DESIGNING AUSTRALIA'S AI SAFETY INSTITUTE

EXPERT SURVEY REPORT

JANUARY
2026

Good Ancestors

Published
5 January 2026

Authors
Emily Grundy, Greg Sadler, Luke Freeman

About Good Ancestors
Good Ancestors is an Australian charity dedicated to improving the long-term future of humanity by providing rigorous, evidence-based, and practical policy recommendations for Australia's biggest challenges. We have been deeply engaged in the AI policy conversation since our creation, working with experts around the world and helping to organise Australians for AI Safety.

Acknowledgments
We thank the 139 respondents who contributed their time and expertise to filling in the survey.

Contact
If you would like to discuss the report or propose further research, please let us know at contact@goodancestors.org.au.

Table of Contents

Executive Summary

Good Ancestors surveyed 139 professionals with expertise in AI safety, governance, and related fields to share their views about the establishment of Australia's AI Safety Institute.

Key findings

Autonomous systems and CBRN top priorities: 85.8% rate the AISI working on autonomous systems as critical or very important, followed by dual-use science/CBRN (79.8%), cyber misuse (81.2%), societal resilience (80.6%), and human influence (80.6%).

Support for effort on catastrophic risks: Only 11.2% of respondents thought the AISI should focus mainly on broader AI risks, with 88.8% indicating the AISI should either take a balanced approach or focus on catastrophic risks.

Evaluations and hardware governance are program priorities: Analysis of open-ended responses identified independent model evaluations and red-teaming as a frequently suggested program area, along with hardware verification and governance.

International connections and focus on catastrophic risks matter most for attracting talent: 67.9% cite strong international AISI network connections as a deal-maker for working at the AISI, followed by leadership focused on catastrophic/frontier risks (64.1%), and a focused mandate on catastrophic/frontier risks (60.3%).

Bureaucracy will deter talent: 90% would be deterred from accepting a role at the AISI by bureaucratic culture that prevents impact. Insufficient funding (under $10 million/year) would deter 54.6%, while being too close to industry / regulatory capture concerns would deter 48.5%.

Australia's unique contribution and middle-power status excites experts: Analysis of open-ended responses indicated Australia's geopolitical positioning as a key factor exciting respondents about the Australian AISI. Respondents highlighted Australia's potential to bridge US-China tensions, build Indo-Pacific capacity, and contribute as a trusted, independent actor in global AI safety.

Substantial funding expected: 77.0% recommend an annual budget of $25 million AUD or more for the AISI to "make a meaningful contribution to AI safety".

Methodology

Good Ancestors surveyed 139 professionals with expertise in AI safety, governance, and related fields between 28 November – 17 December, 2025. Almost two-thirds of the sample (64.0%) had 2+ years of professional experience in AI safety, security, governance, or related work.

Respondents' areas of expertise included AI safety research (51.8%), policy and governance (46.0%), technical AI/ML (40.3%), operations/management (33.8%), government and public sector (25.9%), and security fields (cyber security, biosecurity, national security; 25.9%). The majority (68.3%) were Australian citizens.

The survey asked about:

- What the AISI should prioritise and how it should operate
- Who might be a strong fit to work there (including the respondent, if interested, or others they would recommend)
- What would attract suitable candidates to the AISI

See Appendix A for the full survey.

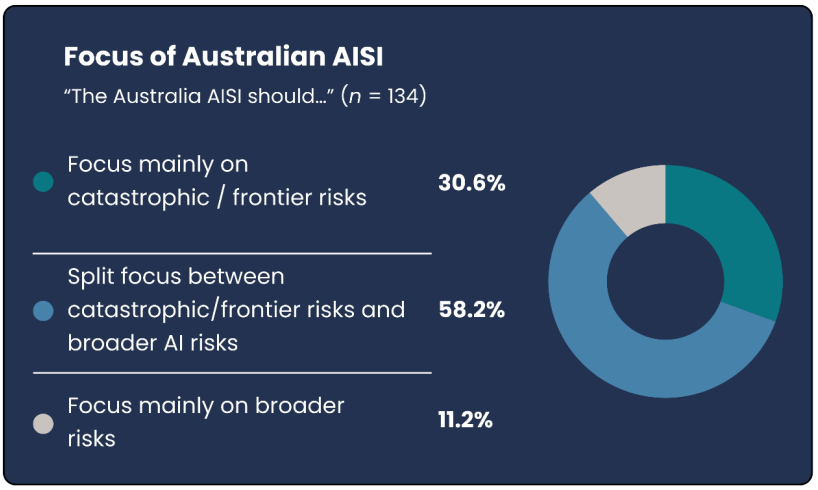
Findings in full

A full breakdown of survey results is provided in Appendix B.

Mission of the AI Safety Institute

Focus on catastrophic/frontier vs broader AI risks

Respondents were asked whether the Australian AISI should focus mainly on catastrophic/frontier risks (e.g., bioweapons, loss of control), broader AI risks (e.g., bias, discrimination, privacy), or split focus evenly. Over half (58.2%) wanted an even split between catastrophic/frontier risks and broader AI risks, while 30.6% preferred a primary focus on catastrophic/frontier risks. Only 11.2% wanted the AISI to focus mainly on broader risks like bias, discrimination, and privacy.

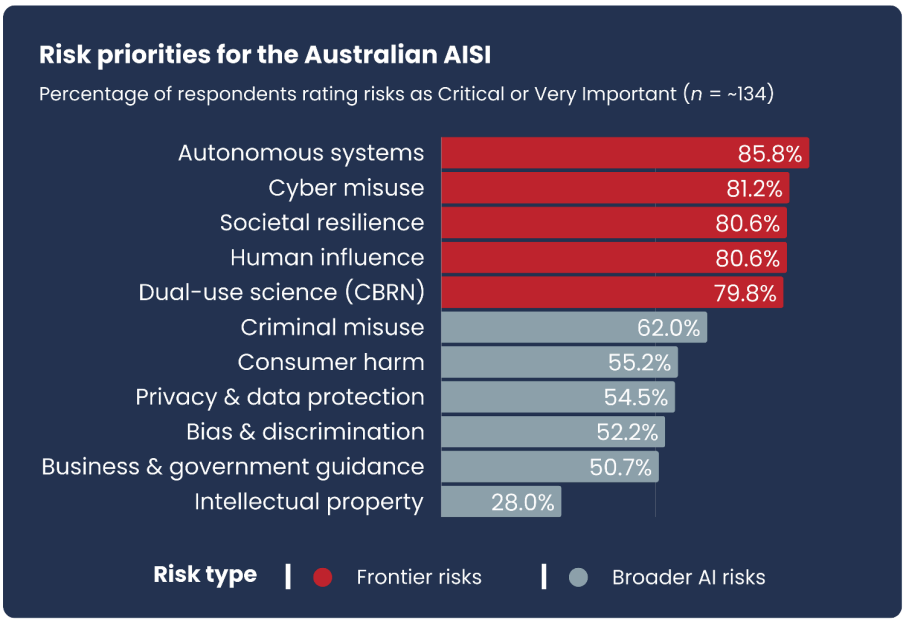


Priority risk areas

Respondents rated how important it was for the Australian AISI to work on various risk areas. **Risk priorities align with catastrophic/frontier focus.**

Addressing autonomous systems was the highest priority, with 85.8% rating it "critical" or "very important". Cyber misuse showed similarly strong support at 81.2%. Dual-use science (CBRN) received 79.8% support in these top categories, while societal resilience and human influence both reached 80.6%.

Broader AI risks were lower priority, like consumer harm (55.2%), privacy and data protection (54.5%), bias and discrimination (52.2%), and intellectual property (28.0%).



Specific programs of work

Respondents were asked what specific programmes of work the Australian AISI should prioritise. Thematic analysis of 107 responses indicated the following:

- **Model evaluations and red-teaming.** Respondents emphasised the need for independent, technically rigorous evaluation environments for pre-deployment testing of advanced models, including testing for dangerous capabilities such as autonomous replication, malicious code generation, and CBRN risks. Respondents noted the value of evaluations conducted in the Australian-context, and real-world testing under degraded conditions.

> *"Evaluate frontier models on benchmarks with Australian contexts, because no one else will do it for us"*

- **Hardware verification and governance.** Respondents identified hardware verification, compute provenance, and supply chain security. This includes developing frameworks for compute attestation, safe procurement standards, hardware security audits, and methods for verifying compute provenance, leveraging Australia's position in critical minerals supply chains and existing organisations (e.g., Data61).

> *"Hardware verification, compute provenance, and supply-chain security, an area of AI safety that remains underdeveloped globally but is critical for long-term governance"*

- **International coordination and regional capacity building.** Respondents highlighted Australia's potential to bridge US-China relations and facilitate international agreements, drawing on its trusted middle-power status and Five Eyes relationships. Responses emphasised building AI safety capacity across the Indo-Pacific and ASEAN region.

> *"Australia's middle power status enables trusted convening role without threatening sovereignty"*

- **Biosecurity and CBRN risks.** Respondents identified dual-use biology, gain-of-function research, DNA synthesis screening, and synthetic pathogen development as potential focus areas for the AISI. Several noted Australia's biosecurity expertise, agricultural pathogen knowledge, and leadership role in the Australia Group as comparative advantages. More broadly, respondents emphasised CBRN (chemical, biological, radiological, nuclear) risks in evaluation contexts.

> *"Australia has world-class biosecurity infrastructure and is a global biomedical leader. Expertise in agricultural pathogens and quarantine systems … position us uniquely to develop evaluation frameworks for AI systems that pose biological design risks."*

- **Technical research and governance development.** Several respondents advocated for interpretability and formal verification research, drawing on Australian strengths in mathematics, theoretical computer science, and formal methods.

> *"Interpretability and formal verification research, a neglected but essential part of safe AI development"*

- **Other programme areas.** Additional priorities mentioned included: deployment monitoring; societal impact assessment including labour market disruption and economic transition; consumer protection leveraging Australia's strong consumer rights traditions; educational programmes for the public and vulnerable groups; defensive acceleration and democratic resilience; and energy-efficient safety compute infrastructure powered by renewable energy.

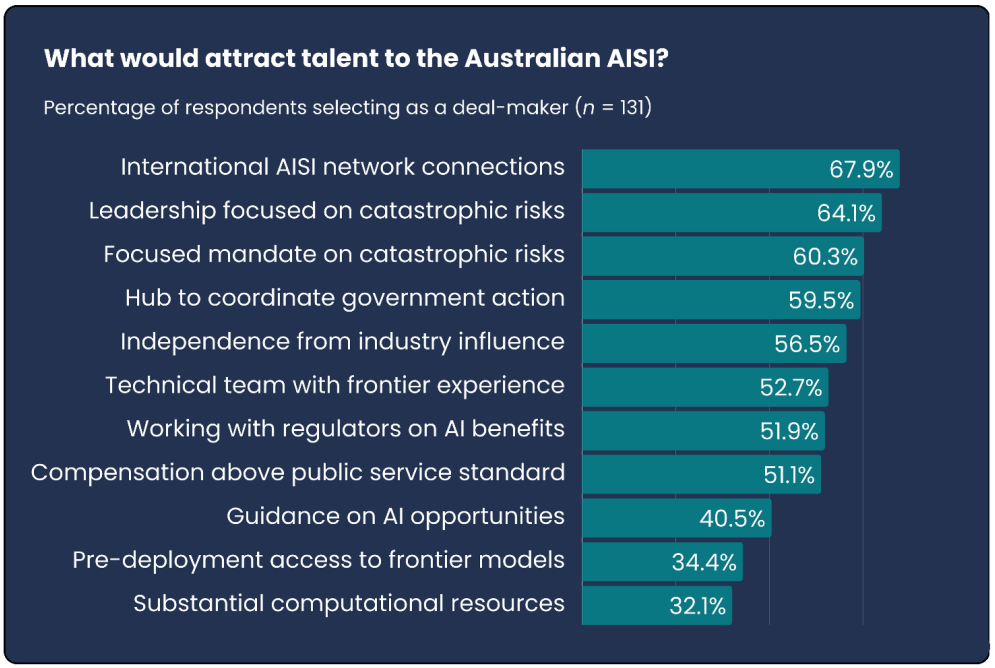
> *"The real risk lives in the weeks and months after deployment, when autonomous systems, agents, and humans co-evolve, drift, collude, and quietly rewire incentives in ways no launch-day impact assessment can see"*

Attracting talent to the AI Safety Institute

Deal-makers

Respondents selected deal-makers—factors that would "strongly pull [them], or someone like [them], toward working at the Australian AISI".

Strong international AISI network connections was the most common deal-maker, with 67.9% selecting this factor. Leadership focused on catastrophic/frontier risks (64.1%) and a focused mandate on such risks (60.3%), were also highly valued.



Respondents were also asked, in an open-text question, what would most attract them, or someone like them, to work at the Australian AISI. Thematic analysis of 112 responses indicated the following:

- **Mission focus and clarity.** Respondents emphasised the importance of a clear, focused mission on catastrophic and frontier AI risks, with credible pathways to impact. Respondents want the Institute to avoid mission drift toward generic AI ethics work or productivity goals that could compromise safety priorities.

> *"Having a strong and clear mission is essential. The organisation needs a well-reasoned theory of change demonstrating how its work genuinely contributes to AI safety"*

- **Leadership quality and technical competence.** Respondents highlighted strong, technically competent leadership as essential, wanting leaders who deeply understand frontier AI risks and can navigate both technical and policy domains. Multiple responses warned against more political appointments, advocating instead for leadership drawn from experts in areas relating to AI, safety, and research.

> *"Most vital: senior leadership who understand frontier AI. I would avoid working there if it's led by a lawyer or policy expert unless they split the role with someone who understands AI and listen to that person often"*

- **Impact and influence.** Respondents wanted the ability to achieve tangible outcomes and avoid "governance theatre". Respondents are seeking genuine influence on policy rather than producing reports without implementation. Many responses emphasised the importance of working on real-world problems with measurable impact, including through international collaborations.

> *"Influence and impact: the opportunity to make a real difference to our ability to navigate and shape the changes ahead."*

- **Independence and decision-making authority.** Respondents emphasised the importance of independence from industry influence and avoiding regulatory capture. Responses mentioned freedom to publish findings, autonomy in research direction, and avoiding bureaucratic constraints.

> *"Independence from Big Tech is paramount. Going forward I expect those companies to try to insinuate themselves as much as possible into AI policy space so a well resourced AISI is really necessary"*

- **Culture and working environment.** Respondents wanted a collaborative, non-hierarchical culture, distinct from traditional public service bureaucracy. They wanted high agency and research autonomy, and the quality of colleagues and intellectual environment matters.

> *"Culture: Collaborative rather than hierarchical, comfortable with uncertainty, valuing practitioners alongside researchers. It needs to be fundamentally different to existing APS hierarchies"*

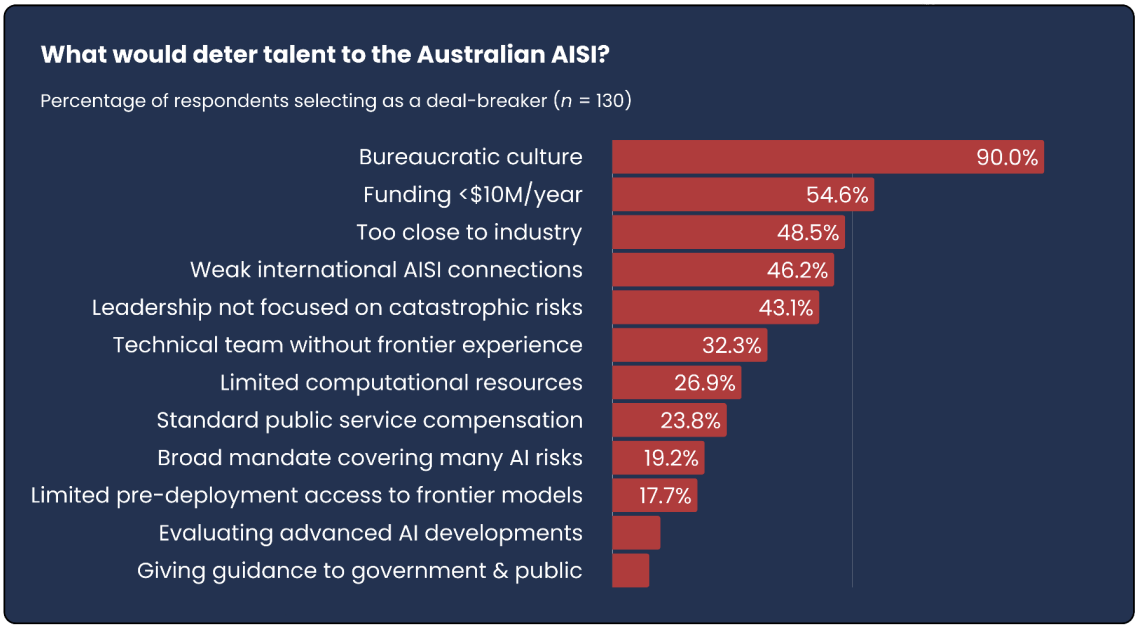
- **Other attractive factors.** Additional factors mentioned included compensation competitive with private sector roles (though often positioned as secondary to mission), access to technical resources including frontier models and compute, strong international network connections (particularly with UK and US institutes), meaningful work programmes on interesting technical problems, opportunities for professional growth and career development, and practical considerations like long-term contracts, and flexible work arrangements.

> *"Strong international connections (AUKUS, Five Eyes, Pacific partners) to ensure global impact"*

Deal-breakers

Respondents selected deal-breakers—factors that would "prevent [them], or someone like [them], from accepting a role at the Australian AISI".

Bureaucratic culture that prevents impact was the strongest deal-breaker, with 90% selecting this option. Funding less than $10 million AUD per year would prevent 54.6% from accepting a role, while being too close to industry or regulatory capture would deter 48.5% of potential candidates.



What excites the community about an Australian AI Safety Institute?

Respondents were asked what about an Australian AISI specifically excited them. Thematic analysis of 111 responses indicated the following:

- **Australia's unique geopolitical position.** Many responses referenced Australia's position as a trusted middle power with access to both major AI nations (US, China) and Indo-Pacific partnerships. This enables Australia to bridge US-China tensions, facilitate international coordination, and be a trusted mediator.

> *"What excites me most about an Australian AISI is its potential to break the US-China binary that dominates AI safety discourse. Australia occupies a unique middle-power position, close enough to major developments to be relevant, distant enough to be genuinely independent, and trusted enough in the Indo-Pacific to convene conversations that neither superpower could."*

- **Filling neglected gaps and doing something different.** Respondents valued the opportunity to pioneer approaches that differ from existing institutes, including addressing neglected areas, not duplicating others, and having a unique contribution.

> *"The prospect of the organisation being bold, innovative and doing something different [to] the existing AISIs, not just being another AISI. That's what I expect from an Australian AISI and that's why I'm excited!"*

- **Global impact and leadership.** Respondents framed the AISI as Australia's opportunity to contribute meaningfully to a global challenge and punch above its weight, similar to its role in energy transition and other areas. Australia's strong safety culture, technical talent, and progressive policy history position it for outsized impact.

> *"Australia punches above its weight in several areas, and Australia is overdue in contributing to AI safety research or shaping global AI policy. Australian AISI will attract international talent and help us to become more technically literate"*

- **Learning from predecessors.** Respondents noted the advantage of not being first to create an AISI, allowing Australia to learn from successes and failures of UK and US institutes (while bringing fresh perspective and energy). The timing provides additional context about AI developments that earlier institutes lacked.

> *"Starting a new AI Safety Institute allows Australia to build upon the successes and shortcomings of other country's AI Safety Institutions."*

- **Opportunity to contribute from Australia.** Multiple respondents expressed excitement about being able to build AI safety careers in Australia rather than relocate overseas, preventing brain drain, and gathering dispersed Australian talent. The AISI removes geographic barriers to meaningful contribution and validates Australia as a serious participant in AI safety.

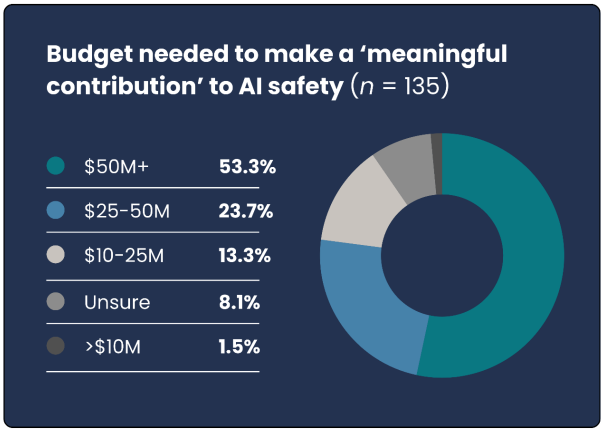
> *"For someone transitioning into AI safety, most opportunities have been concentrated in the USA and UK, making meaningful contribution feel geographically gated. An Australian AISI removes that barrier."*

- **Other factors.** Additional themes included Australia-specific strengths (biosecurity expertise, safety culture, political stability, consumer protection traditions); government engagement and policy impact (providing technical expertise to inform decisions); regional Asia-Pacific leadership (ASEAN partnerships, Pacific Islands support, addressing regional harms); practical real-world safety focus (deployment conditions, protecting vulnerable populations, addressing immediate harms like fraud networks); building Australian AI safety ecosystem (creating hub for collaboration, fostering talent, connecting academia-industry-government); and the signal that Australia is taking AI safety seriously and creating opportunities in the field.

> *"I'm also excited about the ecosystem effects. An AISI can foster more organisations working on AI safety in Australia and create pathways for world-class institutions like CSIRO to contribute their expertise more easily"*

Annual budget recommendations

Respondents supported substantial funding for the Australian AISI. Over half (53.3%) recommended over $50 million AUD per year for the AISI to "make a meaningful contribution to AI safety". Only 13.3% recommended $10-25 million.



Appendix A: Survey questions

Would you like to be connected with AISI opportunities?

Select all that apply, or leave blank

☐ Yes - please suggest me as a potential candidate if suitable. (Note: We may share basic information such as your name and profile (e.g., LinkedIn) with Government, but will NOT share your survey responses.)

☐ Yes - please alert me to opportunities as they become available

Full name

Required if you'd like to be connected with AISI opportunities, otherwise optional (though very helpful for us to understand who completed the survey)

[Open text]

Email

Required if you'd like to be connected with AISI opportunities, otherwise optional

[Open text]

LinkedIn or CV URL

[Open text]

Current role/organisation

[Open text]

What's your area of expertise?

Select all that apply

☐ Technical AI/ML

☐ AI safety research

☐ Policy and governance

☐ Government and public sector

☐ Operations/management

☐ Security (cybersecurity, biosecurity, or national security)

☐ Other:

Do you have 2+ years of professional experience in AI safety, security, governance, or related work?

For example:

- Researchers studying AI safety, alignment, capabilities, or governance

- ML engineers or researchers working on frontier AI systems

- Practitioners implementing AI safety measures or evaluations

- Policy professionals working on AI regulation or standards

- Technical staff at AI labs, research organisations, or safety institutes

- Operations/program management at AI safety organisations

○ Yes

○ No

What is your connection to Australia?

Select all that apply

☐ Australian citizen

☐ Australian permanent resident

☐ Currently living in Australia

☐ Previously lived in Australia

☐ Family connections to Australia

☐ Citizen of US, Canada, UK, or New Zealand

☐ No direct connection (still valuable to hear from you!)

☐ Other:

What should an Australian AISI do?

This section asks what an Australian AISI should prioritise and what would make it most impactful (all questions optional).

The Australia AISI should…

○ Focus mainly on catastrophic/frontier risks (e.g., bioweapons, loss of control)

○ Split focus evenly between catastrophic/frontier risks and broader AI risks

○ Focus mainly on broader AI risks (e.g., bias, discrimination, privacy)

How important is it for the Australian AISI to work on:

[Table format with rows for each risk area and columns: Not important | Somewhat important | Very important | Critical]

- Cyber misuse - AI-assisted cyberattacks

- Dual-use science (CBRN) - chemical, biological, radiological, nuclear weapons

- Criminal misuse - fraud, scams, deepfakes for crime

- Autonomous systems - loss of control, misaligned AI agents

- Societal resilience - widespread AI deployment impacts

- Human influence - manipulation, persuasion, deception

- Bias and discrimination - unfair treatment in high-stakes decisions

- Privacy and data protection - safeguarding personal information

- Intellectual property - copyright and training data rights

- Consumer harm - deceptive or unsafe AI products

- Business and government guidance - supporting responsible AI deployment

To make a meaningful contribution to AI safety, the Australian AISI should have an annual budget of:

For context: Canadian AISI has ~$11 million AUD/year, UK AISI has ~$130 million AUD/year

○ Less than $10 million AUD/year

○ $10-25 million AUD/year

○ $25-50 million AUD/year

○ $50-100 million AUD/year

○ Over $100 million AUD/year

○ Unsure

What specific programs of work should the Australian AISI prioritise?

Consider what approaches to AI safety (e.g. evaluations, interpretability, hardware verification) might be neglected by other institutions, need more global resources, or where Australia has particular comparative advantage (e.g., geopolitical positioning, technical expertise, trusted relationships).

[Open text]

What about an Australian AISI specifically excites you?

[Open text]

Suggesting potential candidates

Please list people you think would be a good fit to work at the Australian AISI.

We're especially interested in people who deeply understand frontier AI risks, have strong expertise and/or relationships in the field, and can navigate government and international networks effectively.

This could include people who might actually want the role, or simply exemplars of what great candidates look like. They don't need to be Australian, though Australian connections are an asset.

Senior roles are most pressing, but all suggestions are helpful. You can also email contact@goodancestors.org.au with suggestions.

[Open text]

Attracting talent to an Australian AISI

This section asks what would attract candidates to work at the AISI (all questions optional).

What would most attract you, or someone like you, to work at the Australian AISI? Please be specific about what matters most.

You could consider factors like mission, leadership, compensation, work programs.

[Open text field]

Which of the following would be deal-makers - things that would strongly pull you, or someone like you, toward working at the Australian AISI?

Select all that apply

☐ Leadership focused on catastrophic / frontier risks

☐ Compensation significantly above standard public service

☐ Focused mandate on catastrophic / frontier risks

☐ Pre-deployment access to frontier models

☐ Technical team with frontier-model experience

☐ Strong international AISI network connections

☐ Substantial computational resources

☐ Independence from industry influence

☐ Working with regulators to safely capture the benefits of AI

☐ Serving as a hub to coordinate government action

☐ Giving guidance on AI opportunities to businesses and the public

☐ Other:

Which would be deal-breakers - things that would prevent you, or someone like you, from accepting a role at the Australian AISI?

Select all that apply

☐ Leadership not focused on catastrophic / frontier risks

☐ Standard public service compensation

☐ Broad mandate covering many AI risks (not focused on catastrophic / frontier risks)

☐ Limited pre-deployment access to frontier models

☐ Technical team without frontier-model experience

☐ Weak international AISI network connections

☐ Limited computational resources

☐ Bureaucratic culture that prevents impact

☐ Too close to industry / regulatory capture concerns

☐ Significant effort evaluating technical developments in advanced AI

☐ Significant effort giving guidance on AI safety to government and the public

☐ Funding less than $10 million AUD/year for the AISI

☐ Other:

Anything else?

Is there anything else you'd like to add?

[Open text]

Appendix B: Survey results

Respondent information

Area of expertise

Question:
"What's your area of expertise?" (Multi-select)

	Respondents (%)
AI safety research	72 (51.8%)
Policy and governance	64 (46.0%)
Technical AI/ML	56 (40.3%)
Operations/management	47 (33.8%)
Government and public sector	36 (25.9%)
Security (cybersecurity, biosecurity, or national security)	36 (25.9%)

Professional experience

Question:
"Do you have 2+ years of professional experience in AI safety, security, governance, or related work?"

	Respondents (%)
Yes	89 (64%)
No	50 (36%)
Total responses	139

Connection to Australia

Question:

"What is your connection to Australia?" (Multi-select)

	Respondents (%)
Australian citizen	95 (68.3%)
Australian permanent resident	8 (5.8%)
Currently living in Australia	27 (19.4%)
Previously lived in Australia	8 (5.8%)
Family connections to Australia	14 (10.1%)
Citizen of US, Canada, UK, or New Zealand	22 (15.8%)
No direct connection (still valuable to hear from you!)	15 (10.8%)

AI Safety Institute focus

Focus on catastrophic/frontier risks vs broad AI risks

Question:
"The Australia AISI should..."

	Respondents (%)
Focus mainly on catastrophic/frontier risks (e.g., bioweapons, loss of control)	41 (30.6%)
Split focus evenly between catastrophic/frontier risks and broader AI risks	78 (58.2%)
Focus mainly on broader AI risks (e.g., bias, discrimination, privacy)	15 (11.2%)
Total responses	134

Risk priority areas

Question:
"How important is it for the Australian AISI to work on:"

	Not important	Somewhat important	Very important	Critical
Cyber misuse	1 (0.8%)	24 (18.0%)	50 (37.6%)	58 (43.6%)
Dual-use science (CBRN)	4 (3.0%)	23 (17.2%)	39 (29.1%)	68 (50.7%)
Criminal misuse	11 (8.2%)	40 (29.9%)	51 (38.1%)	32 (23.9%)
Autonomous systems	1 (0.7%)	18 (13.4%)	27 (20.1%)	88 (65.7%)
Societal resilience	1 (0.7%)	25 (18.7%)	54 (40.3%)	54 (40.3%)
Human influence	4 (3.0%)	22 (16.4%)	54 (40.3%)	54 (40.3%)
Bias and discrimination	23 (17.2%)	41 (30.6%)	43 (32.1%)	27 (20.1%)
Privacy and data protection	19 (14.2%)	42 (31.3%)	43 (32.1%)	30 (22.4%)
Intellectual property	38 (28.8%)	57 (43.2%)	23 (17.4%)	14 (10.6%)
Consumer harm	18 (13.4%)	42 (31.3%)	44 (32.8%)	30 (22.4%)
Business and government guidance	20 (15.2%)	45 (34.1%)	39 (29.5%)	28 (21.2%)

Attracting talent to the AI Safety Institute

Deal-makers

Question:
"Which of the following would be deal-makers - things that would strongly pull you, or someone like you, toward working at the Australian AISI?" (Multi-select, *n* = 131)

	Respondents (%)

Strong international AISI network connections	89 (67.9%)
Leadership focused on catastrophic/frontier risks	84 (64.1%)
Focused mandate on catastrophic/frontier risks	79 (60.3%)
Serving as a hub to coordinate government action	78 (59.5%)
Independence from industry influence	74 (56.5%)
Technical team with frontier-model experience	69 (52.7%)
Working with regulators to safely capture the benefits of AI	68 (51.9%)
Compensation significantly above standard public service	67 (51.1%)
Giving guidance on AI opportunities to businesses and the public	53 (40.5%)
Pre-deployment access to frontier models	45 (34.4%)
Substantial computational resources	42 (32.1%)

Deal-breakers

Question:
"Which would be deal-breakers - things that would prevent you, or someone like you, from accepting a role at the Australian AISI?" (Multi-select, *n* = 130)

	Respondents (%)
Bureaucratic culture that prevents impact	117 (90%)
Funding less than $10 million AUD/year	71 (54.6%)
Too close to industry/regulatory capture concerns	63 (48.5%)
Weak international AISI network connections	60 (46.2%)
Leadership not focused on catastrophic/frontier risks	56 (43.1%)
Technical team without frontier-model experience	42 (32.3%)
Limited computational resources	35 (26.9%)
Standard public service compensation	31 (23.8%)
Broad mandate covering many AI risks	25 (19.2%)
Limited pre-deployment access to frontier models	23 (17.7%)
Significant effort evaluating technical developments in advanced AI	13 (10.0%)
Significant effort giving guidance on AI safety to government and the public	10 (7.7%)

Annual budget recommendations

Question:
"To make a meaningful contribution to AI safety, the Australian AISI should have an annual budget of:"

	Respondents (%)
Less than $10 million AUD/year	2 (1.5%)
$10-25 million AUD/year	18 (13.3%)
$25-50 million AUD/year	32 (23.7%)
$50-100 million AUD/year	32 (23.7%)
Over $100 million AUD/year	40 (29.6%)

Unsure	11 (8.1%)
Total responses	135