



Australian Government
Department of Industry,
Science and Resources

Ministerial Submission
MS25-001574

FOR ACTION – Establishing the Artificial Intelligence Safety Institute

TO: Minister for Industry and Innovation; Minister for Science
CC: Assistant Minister for Science, Technology and the Digital Economy

TIMING

Important: by 8 December 2025 to progress work to establish the AI Safety Institute.

Recommendation

• That you note the department’s approach to establishing the AI Safety Institute (see Attachment A).

Noted / Please discuss

s 47C

Noted / Please discuss

Signature	Date: / /2025
-----------	---------------------

KEY POINTS

- s 34(3) s 47C
- The Albanese Government will announce its decision to establish the AISI on 25 November 2025. This will reaffirm Australia's commitment to collaborate with international partners on AI safety as a signatory to the Bletchley Declaration, the Seoul Declaration, and the Paris Statement.
- The National AI Plan is scheduled for release on 2 December 2025. The National AI Plan will announce the Government's position on the regulation of AI with the establishment of the AISI a key element of Australia's regulatory approach to AI.
- The department is well-progressed in its planning for the AISI, including structure and governance arrangements, procurement of information, communications and technology (ICT) systems, and recruitment s 47C

Announcing the Government's decision to establish an Australian AISI

s 34(3)

- - and advice from your office.
- At your office's request, we provided a list of stakeholders with potential interest in the AISI, a draft media release and draft talking points on 20 and 21 November 2025 to support the announcement of the AISI.
- The announcement will take place in AI Week and coincides with the Australian Industry Group's (Ai Group) 'AI in Action-Shaping the Future of Australian Industry' event at the National Press Club.

s 47C

s 47C

Approach to establishing the AISI

- The AISI Establishment Plan Overview (see [Attachment B](#)) highlights the priority actions planned to operationalise the AISI. This includes organisational design, staffing and recruitment, product development, ICT (testing) capability, stakeholder engagement, and communications.
- Some elements of the AISI are in place, including foundational elements of its upstream risk and downstream harm monitoring functions, as well as its international AI safety engagement function. Remaining components will become operational as procurement of ICT systems and merit-based recruitment are completed. These processes will take several months, noting that progress will slow down over the December–January holiday period. Some of the more complex elements, such as the commencement of in-house testing capability for frontier AI, will be operational by 30 April 2026.
 - Recruitment for AISI positions will follow established procedures for APS employees and, once employed, they will be bound by the requirements that apply to all APS employees, including in relation to security and conflicts of interest.

s 33(a)(i)

- ICT procurement will be undertaken in accordance with the rules and guidance set out by the Digital Transformation Agency. Processes will adhere to the requirements of the *Public Governance, Performance and Accountability Act 2013* (PGPA Act) and the Protective Security Policy Framework. Technology and Digital Policy Division is working closely with CIO Division on this process.
- Engagement with international partners, industry, academia, and across government to build awareness of the AISI will commence immediately. This will be strengthened by Australia's participation in the upcoming meeting of the International Network of AI Safety Institutes on 4–5 December 2025, held on the sidelines of the Neural Information Processing Systems (NeurIPS) conference in San Diego, USA.

Progressing priority AI safety workstreams

- The department will immediately progress several priority AI safety workstreams to maintain momentum between the announcement of the AISI and its official commencement.
 - The department will continue work with the International Network of AI Safety Institutes, deliver AI safety research, and monitor emerging AI capabilities and risks.

s 47C

s 33(a)(i)

s 47C

s 47C

CONSULTATION

- National AI Centre; Chief Information Officer Division; Chief Operating Officer Division.

Clearance Officer

Simon Writer

Head of Division

Technology and Digital Policy Division

Mob: s 22(1)(a)(ii)

Clearance Date: 24 November 2025

Contact Officers:

s 22(1)(a)(ii)

Manager, AI Law Design and Policy

s 22(1)(a)(ii)

Manager, AI Safety and Programs

s 22(1)(a)(ii)

ATTACHMENTS

s 47C

B. AISI Establishment Plan Overview

s 33(a)(i)

ATTACHMENT B

AI SAFETY INSTITUTE ESTABLISHMENT PLAN OVERVIEW

Objective

Establish the AI Safety Institute (AISI) within the department, delivering a trusted, technically rigorous capability for monitoring, testing, and facilitating information sharing on emerging AI capabilities, risks and harms.

Timeline & Key Milestones

AISI activities: From set-up to launch	Dependencies / Risks
Week 1 - 24 November	
<i>AISI public announcement 25 November</i>	
- Brief Minister and Assistant Minister on AISI set up approach - Establish AISI establishment Project Board - Establish AISI project management team - Develop AISI establishment plan - Recruitment planning (including SES B1) - AISI talking points (including holding lines ahead of announcement) - Prioritise BAU and AISI activities	s47C
Month 1 - December	
<i>National AI Plan announcement 2 December</i>	
- s47C - Targeted external SES engagement on AISI announcement - Co-design AISI deliverables	s47C
s47C	
- Begin onboarding technical and non-technical staff - Set up ICT infrastructure - Establish AI testing capability - s47C	Recruitment interest Technical onboarding IT procurement Collaboration across DISR (e.g. CIOD, HR, TDPD)
s47C	
- Finalise onboarding of technical staff - Operationalise AI testing capability	ICT operationalisation and security risks s47C

Key Activities

- Governance
 - Establish the AISI establishment Project Board.
 - Define scope, objectives, risk, assumptions, and key performance indicators.
 - Design monitoring and evaluation framework.
 - Confirm budget, resources, roles and responsibilities.
 - Identify support roles across HR, IT, Communications, and Property.
- Operationalisation
 - Recruit technical and non-technical staff, prioritising technical experts to begin assessing the risks and capabilities of frontier AI models.
 - Recruit SES Band 1
 - Procure and implement ICT.
- Engagement and communication
 - Align engagement and communication activities with the implementation of the National AI Plan.

s47C



Australian Government
Department of Industry,
Science and Resources

Ministerial Submission

MS26-000076

FOR INFORMATION - Update on establishing the AI Safety Institute Branch

TO: Assistant Minister for Science, Technology and the Digital Economy

CC: Minister for Industry and Innovation; Minister for Science

TIMING

Routine: 23 February 2026

Recommendation

- That you **note** the direction of, and progress made toward, establishing the Australian AI Safety Institute as a new branch within the Department of Industry, Science and Resources (DISR), with operational capability expected from April 2026.

Noted / Please discuss

Signature

Date: / /2026

KEY POINTS

- In the 2025-26 MYEFO, the department received \$29.8 million over 4 years from 2025-26 to establish an AI Safety Institute function, and a further \$7.9 million ongoing from 2029-30 to sustain it. This is a central element of the National AI Plan's "Keeping Australians Safe" pillar. This submission provides an update on progress toward establishing the AI Safety Institute (see [Attachment A](#)).

Objectives, functions and structure

- The AI Safety Institute will support the Government, agencies and regulators to maintain safety measures, laws and regulatory frameworks that keep pace with rapid technological change.
- The AI Safety Institute will be constituted as a branch in Technology and Digital Policy Division. It will pursue 3 strategic objectives and perform 3 core functions that support coordinated action to reduce avoidable harm from advanced AI, enable better deployment decisions and shape national and international AI governance.

- **Objective 1:** Enable government to understand emerging developments in AI and their potential implications.
 - **Function 1:** The AI Safety Institute will monitor and analyse emerging AI capabilities, risks and real-world impacts, as well as emerging technical, legal, regulatory and institutional governance approaches to risk mitigation. This will include technical evaluations of advanced AI systems and collaboration with the International Network for Advanced AI Measurement, Evaluation and Science.
- **Objective 2:** Support government agencies and regulators to respond to emerging AI-related risks and harms in a coherent and coordinated way.
 - **Function 2:** The AI Safety Institute will share insights and evidence on emerging AI capabilities, risks, harms and mitigation strategies to support informed decision-making across government. It will track how responsible agencies and regulators implement mitigation actions and help ensure no risks or harms are overlooked.
- **Objective 3:** Shape approaches to the safe development, deployment and governance of advanced AI systems in line with Australia's interests.
 - **Function 3:** The AI Safety Institute will engage domestically and internationally to advance understandings of, and effective approaches to AI safety and governance and embed learnings and best practice across the system. This will include providing AI safety insights and AI risk mitigation advice for the National AI Centre (NAIC) to translate into accessible industry guidance and public information.
- To deliver these objectives and perform these functions, the AI Safety Institute will be established with a four-team structure (further details at [Attachment B](#)).
 - The *Emerging AI Risks* team will identify risks early by tracking AI developments, upstream emerging risks, and downstream AI harms.
 - The *Technical AI Safety Evaluation* team will provide hands-on technical evaluation and safety-related research on advanced AI systems to identify emerging capabilities, risks and failure modes.
 - The *AI Risk Advisory* team will provide advice to agencies and regulators on best-practices to shape safe AI development, deployment and governance. It will lead domestic and international expert engagement, as well as engagement with the National Intelligence Community.
 - The *AI Regulatory Engagement* team will focus on the AI Safety Institute's relationships across the public sector and manage intake, triage and flow of requests and relevant information between the AI Safety Institute and public sector stakeholders. It will track implementation of mitigation actions taken by responsible regulators and agencies and monitor domestic and international legal and regulatory developments.

- The AI Safety Institute is not a compliance or enforcement body, has no statutory powers and will not offer a public advisory service.
- The AI Safety Institute will drive cross-government engagement on AI safety issues, guidance and technical information sharing. The NAIC will be the primary channel for disseminating the AI Safety Institute's insights on advanced AI risks and safety-relevant capabilities to industry and the public (see [Attachment C](#)). This leverages the NAIC's established networks and communication platforms, ensuring information reaches relevant stakeholders efficiently and consistently.

Progress toward operationalising the AI Safety Institute

s 33(a)(i), s 47C

s 33(a)(i), s 47C

s 47E(d)

s 33(a)(i), s 47C

DATA REFERENCED

- Nil

CONSULTATION

- National AI Centre, Chief Information Officer Division.

Clearance Officer

Simon Writer

Head of Division

Technology and Digital Policy Division

s 22(1)(a)(ii)

Clearance date: 2 February 2026

Contact Officer

Chelle Adamson

A/g General Manager, AI Governance

s 22(1)(a)(ii)

ATTACHMENTS

- A. Progress towards the AI Safety Institute's establishment
- B. AI Safety Institute branch structure and activities
- C. How the AI Safety Institute and National AI Centre will work together
- D. AI Safety Institute workflows
- E. Key messages for information sessions
- F. Engagement plan and information sessions