AUSTRALIA

Preparing for human-centric use of artificial intelligence

Emma Schleiger, Qinghua Lu, Stefan Harrer, Liming Zhu
Commonwealth Scientific and Industrial Research Organisation

Key takeaways

- Ethical principles and human-centric approaches to AI are informing Australia's emerging framework for AI governance. The number of tertiary education offerings for AI have increased in Australia and are complemented by an initiative to attract and train job ready AI specialists.
- While active programs to enhance diversity in Australia›s STEM workforce exist, they are not specifically tailored to address AI. Additionally, there is a recognized need to enhance ethical competence and raise awareness of human rights in AI-related scientific endeavours. However, more customized resources for the science sector are required.
- Other challenges remain to be addressed such as the high-performance and data computing infrastructure needed for AI and AI-enabled science and the implementation of FAIR and CARE data principles.

Australia's government, scientific organizations and universities are exploring the preparedness of the national science system to capture the opportunities and mitigate the risks of AI to accelerate scientific discovery. For example, the national science agency, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), released the report *Artificial Intelligence for Science – Adoption Trends and Future Development Pathways* (Hajkowicz et al., 2022). It examines the impact of AI on science and the imperative for research organizations to invest in mechanisms to harness the benefits and mitigate the risks of these technologies. The report outlines six future development pathways to enable the transition, including hardware and software upgrades, data capability uplift, improved education and training, the development of human-centred AI, improved workforce diversity and ethical capability. Organizations throughout Australia's national science system have begun expanding their capacity for AI uplift in these areas with recent research initiatives, activities, programs and guidelines. However, challenges remain to be addressed.

Hardware and software

Scientific organizations seeking to uplift their AI capability must make decisions about hardware, software and computational infrastructure upgrades. The Australian Academy of Science recently held a national roundtable to discuss the Australian science sector's future supercomputing needs. The group highlighted the need for a national strategy and an exascale computing facility to secure Australia's sovereign research capability and enable science to meet national and regional priorities into the future (Australian Academy of Sciences, 2023).

Data

Future AI capability uplift also requires investment in high-quality data which is fit for purpose, provenance assured, validated, up to date and ethically obtained. The Australian government is leading by example through its Data and Digital Government Strategy (Government of Australia, 2023). This initiative focuses on adopting best-practice approaches to data collection, management and use to become a data-driven organization.

In conjunction with the increasing use of AI, it is essential for Australia to better implement the FAIR (Findable, Accessible, Interpretable and Reusable) and CARE (Collective benefit, Authority to control, Responsibility and Ethics) data principles. These and other principles and practices from open science, the Indigenous Data Sovereignty movement and participatory data stewardship all provide critical guidance for the creation, use and management of the data that will underpin AI in Australia's science system.

Education, training and capability

There is an imperative for education, training and capability uplift across the science sector and into lifelong education. The number of tertiary AI courses offered in Australia almost doubled between 2020 and 2023, providing greater educational opportunities (37 offerings in 2020, 69 in 2023) (OECD, 2024).

The Australian Human Rights Commission (2023) has recommended that 'professional development and training be provided to teachers' and 'schools should introduce comprehensive digital literacy programs to provide students with the skills needed to engage with generative AI tools in a responsible and ethical way'.

In 2021, AUD 24.7 million was invested in establishing CSIRO's Next Generation AI Graduates Program to attract and train job-ready AI specialists in Australia (CSIRO, 2021). Currently, more than a thousand CSIRO researchers are working on a diverse range of AI and data science projects (CSIRO, a).

In December 2023, CSIRO' researchers published the world's first responsible AI book for practitioners (Lu, 2023a). The book introduces Responsible AI Pattern Catalogue, featuring over sixty best practices for different types of stakeholders. It also contains case studies from industry partners on how they have applied these best practices.

Human-centric artificial intelligence

Human–AI collaboration and human-centric AI is designed and implemented to ensure humans can work effectively with AI and benefit from the complementary strengths of humans and AI systems to carry out tasks to higher standards than either can achieve alone. In 2023, Australia signed the Bletchley Declaration affirming that AI should be designed, developed and deployed in a human-centric, responsible and trustworthy manner. Australia then signed the Seoul Declaration in May 2024, which builds upon the Bletchley Declaration. Signatories of the Seoul Declaration committed to a shared understanding of the opportunities and risks posed by AI.

CSIRO's collaborative intelligence (CINTEL) program of work is developing the science and technology to ensure AI systems support humans to solve scientific challenges, such as highly labour-intensive tasks like genome annotation (CSIRO, b). Annotation uses the genome sequence to create biological phenotypes critical for increasing crop yields through selective breeding. The group is developing a scalable approach involving collaboration between a domain expert and AI that will allow for accurate and timely annotation of genomes.

CSIRO's Science Digital – AI for Science program of work is developing an agentic AI platform that allows scientist users to integrate AI agents into scientific discovery workflows. AI agents will assist and leverage the work of human scientists from hypothesis generation to experimental design to outputs analysis. The AI agent platform will evolve into a user and developer ecosystem where scientists can build their own AI agents alongside AI agents built by others. Moreover, CSIRO has entered a strategic partnership with Google to promote the role of AI in Science, to educate and upskill scientists on safe and responsible use of AI, and to nurture a community of practice in AI for Science.

Gender, ethnic and cultural diversity

The AI workforce lacks gender, ethnic and cultural diversity, which limits the quality of outcomes. Improving this will contribute to an uplift in AI capability within research organizations.

The Government of Australia's (2020) *Advancing Women in STEM Strategy Action Plan 2020* provides a national, coordinated approach to achieving sustained increases in gender equity in science, technology, engineering and mathematics (STEM). Currently, only 20% of AI and computer science PhD students are female in Australia, however it is notable that 44% of the new AI talent entering the market in Australia are female, placing Australia amongst the highest gender diversity in the world for new talent (Randstad, 2024).

Programs such as Deadly Science (Deadly Science) and the Indigenous STEM Education Project (CSIRO, 2021) seek to support and engage Aboriginal and Torres Strait Islander students in science- and STEM-related careers. Between 2014 and 2021, the Indigenous STEM Education Project reached over 23,000 participants in 603 schools, and Deadly Science has delivered 7,500 boxes of science resources to over 800 schools.

CSIRO's Responsible AI Pattern Catalogue (Lu et al., 2023b) and Diversity and Inclusion (Zowghi and da Rimini, 2023) in AI Guidelines were highlighted in the National Framework for the Assurance of AI in Government. The same works are integrated into the digital pathway co-developed with the National AI Centre and will be accessible as interactive tools to all Australian businesses.

Ethical capability

Evolving standards and regulation of the design and implementation of AI require investment in ethical capability – including technology, skills and cultures. In support of responsible innovation, the Australian government has produced a framework of eight ethics principles to ensure AI is safe, secure and reliable (Dawson et al., 2019; DISR, a). This was followed by

the 2023 discussion paper *Safe and Responsible AI in Australia* (DISR, 2023) to support responsible AI practices and increase community trust and confidence through consultative government responses. The Australian government's January 2024 interim response to the consultation identified a range of legal, regulatory and governance measures that are needed to ensure AI is designed, developed and deployed safely and responsibly (DISR, 2024).

Following this work, the National AI Centre (NAIC) has developed the first iteration of the Voluntary AI Safety Standard, with 10 voluntary guardrails aimed to help organisations develop and deploy AI systems safely and reliably (DISR, 2024). The government is also considering options for mandatory approaches and released a set of ten mandatory guardrails for AI in high-risk settings for consultation in September 2024 (DISR, 2024). Within their proposal, the Australian Government has sought advice on the proposed guardrails themselves and the three different regulatory options available to the Australian Government to mandate the proposed guardrails, including the introduction of a cross-economy Australian AI act.

CSIRO's Responsible Innovation Future Science Platform is a program of research that systematically and scientifically assesses the risks, benefits and uncertainties of future science and technology. Meanwhile, the Australian Human Rights Commission (2021) recommends that 'professional accreditation bodies for [STEM] should introduce mandatory training on human rights by design as part of continuing professional development'. However, no framework or strategies are in place for such upskilling in the science sector, and very few professional accreditation bodies exist.

CSIRO Responsible AI Team collaborated with Alphinity Investment Company on the world's first Responsible AI Framework using the ESG lens (Alphinity and CSIRO, 2024), involving 26 ASX and internationally listed companies. The work was widely reported internationally, with many investor roundtables discussing their use of the framework and potential licensing for external use.

Other challenges

As well as impacting how science is done, AI may impact how science is administered, governed, funded and assessed. Australia's research councils, the Australian Research Council and the National Health and Medical Research Council, have created policies to account for the role of generative AI in their grant processes (ARC, 2023; NHMRC, 2023). The use of generative AI is prohibited in assessing applications to preserve the confidentiality and integrity of the process. For applicants, the policies note the potential benefits and need for caution in using AI but do not list any specific restrictions on the use of AI by applicants.

References

ARC. 2023. *Policy on Use of Generative Artificial Intelligence in the ARC's Grants Programs*. Canberra, Australian Research Council. https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/artificial-intelligence

Australian Academy of Science. 2023. 'Is Australia ready for our supercomputing future?' *News and Media Releases*, 29 November. Canberra: Australian Academy of Science. https://www.science.org.au/news-and-events/news-and-media-releases/is-australia-ready-for-our-supercomputing-future (2023).

Australian Human Rights Commission. 2021. *Human Rights and Technology Final Report.* Sydney: Australian Human Rights Commission. https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology

Australian Human Rights Commission. 2023. *Utilising Ethical AI in the Australian Education System: Submission to the Standing Committee on Employment, Education and Training*. Sydney: Australian Human Rights Commission. https://humanrights.gov.au/our-work/legal/submission/utilising-ethical-ai-education-system

CSIRO. a. 'Artificial Intelligence'. *CSIRO*. Canberra, Commonwealth Scientific and Industrial Research Organisation. https://www.csiro.au/en/research/technology-space/ai

CSIRO. b. *Collaborative Intelligence Future Science Platform*. Canberra, Commonwealth Scientific and Industrial Research Organisation. https://research.csiro.au/cintel/.

CSIRO. 2021 c. 'Indigenous STEM Education Project'. *CSIRO*. Canberra, Commonwealth Scientific and Industrial Research Organisation. https://www.csiro.au/en/education/programs/indigenous-stem-education-project.

CSIRO. 2021. '2021-22 Federal Budget Statement'. *CSIRO*, 12 May. Canberra, Commonwealth Scientific and Industrial Research Organisation. https://www.csiro.au/en/news/All/News/2021/May/2021-22-Federal-Budget-Statement.

Alphinity Investment Management (Alphinity) and Commonwealth Scientific and Industrial Research Organisation (CSIRO), The intersection of Responsible AI and ESG: A Framework for Investors, CSIRO, 2024. https://www.csiro.au/en/research/technology-space/ai/responsible-ai/rai-esg-framework-for-investors

Dawson, D. et al. 2019. *Artificial Intelligence: Australia's Ethics Framework. A Discussion Paper.* Canberra, Commonwealth Scientific and Industrial Research Organisation. https://www.csiro.au/en/research/technology-space/ai/AI-Ethics-Framework.

Deadly Science. https://deadlyscience.org.au/about-us/.

DISR. a. *Australia's AI Ethics Principles*. Canberra, Department of Industry, Science and Resources. https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles.

DISR. 2023. *Safe and Responsible AI in Australia: Discussion Paper.* Canberra, Department of Industry, Science and Resources. https://consult.industry.gov.au/supporting-responsible-ai.

DISR. 2024. *Safe and Responsible AI in Australia Consultation: Australian Government's Interim Response*. Canberra, Department of Industry, Science and Resources. https://consult.industry.gov.au/supporting-responsible-ai.

DISR. 2024. *Voluntary AI Safety Standard.* Canberra, Department of Industry, Science and Resources. https://www.industry.gov.au/publications/voluntary-ai-safety-standard

DISR. 2024. *Proposals Paper for Introducing Mandatory Guardrails for AI in High-Risk Settings*. Canberra, Department of Industry, Science and Resources. https://consult.industry.gov.au/ai-mandatory-guardrails

Government of Australia. 2020. *Advancing Women in STEM 2020 Action Plan*. https://www.industry.gov.au/publications/advancing-women-stem-strategy

Government of Australia. 2023. *Data and Digital Government Strategy*. https://www.dataanddigital.gov.au/ (Accessed 21 February 2024).

Hajkowicz, S. et al. 2022. *Artificial Intelligence for Science – Adoption Trends and Future Development Pathways*. Canberra, Commonwealth Scientific and Industrial Research Organisation. https://www.csiro.au/-/media/D61/AI4Science-report/AI-for-Science-report-2022.pdf

Lu, Qinghua, Zhu, Liming, Whittle, Jon, Xu, Xiwei. 2023. *Responsible AI: Best Practices for Creating Trustworthy AI Systems*. Addison-Wesley Professional.

Lu, Qinghua, Liming Zhu, Xiwei Xu, Jon Whittle, Didar Zowghi, and Aurelie Jacquet. "Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering." *ACM Computing Surveys* 56, no. 7 (2024): 1-35.

NHMRC. 2023. *Policy on Use of Generative Artificial Intelligence in Grant Applications and Peer Review*. Canberra: National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/resources/policy-use-generative-artificial-intelligence.

OECD. 2024. 'AI courses in English by degree type'. *OECD.AI Policy Observatory*. https://oecd.ai/en/data?selectedArea=ai-education&selectedVisualization=ai-courses-by-degree-type (Accessed 21 February 2024).

Parliament of Australia. 2024. *Inquiry into the use and governance of artificial intelligence systems by public sector entities*. Canberra, Parliament of Australia. https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/Public_Accounts_and_Audit/PublicsectoruseofAI

Randstad Enterprise. 2024. *2024 In-Demand Skills Dashboard, AI & Automation*. https://enterprise.randstadsourceright.com/in-demand-skills-ai

Zowghi, Didar and da Rimini, Francesca. 2023. 'Diversity and Inclusion in Artificial Intelligence', in Lu Q et al. (eds) *Responsible AI: Best Practices for Creating Trustworthy AI Systems*, Addison-Wesley, Sydney.