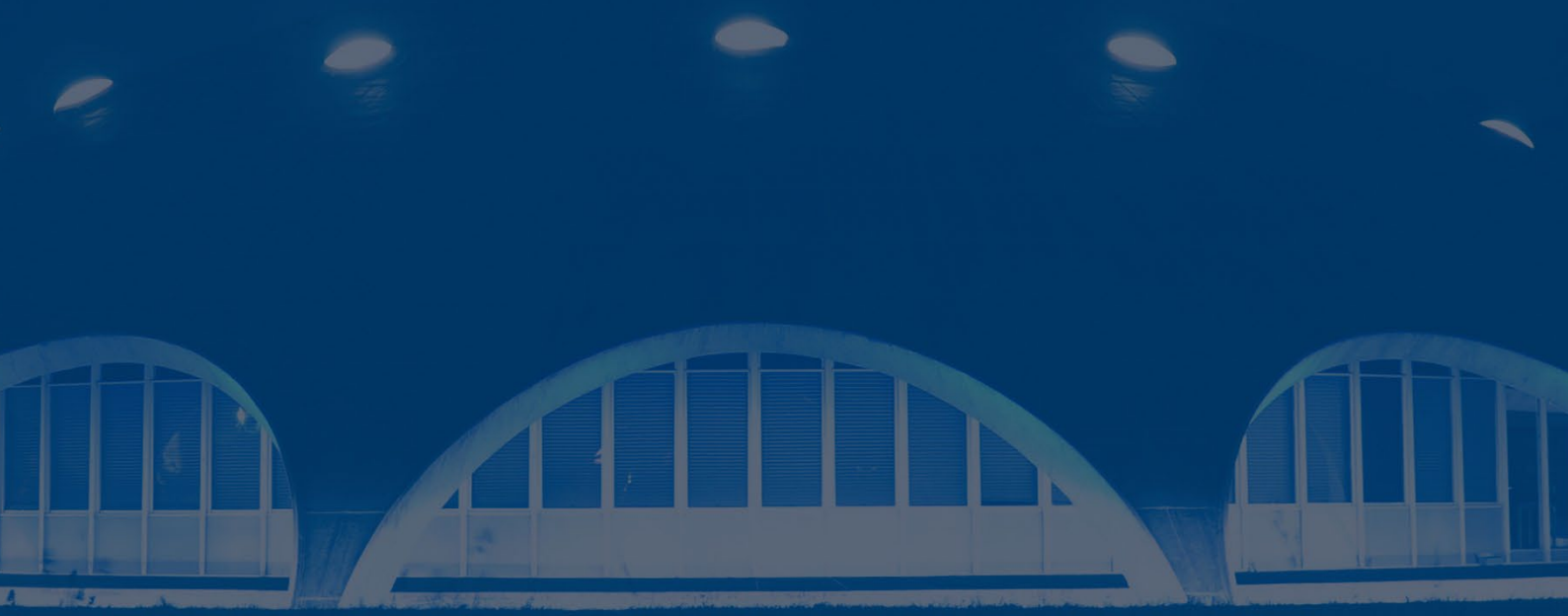
Australian Academy of Science

Is Australian science ready for AI?

Systems to disseminate knowledge

December 2025



'Is Australian science ready for AI?' is a series of discussion papers that explore the preparedness of the Australian science sector for AI advances.

Are our systems for disseminating knowledge AI-ready?

Generative AI can supercharge manuscript preparation, but it can also drown publishers in a tidal wave of submissions, further straining an already fragile peer-review system.

AI poses risks for scientists and scientific publishing, mainly due to the system's dependence on profit-driven models. Publishers are increasingly experimenting with providing expedited peer reviews or licensing agreements that give large language model (LLM) developers access to scientific data. At the same time, AI tools are being used within the peer review process itself, raising concerns about data privacy, consent, and how unpublished scientific work is being handled within these systems.

This paper explores how generative AI is accelerating scientific knowledge dissemination while addressing the challenges it brings to the scientific publishing ecosystem.

Generative AI is accelerating manuscript preparation

AI has the potential to revolutionise scientific publishing by increasing the efficiency of research processes. It could also aid in faster and more thorough preparation of systematic reviews, metric analyses, and scientific evaluation.[1]

Large language models can rapidly summarise, synthesise and reveal novel connections within vast quantities of research literature.

One of the benefits of generative AI is assisting researchers with time-consuming tasks such as preparing manuscripts for publication. However, scientific publishing relies on the unpaid labour of peer reviewers, who may now, because of generative AI, be facing an even greater volume of publications to be reviewed. This could have flow-on effects on how long research takes to be published.

Additionally, generative AI may further encourage a process called 'salami slicing', where scientists publish slices of their work in separate papers to generate track record, rather than publishing it together as a single comprehensive paper. This can be supported by AI's rapid generation of text, including circumnavigating the detection of self-plagiarism by rephrasing text.

Scientific publishing and the AI challenge

While generative AI tools can improve the quality of research, expedite the process of conducting literature reviews and drafting journal articles, and increase efficiency, they also introduce significant challenges. These include the risk of incorporating misinformation and disinformation into scientific literature, as well as replicating and reusing published content without the knowledge, permission or proper attribution of authors or publishers.[2]

In this context, scientific publishing faces two major challenges: safeguarding the integrity of the scientific record by identifying and preventing misinformation and disinformation, and ensuring proper recognition and protection of authorship.

The traditional scientific dissemination system, predominantly dependent on journal-based models, is not adequately prepared for the transformative capabilities of AI.

Scientific publishing has long been monetised by publishers offering exclusivity and control over scientific knowledge. As generative AI undermines the value of traditional access and paywall models, this disruption could fundamentally shift how knowledge is validated, curated, applied and monetised.

Disruptions to scientific publishing can have broad impacts on the science system

The disruption to scientific publishing caused by AI has far-reaching effects on the Australian science system.

- AI could undermine the standing of the unit of the scientific paper fundamental to the record of science, thus affecting systematic reviews in fields including but not limited to health and environmental science.
- AI tools could be deliberately prompted to produce citation-heavy outputs, undermining the use of citation metrics in funding allocation across the science system, such as in formulas for research block grants to universities or metrics used within institutions and science agencies.
- Accurately measuring the quality and impact of science at an individual, institutional or national level could become more difficult.
- A flood of AI-influenced papers will confound processes like the assessment of grant proposals and promotion applications that rely, in part, on publishing track record.
- Use of AI systems can homogenise the style and directions of inquiry in scientific research, risking diversity of thought and potentially supressing novel ideas.

Ownership of journal articles and associated data by scientists

The rapidly growing popularity of AI tools is driving a worldwide race for vast quantities of data to train these tools, particularly large language models. This has made data held by publishers in the form of text that they own a commodity in rising demand. Large companies creating these AI tools have begun purchasing rights to data from publishers to train their models.

As generative AI continues to grow in popularity and be used by more scientists, the current publishing model could have its weaknesses exacerbated or improved.

On the one hand, a surge in quickly produced articles may cause delays in the review process, and researchers will need to pay to have their work published and see their work monetised by being sold as data to train AI models by publishers. On the other hand, AI may offer access to scientific knowledge long hidden behind paywalls, making it more accessible to the general public and policymakers, where it can be used to support evidence-informed decision making.

With thoughtful regulation and licensing models, AI systems could be required to recognise authors, could extend the reach and impact of science including to non-scientific audiences, and could reduce the cost of publication to researchers and research organisations. This could create an opportunity to reimagine the publishing model, where researchers could be content creators who are paid, rather than paying others to publish their work.

Academic publishers are responding to the use of these tools

Major academic publishers have released policy statements that have restricted the use of AI-generated text in publications, prohibited the inclusion of AI as authors, and set out guidelines for acknowledging where the text came from.

These policies, which focus on the authorship and creation of research outputs and how they are evaluated (publication and accountability), may be increasingly important to the management and maintenance of functional and robust science.

AI may aid science communication and democratise access to science

Science communication is the practice of the public communication of individual scientists, strategic communication of scientific organisations, science journalism, and other forms of public communication related to science.[3]

AI has the potential to remove barriers in science communication by democratising the dissemination of knowledge, overcoming language barriers that frequently restrict broader international participation. This could open new opportunities for scientific collaboration for Australia, where language barriers have prevented collaboration in the past.

Generative AI tools could also increase public understanding of scientific concepts and discoveries by providing reliable information and fostering dialogue between scientists and the public through:

- streamlining and automating the process of generating scientific communication content
- summarising scholarly publications and findings in an accessible manner or answering specific questions from a non-scientific perspective
- broadly democratising science by providing access to information.

On the other hand, the hallucination effect of generative AI raises questions about the accuracy of outputs science communicators may use. While newer versions of GPT and other AI tools have made improvements, it is essential to note that generative AI has limitations due to biases in training data and output modalities.

Synthetic training data – data created using algorithms rather than scientific sources – used to train AI models, can also feed into the hallucination effect, as scientists cannot generate original datasets rapidly enough to match AI training demands.

There are concerns that AI-powered language models could generate an overwhelming amount of information, leading to a pollution of reliable knowledge. The addition of mis- or disinformation could exacerbate this problem, and attempts to bypass built-in restrictions have already been observed. This could result in widespread falsehoods, especially in a communication ecosystem where scientific authority is contested.

The importance of data stewardship and sovereignty

AI research and development currently depends on open science because the sharing of datasets is required to continue to train AI models and enable large-scale data analysis.

ChatGPT, for example, has been trained on publicly available data including books, websites and articles. Australians are producing unprecedented amounts of data, yet Australia does not have an open science strategy and frameworks, and those that exist, such as the UNESCO Recommendation on Open Science, do not mention AI.

The pursuit of big data is transitioning to the pursuit of better-quality data. High-quality data is essential if AI systems are to make accurate predictions and build reliable and reproducible outcomes – processes that are fundamental to science.[4]

However, access to high-quality scientific data is increasingly shaped by commercial agreements such as licensing deals between publishers and AI developers, raising concerns about transparency, equitable access and the future of open science.

In 2021, the Academy warned that despite the increasing importance of data in research and public policy, Australia has no coherent national research data policy.[5]

Ensuring that data used to train AI models is accessible, interoperable, and reproducible, and dealing with the scientific challenges of corrupted data, remain significant challenges.

Australia's privacy protection framework takes strict approaches to sharing publicly held data while lacking the same level of control over privately held data.[4] Public sector data sharing is governed by various legislative frameworks prioritising privacy, proprietary control and informed consent. While some scientific sectors adhere to FAIR principles for reliable and transparent data, no national open science policy governs the ethical use of citizens' data across all sectors.

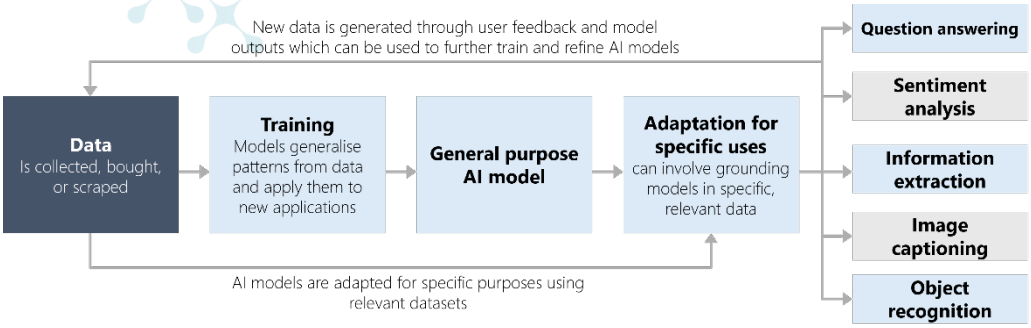


Figure 1. Data is both an input and an output of AI models. Adapted from the Productivity Commission (2024).[4]

Coupled with this is the need to respect Indigenous data sovereignty with data protocols like CARE principles that provide a framework for researchers looking to engage with Indigenous Knowledges and Indigenous Knowledge Holders.

Reviewers

Expert reviewers

Emeritus Professor Deborah Bunker, The University of Sydney Business School

Dr Ehsan Nabavi, Centre for the Public Awareness of Science, Australian National University

Dr Emma Schleiger, University of the Sunshine Coast

Professor Toby Walsh FAA FTSE, UNSW AI Institute

Dr Danny Kingsley, Deakin University

Adjunct Professor Ginny Barbour, Queensland University of Technology

Acknowledgements

The production of this discussion paper was supported by staff of the Australian Academy of Science, with Dr Sage Kelly, Dr Hayley Teasdale, Dr Ben Swinton-Bland, Alexandra Lucchetti, Dr Rakshanya Sekar, and Chris Anderson gratefully acknowledged. Edited by Robyn Diamond.

References

1. Atkinson, C. F. Cheap, quick, and rigorous: artificial intelligence and the systematic literature review. *Soc Sci Comput Rev* 42, 376–393 (2024).

2. Heath, N. Authors outraged to discover Meta used their pirated work to train its AI systems. *ABC* (2025).

3. Schäfer, M. S. The notorious GPT: science communication in the age of artificial intelligence. *Journal of Science Communication* 22, Y02 (2023).

4. Productivity Commission 2024. *Making the most of the AI opportunity: AI raises the stakes for data policy, Research Paper 3.* www.pc.gov.au (2024).

5. Australian Academy of Science. *Advancing data-intensive research in Australia.* (2021).