altered or compromised, since AI capability depends on continuous retraining and cloud infrastructure governed beyond Australia's control. Sovereignty in the AI era therefore requires more than platform ownership; it requires control over the systems that define military capability itself, sovereign compute, sovereign model development, secure integration with Australian manufactured hardware and sensors, and a domestic workforce capable of designing, validating and securing these systems.

1.3 Strategic Vulnerabilities in Foreign Governed AI

The integration of AI across defence capability creates a new category of strategic vulnerability. These vulnerabilities are not discrete; they interact and compound, producing systemic exposure across the defence organisation. Geopolitical competition is driving a structural divergence in global technology networks, with AI models, semiconductor supply chains, cloud infrastructure and data governance frameworks increasingly shaped by competing blocs with distinct political and legal priorities.[10] In this environment, Australia's reliance on foreign governed AI systems exposes defence capability to geopolitical, legal, operational and technical risks that cannot be mitigated through traditional procurement, alliance structures or risk management frameworks.

Australia faces geopolitical exposure because frontier AI capability is concentrated in jurisdictions whose political priorities may diverge from our own. Export controls, national security directives or political decisions in the US or China can restrict access to models, compute or cloud infrastructure with immediate operational consequences. These decisions are made according to their national interests, not Australia's, and in a deteriorating strategic environment access to critical AI capability may be curtailed without warning. Foreign governments can compel access to data or model outputs processed through foreign owned infrastructure, even when hosted in Australia, creating a structural exposure that cannot be mitigated through technical measures alone. For defence, intelligence, and critical infrastructure data, this represents an unavoidable sovereignty risk.

Australia also faces supply chain opacity. Frontier AI systems are built on globally dispersed supply chains consisting of closed training data, opaque model design, and complex integration pathways. Australia has limited visibility into how these systems are trained, secured or updated. The opacity of frontier AI systems increases this risk, with models operating as 'black boxes' with limited transparency regarding training data, decision processes or embedded vulnerabilities. This means Australia cannot independently verify their integrity or detect manipulation.[11] This makes it difficult to assess risk, verify compliance with Australian law or ensure that systems have not been compromised.

These structural dependencies create operational vulnerability. If access to frontier models or cloud platforms is restricted, whether through political decision, commercial reprioritisation or geopolitical crisis, Australia may lose access to capabilities that defence operations have come to rely on. This dependence undermines operational continuity and strategic autonomy, particularly in fast paced or contested environments where access to compute, data, and model updates is vital.

Finally, foreign governed AI systems introduce heightened risk of exploitation by adversaries. When Australia cannot independently audit or verify the models it relies on, it becomes more vulnerable to data poisoning, model manipulation or targeted exploitation. These vulnerabilities may be exploited to degrade capability, manipulate outputs or compromise systems in ways that are difficult to detect or attribute. The inability to assure the integrity of foreign governed models represents a direct sovereignty risk.

Australia's strategic environment demands a defence enterprise capable of operating independently, securely and reliably under conditions of geopolitical tension, contested access and technological disruption. As AI becomes the analytical engine of defence capability, sovereign control over the systems that enable sensing, analysis, decision-making and operational coordination becomes essential. Sovereign capability in AI is not about autarky or technological isolation; it is about ensuring that the systems underpinning national security operate under Australian authority, within Australian law and in alignment with Australia's strategic interests.

Meeting this challenge requires a fundamental shift in how Australia conceptualises defence preparedness. It is no longer sufficient to procure platforms, integrate foreign systems or rely on alliance-based access to critical technologies. The foundations of capability now lie in compute infrastructure, data governance, model assurance, integration capacity and workforce pipelines. These are industrial base questions, not niche technology questions. They require long-term investment, institutional coordination and a clear understanding of the mechanisms through which foreign dependency creates strategic vulnerability.

The following section sets out the components of sovereign defence AI capability and the industrial foundations required to support them.

1.4 Building Sovereign Defence AI Capability

Sovereign capability in defence AI does not require Australia to build frontier models from scratch. It requires the ability to host, govern, audit and adapt AI systems in ways that ensure operational continuity, security and alignment with national interests. This capability rests on the following industrial foundations.

Sovereign compute and hosting is the first and most fundamental. Defence requires the ability to train, fine-tune, host and operate AI models within infrastructure that is physically located in Australia, governed by Australian law and secured to defence grade standards. Without it, Australia remains dependent on foreign cloud platforms whose availability, pricing and security conditions are determined by external actors. Sovereign compute is essential for hosting operational models, processing classified data, conducting model assurance and adversarial testing, and maintaining continuity if foreign access is restricted. It requires investment in secure data centres, high-performance computing clusters and specialised hardware, alongside governance frameworks that insulate these facilities from foreign legal claims.

Sovereign data governance in the defence context requires the most rigorous application of data sovereignty principles. Defence sensitive data, ISR outputs, autonomous-system data, logistics and supply chain data, command-and-control records, and model training datasets collected from classified operations, must be stored and processed exclusively within Australian governed infrastructure, subject to classification level controls, need-to-know access frameworks and operational security requirements that exceed civilian critical infrastructure standards. This is not simply a privacy or compliance requirement, it is an operational necessity. Data that cannot be governed under Australian authority cannot underpin systems that operate under Australian command. These defence data governance requirements inform the broader tiered data sovereignty classification architecture developed in Pillar Two, where the critical infrastructure baseline extends and generalises from the defence context, but the defence floor is higher, and the consequences of breach are more immediate.

Model assurance and verification is another foundation. AI systems used in defence must be reliable, secure and predictable under all conditions, free from vulnerabilities or backdoors, resilient to manipulation, and compliant with Australian legal and ethical standards. This requires access to model weights, the capacity to conduct

adversarial testing, and specialised expertise in machine learning security and red-teaming. As previously established, foreign governed models often operate as opaque systems whose training data, update cycles and internal mechanisms cannot be independently verified, creating unacceptable risk in defence contexts. Continuous assurance, supported by institutional capacity within Defence, ASD and DISR, is therefore essential.

The National Model Assurance Facility is the institutional vehicle through which sovereign model assurance is operationalised across all three pillars of this report. It would not be a regulatory body, but a standing technical centre capable of conducting the work that Defence, regulators and operators cannot perform independently: adversarial testing, continuous verification of model behaviour, training data auditing, red-team evaluation under classified conditions, and production of legally admissible assurance findings. In the defence context specifically, it is the mechanism through which Australia verifies that AI systems underpinning ISR, autonomous platforms, logistics, cyber defence and command-and-control behave as intended and remain reliable when foreign governed alternatives are unavailable or compromised. The facility operates as a continuous verification infrastructure rather than a one-off certification checkpoint, updating its testing methodologies as models are retrained, updated and exposed to new operational conditions. It is vital that it is sovereign, as a model assurance facility that relies on foreign governed compute, foreign tools or foreign personnel cannot provide the independent verification that sovereign assurance requires.

Defence specific model development is the next foundation, addressing areas where foreign governed systems pose unacceptable risk. These models do not need to match frontier capabilities, they just need to be reliable, secure and fit for purpose across ISR analysis, autonomous-system navigation, logistics optimisation, cyber defence automation and command-and-control decision support. Developing them requires sovereign compute, secure data pipelines and close collaboration between Defence, industry and research institutions. The strategic advantage being that Australia governs how models are trained, updated and deployed; models can be hardened and tested under classified conditions; and capability is maintained even if access to foreign systems is restricted.

Integration and systems engineering is the fifth foundation. Sovereign capability requires more than just models and compute, it requires the ability to integrate AI systems with defence platforms, sensors and networks. Integration is where AI becomes operational capability, and it is where foreign dependency is most acute. Reliance on foreign integrators risks embedding vulnerabilities deep within defence architecture. Sovereign integration capacity ensures Australia can adapt AI systems to local platforms, secure data flows across domains and update systems without foreign approval. It cannot be imported, it requires long-term investment in domestic firms, workforce development and institutional partnerships.

Workforce and institutional capacity are the final foundation. Sovereign defence AI capability ultimately depends on machine-learning engineers, systems engineers, cybersecurity specialists, data governance experts and AI assurance specialists. These roles then provide the institutional capacity needed within Defence, DISR and ASD to govern, procure and oversee AI systems effectively. Education and DEWR play a crucial enabling role in shaping national skills pipelines, vocational and tertiary pathways, as well as workforce development programs. A central institutional vehicle for this effort would be a National AI Skills Academy, bringing together organised labour, research institutions and defence and critical infrastructure operators to embed unions within training design and delivery. This tripartite model ensures programs reflect the conditions in which workers actually operate AI systems, facilitates ongoing feedback between workplaces and training providers, and produces a workforce

pipeline that is both technically capable and institutionally grounded. Without this workforce and institutional capability, sovereign compute and model assurance cannot be operationalised. The workforce dimension of sovereign defence AI capability is not simply a supply side question of producing enough engineers. It is a sovereignty question. When the technical knowledge required to understand, audit and adapt a defence AI system resides with foreign providers rather than in an Australian institution, operational sovereignty is titular rather than legitimate. A Defence AI system is sovereign only if Australian personnel can govern it, not simply operate it. This distinction, between operation and governance, between using a system and understanding it well enough to verify, adapt and secure it, is the human capability aspect of the broader sovereignty principle. It is why the workforce foundations identified in this report are industrial base requirements and why union participation in training design, through the National AI Skills Academy, is a governance mechanism rather than a stakeholder concession.

1.5 AUKUS AI Capability Integration

AUKUS offers Australia its most significant near-term pathway in closing the sovereign defence AI capability gap. The partnership's Pillar II advanced capabilities stream provides a structured framework through which Australia can access allied technical expertise, co-develop sovereign defence AI systems and build the domestic industrial and workforce foundations required to sustain them. But realising these benefits requires intentional strategy. AUKUS integration is a mechanism for accelerating sovereign capability, not a substitute for it. In the absence of domestic compute, assurance infrastructure and a trained workforce, allied technology transfers risk deepening operational dependence rather than building strategic independence.

An immediate priority is aligning Australia's sovereign AI requirements with Pillar II's collaborative development pipeline, identifying the defence AI systems across ISR analytics, autonomous platform navigation, logistics optimisation, cyber defence automation and decision support where co-development can accelerate domestic capability while guaranteeing genuine Australian ownership of intellectual property, training data and assurance rights. Co-development must embed Australian organisations at the model development layer, not as downstream integrators. Technology transfer that places Australian firms inside allied AI development programs, rather than receiving finished systems, is the only form of cooperation that builds lasting sovereign capability.

Interoperability standards are another crucial priority. As autonomous and AI enabled systems are deployed alongside allied platforms, Australia must be able to verify the integrity of shared systems, audit their behaviour under Australian legal standards and maintain operational independence if allied access is disrupted. This requires shared assurance protocols within AUKUS, covering agreed frameworks for model verification, adversarial testing and data sovereignty protections across all Pillar II AI systems. Australia should drive the design of these protocols rather than accepting allied standards passively, ensuring that interoperability does not become a means for the same external dependency that sovereign capability is designed to avoid.

The Five Eyes intelligence sharing architecture presents a parallel opportunity. AI enabled ISR systems increasingly depend on shared computational infrastructure and model training pipelines. Australia must negotiate access arrangements that preserves sovereign control over Australian data, ensures model outputs affecting Australian operations are auditable under Australian law, and prevents allied cloud or compute dependencies from becoming a sovereignty risk.

AUKUS also positions Australia as a potential regional provider of AI governance and assurance

capability. Democratic partners across the Indo-Pacific, seeking alternatives to both US commercial dependence and Chinese platform risks, represent a genuine market for Australia's emerging forensic AI audit capability, sovereign cloud infrastructure and governance aligned AI systems. By anchoring its AUKUS investments to exportable governance standards and interoperable assurance frameworks, Australia can convert the sovereignty agenda outlined in this report into a source of regional strategic influence, establishing itself not merely as a capable AUKUS partner but as the trusted AI governance anchor of the democratic Indo-Pacific.

1.6 Implementation Pathway

Roles and responsibilities are assigned as follows:

Defence defines capability requirements, mandates sovereign hosting and model assurance across ISR, autonomous platforms, logistics, cyber defence and command-and-control, and aligns procurement with sovereign capability requirements. Defence also leads development of defence specific models in priority domains where foreign governed systems pose unacceptable risk.

ASD establishes verification and threat-resilience standards, develops AI specific cybersecurity frameworks, leads the national AI red-team capability, and develops the AI security expertise required to participate in allied red-teaming and threat testing programs under AUKUS Pillar II.

DISR develops sovereign compute, establishes the national model assurance facility, builds domestic integration capacity through long-term partnerships with Australian firms, and coordinates workforce pipelines with universities, Technical and Further Education institutions (TAFEs) and industry. DISR must also align sovereign compute and model-assurance investments with the technical requirements of Pillar II programs.

Finance embeds sovereign capability weighting into defence procurement, mandates open standards and auditability rights, supports joint procurement mechanisms, and ensures contestability through independent evaluation mechanisms and challenge-based procurement for emerging AI capabilities.

Home Affairs aligns sovereign AI with national cybersecurity and critical infrastructure obligations under the SOCI Act, leads supply chain assurance and foreign interference risk assessment, and ensures sovereign AI infrastructure is integrated into national resilience planning.

AGD develops the legal architecture for the national autonomous systems assurance regime, including human authorisation requirements for lethal functions, testing aligned with international legal obligations and certification standards, and a liability framework allocating responsibility across developers, integrators and operational commanders.

DFAT leads Australia's engagement via Five Eyes AI workstreams, negotiates access arrangements that preserve sovereign control over Australian sourced data, and ensures model outputs affecting Australian operations are auditable under Australian law.

Education and DEWR develop workforce pathways in defence AI, cybersecurity and systems engineering, and coordinate with Defence and DISR to ensure training pipelines meet operational requirements, including clearance ready pathways and the National AI Skills Academy.

PM&C provides the strategic spine: coordinating whole-of-government sovereign AI investment decisions, driving alignment across Defence, ASD, DISR, Finance, Home Affairs, Education and DEWR, maintaining a

cross-portfolio register of AI dependencies, vulnerabilities and emerging threats, and ensuring sovereign capability is treated as a national power requirement rather than a portfolio specific initiative.

Implementation proceeds across three phases:

Phase 1 (0–12 months): Foundations

- Defence mandates sovereign hosting requirements for all new AI systems.
- ASD drafts high-risk AI classification criteria and continuous-verification standards.
- Finance introduces sovereign capability weighting into defence procurement.
- DISR establishes the interim secretariat for whole-of-government sovereign AI coordination.
- Home Affairs aligns sovereign AI requirements with SOCI Act obligations.
- AGD scopes the legal architecture for a national autonomous systems assurance regime, including human authorisation requirements for lethal functions and a liability framework allocating responsibility across developers, integrators and operational commanders.
- DFAT initiates a dedicated AI sovereignty workstream within the Five Eyes governance structure, aligned with AUKUS Pillar II standards, to negotiate access arrangements that preserve sovereign control over Australian-sourced data.
- PM&C establishes whole-of-government coordination mechanisms and commissions the design of a cross-portfolio AI dependency and vulnerability register, defining scope, escalation protocols and agency responsibilities.
- Defence, ASD and DISR establish an AUKUS Pillar II alignment framework to map sovereign AI requirements against allied development pipelines.

Phase 2 (1–3 years): Capability Build

- Defence begins phased migration of high-risk workloads into sovereign compute environments and embeds Australian personnel inside AUKUS Pillar II AI development programs at the model development layer, securing sovereign access to training data, model weights and assurance rights.
- Defence commences development of defence-specific models in priority domains – ISR analytics, logistics optimisation, cyber-defence automation and decision support – where foreign governed systems pose unacceptable risk.
- ASD operationalises defence AI assurance, including red-team testing and periodic recertification.
- DISR commences sovereign cloud and compute procurement and establishes the national model assurance facility, with institutional capacity to conduct adversarial testing, continuous verification and model behaviour auditing.
- Education and DEWR launch targeted defence AI and cyber workforce pathways, including the National AI Skills Academy, and a national stackable micro-credential framework designed for portability, regional access and progressive skill development.
- Home Affairs integrates sovereign AI requirements into national cyber and critical infrastructure frameworks.
- DFAT operationalises the Five Eyes AI sovereignty workstream, negotiating data access arrangements and ensuring model outputs affecting Australian operations are auditable under Australian law.

- AGD, Defence and ASD develop and pilot the national autonomous-systems assurance regime, including IHL-aligned testing and certification standards.

- PM&C drives cross-government alignment on sovereign capability standards and investment decisions and operationalises the cross-portfolio AI dependency and vulnerability register.

Phase 3 (3–7 years): Scale-Up

- DISR expands sovereign compute capacity to meet long-term defence demand.

- Defence embeds sovereign assurance standards across all strategic procurement.

- Defence, AGD and ASD operationalise the national autonomous systems assurance regime across all deployed autonomous and semi-autonomous platforms.

- Australia operationalises AUKUS aligned assurance protocols and interoperable AI standards, supporting export-ready governance aligned AI capability across the Indo-Pacific.

- PM&C ensures whole-of-government alignment, Cabinet level oversight and continuous operation of the cross-portfolio AI dependency and vulnerability register.

1.7 Implementation Risks and Mitigation

The transition to sovereign defence AI capability will encounter structural risks across compute, supply chains, procurement, workforce and governance. These risks are not isolated and will interact and compound across the defence enterprise. Mitigation requires coordinated action across Defence, ASD, DISR, Finance, Home Affairs, Education, DEWR and PM&C.

Risk	Mitigation
Compute shortfall – Australia lacks sovereign compute capacity to host and assure defence grade AI at scale.	DISR led phased procurement of sovereign compute; allied secure cloud arrangements as interim measures; long-term demand signals to enable domestic providers to scale.
Procurement lock-in – sovereign capability preferences may entrench incumbents or create proprietary dependencies, undermining contestability.	Finance mandates open standards, access to model weights, independent evaluation mechanisms, and challenge-based procurement for emerging AI capabilities.
Workforce shortage – insufficient cleared AI engineers, systems integrators, cybersecurity specialists and model-assurance experts to sustain sovereign capability.	Education and DEWR expand tertiary and vocational pathways, targeted scholarships, skilled-migration programs, and clearance ready training pipelines aligned with Defence and ASD requirements.
Assurance capacity gap – demand for continuous verification, adversarial testing and AI specific cybersecurity will exceed current institutional capability.	ASD expands AI security expertise; DISR develops sovereign assurance infrastructure; Home Affairs integrates AI specific requirements into critical infrastructure regulation.

Risk	Mitigation
Supply chain opacity – limited visibility into vulnerabilities embedded in globally distributed hardware, proprietary model architectures and opaque update cycles.	Home Affairs, ASD and DISR jointly implement AI specific supply chain assurance requirements and foreign interference risk assessments, supported by sovereign hosting mandates.
Fragmented governance – without central coordination, sovereign capability efforts may diverge across Defence, ASD, DISR, Finance, Home Affairs, Education and DEWR.	PM&C coordinates whole-of-government sovereign AI capability, enforces alignment across portfolios, and maintains a cross-portfolio register tracking AI dependencies, vulnerabilities and emerging threats.

AUKUS interoperability introduces a dual sided risk, while Pillar II cooperation can accelerate sovereign capability, if poorly structured it can entrench new forms of dependency on allied compute, modelassurance pipelines or proprietary architectures. All AUKUS AI collaboration must be tied to sovereign hosting, access to model weights, Australian assurance rights and domestic workforce participation. The objective is interoperability, not harmonisation – allied assurance evidence contributes to without substituting for Australian verification requirements. Interoperability without subordination is the assurance expression of the broader sovereignty principle: engagement with allied capability accelerates Australian capability but does not replace it.

Autonomous systems governance introduces a distinct legal and ethical risk category. As AI enabled autonomous and semiautonomous platforms become operational, Australia must demonstrate compliance with international humanitarian law (IHL) and maintain clear human oversight obligations. Reliance on opaque, foreign governed models for targeting, navigation or threat assessment increases the risk that Australia cannot demonstrate how such systems make decisions, whether they satisfy distinction and proportionality requirements, or who bears responsibility when autonomous systems cause harm. Defence, AGD and ASD should establish a national autonomous systems assurance regime incorporating mandatory human authorisation requirements for lethal functions, IHL aligned testing and certification standards, and a liability architecture that allocates responsibility across developers, integrators and operational commanders.

Pillar Two: AI and Critical Infrastructure Resilience

2.1 AI as a Systemic Infrastructure Risk

AI is transforming critical infrastructure by reshaping how systems operate, coordinate and respond under stress. It now automates decision-making across energy, water, transport and telecommunications networks; enables real-time coordination across geographically dispersed assets; and accelerates both defensive and offensive cyber threat activities. These shifts create tightly linked, data dependent systems whose performance and resilience increasingly rely on machine learning models rather than traditional software.

This transformation introduces a new front of vulnerability. Infrastructure systems that once behaved predictably now rely on machine learning models that infer patterns from historical data, adapt to changing conditions, make predictions under uncertainty and update continuously based on new inputs. This flexibility may enable efficiency and optimisation, but it also introduces new means of failure that are difficult to anticipate, detect or mitigate, particularly when the underlying models are opaque, externally governed or sensitive to data quality, model drift and manipulation.

AI also enables the rise of tightly coupled infrastructure networks. These systems depend on continuous data flows, reliable connectivity, consistent model performance and secure integration across domains. When AI systems are hosted on foreign cloud platforms, Australia becomes dependent on the availability, security and governance of those platforms. A foreign political decision, commercial outage or cyber incident affecting a cloud provider can disrupt essential services across multiple sectors simultaneously.

At the same time, AI is accelerating the speed, scale and sophistication of cyber threats. Adversaries already use AI enabled tools to automate reconnaissance, identify vulnerabilities and execute intrusions at machine speed, overwhelming traditional perimeter-based security models. They can generate highly targeted phishing campaigns, craft malware that adapts to defensive measures, manipulate data streams feeding critical systems and conduct largescale coordinated attacks.[12] Defensive systems must evolve accordingly, relying on AI enabled anomaly detection, behavioural analysis and automated response. Yet these defensive systems themselves become targets, if adverse actors compromise the models underpinning cyber defence, they can blind detection systems, manipulate alerts or trigger false positives that disrupt operations.

Zero-trust architectures, requiring continuous authentication, strict access controls and realtime verification, are one necessary response to AI enhanced threats.[13] But zero-trust alone cannot address the structural vulnerabilities created by foreign governed cloud platforms, extraterritorial data exposure and opaque model update cycles. When the AI systems that underpin critical infrastructure are governed by foreign jurisdictions, Australia cannot guarantee their integrity, availability or alignment with national interests in a crisis.

Addressing AI enhanced infrastructure risk therefore requires sovereign industrial capability, encompassing domestic compute to host and operate critical models; domestic data governance to ensure sensitive operational data remains under Australian authority; domestic model assurance to verify integrity and detect manipulation; and domestic integration capacity to embed secure systems across critical sectors. AI has become a systemic infrastructure risk not because it is unreliable, but because it introduces new dependencies, new means of failure and new

forms of exposure that can only be mitigated through sovereign control of the systems that underpin national resilience.

2.2 The New Vulnerability Surface

AI introduces vulnerabilities across every layer of critical infrastructure, and these vulnerabilities do not exist in isolation. They interact, compound and spread across sectors, creating a systemic exposure that traditional cybersecurity models were never designed to manage.

At the model layer, AI systems fail in ways traditional software never could. Machine learning models can drift as conditions change, degrade as data distributions shift, or generalise incorrectly in ways that are difficult to predict. They are susceptible to data poisoning, where training inputs are manipulated; adversarial prompts, where small disruptions cause misclassification; and subtle behavioural failures that emerge only under operational stress.[14] When these models are governed by foreign jurisdictions, Australia cannot independently verify their robustness or detect hidden vulnerabilities embedded in unclear update cycles.

These risks extend into the data pipelines that feed AI systems. Critical infrastructure increasingly depends on continuous data flows for real-time optimisation and anomaly detection. If adversaries spoof, inject or corrupt data streams, they can cause models to behave unpredictably or make unsafe decisions.[15] Foreign governed cloud platforms amplify this exposure, as operational data may transit through jurisdictions with different legal frameworks, creating opportunities for foreign access, compelled disclosure or interference.[16]

Cloud infrastructure itself is a major point of vulnerability. Most AI systems used in essential services are hosted on foreign cloud platforms subject to extraterritorial legal authority.[17] These platforms are also single points of failure, with outages, cyber incidents or political decisions affecting a cloud provider, that can disrupt multiple sectors simultaneously. When the infrastructure that underpins critical services is concentrated in a handful of foreign controlled environments, Australia inherits their risks.

The supply chain that supports AI systems adds another layer of opacity. Hardware components, software libraries, model weights and training data are sourced globally, often from jurisdictions with varying security standards. Australia has limited visibility into these supply chains, making it difficult to assess sources, detect tampering or ensure compliance with national security requirements.[18]

Finally, AI accelerates the integration of critical infrastructure systems, creating tightly coupled networks where failures cascade rapidly. Energy grids depend on telecommunications networks; water systems depend on energy; transport systems depend on both. AI intensifies these interdependencies by coordinating operations across domains in real-time. A vulnerability in one sector can spread across others, turning a localised failure into a multisector disruption.[19]

Together, these layers form a new vulnerability surface, one defined not by isolated technical weaknesses but by interconnected, AI dependent systems whose behaviour, security and resilience increasingly rely on models, data flows and cloud platforms that Australia does not control. Sovereign capability in compute, data governance, model assurance and integration, is now essential to managing this systemic risk.

These vulnerabilities are amplified by Australia's structural dependence on foreign governed AI systems and cloud infrastructure. This dependence creates a sovereignty problem, where Australia cannot guarantee the availability, integrity or security of the systems that underpin essential services.

The sovereignty problem operates through three mechanisms:

- foreign political decisions can restrict

access to models, compute or cloud services

- foreign legal jurisdictions can compel access to data or model outputs
- foreign commercial priorities can reshape capability through pricing changes, service deprecation or product redesign

Essential services must remain available under all conditions, including geopolitical tension, commercial disruption and cyberattack, yet when the systems that support these services are governed by external authorities, continuity cannot be assured.

AI is not entering Australia's critical infrastructure evenly, but the direction of travel is consistent. Essential services are becoming more automated, more data driven, more interconnected and more dependent on AI enabled optimisation. This creates efficiency and resilience benefits, but it also introduces new vulnerabilities amplified by foreign dependency.

2.3 Sector Specific Dependencies and Vulnerabilities

AI is now embedded across every major sector of Australia's critical infrastructure, shaping how essential services are produced, distributed and managed. While the applications differ, the pattern is consistent. Sectors are becoming more automated, more data driven, more interconnected and more dependent on AI enabled optimisation. This creates efficiency and other benefits, but it also introduces new vulnerabilities that are magnified when the underlying models, cloud platforms or data pipelines are governed externally. The following sectoral dependencies demonstrate how AI driven optimisation can become a point of systemic exposure.

In energy systems, AI is now indispensable for forecasting demand and generation, optimising grid stability, managing distributed energy resources, detecting anomalies in transmission networks and coordinating storage and dispatch.[20] These functions depend on continuous data flows and reliable model performance. Vulnerabilities emerge when optimisation models are hosted on foreign cloud platforms, update cycles are controlled externally, training data includes foreign sourced data or algorithms are opaque and proprietary. A disruption in AI enabled forecasting or balancing systems could trigger instability across the national electricity market, particularly during heatwaves, storms or peak demand periods.[21]

Water and wastewater systems rely on AI for leak detection, pump optimisation, chemical dosing, predictive maintenance and contamination monitoring.[22] These systems depend on dense sensor networks and machine learning models that interpret real-time data. Vulnerabilities arise when adverse actors manipulate sensor inputs, foreign governed models misclassify anomalies, cloud outages disrupt optimisation tools or data pipelines are compromised. Because water systems are tightly coupled with energy and telecommunications, disruptions can cascade across sectors.

In transport networks, AI underpins traffic signal optimisation, rail scheduling, aviation routing, port logistics and freight coordination. These systems rely on foreign hosted optimisation models, autonomous system components that receive external updates and data streams that can be manipulated to create congestion or misrouting. Cloud outages can disrupt routing or scheduling, paralysing supply chains, delaying emergency services and undermining economic activity.[23] Ports and freight hubs are particularly exposed because they depend on AI for container tracking, crane automation and logistics coordination.

Telecommunications networks form the backbone of all other critical infrastructure sectors. AI is used to optimise routing, detect anomalies, manage congestion, allocate spectrum and support 5G architectures.[24] These systems are highly automated

and deeply integrated with cloud platforms. Vulnerabilities arise when network management AI is foreign governed, model updates are externally controlled, adverse actors exploit AI enabled routing algorithms or cloud outages disrupt network management functions. A disruption in telecommunications cascades across energy, water, transport, finance, health and emergency services.[25]

In financial systems, AI supports fraud detection, algorithmic trading, credit risk modelling, anti-money laundering analysis and realtime settlement and clearing.[26] These systems operate at high speed and high volume, making them sensitive to misclassification, data manipulation or cloud outages. If foreign governed models misclassify transactions, if data streams are corrupted or if AI enabled fraud tools are compromised, the result can be liquidity shocks, market instability or loss of public confidence.

Health systems increasingly rely on AI for diagnostic imaging, hospital triage, supply chain management, electronic medical records and predictive analytics for emergency demand.[27] These systems are uniquely sensitive because disruptions have immediate human consequences. Vulnerabilities arise when diagnostic models are foreign governed, cloud outages disrupt hospital operations, adversarial manipulation affects triage algorithms or supply chain optimisation tools fail during crises. During natural disasters, pandemics or mass casualty events, foreign dependency creates unacceptable risk.

In food supply chains, AI is used for crop forecasting, logistics optimisation, coldchain management, inventory planning and supermarket distribution.[28] These systems depend on real-time data from farms, freight networks, ports and retail systems. Vulnerabilities arise when optimisation models are foreign hosted, routing algorithms depend on external cloud services, or adversaries manipulate data to disrupt distribution. Because food supply chains rely on transport, energy and telecommunications, disruptions can rapidly escalate into shortages, price shocks or regional supply failures.

Emergency services management increasingly depends on AI for fire behaviour modelling, flood prediction, resource allocation, dispatch optimisation and situational awareness platforms that integrate data from satellites, sensors and public reporting.[29] These systems must remain available during crises, precisely when cloud platforms are most likely to be stressed or targeted. Vulnerabilities arise when predictive models are foreign governed, data pipelines are disrupted, or cloud outages degrade command-and-control systems. A failure in AI enabled emergency management tools can delay response, misallocate resources or reduce situational awareness during life threatening events.

Across all sectors, these dependencies create crosssector cascading risks. Energy and telecommunications form a mutually dependent pair: energy systems rely on telecommunications for sensor data, control signals and grid balancing algorithms, while telecommunications rely on energy for base station power, data centre operation and network management systems. A failure in one rapidly destabilises the other. Transport systems underpin national supply chains, meaning disruptions in AI enabled routing, port automation or freight optimisation can paralyse logistics and economic activity. Water systems support health and public safety, so failures in AI enabled contamination monitoring or pump optimisation can cascade into hospital disruptions and emergency services strain. Financial systems depend on telecommunications and energy, meaning failures in AI enabled trading, settlement or fraud detection can impair government services and undermine public confidence.

The pattern is clear: AI enabled optimisation has become essential to the functioning of modern infrastructure, but it has also become a point of systemic vulnerability. When the models, data flows or cloud platforms that underpin these systems are governed by foreign jurisdictions, Australia

cannot guarantee their availability, integrity or alignment with national interests during periods of stress. Sovereign capability, in compute, data governance, model assurance and integration, is now essential to ensuring that critical infrastructure remains resilient under all conditions.

2.4 The Structural Nature of the Risk

The risks outlined above are not the result of poor cybersecurity, inadequate regulation or insufficient investment. They are structural. They arise because AI is becoming the operational underpinning of critical infrastructure; because the models, data pipelines and optimisation tools that underpin essential services are largely hosted on foreign cloud platforms; because foreign governments can compel access or restrict availability; because adverse actors can exploit vulnerabilities in models and data flows; and because tightly coupled systems amplify the impact of disruptions across sectors. Each of these mechanisms reflects a single underlying condition: the systems that govern Australia's essential services are not governed by Australia.

This creates a sovereignty problem where Australia cannot guarantee the continuity, integrity or security of essential services when the systems that support them are governed by external authorities. As energy, water, transport, telecommunications, finance, health, food supply chains and emergency services become more automated, more interconnected and more dependent on AI enabled optimisation, the question of who controls the underlying infrastructure becomes central to national resilience.

The economic dimensions of this dependency compound the strategic ones. Data centres are frequently cited as evidence of domestic digital investment, yet the value generated by that investment does not remain in Australia. Analysis of domestic data centre economics suggests that for every $100 invested, as much as $70-$80 leaves the country almost immediately through imported hardware, proprietary software licences, and external service provision controlled by foreign firms.[30] This mirrors a structural pattern Australia has navigated across earlier industrial eras, supplying land, energy, and capital while the higher-value components of production, design, intellectual property, system governance, accumulate offshore. In the AI era, the equivalent of raw material export is data that is generated domestically, processed through foreign governed infrastructure, and returned as a service over which Australia retains no sovereign control. Sovereign compute investment addresses this leakage only if it is designed from the outset to retain operational control, intellectual property, and governance authority within Australian jurisdiction, not merely to provide physical hosting while the value chain remains external.

The economic pattern and the strategic risk share a common cause, Australia's regulatory and governance frameworks were designed for a world in which physical assets and the systems operating them were distinct. That distinction no longer holds. Algorithms are no longer discrete software tools operating in support of infrastructure, they are infrastructure. The systems that balance the national electricity grid, coordinate freight logistics, manage hospital triage, govern financial settlement, and direct emergency response are performing functions indistinguishable in their national importance from the physical assets that regulation has long treated as critical infrastructure. Yet the legal and regulatory frameworks governing those physical assets, licensing requirements, performance standards, resilience obligations, public interest conditions, do not extend to the algorithmic systems that increasingly determine how those assets behave. Recognising algorithms as infrastructure is not a theatrical move, it is a precondition for effective regulation. Legal authority without the ability to reach the systems actually making operational decisions is not sovereignty, it is the appearance of oversight while control resides elsewhere. Extending infrastructure regulation to encompass the AI systems embedded

within essential services is the foundation on which sovereign assurance requirements, sector regulator mandates, and enforceable remediation obligations must be built.

The vulnerabilities identified above therefore reflect that essential services increasingly rely on AI systems that Australia does not fully control, governed by frameworks that were not designed to reach them.

2.5 The Sovereign Capability Architecture for Critical Infrastructure

Securing Australia's critical infrastructure in an AI enabled world requires more than technical upgrades or incremental regulatory reform. It demands a structural shift in how Australia conceptualises resilience, risk and sovereignty. The vulnerabilities outlined in the previous section cannot be mitigated through isolated initiatives. They require a coordinated capability architecture built on five mutually reinforcing foundations: sovereign compute, sovereign data governance, sovereign model assurance, domestic integration capacity and sovereign workforce and institutional capability.

Sovereign compute is the foundation of sovereign AI capability. It refers to the ability to host, operate and secure AI systems within infrastructure that is physically located in Australia, governed by Australian law and insulated from foreign jurisdictional claims. Critical infrastructure sectors rely heavily on cloud-based AI services, typically hosted on foreign platforms subject to foreign legal frameworks. Without sovereign compute, Australia remains dependent on these platforms for hosting operational models, processing sensitive data, conducting model assurance and adversarial testing, and maintaining continuity of service if foreign access is restricted. Building sovereign compute capacity requires secure Australian governed data centres, high-performance computing clusters optimised for AI workloads, sovereign cloud environments for essential services, governance frameworks that insulate these facilities from foreign legal claims, and resilience and failover capacity across jurisdictions. Sovereign compute is not a luxury; it is a prerequisite for ensuring that essential services remain available under all conditions.

Sovereign data governance ensures that operational data remains under Australian jurisdiction; that data pipelines are secure and resistant to tampering; that training and inference data is protected from foreign access; and that data flows comply with Australian privacy, security and critical infrastructure laws. Even where critical infrastructure operators store data in Australian facilities and contract with providers under Australian law, the underlying platforms and model hosting environments may remain governed by foreign legal frameworks that can compel disclosure, access or modification without Australian consent or recourse. This is the data dimension of the cloud hostage condition. Foreign governments can compel access to sensitive operational patterns regardless of where data is physically stored. Mandatory data sovereignty requirements, including tiered classification of sensitive datasets, enforceable restrictions on offshore access, and explicit carve-outs for AUKUS and Five Eyes data sharing obligations, are the mechanism through which Australia breaks this dependency at the legal layer. Technical measures such as encryption and access controls reduce risk but cannot substitute for jurisdictional control.

Sovereignty in data governance is ultimately a legal and institutional question, not a technical one. To make this principle operational, a tiered classification system is required. At minimum, three tiers are necessary:

- **Tier 1 (Sovereign Mandatory) –** data whose offshore access or processing would create unacceptable national security or critical infrastructure risk, including operational technology data from energy, water and telecommunications systems, defence

sensitive logistics and supply chain data, biometric and identity data held by government agencies, and electoral administration systems. This data must be stored and processed exclusively within Australian governed infrastructure and must never transit foreign legal jurisdictions.

- **Tier 2 (Sovereign Preferred) –** data that creates material sovereignty risk if governed offshore but for which transition timelines and cost constraints may require temporary arrangements; sovereign hosting is mandated within defined timeframes, and offshore access must be notified and audited.

- **Tier 3 (Standard Compliance) –** data governed by existing privacy and security frameworks, for which current protections are adequate and sovereign hosting is not mandated.

Two exceptions must be built into this system. First, an AUKUS and Five Eyes carve-out must ensure that mandatory data sovereignty requirements do not inadvertently restrict intelligence sharing obligations and operational data exchange with Australia's closest allies. Data exchanged under AUKUS Pillar II co-development programs, Five Eyes intelligence frameworks or bilateral defence arrangements should be governed by agreed standards rather than general sovereign hosting mandates, provided those standards meet minimum assurance requirements certified by ASD. Second, a transitionary pathway must be built into the classification regime, recognising that many essential services operators currently lack the sovereign infrastructure to achieve immediate Tier 1 compliance. Phased transition timelines,12-month foundations, 36-month capability build and 7-year full compliance goals, must be matched by sovereign compute investment to ensure that mandates do not outrun domestic infrastructure capacity.

Sovereign model assurance is the process through which Australia verifies that AI systems used in critical infrastructure are reliable, secure and predictable under all conditions. It requires the ability to audit model behaviour, test for vulnerabilities, conduct adversarial evaluation, verify training data origin, monitor for model drift and ensure compliance with Australian law and safety standards. Foreign governed models often operate as black boxes, in which training data is opaque, update cycles externally controlled, internal mechanisms are independently unverifiable, creating unacceptable risk in critical infrastructure contexts. Sovereign model assurance requires access to model weights and architecture; the ability to conduct red-teaming and adversarial testing; continuous monitoring of model performance; sector specific assurance standards; and institutional capacity within both regulators and operators. Model assurance is not a one-off process, AI systems evolve through updates, retraining and exposure to new data, making continuous assurance essential to maintaining reliability and security over time.

Domestic integration capacity is where sovereign capability becomes operational. Sovereign compute, data governance and model assurance are necessary but insufficient without the ability to integrate AI systems with the physical infrastructure that delivers essential services. Integration requires systems engineering expertise, secure data pipelines, interoperability standards, platform specific adaptation and continuous testing and validation. Foreign dependency is most acute at this layer, with reliance on foreign integrators or foreign governed systems risks embedding vulnerabilities deep within critical infrastructure architecture. Domestic integration capacity ensures Australia can adapt AI systems to local infrastructure, maintain operational independence, secure data flows across sectors, update systems without foreign approval and respond rapidly to emerging threats. Building this capacity requires long-term investment in Australian firms, workforce development and institutional partnerships.

Workforce and institutional capacity is the final foundation. Critical infrastructure resilience

requires machine learning engineers, cybersecurity specialists, systems engineers, data governance experts, AI assurance specialists and sector specific AI practitioners, as well as institutional capacity within regulators and operators to govern, procure and oversee AI systems effectively. Institutional capacity requires governance frameworks for AI procurement and deployment, standards for model assurance and security, mechanisms for cross-sector coordination and long-term investment strategies aligned with national resilience needs. Without workforce and institutional capacity, sovereign compute and model assurance cannot be realised. Sovereign capability is ultimately a human capability, the ability of Australian institutions and experts to build, secure and govern the systems that underpin essential services.

The five foundations outlined above are interdependent. Investment decisions made in one area affect the others and procurement conditions set in one portfolio shape capability outcomes across all three pillars. Coordinating this architecture requires a standing governance institution with the mandate and authority to align investment, set standards and enforce obligations across government, industry and the workforce simultaneously. The National AI Sovereignty Compact is that institution. Unlike standard interdepartmental committees or industry advisory bodies, the Compact is a tripartite body of government, industry and unions, whose design reflects a structural argument, that sovereign AI capability is built on industrial foundations, and industrial foundations cannot be governed effectively without the active participation of the workforce that operates them. A coordination mechanism confined to government and industry would recreate accountability deficits that has allowed AI adoption in banking, telecommunications and public administration to hollow out domestic technical expertise while concentrating knowledge with foreign providers. Unions participate not as stakeholder representatives but as co-designers of deployment conditions, upholders of worker data rights and guarantors of democratic accountability at the workplace level, converting sovereign capability obligations from regulatory requirements into enforceable industrial rights.

The Compact's coordinating functions would operate across four departments. It aligns investment decisions across PM&C, DISR, Finance and Defence, ensuring that sovereign compute, model assurance infrastructure, workforce pipelines and R&D programs are sequenced and funded as a connected capability agenda rather than competing portfolio priorities. It sets and monitors sovereign capability standards, the minimum requirements for sovereign hosting, model assurance, data governance and domestic integration capacity that apply across defence, essential services and critical democratic systems. It would govern procurement conditions, working with Finance to ensure that providers deploying AI in critical infrastructure demonstrate compliance with sovereign capability and labour standards as a condition of Commonwealth procurement eligibility. And it oversees deployment governance, through advance consultation obligations, workforce impact assessment requirements and collective bargaining frameworks through which AI adoption is managed equitably across sectors facing significant disruption. PM&C chairs the Compact and is responsible for ensuring its decisions are binding across portfolios rather than advisory. DISR, Defence, Home Affairs, Finance, Education and DEWR hold primary portfolio responsibilities within it, and industry and union representatives participate on equal terms in standard setting and investment sequencing decisions.

2.6 Sovereign AI Research and Development

Sovereign capability cannot be sustained through procurement alone. The systems that underpin national resilience, will continue to evolve in complexity and capability. If Australia relies entirely on foreign models and platforms to

operate these systems, procurement of existing foreign technology simply defers, rather than resolves, the sovereignty problem. Over time, the capability gap between states that develop AI and states that consume it will widen. A credible sovereign AI strategy must therefore include a research and development investment architecture capable of producing, adapting, and continuously improving the AI systems on which essential services depend.

Australia does not need to compete with the United States or China in frontier general purpose model development. The resource requirements for that competition, hundreds of millions of dollars in compute, training data at global scale, and engineering workforces of a size Australia cannot rapidly assemble, are beyond realistic reach in the near term. The strategic objective is more targeted, with Australia needing to develop sovereign capability in the specific AI applications where dependence creates unacceptable national risk, and to build the research and institutional infrastructure that allows Australia to govern, adapt, and improve AI systems across the lifecycle of their deployment in critical functions.

Domain specific AI systems represent the most immediately actionable R&D priority. In energy, this means sovereign models for grid balancing, distributed energy resource management, and demand forecasting, systems that can be trained on Australian grid data, maintained under Australian jurisdictional control, and updated without foreign approval. In water, it means sovereign anomaly detection and treatment optimisation systems designed for Australian infrastructure conditions. In defence, it means the ISR analytics, logistics models, and cyber defence automation tools identified in Pillar 1. These systems do not require frontier compute capabilities, instead they require domain expertise, sovereign data pipelines, and institutional commitment. Australia already possesses the research talent to develop them. What is missing is the coordinated funding architecture, sovereign compute access, and public sector demand signal required to convert that talent into operational capability.

The National Reconstruction Fund and the Defence Industry Cooperative Research Centres program offer existing mechanisms for channelling long-term capital into sovereign AI R&D. Both instruments should be expanded to fund AI capability development in critical infrastructure domains, with sovereign capability weighting built into grant criteria, requiring recipients to demonstrate Australian jurisdictional control over model training, data origin, intellectual property, and deployment infrastructure. This ensures that public R&D investment generates sovereign capability rather than subsidising foreign platforms.

University and research institution partnerships are a second essential pillar. Australia's universities, particularly those with established AI research capabilities, are underutilised as sovereign capability assets. A national sovereign AI research network, coordinating investment across the Commonwealth Scientific and Industrial Research Organisation (CSIRO), university AI institutes, and the Defence Science and Technology Group (DSTG), would create the institutional spine for long-term R&D. This network should be focused on national priority capability areas, operate on sovereign compute infrastructure, maintain open publication standards where security permits, and feed directly into the workforce pipelines. Research that produces both operational capability and trained personnel is a compounding return on investment, as the goal is not research papers but sovereign practitioners.

Intellectual property and data governance must be embedded in R&D funding conditions from the outset. Commonwealth funded AI research should be subject to requirements that retain ownership of model weights, training data, and architectural innovations within Australia. Where partnerships with international firms are necessary, agreements must preserve Australia's right to audit, adapt, and redeploy model outputs without seeking foreign permission. The pattern identified, where domestic investment generates returns that flow offshore

through imported IP and foreign controlled platforms, must not be replicated within the public R&D system.

Finally, sovereign R&D investment must be explicitly connected to the model assurance infrastructure described in section 2.5. Research that produces models which cannot be independently audited, stresstested against adversarial conditions, or continuously verified is not sovereign capability, it is an externally governed capability located offshore. The national model assurance facility recommended in 1.4 should function as both a verification body and a technical centre for R&D on assurance methodologies, developing the tools and standards required to verify AI behaviour not just at deployment, but continuously over the lifecycle of systems embedded in national infrastructure.

2.7 AI, Economic Transition, and the Labour Dimension of Sovereignty

Sovereign AI capability is an industrial policy agenda, and industrial policy is inseparable from labour. The workforce that designs, builds, operates, secures, and governs the AI systems described in this report is the human core of sovereignty. Without deliberate intervention, AI adoption in Australia risks following a pattern already visible in banking and telecommunications, where efficiency gains at the enterprise level paired with the offshoring of technical functions, is eroding domestic operational expertise and hollowing out of the skills base that sovereign capability depends on. A sovereign AI agenda must therefore treat workforce agency not as a social objective to be pursued alongside the economic strategy, but as a fundamental element of it.

Australia's major service industries illustrate how AI deployment, outsourcing, and offshoring interact to reshape workforce capability and operational sovereignty. In banking, institutions including the Commonwealth Bank and Bendigo Bank have automated substantial components of back-office processing, customer service triage, and compliance workflows.[31] These transitions have been accompanied by increased reliance on external technology providers and offshore service centres, reducing the stock of domestic technical expertise available to govern, audit, and secure these systems. When the technical knowledge required to understand how a system operates resides with a foreign provider rather than in the employing institution, accountability becomes nominal rather than substantive. Regulatory oversight becomes dependent on provider cooperation, and when Australia cannot modify, audit or replace systems without external authorisation, the operational dimension of sovereignty diminishes.

Telstra presents a parallel case in telecommunications.[32] The integration of AI into network management, diagnostics, and customer support systems has been accompanied by the expansion of offshore technical teams and managed service providers. Critical operational functions, the knowledge of how networks behave, where vulnerabilities lie, and how anomalies are detected, increasingly reside outside Australia's workforce. These are not minor administrative functions; they are the operational intelligence on which national communications resilience depends. The pattern is consistent, AI adoption that does not embed domestic technical expertise produces, at the system governance level, the same dependency that sovereign compute and data localisation are designed to address at the infrastructure level.

These cases have direct implications for sovereign AI strategy. The policy logic that motivates investment in sovereign compute, retaining operational control over the infrastructure that underpins essential services, applies equally to workforce capability. A data centre is sovereign only if Australian workers can operate, secure, and govern it. A sovereign model assurance facility is only as sovereign as the assurance experts staffing it. The workforce pipeline is not supporting

infrastructure for the sovereignty agenda, it is the sovereignty agenda made operational.

Several structural dynamics require deliberate policy intervention. First, AI adoption in large enterprises tends to be accompanied by outsourcing and offshoring of technical functions, particularly where domestic skills are scarce, and foreign providers offer immediate capability at lower cost. Without counterbalancing incentives, this pattern will deepen as AI adoption accelerates, concentrating technical capability in a small number of foreign firms and reducing Australia's ability to govern, audit, and adapt its own critical systems. Procurement conditions, skills investment, and enterprise level industrial agreements are the levers through which government can redirect this incentive structure.

Second, automation displaces employment in ways that are sectorally and regionally concentrated. The workers most exposed to displacement across customer service, administrative processing, and routine data tasks, are often in regions where alternative employment options are limited and where the skills transition to AI adjacent roles requires ongoing support. The economic and social costs of poorly managed transition concentrate in communities that are already exposed to structural economic change. A sovereign AI agenda that generates capability gains while simultaneously generating concentrated community level harm is not a national success; it is a redistributive failure with political and social cohesion consequences.

Third, the quality of AI deployment, whether it augments human capability or simply substitutes labour, has long-term implications for sovereign capability. Deployment models that remove workers from complex operational roles degrades the organisational knowledge base that high-quality governance, incident response, and system adaptation require. Augmentation oriented deployment, where AI handles routine processing while workers are redeployed to higher order governance, oversight, and adaptation functions, creates more durable capability than labour replacement deployment models, as it retains and develops the human expertise that sovereign systems ultimately rely on.

Giving effect to these principles requires a coordinated policy architecture. Industrial agreements that include advance notice of AI driven organisational change, rights to explanation for automated decisions affecting employment, commitments to retraining and redeployment, and workforce impact assessment requirements are the workplace level mechanism through which transition is managed equitably. The National AI Sovereignty Compact must include explicit labour standards, not as a concession to union stakeholders but as a governance requirement for sovereign capability. Enterprises deploying AI in critical infrastructure should be required, as a condition of sovereign procurement eligibility, to demonstrate that deployment models meet minimum standards for human oversight, workforce impact assessment, and worker agency.

Regionally, the sovereign AI agenda creates genuine economic opportunity. AI enabled industries, logistics, energy infrastructure, advanced manufacturing, precision components, data centre operations, are distributed across regional Australia. The workforce development investments required to build sovereign capability, TAFE-delivered micro-credentials, apprenticeship and early career pathways, mid-career reskilling programs, can be specifically designed to service regional labour markets, converting the sovereignty investment into a mechanism for regional economic diversification and resilience.

2.8 Implementation Pathway

The implementation pathway for sovereign AI capability in critical infrastructure rests on four principles: authority must be explicit; sovereignty must be embedded in core infrastructure; industrial strategy must align with resilience needs; and continuity must be guaranteed even if foreign

access is restricted.

Roles and responsibilities are assigned as follows:

Home Affairs leads under the SOCI Act, establishing AI sovereignty standards for essential services, mandating foreign dependency assessments, developing sector specific resilience plans, building a national AI incident reporting framework, and coordinating cross-sector exercises that simulate AI enabled disruptions.

ASD extends its mandate to include model assurance standards, adversarial testing, continuous monitoring and AI specific cybersecurity frameworks. ASD develops the national AI red-team capability and certifies sovereign compute environments to defence grade security standards.

DISR leads development of sovereign compute infrastructure, establishes the national model assurance facility, builds domestic integration capacity through long-term partnerships with Australian firms. DISR leads the sovereign AI R&D investment framework, aligning National Reconstruction Fund (NRF) and Defence Industry Cooperative Research Centres (CRC-Defence) funding with critical infrastructure priority domains and embedding sovereign data and sovereign IP conditions into all Commonwealth AI R&D grants, and coordinates workforce pipelines with Education, DEWR and sector regulators. Where existing authority is insufficient, DISR's mandate should be expanded through Cabinet to explicitly include sovereign AI capability for critical infrastructure.

Finance reforms procurement rules so that sovereign hosting, model assurance and domestic integration capacity become default requirements for critical infrastructure digital systems. Strategic procurement reform is a primary market shaping instrument, embedding local capability outcomes, supply chain security requirements and contestability safeguards as default conditions of Commonwealth digital procurement. The trade law dimensions of this mandate are a shared responsibility of Finance, AGD and DFAT, who must assess World Trade Organization (WTO) Government Procurement Agreement (GPA) obligations and invoke national security exceptions as required before procurement conditions are finalised.

AGD insulates sovereign compute, cloud and data environments from foreign legal claims, clarifies the limits of extraterritorial legislation including the CLOUD Act and China's National Security Law, and develops legal protections for sovereign hosting environments. Where sovereign capability preferences cannot be fully insulated from trade obligations through existing exceptions, AGD and DFAT should develop a trade law reform agenda securing permanent legal space for the national AI sovereignty agenda.

Education and DEWR build the national skills pipeline through expanded tertiary and vocational programs in machine learning, cybersecurity, systems engineering and AI assurance, national accreditation standards for AI security professions, and coordination with DISR and sector regulators to align training with operational needs.

Sector regulators (Australian Energy Regular, ACMA, Australian Prudential Regulation Authority, etc.) and state-based water and transport regulators, embed sovereignty requirements into operational standards, mandate sovereign hosting for essential AI systems, require model assurance certification for critical functions, and ensure operators maintain fallback modes and manual override capabilities.

PM&C provides the strategic spine, coordinating whole-of-government alignment, driving investment decisions, and ensuring sovereign capability is treated as a national resilience requirement rather than portfolio specific initiatives.

Cross-government coordination must include a national AI sovereignty framework defining roles and standards; a whole-of-government investment strategy for sovereign compute, model assurance and integration capacity; shared

security and assurance standards for AI systems used in essential services; joint procurement mechanisms; and a national AI risk register tracking dependencies, vulnerabilities and emerging threats. PM&C leads this coordination effort.

Implementation proceeds in three phases.

Phase 1 (0–12 months): Foundations

- Home Affairs identifies sovereign data classes and priority sectors requiring mandatory data sovereignty controls under the SOCI Act.
- ASD drafts AI assurance and zero-trust requirements for essential service operators.
- Finance reforms procurement rules to incorporate sovereign capability weighting for critical infrastructure digital systems and embeds sovereign capability requirements into all Commonwealth AI R&D grants.
- DISR begins design of sovereign compute and cloud procurement for essential services and establishes a national sovereign AI R&D investment framework, aligning NRF, CRC-Defence and existing research programs with critical infrastructure domains.
- AGD develops legal protections for sovereign compute and cloud environments, clarifying exposure to foreign extraterritorial claims.
- Education and DEWR begin planning national AI workforce pathways for critical infrastructure sectors.
- PM&C establishes whole-of-government coordination settings and convenes a cross-portfolio steering group (DISR, Defence, ASD, Education, CSIRO) to define national priority application domains and coordinate early investment sequencing.
- ASD and CSIRO begin scoping technical requirements for integrating sovereign R&D outputs into the national model assurance facility.
- Universities and CSIRO establish a preliminary network of sovereign AI research points, identifying existing strengths and gaps in domain specific capability.

Phase 2 (1–3 years): Capability Build

- Home Affairs introduces mandatory data sovereignty requirements for high-risk sectors.
- DISR commences sovereign compute procurement across energy, water, transport, telecommunications, finance and health, and funds multi-year R&D programs for domain specific sovereign AI systems with mandatory sovereign data and sovereign IP conditions.
- ASD operationalises continuous verification, audit and adversarial testing for high-risk AI systems in essential services and develops and pilots' assurance methodology R&D programs including adversarial testing tools and model behaviour auditing frameworks.
- Sector regulators begin embedding sovereignty requirements into operational standards.
- Education and DEWR launch critical infrastructure AI workforce pathways through the National AI Workforce Compact.
- AGD finalises legislative and regulatory safeguards for sovereign data and compute environments.
- Universities and CSIRO operationalise the national sovereign AI research network on sovereign compute infrastructure, producing workforce pipelines.
- Finance and DISR jointly implement co-in-

vestment mechanisms prioritising Australian firms capable of delivering sovereign AI systems for critical infrastructure operators.

Phase 3 (3–7 years): Scale-Up

- DISR expands sovereign compute, cloud and data infrastructure to meet long-term national demand and scales sovereign AI R&D funding into a permanent national program anchored in sovereign compute and long-term industry partnerships.

- Sector regulators embed sovereign assurance standards across all strategic procurement and operational compliance frameworks.

- ASD and the national model assurance facility integrate R&D outputs into continuous verification pipelines, ensuring all publicly funded models meet sovereign assurance standards across their lifecycle.

- Home Affairs and DISR support the growth of domestic systems integration capability across essential services.

- Defence, DISR and CSIRO support the emergence of export ready sovereign AI systems, positioning Australia as a provider of governance aligned AI capability to Indo-Pacific partners.

- PM&C embeds sovereign AI R&D into long-term national capability planning, ensuring alignment with industrial strategy, workforce development and critical infrastructure resilience.

2.9 Implementation Risks and Mitigation

Implementing sovereign AI capability across critical infrastructure will encounter structural and political constraints that must be anticipated and managed. Mitigation requires coordinated action across Home Affairs, ASD, DISR, Finance, AGD, Education, DEWR and PM&C.

R&D fragmentation and foreign dependency pose a distinct structural threat that warrants separate treatment. Without a coordinated national R&D architecture, publicly funded research may default to foreign governed platforms, offshore compute/international IP arrangements that transfer ownership of model weights, training data or architectural innovations outside Australian jurisdiction. This creates a long-term dependency loop, in which even when models are developed domestically, the underlying infrastructure, IP rights or update pathways remain externally controlled. Mitigation requires embedding sovereign capability conditions into all Commonwealth AI R&D funding, mandating Australian jurisdictional control over data, model weights and deployment infrastructure, and ensuring all research outputs are compatible with the national model assurance facility.

Risk	Mitigation
Cloud provider resistance – mandatory data sovereignty requirements will face commercial and diplomatic pushback from major foreign cloud providers.	Home Affairs leads a risk tiered regulatory approach with transition periods, narrowly defined exemptions and clear criteria for high-risk workloads that must be hosted on sovereign infrastructure.
Compute shortfall – Australia cannot immediately host all high-risk workloads on sovereign infrastructure.	DISR adopts phased procurement prioritising the most sensitive sectors first; allied secure cloud arrangements serve as interim measures until domestic capacity scales.
Procurement distortion – sovereign capability preferences may entrench incumbents, reduce contestability or lock operators into immature domestic offerings.	Finance enforces open technical standards, independent evaluation mechanisms and transparent performance benchmarks to ensure sovereign capability strengthens resilience rather than creating protected monopolies.
Regulatory capacity gap – sector regulators lack the technical expertise to audit AI systems, assess model assurance claims or evaluate sovereign hosting compliance.	ASD staff deployments, shared national assurance capability and coordinated uplift programs build long-term regulatory expertise across energy, water, transport, telecommunications, finance and health.

Pillar Three: AI, Democratic Sovereignty, Institutional Integrity, and Social Cohesion

3.1 AI and the Transformation of the Information Environment

AI is reshaping the information environment in ways that alter how democratic societies form opinions, interpret events and understand themselves. It automates the production of persuasive content at scale, generating text, images, audio and video that increasingly resemble authentic human communication.[33] It personalises information flows through recommendation systems that tailor content to individual preferences, beliefs and emotional states,[34] as well as accelerating the speed and reach of information, enabling platforms to spread content across networks in seconds, shaping public perception long before institutions can respond.[35]

These transformations create profound vulnerabilities for democratic systems. AI enabled content production lowers the cost of misinformation and disinformation, allowing foreign and domestic actors to generate persuasive narratives, impersonate public figures, manipulate public sentiment and overwhelm public debate.[36] Synthetic media further erodes confidence in authentic evidence, enabling reputational attacks, political deception and the 'liar's dividend,' where genuine information can be dismissed as fake.[37]

Personalisation compounds these risks by fragmenting the public sphere. AI driven recommendation systems reinforce confirmation bias, amplify emotionally charged content and create tightly bounded echo chambers.[38] Political actors can microtarget messages, foreign actors can exploit social divisions, and extremist content can be spread to vulnerable audiences.[39] As individuals occupy increasingly separate informational worlds, democratic debate becomes more difficult, institutional legitimacy weakens and social cohesion erodes.[40]

AI also accelerates the speed and scale of information operations. Adversaries can automate bot networks, target specific demographic groups, exploit social fractures and saturate public discourse with rapidly mass-produced content.[41] These operations can distort public perception, influence elections and destabilise civic trust. The speed of AI driven information flows outpaces the ability of governments, media and civil society institutions to respond, verify or contextualise emerging narratives.[42]

These vulnerabilities are intensified by structural dependencies. The platforms that mediate public debate are overwhelmingly governed by foreign companies operating under foreign legal frameworks. Their incentives, commercial priorities and regulatory obligations do not necessarily align with Australia's democratic interests.[43] In several jurisdictions, these firms have resisted regulatory oversight and, when challenged, have mobilised the political authority of their home governments.[44] This dynamic places core elements of Australia's democratic information environment under external control, creating a sovereignty problem, where democratic resilience is increasingly shaped by decisions made outside Australia's jurisdiction.

Together, these dynamics represent a systemic transformation of the information environment. AI is not imply a new tool within this environment; it is reshaping the environment itself, altering how information is produced, distributed, personalised and weaponised. Understanding this transformation is essential to diagnosing the sovereignty risks that now confront Australia's democratic institutions.

3.2 The Sovereignty Problem: Institutions, Legitimacy, and Social Cohesion

Australia's democratic sovereignty problem operates through three intertwined mechanisms. Foreign political decisions shape platform governance, as overseas governments can influence platform policies through regulation, political pressure or national security directives. These decisions affect content moderation, algorithmic transparency, data access, platform availability and information security practices. Australia has limited influence over these choices, even when they directly affect domestic democratic processes.[45]

Foreign commercial incentives further shape information flows. Platforms optimise for engagement rather than democratic integrity, prioritising content that maximises attention, including misinformation, outrage riven material, polarising narratives and emotionally charged content. These incentives distort public discourse, widen division and weaken the informational foundations on which democratic debate depends.

Foreign legal jurisdiction increases these vulnerabilities. Data used for personalisation, targeting and content moderation is often stored or processed in jurisdictions with different legal frameworks. This creates exposure in which foreign governments can compel access, platform decisions may prioritise compliance with foreign law, and Australia cannot independently audit algorithms, data flows or model governance practices. Together, these mechanisms create a form of dependency incompatible with democratic sovereignty.[46]

These structural exposures directly affect institutional legitimacy. Democratic institutions depend on public trust, yet AI enables misinformation that undermines confidence in public authorities, accelerates narratives that delegitimise officials and creates uncertainty about the authenticity of evidence. Synthetic media can fabricate misconduct, impersonate public figures and erode trust in authentic communication.[47] False information spreads faster than corrections, coordinated campaigns can overwhelm institutional responses and public officials may be targeted with precision that exploits personal vulnerabilities or community divisions. Traditional communications strategies are insufficient in an environment where AI enabled attacks on legitimacy operate at machine speed.

The risks associated with largescale data aggregation and AI assisted profiling are no longer hypothetical. The exposure of Zhenhua Data's operations, in which a Chinese technology firm systematically collected detailed information on politicians, business leaders, journalists, union representatives, judges and defence personnel across multiple countries, demonstrated how open source data, behavioural analytics and relational mapping can be fused to generate strategic intelligence on democratic institutions.[48] By aggregating location data, financial information, social networks and behavioural patterns, hostile actors can identify key personnel, map institutional relationships and pinpoint vulnerabilities across public and private sectors. These capabilities move beyond passive surveillance into active strategic targeting, enabling more precise forms of coercion, blackmail and institutional disruption than traditional espionage. For Australia, the Zhenhua case illustrates the sovereignty problem at the heart of this section, the information environment that sustains democratic institutions, the networks of trust, authority and communication through which they operate, is itself a target. Foreign governed data platforms and AI systems are not passive tools for influence operations, they are the infrastructure that makes strategic profiling at scale possible. Protecting institutional legitimacy therefore requires sovereign control over the data architectures and AI systems through which Australia's democratic life is mediated.

AI also places new pressures on social cohesion, the foundation of democratic stability. Cohesion depends on shared norms, mutual trust, tolerance

of disagreement and a sense of common purpose. AI driven personalisation amplifies polarising content, reinforces ideological echo chambers and enables targeted manipulation of vulnerable groups. Extremist narratives can spread rapidly, while synthetic media and personalised manipulation undermine trust in authentic communication. These pressures fragment society, weaken democratic resilience and create openings for foreign interference.

Australia's democratic information environment operates under what this report describes as a platform hostage condition, in which algorithms that rank information, the systems that detect synthetic content, and the infrastructure that distributes public discourse are governed by foreign companies operating under foreign legal frameworks, optimised for foreign commercial priorities, and subject to foreign political authority. Like the cloud hostage condition in infrastructure and defence, this dependency is not a failure of policy, it is the predictable consequence of building democratic communication on foreign governed platforms. The distinction matters as in the cloud hostage condition, the risk is operational disruption, whereas in the platform hostage condition, the risk is institutional delegitimisation. Sovereign information infrastructure addresses the latter, but only if it is designed from the outset as a genuine industrial capability, not a technical patch applied to a foreign governed system

3.3 Sovereign Information Infrastructure

Australia's information environment is mediated by platforms and AI systems overwhelmingly governed offshore. This creates structural vulnerabilities where foreign political decisions, commercial incentives and legal frameworks shape the flow of information within Australia's democracy. To minimise this exposure, Australia must develop sovereign information infrastructure, systems that ensure the integrity, availability and security of public interest information, independent of foreign governed platforms.

Sovereign information infrastructure includes Australian governed content distribution systems for emergency communication, public health messaging and civic information; secure domestic hosting environments for public interest digital services; sovereign data pipelines for election information, government communication and civic education; and public interest algorithms designed to prioritise accuracy, reliability and democratic values. These systems do not replace global platforms but complement them by ensuring that essential democratic functions do not depend on foreign governed infrastructure.

A core component is the maintenance of sovereign public interest information channels. Emergency alerts, public health communication, election information, institutional announcements and civic education content must be sovereignly hosted, resilient to cyberattack, insulated from foreign jurisdiction, accessible across platforms and trusted by the public. Public interest communication cannot rely solely on commercial platforms whose incentives prioritise engagement over accuracy.

Sovereign hosting is equally essential for critical civic systems. Election administration, parliamentary communication, public service announcements and government digital services must operate on infrastructure governed by Australian law. Foreign governed cloud platforms introduce unavoidable exposure to foreign legal claims and political decisions, creating risks that cannot be mitigated through contractual or technical safeguards alone.

As synthetic media becomes more sophisticated, sovereign capabilities in content verification becomes a foundational element of information infrastructure. Verification systems must allow citizens to distinguish authentic communication from synthetic manipulation, enable institutions to respond quickly to fabricated content, allow public officials to authenticate their statements,

support media organisations in verifying evidence and protect election administrators from high impact disinformation events. This requires national digital signature infrastructure for public officials and institutions, AI enabled detection tools for synthetic media, public facing verification portals, crossplatform monitoring systems for coordinated manipulation and rapid response mechanisms for disinformation incidents.

Authenticity infrastructure is central to this capability. Public officials must be able to authenticate statements, videos, audio recordings and public announcements through systems like cryptographic signatures, secure identity verification systems and tamper resistant communication channels. Without this, impersonation and synthetic attacks can and will undermine institutional credibility.

Australia must also maintain sovereign capability to detect deepfake videos, synthetic audio, AI generated images, fabricated documents and coordinated synthetic identity networks. Detection cannot rely solely on foreign platforms, which may prioritise commercial or foreign policy considerations. Sovereign information infrastructure therefore requires domestic technical capability, secure hosting, public interest algorithms and national verification systems that operate under Australian authority.

Data critical to civic function, electoral administration records, public official authentication credentials, emergency communication systems and government digital service infrastructure, falls within Tier 1 (Sovereign Mandatory) of the data classification framework established in Pillar Two. Platform mediated data, the content, targeting parameters and behavioural data flowing through foreign recommendation systems, presents a different challenge, as it is generated domestically but processed and governed offshore, placing it structurally outside Australian jurisdictional reach. The platform hostage condition operates exactly at this layer. Sovereign information infrastructure addresses the Tier 1 dimension, while platform governance addresses the offshore processing dimension. Both are required, and neither is a substitute for the other.

3.4 Domestic Governance and Civic Resilience Capacity

Technical infrastructure alone cannot secure Australia's democratic information environment. Sovereign capability also requires domestic governance capacity, the ability to set standards, enforce compliance and oversee the AI systems that shape public discourse and public administration. This governance requirement operates across two distinct domains that must be addressed separately, these being the regulation of private platforms and their AI systems, and the governance of AI in public sector decision making. These domains share underlying principles, transparency, accountability, auditability and meaningful avenues for redress, but involve different actors, different legal frameworks and different regulatory bodies.

Both areas share a common enforcement gap. In platform governance, ACMA's existing powers allow it to request information from platforms but not to independently audit algorithmic systems, access model architectures or compel technical evidence of compliance.[49] The voluntary disinformation code has demonstrated this clearly, as ACMA's own oversight reports have found, platforms self-report on their own terms, against metrics they have largely designed themselves, without consistent Australian trended data or independent verification of whether code commitments are being met.[50] The Misinformation and Disinformation Bill 2024, which would have introduced enforceable co-regulatory obligations, was abandoned before passing.[51] In public sector governance, the enforcement gap is equally evident. Despite the Robodebt Royal Commission's 2023 recommendations for legislative reform and an independent audit body, no such framework

has been legislated, and the pattern has already repeated, with the Commonwealth Ombudsman finding that DEWR unlawfully terminated income support payments for nearly a thousand jobseekers using an incorrectly coded automated system.[52] In both cases, regulatory intent without independent technical verification is not accountability, it is the appearance of oversight

Platform governance addresses the recommendation algorithms, content ranking systems, synthetic content detection tools and political advertising targeting mechanisms operated by private technology companies. ACMA is the primary regulatory body in this context, and its mandate must be extended to require platforms to disclose how recommendation systems operate, provide visibility into targeting practices, allow independent audit of high-risk algorithms, and share data on coordinated manipulation. Influence operation reporting must be mandatory, timely and enforceable, requiring platforms to detect coordinated inauthentic behaviour, report foreign interference attempts and provide early warning during crises. Without enforceable reporting obligations, Australia remains dependent on voluntary disclosures shaped by foreign commercial and political incentives.

Public sector automated decision-making addresses AI systems deployed by government agencies to make or inform administrative decisions affecting individuals across welfare, migration, law enforcement, tax compliance, housing and health. The regulatory challenge is fundamentally different, as the actors are public authorities whose decisions carry legal force, and the affected parties are citizens entitled to procedural fairness, explanation and appeal. The Robodebt scheme illustrated what happens when automated decision-making in high-stakes public administration operates without adequate human oversight, explainability or accountability. The governing framework must require human-in-the-loop mechanisms that function as substantive safeguards rather than formalities; mandatory explainability in plain language at the point of decision; independent auditing of model performance, bias and outcomes; and enforceable appeal and redress rights. AGD should lead this effort, with oversight shared with the Commonwealth Ombudsman, the Office of the Australian Information Commissioner (OAIC) and relevant sector regulators.

Domestic enforcement capability is the backbone of both governance functions. ACMA, AGD, the Ombudsman and the OAIC must maintain technical expertise capable of auditing AI systems independently, not simply receiving self-reported disclosures. Without that capacity, enforcement is dependent on the cooperation of regulated entities. France's Pôle d'Expertise de la Régulation Numérique (PEReN) provides a relevant institutional model. Under the PEReN design, sectoral regulators retain legal mandate and enforcement authority while a dedicated forensic body conducts the technical work, auditing closed systems, identifying algorithmic bias and producing legally admissible findings. This separation of legal authority from technical execution is the governing design principle. Australia currently lacks an equivalent body. Establishing a sovereign AI forensic capability modelled on PEReN, anchored within the national assurance architecture, would convert regulatory intent into demonstrable, enforceable accountability across both platform governance and public sector oversight.

Additionally, democratic resilience institutions can complement governance capacity by strengthening society's ability to resist manipulation and maintain trust across the steady state of the information environment. Media literacy programs equip citizens to identify misinformation, recognise synthetic media and navigate polarised environments. Community based initiatives build trust networks, counter polarisation and strengthen crisis communication. These civic foundations depend on a workforce capable of detecting synthetic media, auditing algorithms, analysing influence operations and conducting platform oversight, supported by cross-agency coordination and long-term public interest governance

frameworks.

AI Reserves would address this disruption. Consisting of a standing capability of trained personnel and pre-configured deployable systems enabling rapid mobilisation in response to cyber incidents, critical infrastructure disruptions or large-scale information operations. Where civic resilience institutions uphold distributed democratic capacity continuously, AI Reserves activate when that capacity comes under severe stress. Drawing on specialist teams in cybersecurity, data analytics, systems engineering and emergency logistics, AI Reserves could support cyber defence, emergency communications and critical infrastructure management at short notice, operating across federal and state jurisdictions. Home Affairs would serve as the lead agency for establishing, maintaining and activating the Reserves, with Defence holding responsibility for classified operational environments. Standards, readiness requirements and activation protocols would be jointly defined by Home Affairs and Defence. DEWR and Education would coordinate the workforce pipelines required to maintain reserve capacity between activations, ensuring that sovereign AI investment converts into operational resilience that activates when it is most needed.

3.5 Trade Unions as Democratic and Civic Resilience Institutions

Trade unions occupy a distinct role within Australia's democratic civic resilience architecture that extends beyond workplace representation. Their significance lies in three capacities that no other institution replicates at scale.

First, unions provide structured resistance to algorithmic manipulation at the workplace level. AI systems increasingly govern how shifts are allocated, performance is assessed, and disciplinary outcomes are determined across the media, telecommunications, logistics and public administration sectors; precisely the industries whose workers operate the infrastructure that mediates democratic communication. Where algorithmic management removes human judgement from decisions affecting workers in these sectors, it simultaneously degrades the operational knowledge and institutional independence on which democratic information infrastructure depends. Unions provide the mechanism through which workers in these sectors can challenge opaque algorithmic governance, advocate for limits on behavioural surveillance, and ensure that consent and transparency obligations are genuine rather than titular, converting those protections into enforceable rights through collective bargaining instead of than leaving them to regulatory discretion.

Second, unions are early warning institutions for foreign interference and platform governance failures. Workers in telecommunications, media and digital platform adjacent industries have direct operational visibility into how content is moderated, how algorithms behave under pressure, and where platform governance obligations are met on paper but not in practice. Structured channels through which workers can report algorithmic anomalies, content manipulation patterns or governance failures to relevant authorities turn that operational knowledge into a civic intelligence function. No regulatory body has equivalent ground level visibility into platform behaviour.

Third, union participation in AI governance at the democratic level ensures that sovereign information infrastructure is accountable to the people who operate it. The platform hostage condition identified in 3.2 is not only a regulatory failure, but also a workplace governance failure. The workers who operate Australia's telecommunications networks, media organisations and public sector communication systems are the human core of democratic information infrastructure. Their knowledge of how these systems actually behave, where vulnerabilities lie and how anomalies manifest is the operational intelligence on which sovereign assurance depends. Embedding that knowledge into the design and governance of

sovereign information infrastructure through the tripartite Compact structure established in 2.5, ensures that democratic AI governance reflects real-world conditions rather than regulatory assumptions about how systems should work.

The role of unions in democratic AI governance is not peripheral to sovereignty. In the information environment as in the industrial base, sovereignty is ultimately a human capability, and unions are the institutional mechanism through which that human capability is organised, sustained and made accountable.

3.6 International AI Governance Engagement

Australia's democratic sovereignty agenda cannot be pursued in isolation. The systems that create sovereign vulnerability are global in reach, and building genuine sovereign capability requires active engagement with the international AI governance architecture, involving shaping global standards, building coalitions of democratic states, and establishing Australia as a trusted provider of governance-oriented AI capability in the Indo-Pacific.

Australia's claim to this role rests on three advantages that neither the United States nor other democratic partners can fully replicate in the region. First, Australia's regulatory independence through its institutional architecture, independent ACMA, a credible OAIC, and a democratic governance tradition insulated from the commercial capture that has constrained US platform regulation, provides a model that regional partners can adopt without inheriting the political baggage of US tech sector dominance. Second, Australia's presence across the Pacific, its established relationships with Pacific Island nations, Association of Southeast Asian Nations (ASEAN) partners and democratic Northeast Asian states, and its Five Eyes and AUKUS memberships give it access to the institutional channels through which governance influence can be operationalised at regional scale. Third, Australia is large enough to build and export sovereign AI governance standards but not so large as to replicate the hegemonic dynamic that makes US and Chinese AI governance models unattractive to smaller states seeking genuine independence. These advantages are not permanent, they require deliberate cultivation through the investments and partnerships outlined in this report.

At the multilateral level, Australia must engage in emerging global AI governance forums, including the AI Safety Summits initiated in the UK and continued through Seoul and Paris, not solely as a participant but as an agenda setter. This means advocating for governance frameworks that centre democratic accountability, platform transparency and the protection of open information environments, and contributing to international standards for algorithmic auditing, synthetic media verification, cross-border influence operation reporting and the interoperability of national AI assurance frameworks. Where international standards are being shaped, Australia's absence means other powers' preferences reign by default.

Regionally, Australia's Indo-Pacific partnerships provide the institutional basis for a democratic AI governance network. Many partners face the same sovereignty vulnerabilities, via foreign governed platforms mediating domestic information environments, AI systems deployed in essential services without adequate audit capability, and democratic institutions exposed to externally enabled influence operations. Australia can address these vulnerabilities directly, sharing governance frameworks, providing technical assistance for AI audit capability and exporting sovereign assurance standards. A regional network of states adopting Australian aligned AI governance standards creates both a security buffer and a market for Australia's emerging sovereign AI capability.

Bilaterally, the EU's AI Act represents the most advanced regulatory framework currently in

operation. Australia should seek formal regulatory cooperation with the European Commission, including mutual recognition of assurance standards, information sharing on platform governance and joint development of synthetic media detection tools, that leverages EU technical capacity without creating new forms of external dependency. Similar cooperation should be pursued with Canada and the United Kingdom.

DFAT must be recognised as a central actor in delivering this agenda. AI governance is now a dimension of foreign policy and strategic competition, not a technical matter confined to domestic portfolios. DFAT should be given an explicit mandate to advance Australia's AI governance interests in bilateral and multilateral forums, coordinate the diplomatic dimensions of the National AI Sovereignty Compact, and develop dedicated in-house AI governance expertise embedded within its regional engagement frameworks across the Pacific Step-up, the Australia–ASEAN relationship and the Quad.

3.7 Implementation Pathway

The implementation pathway rests on four principles: authority must be explicit; sovereignty must be embedded in democratic infrastructure; resilience must be systemic across institutions, platforms, communities and citizens; and continuity must be guaranteed even if foreign platforms restrict access or become compromised.

Roles and responsibilities are assigned as follows:

Home Affairs establishes AI specific information integrity standards, mandates risk assessments for AI enabled political communication tools, develops a national influence operations framework, coordinates cross platform incident response and maintains a national synthetic media alert system.

ASD defines security and verification standards for civic critical AI systems, develops sovereign synthetic media detection capability, monitors AI enabled influence operations, supports platform governance regulators with technical expertise and ensures sovereign information infrastructure meets national security standards.

DISR develops sovereign information infrastructure platforms for public interest communication, establishes a national content verification facility, supports domestic development of public interest algorithms, builds platform governance tools for independent audit, and develops workforce pipelines in AI governance, information security and civic technology engineering. Where existing authority is insufficient, DISR's mandate should be expanded through Cabinet to explicitly include sovereign AI capability for democratic integrity.

ACMA mandates transparency requirements for AI enabled political advertising, requires platforms to report influence operations, establishes algorithmic transparency standards and audits platform compliance.

The AEC hosts election critical systems on sovereign infrastructure, maintains authenticity infrastructure for election communication, develops rapid response protocols for synthetic media incidents and collaborates with ASD and Home Affairs on influence operation detection.

DFAT leads Australia's engagement in multilateral AI governance forums, negotiates bilateral cooperation agreements on AI assurance and platform governance, coordinates technical assistance programs for Indo-Pacific partners, and embeds AI sovereignty objectives into regional engagement strategies including the Pacific Step-up, ASEAN engagement and Quad partnerships.

Education and DEWR deliver national media literacy and civic resilience programs and build workforce pipelines in AI governance and civic technology engineering.

Media organisations and civil society integrate sovereign content verification tools into editorial workflows, provide media literacy training, monitor

localised influence operations and act as trusted intermediaries during crises.

PM&C provides the strategic spine: coordinating whole-of-government alignment, defining authority for international AI governance engagement, ensuring DFAT's mandate is integrated into national security and economic sovereignty planning, and leading cross-government coordination on democratic information sovereignty.

Implementation proceeds in three phases.

Phase 1 (0–12 months): Foundations

- PM&C commissions the National AI Capability and Security Strategy with a dedicated chapter on democratic integrity.
- ACMA drafts transparency and disclosure requirements for political uses of AI.
- AEC begins designing sovereign hosting for election critical systems.
- ASD drafts synthetic media detection and authenticity verification standards.
- DFAT establishes an International AI Governance Coordination Unit and maps Australia's sovereign assurance standards against EU, UK, Canadian, Japanese, South Korean and Indian frameworks to identify areas for early mutual recognition agreements.
- ACMA and the AEC initiate formal cooperation channels with international regulators on platform transparency, cross-border influence operation reporting and synthetic media detection.
- PM&C defines whole-of-government authority for international AI governance engagement, integrating DFAT's mandate into national security and economic sovereignty planning.

Phase 2 (1–3 years): Capability Build

- AEC migrates election critical systems into sovereign environments.
- ACMA operationalises platform transparency requirements and political AI auditing.
- ASD stands up national synthetic media detection and rapid response capability.
- Education and DEWR launch national media literacy and civic resilience programs.
- DFAT negotiates bilateral and multilateral cooperation agreements on AI assurance, platform governance and synthetic media verification with the EU, UK, Canada, Japan, South Korea and India.
- ASD and DISR participate in allied adversarial testing and model verification programs, ensuring Australian assurance standards are interoperable while preserving sovereign control over Australian sourced data.
- DFAT and DISR coordinate technical assistance programs for Indo-Pacific partners, supporting regional AI audit capability, sovereign cloud alignment and governance frameworks.
- ACMA and the AEC operationalise cross-border transparency mechanisms with democratic regulators, enabling shared detection of influence operations and coordinated responses to platform level manipulation.

Phase 3 (3–7 years): Scale-Up

- DISR expands sovereign information infrastructure capacity for public interest communication.
- ACMA and the AEC embed sovereign assurance standards across all civic critical systems.

- DFAT leads establishment of a regional democratic AI governance network across the Indo-Pacific, anchored in Australian assurance standards and public interest AI governance frameworks.

- ASD and ACMA integrate Australia's assurance and platform governance standards into multilateral interoperability frameworks, enabling shared verification of high-risk AI systems across democratic partners.

- PM&C matures cross-government governance, annual capability reporting and long-term alignment with the National AI Sovereignty Compact.

3.8 Implementation Risks and Mitigation

Implementing sovereign capability in the democratic information environment will encounter structural risks that must be anticipated and managed. Mitigation requires coordinated action across Home Affairs, ASD, DISR, ACMA, the AEC, DFAT, Education, DEWR and PM&C.

Risk	Mitigation
Regulatory capacity gap – ACMA and the AEC lack the technical expertise to audit high-risk algorithms, verify platform disclosures or assess AI enabled political communication independently.	ASD led placements, shared assurance infrastructure and long-term investment in specialist regulatory expertise convert regulatory intent into demonstrable, enforceable accountability.
Platform resistance – transparency, disclosure and audit requirements will face pushback from major platforms with misaligned commercial incentives, and firms that have historically mobilised foreign government authority when challenged.	Phased implementation, alignment with emerging international standards and clear domestic enforcement mechanisms ensure transparency obligations are credible and durable.
Synthetic media trust fragility – high-impact fabrications targeting public officials or electoral processes may erode trust before verification systems can intervene.	ASD and PM&C establish rapid response protocols, national authenticity infrastructure and coordinated communication channels capable of containing synthetic media incidents at speed.
International governance fragmentation – domestic regulatory capacity cannot fully mitigate vulnerabilities originating offshore or embedded in foreign governed platforms.	DFAT leads coordinated engagement with democratic partners to shape interoperable assurance standards, influence operation reporting frameworks and cross-border transparency obligations, ensuring global governance gaps do not become domestic sovereignty risks.

Capability or Dependency: Australia's Choice in the AI Era

Artificial intelligence has become a structural determinant of national power. It shapes how states defend themselves, how economies coordinate, how infrastructure operates and how democratic institutions maintain legitimacy. In this environment, sovereignty is no longer defined by territorial control, military capability or economic scale alone. It is defined by whether a nation can design, govern and operate the computational systems that now underpin every domain of national life.

Australia's current position is one of structural dependency. The systems that increasingly determine national resilience, the cloud platforms hosting essential services, the AI models supporting defence and intelligence, the data pipelines sustaining critical infrastructure, the algorithms mediating public discourse, are governed by foreign jurisdictions, foreign commercial priorities and foreign political decisions.

This report has argued that addressing this condition requires a national capability agenda grounded in industrial strategy, not technology policy. Sovereign compute, sovereign data governance, sovereign model assurance, domestic integration capacity and workforce pipelines are not optional enhancements, they are the industrial foundations on which national resilience in the AI era depends. Without them, Australia cannot guarantee operational continuity in a crisis, cannot secure essential services against foreign governed disruption, and cannot protect the integrity of its democratic institutions from externally enabled manipulation.

The ten recommendations in this report form a coherent and coordinated response. A National AI Capability and Security Strategy sets the strategic frame. A National AI Sovereignty Compact aligns government, industry and labour behind shared capability objectives. Mandatory data sovereignty requirements, sovereign cloud and compute, and a national AI assurance framework address the infrastructure and verification foundations. Strategic procurement reform and a national workforce pipeline build the domestic industrial and human capacity on which all other capability depends. Platform governance, public sector accountability, and AI risk management address the democratic and accountability dimensions. A sovereign AI R&D program and an economic transition and labour compact ensure that the capability built is both technically durable and broadly shared.

The coordination architecture assigns clear responsibilities across PM&C, ASD, DISR, Finance, Defence, Home Affairs, AGD, DFAT, Education, DEWR, ACMA and the AEC. It ensures that sovereign AI capability is developed, governed and deployed in a way that is strategically aligned, operationally secure and democratically accountable.

Australia stands at a decisive moment. The choices made in the coming years will determine whether we remain a downstream consumer of foreign AI systems, subject to external political decisions we cannot influence and cannot escape, or whether we build the sovereign industrial capability required to defend our interests, strengthen our institutions and secure our democratic future. Sovereignty in the AI era is not isolation. It is the ability to operate on our own terms, under our own laws, in alignment with our national interest. This report sets out the pathway. The question is whether Australia will take it.

Policy Recommendations

The following recommendations translate the report's analysis into a coordinated national capability agenda.

Recommendation 1: National AI Capability and Security Strategy

That the Australian Government adopt a whole-of-government National AI Capability and Security Strategy, led by PM&C with Defence, Home Affairs, DISR and Finance, treating AI as strategic infrastructure. The Strategy must set measurable, time-bound capability targets across sovereign compute, data governance, workforce, model assurance and supply chain resilience, with annual public reporting to Cabinet and a designated accountability framework specifying which agency is responsible for each outcome. PM&C is responsible for ensuring this Strategy functions as a live governance instrument: annual capability reporting to Cabinet, portfolio level assessments before each budget process, and a biennial independent review of the National AI Risk Register against which progress is measured.

Recommendation 2: National AI Sovereignty Compact

That the Australian Government establish a National AI Sovereignty Compact as a standing tripartite body of government, industry and unions, to coordinate investment, standards, procurement settings, workforce pipelines, and deployment governance for sovereign AI capability across defence, critical infrastructure, and the democratic information environment. Union participation must extend beyond workforce representation to include co-design of deployment conditions, custodianship of worker data rights, and democratic accountability. Collective bargaining frameworks should be recognised as a legitimate mechanism for operationalising AI governance at the workplace level, and industrial agreements incorporating rights to explanation, advance notice of technological change, and independent review processes are part of the Compact's delivery architecture, not peripheral to it.

Recommendation 3A: Mandatory Data Sovereignty Requirements

That the Australian Government, led by Home Affairs under the Security of Critical Infrastructure Act and in partnership with AGD, introduce mandatory data sovereignty requirements for high-risk sectors, incorporating a tiered classification framework distinguishing at minimum between defence classified, critical infrastructure operational, and sensitive government data; domestic jurisdictional control over sensitive datasets; and enforceable restrictions on offshore access, including where 'localised' data remains subject to foreign legal compulsion. The framework must provide explicit carve-outs and protocols for data sharing obligations under AUKUS and Five Eyes arrangements that preserve allied interoperability without undermining domestic sovereignty protections.

Recommendation 3B: Sovereign Cloud and Compute Program

That the Australian Government, led by DISR in coordination with Defence and Home Affairs, co-invest in trusted domestic cloud and high-performance compute capacity to support sovereign AI development and deployment, prioritising Australian operational control, security assurance, and long-term compute supply resilience for strategic workloads across defence and essential services. The program must define capacity benchmarks calibrated against modelled demand, use allied secure cloud arrangements as an interim measure while domestic capacity scales, and publish transition timelines for migrating high-risk workloads off foreign governed platforms.

Recommendation 3C: National AI

Assurance Framework

That the Australian Government, led by ASD in coordination with Defence, DISR and Home Affairs, establish a national AI assurance framework for high-risk digital and AI enabled systems across government, defence and critical infrastructure, incorporating mandatory sovereign assurance requirements, continuous verification, independent red-teaming, periodic recertification, and enforceable remediation obligations. Zero-trust security principles should be adopted as one component of the broader assurance architecture. The framework must achieve interoperability with allied and international standards including the National Institute of Standards and Technology (NIST) and European Union (EU) AI Act, and AUKUS Pillar II assurance standards, to ensure Australian firms operating across jurisdictions do not face conflicting compliance obligations.

Recommendation 4: Strategic Procurement Reform

That the Australian Government use defence and public procurement as a market shaping tool to expand domestic AI, advanced manufacturing and systems integration capability, including explicit local capability outcomes, supply chain security requirements, and pathways for Australian firms to scale. Finance must enforce open technical standards, independent evaluation mechanisms and transparent performance benchmarks to ensure contestability. The Government should seek legal advice on navigating WTO and free trade agreement obligations and publish guidance on the permissible scope of local capability preferences under Australia's international trade commitments.

Recommendation 5A: Immediate AI Workforce Response (0–2 Years)

That the Australian Government implement an immediate response to acute sovereign AI workforce shortages through: targeted skilled migration pathways for cleared AI engineers, systems integrators and model assurance specialists; industry partnership programs that embed Australian workers in sovereign capable domestic firms; and streamlined security clearance pipelines for defence and intelligence AI roles. This response should be developed with Defence, ASD and Home Affairs and be operational within 12 months.

Recommendation 5B: Long-Term AI Workforce and Skills Pipeline (2–7 Years)

That the Australian Government, led by Education and DEWR in alignment with DISR and Defence, establish a structural national AI skills and workforce pipeline including universal AI literacy programs, portable training access, expanded vocational micro-credentials, and targeted defence and cyber pathways developed with unions, industry and education providers. The pipeline should be scaled and costed against modelled long-term demand and include national accreditation standards for AI security professions. The pipeline should be delivered through a National AI Skills Academy, bringing together organised labour, research institutions, and defence and critical infrastructure operators to embed unions within training design and delivery. Micro-credentials should adopt a stackable architecture, portable across jurisdictions, nationally standardised, and accessible in regional areas, allowing workers to build expertise progressively without requiring full-qualification enrolment at each stage.

Recommendation 6A: Platform Governance and Transparency Framework

That the Australian Government establish a statutory platform governance framework, administered by ACMA, requiring platforms to disclose how recommendation algorithms operate; report on influence operations and coordinated inauthentic behaviour; enable independent audit of high-risk algorithms; and share relevant data with Australian authorities. Enforcement mechanisms must be credible, phased implementation should align with emerging international standards, and the framework should include provisions

for platforms that mobilise foreign government authority to resist Australian regulatory oversight.

Recommendation 6B: Public Sector AI Accountability Framework

That the Australian Government implement a National Framework for Public Sector AI Accountability, administered by AGD, incorporating mandatory disclosure and auditing of AI assisted government decision-making, strengthened appeal rights for individuals subject to automated decisions, enforceable transparency requirements for AI use in politically sensitive contexts, and periodic independent review of automated decision systems across Commonwealth agencies. As a further component, that the Australian Government establish a Sovereign AI Forensic Capability, a dedicated public sector body modelled on France's PEReN, providing independent technical expertise to audit algorithmic systems on behalf of ACMA, the AEC, and sectoral regulators. Sectoral regulators retain legal mandate and enforcement authority; the forensic body conducts the technical work: auditing closed systems, identifying bias, and producing admissible findings. This separation of legal authority from technical execution is the governing design principle. The body should align with Group of Seven (G7) forensic standards to enable regulatory tooling and audit methodology sharing with allied democratic partners.

Recommendation 6C: National AI Risk Register

That the Australian Government establish a National AI Risk Register, owned by PM&C and technically maintained by ASD, to track cross-government dependencies, vulnerabilities and emerging threats from the use of foreign governed AI systems. The register must map exposure across defence, intelligence, critical infrastructure and public administration; identify high-risk dependencies on foreign models, cloud platforms and supply chains; define escalation protocols specifying risk thresholds that trigger Cabinet briefing; and integrate with existing SOCI sector risk registers. PM&C must publish an annual declassified summary of Australia's AI sovereignty posture.

Recommendation 7: AUKUS AI Capability Integration

That the Australian Government develop a dedicated AUKUS AI Capability Integration Strategy, coordinated by Defence and PM&C, to ensure Australia's sovereign AI capability agenda is aligned with and mutually reinforcing of AUKUS Pillar II advanced capabilities work. The strategy should:

- Embed Australian entities at the model development layer of Pillar II AI programs, not merely as downstream integrators, securing sovereign ownership of intellectual property, training data and assurance rights.

- Develop shared assurance protocols within AUKUS for model verification, adversarial testing and data sovereignty protections, with Australia driving the design of these standards rather than accepting allied defaults.

- Negotiate Five Eyes access arrangements that preserve sovereign control over Australian sourced data and ensure model outputs affecting Australian operations are auditable under Australian law.

- Align DISR sovereign compute and model assurance investments with Pillar II technical requirements and embed AUKUS integration requirements into Defence capability planning from the outset.

- Position Australia as a regional provider of governance aligned AI capability, exporting assurance standards and sovereign cloud frameworks to Indo-Pacific democratic partners.

Recommendation 8: International AI Governance Engagement Strategy

That the Australian Government develop and

resource an International AI Governance Engagement Strategy, led by DFAT in coordination with DISR, Home Affairs and PM&C, to ensure Australia actively shapes international AI norms rather than responding to them. The strategy should:

- Pursue leadership roles in key multilateral governance forums, including the Organisation for Economic Co-operation and Development (OECD) AI Policy Observatory, Group of Twenty (G20) AI Principles, the United Nations Advisory Body on AI, and the Global Partnership on AI.

- Negotiate equivalence or mutual recognition arrangements with the EU AI Act and aligned jurisdictions to reduce compliance fragmentation for Australian firms.

- Advocate for international norms on AI sovereignty, data compulsion limits – including constraints on extraterritorial legislation such as the CLOUD Act – and platform governance transparency that reflect Australia's national interest.

- Use bilateral trade and technology agreements to build AI supply chain resilience and diversify access to frontier capability beyond current US concentrated dependencies.

- Develop a regional democratic AI governance network across the Indo-Pacific, anchored in Australian assurance standards and audit capability, with DFAT embedding AI sovereignty objectives into the Pacific Step-up, ASEAN engagement, and Quad partnerships.

Recommendation 9: Sovereign AI Research and Development Investment

That the Australian Government establish a National Sovereign AI R&D Program, coordinated by DISR and the National AI Centre, to build applied research, evaluation and pre-competitive development capacity underpinning long-term sovereign capability. The program should:

- Fund applied AI research in priority sovereign capability domains, defence AI, critical infrastructure resilience, content verification, and model assurance, through CSIRO, university research centres and the Defence Science and Technology Group.

- Establish a national AI evaluation and benchmarking capability to independently assess the safety, security and fitness-for-purpose of AI systems used in government and critical infrastructure, reducing dependence on vendor self-assessment.

- Support pre-competitive R&D partnerships between industry, universities and government that generate intellectual property remaining under Australian ownership.

- Target domain specific applied capability, evaluation infrastructure, and the human capital required to adapt, audit and operate AI systems in sovereign contexts, not frontier general-purpose model development.

Recommendation 10: AI Economic Transition and Labour Compact

That the Australian Government, through the National AI Sovereignty Compact (Recommendation 2) and led by DEWR, develop a dedicated AI Economic Transition and Labour Compact to manage the workforce and social licence dimensions of AI adoption. The Compact should:

- Establish sectoral transition frameworks, developed with unions, industry and government for industries facing significant AI driven disruption, including financial services, telecommunications, logistics and public administration.

- Require large employers deploying AI systems with material workforce impacts to notify and consult with affected workers and their representatives before deployment.

- Fund portable retraining entitlements for workers displaced by AI adoption, accessible regardless of employer or industry sector.

- Build the social licence for sovereign AI capability by demonstrating that adoption benefits are broadly shared and that workers have a genuine voice in how AI is deployed in their workplaces.

Endnotes

1 Kelley M. Sayler, *Pentagon-Anthropic Dispute over Autonomous Weapon Systems: Potential Issues for Congress*, CRS Insight No. IN12669 (Washington, DC: Congressional Research Service, April 21, 2026), accessed May 21, 2026, https://www.congress.gov/crs-product/IN12669

[1] Kevin Honglin Zhang, *Geoeconomics of US–China Tech Rivalry and Industrial Policy, Asia and the Global Economy* 4, no. 2 (2024): 100098, https://doi.org/10.1016/j.aglobe.2024.100098

2 *Data Beyond Borders: Australian Data Stored in Non-Australian Cloud Environments*, Australian Cyber Security Magazine, 2024, https://australiancybersecuritymagazine.com.au/data-beyond-borders-australian-data-stored-in-non-australian-cloud-environments/

The Chinese Communist Party (CCP): A Quest for Data Control, Center for Internet Security, 2023, https://www.cisecurity.org/insights/blog/the-chinese-communist-party-ccp-a-quest-for-data-control

3 Australian Strategic Policy Institute, *The Strategist*, "AI is Changing Indo-Pacific Naval Operations," by Jihoon Yu, 23 April 2025, https://www.aspistrategist.org.au/AI is-changing-indo-pacific-naval-operations/

4 Alex Vakulov, "Embedded AI in Military Drones Is Redefining Autonomy and Operations," *Institute for Defense and Government Advancement (IDGA)*, accessed May 21, 2026, https://www.idga.org/government-defense-it-communications/articles/embedded-AI in-military-drones-is-redefining-autonomy-and-operations

5 Maj. Sharlene Tilley, "Smart Logistics: Navigating the AI Frontier in Sustainment Operations," *The United States Army*, October 17, 2024, https://www.army.mil/article/280377/smart_logistics_navigating_the_ai_frontier_in_sustainment_operations

6 Svenja Kirsch and Bethan Saunders, *Addressing Russian and Chinese Cyber Threats: A Transatlantic Perspective on Threats to Ukraine and Beyond*, Belfer Center for Science and International Affairs, Harvard Kennedy School, May 2023, https://www.belfercenter.org/sites/default/files/pantheon_files/files/publication/HKS_RussianChineseCyberThreats_230526.pdf

7 Alessandro Arduino, *The Cyber Arsenal: North Korea's Silent Power*, Istituto per gli Studi di Politica Internazionale (ISPI), 2026, https://www.ispionline.it/en/publication/the-cyber-arsenal-north-koreas-silent-power-229242

Cyber Operations as Iran's Asymmetric Leverage, The Soufan Center, IntelBrief, 17 March 2026, https://thesoufancenter.org/intelbrief-2026-march-17/

8 Maaike Roelofs and Karel van den Bosch, *Intelligent Decision Support and Meaningful Human Control in Multi-Domain Operations* (Brussels: NATO Science and Technology Organization, 2024), https://www.sto.nato.int/document/intelligent-decision-support-and-meaningful-human-control-in-multi-domain-operations/

9 Matthew Jones, "The Profession of Arms in an AI Enabled World," *Australian Army Journal* (Early Access), Australian Army Research Centre, DOI: 10.61451/220105, accessed May 21, 2026, https://researchcentre.army.gov.au/library/australian-army-journal-aaj/early-access-articles/profession-arms-AI enabled-world

10 Kevin Honglin Zhang, *Geoeconomics of US–China Tech Rivalry and Industrial Policy, Asia and the Global Economy* 4, no. 2 (2024): 100098

11 Akintunde Akinbo, *Demystifying the Black Box: Responsible AI Governance Without the Jargon – Why This Conversation Matters*, SSRN Working Paper, 17 March 2026 (written 26 July 2024), https://ssrn.com/abstract=6178958

12 Google Cloud, "Distillation, Experimentation and Integration: AI Adversarial Use," *Google Cloud Blog*, February 2026, https://cloud.google.com/blog/topics/threat-intelligence/distillation-experimentation-integration-AI adversarial-use

13 Akharchaf, Youssef, *Zero Trust Architecture in Cloud Security: Challenges and Implementation*, 2025, https://doi.org/10.13140/RG.2.2.26671.04000

William Yeoh, et al., *Zero Trust Cybersecurity: Critical Success Factors and a Maturity Assessment Framework*, Computers & Security 133 (2023): 103412, https://doi.org/10.1016/j.cose.2023.103412

14 Shantanu Pal, Geeta Yadav, Zahra Jad-

idi, Ahsan Habib, Md. Palash Uddin, Chandan Karmakar, and Sandeep Shukla, "Vulnerabilities in Machine Learning for Cybersecurity: Current Trends and Future Research Directions," *Journal of Information Security and Applications* 96 (2026): 104269, https://doi.org/10.1016/j.jisa.2025.104269

15 Australian Cyber Security Centre, *Artificial Intelligence and Machine Learning: Supply Chain Risks and Mitigations* (Canberra: Australian Government, first published 16 October 2025, updated 5 March 2026), https://www.cyber.gov.au/business-government/secure-design/artificial-intelligence/artificial-intelligence-and-machine-learning-supply chain-risks-and-mitigations

16 Andrew Horton, "Mitigating Australia's Cloud-Computing Risks Is Still Work in Progress," *The Strategist* (Australian Strategic Policy Institute), 28 October 2024, https://www.aspistrategist.org.au/mitigating-australias-cloud-computing-risks-is-still-work-in-progress/

17 Ibid

18 Luke Eason, "Is Foreign Interference a Blind Spot in Your Technology Supply Chain?," *BDO Australia*, 18 August 2025, https://www.bdo.com.au/en-au/insights/digital-technology/is-foreign-interference-a-blind-spot-in-your-technology-supply chain

19 Australian Computer Society, "A Single System Failure Can Take Down an Industry," *Information Age*, 3 December 2025, https://ia.acs.org.au/article/2025/a-single-system-failure-can-take-down-an-industry.html

20 Persistence Market Research. (2026). *AI in energy distribution market size, share, and growth forecast, 2026–2033.*

21 RAND Corporation, "The U.S. Power Grids Need AI Disclosure Requirements," *RAND Commentary*, June 2025, https://www.rand.org/pubs/commentary/2025/06/the-us-power-grids-need-AI disclosure-requirements.html

22 Antonio Moreno-Rodenas, Koen Verbist, Annelies Mertens, et al., *Applications of Artificial Intelligence for Water Management*

23 'AWS Outage Exposes Global Supply Chain Vulnerabilities' on Supply Chain World Magazine (21 October 2025) https://scw-mag.com/news/aws-outage-exposes-global-supply chain-vulnerabilities/

24 Asad Alam, Asif Umer, Insaf Ullah, and Ahmed Alsayat, "AI Enabled Cybersecurity Framework for Future 5G Wireless Infrastructures," *Scientific Reports* 16, no. 1 (2026): 7055, https://doi.org/10.1038/s41598-026-37444-8

[34] Australian Communications and Media Authority, *AI in Telecommunications* (Canberra: Australian Communications and Media Authority, April 2026),

25 Wiggins, J. (2025, April 22). It's like a marriage inside Telstra and Accenture's AI reinvention. Australian Financial Review. https://www.afr.com/companies/telecommunications/it-s-like-a-marriage-inside-telstra-and-accenture-s-AI reinvention-20250422-p5ltf1

26 Commonwealth Bank of Australia. (2026, February 5). *CommBank releases Australian-first report outlining how it is adopting AI.* https://www.commbank.com.au/articles/newsroom/2026/02/cba-approach-to-adopting-AI report-announcement.html

27 Varnosfaderani, S. M., & Forouzanfar, M. (2024). *The role of AI in hospitals and clinics: Transforming healthcare in the 21st century.* Bioengineering, 11(4), 337.

28 Elisa Industriq. (2026, February 3). *The future of food supply chains: Innovation beyond 2026.*

29 Workplace Emergency Management. (2025, October 22). *How AI is transforming emergency response in Australia.*

30 *Sydney Morning Herald, For every $100 in data centres, $80 leaves Australia almost immediately*, 31 March 2026, available at: https://www.smh.com.au/technology/for-every-100-in-data-centres-80-leaves-australia-almost-immediately-20260331-p5zkb5.html

31 Knight, E. (2026, April 9). The race to the bottom on AI job cuts has already started. The Sydney Morning Herald.

[34] Eyers, J. (2026, February 23). CBA sets out plan to retrain workers in AI age and as job cuts loom. Australian Financial Review.

32 Wiggins, J. (2025, April 22). It's like a marriage inside Telstra and Accenture's AI reinvention. Australian Financial Review.

33 Romanishyn, A., Malytska, O., & Goncharuk, V. (2025). *AI driven disinformation: Policy recommendations for democratic resilience.* Frontiers in Artificial Intelligence. https://doi.org/10.3389/frai.2025.1569115

34 Guimaraes e Silva, L. C. (2025). *Deceptive influencing as a harm to democratic politics: Warding off political manipulation through micro-targeting with a limited right not to be profiled.*

SSRN. https://doi.org/10.2139/ssrn.5188504
35 Australian Strategic Policy Institute. (2025, February 6). *Safeguarding Australian elections: Addressing AI enabled disinformation* [Video]. YouTube. https://www.youtube.com/watch?v=t-Sa8tG1vgUo
36 Romanishyn, A., Malytska, O., & Goncharuk, V. (2025). *AI driven disinformation: Policy recommendations for democratic resilience*. Frontiers in Artificial Intelligence.
37 Dan, V. (2025). *Deepfakes as a democratic threat: Experimental evidence shows noxious effects that are reducible through journalistic fact checks*. The International Journal of Press/Politics. Advance online publication. https://doi.org/10.1177/19401612251317766
38 Australian Communications and Media Authority. (2020). *Misinformation and news quality on digital platforms in Australia: A position paper to guide code development*. https://www.acma.gov.au/sites/default/files/2020-06/Misinformation%20and%20news%20quality%20position%20paper.pdf
39 Guimaraes e Silva, L. C. (2025). *Deceptive influencing as a harm to democratic politics: Warding off political manipulation through micro-targeting with a limited right not to be profiled*.
40 Romanishyn, A., Malytska, O., & Goncharuk, V. (2025). *AI driven disinformation: Policy recommendations for democratic resilience*. Frontiers in Artificial Intelligence.
41 Australian Strategic Policy Institute. (2025, February 6). *Safeguarding Australian elections: Addressing AI enabled disinformation*
42 United Nations Development Programme (UNDP). (2025). *AI for the next generation of public services*. Retrieved from https://www.undp.org/sites/g/files/zskgke326/files/2025-12/AI for-the-next-generation-of-public-services.pdf
43 Australian Communications and Media Authority. (2020). *Misinformation and news quality on digital platforms in Australia: A position paper to guide code development*. https://www.acma.gov.au/sites/default/files/2020-06/Misinformation%20and%20news%20quality%20position%20paper.pdf
44 SOMO. (February 2026). *Australia: Big Tech weaponises its platforms against the government*. https://www.somo.nl/australia-big-tech-weaponises-its-platforms-against-the-government/
45 Ibid
46 Australian Communications and Media Authority. (2020). *Misinformation and news quality on digital platforms in Australia: A position paper to guide code development*.
47 Dan, V. (2025). *Deepfakes as a democratic threat: Experimental evidence shows noxious effects that are reducible through journalistic fact checks*.
48 Hurst, D., Kuo, L., & Graham-McLay, C. (2020, September 14). *Zhenhua Data leak: personal details of millions around world gathered by China tech company*. The Guardian. https://www.theguardian.com/world/2020/sep/14/zhenhua-data-full-list-leak-database-personal-details-millions-china-tech-company
49 ACMA, *New ACMA Powers to Combat Misinformation and Disinformation*, Department of Infrastructure, 2023
50 ACMA, *Third Report on Digital Platforms' Efforts Under the Australian Code of Practice on Disinformation and Misinformation*, September 2024
51 Rizzi et al., "Combatting Mis and Disinformation in the Digital Age," *Australasian Journal of Information Systems*, 2025
52 Koziol, "After the Robodebt Disaster, We Risk Making the Same Mistakes," *The New Daily*, April 2026; and Australian Public Law, "Regulating Automated Government Decision-Making," November 2025

John Curtin Research Centre Governance

BOARD

Sam Almaliki (Chair)
Hon. Annastacia Palaszczuk AC
Priya Brown
Gerard Dwyer
Lewis Hamilton
Jenny Hill
Dr Shireen Morris
Kosmos Samaras
Stacey Schinnerl
Sebastian Zwalf
Dr Henry Pinskier (immediate past chair)

ADVISORY COUNCIL
Professor the Hon Geoff Gallop AC | Former Western Australian Premier (Co-Chair)
Hon Morris Iemma | Former NSW Labor Premier
Hon Kim Carr | Former Labor Senator
Hon Stephen Conroy| Former Labor Senator
Dr Mike Kelly AM | Former Labor MP
Senator Deborah O'Neill
Hon Peter Khalil MP
Hon Daniel Mulino MP
Michael Easson AM | Former Secretary NSW Labor Council
Geoff Fary | Former Assistant Secretary, Australian Council of Trade Unions
Josh Peak | Secretary, SDA South Australia & Northern Territory
Helen Cooney | Shop, Distributive and Allied Employees' Association
Ari Suss | Linfox Advisory
Lord Maurice Glasman | UK Labour Party
Jon Cruddas MP | UK Labour Party