

**Australian Academy of Science submission on the
National Digital Research Infrastructure (NDRI) Investment Plan consultation**

The Academy welcomes the development of the National Digital Research Infrastructure (NDRI) Investment Plan to provide further detail as to how the vision for the future of Australia's research infrastructure will be realised.

The Academy recommends the NDRI Investment Plan includes strategic investments to deliver:

- Exascale high-performance computing and data infrastructure to accelerate Australian research, integrating artificial intelligence, quantum computing, and cloud computing.
- National data infrastructure and services that support data discovery, access, and curation and ensure rich metadata for all research sectors, with appropriate data security.
- A nationally coordinated digitalisation program for research collections to deliver broader access and preservation of nationally significant collections as a resource for conservation and science.
- Programs in schools, tertiary institutions and workplaces to upskill Australia's workforce in current and future NDRI capabilities.

[A strategic and ambitious plan to build Australia's national capacity in high-performance computing and AI](#)

For Australia to be competitive in science and technology, targeted investment in next-generation high-performance computing and data infrastructure, including exascale capability, is essential.

As scientific research becomes increasingly complex and integrates advances in AI, the demand for advanced computational power is growing exponentially. Exascale computing will provide the computational power and capacity to tackle large-scale challenges across domains such as earth system science data, climate science, genomics, modelling and advanced manufacturing. Without this capability, Australia's ability to innovate and address emerging societal challenges will be limited as researchers struggle to keep pace with global advances. To build our national capacity, the following activities should be considered as a priority:

1. **Establishing a Tier-0 exascale supercomputing facility in the Asia-Pacific region, hosted in Australia, would secure the nation's leadership in AI and HPC.** This would position Australia as a regional hub for cutting-edge research, catalysing new opportunities for scientific breakthroughs, economic growth, and technological advances.
2. **Integrating high-resolution national datasets with accessible HPC platforms** would unlock new opportunities in data-intensive research and AI by ensuring seamless modelling, analysis, and real-time insights. Co-locating data and computing resources is essential for driving progress in critical areas such as environmental monitoring, precision medicine, and energy systems. This approach reduces bottlenecks associated with data transfer, enhances research productivity, and strengthens Australia's capacity to deliver impactful scientific outcomes at both national and global scales.
3. **Investments in specialised research groups to develop expertise in emerging technologies**, such as AI and quantum computing, would strengthen Australia's capability to explore new scientific frontiers and sustain innovation.
4. **Participation in international collaborations on pre-exascale and exascale systems** would align Australia with global best practices and data standards, facilitating seamless integration into the international research ecosystem. These partnerships would enhance local expertise and ensure Australian researchers have access to the latest advances in HPC environments.

Strengthening Australian research through secure and inclusive data infrastructure

Secure, accessible, inclusive and well-integrated national data infrastructures and services are a priority under the NDRI. Investing in these systems and services is essential to supporting data discovery, access, and interoperability across research sectors and domain-specific repositories while striking an appropriate balance between research openness and data security. Such infrastructure will promote the sharing of research progress and outputs, accelerating Australia's research capabilities.

To drive this vision, specific investments under the NDRI plan could include:

- Sustained support for **a national vocabulary infrastructure** where researchers, NCRIS facilities and government bodies can access and assess semantic standards endorsed by authoritative bodies to promote consistency across research data. Research vocabularies communicate what data are by defining agreed meanings within or across domains, enabling data analysis and aggregation.¹
- **Developing an integrated and standardised observational data system, co-located with national computing facilities**, to support research through access to national and global reanalyses and satellite data essential for areas such as climate change, AI and machine learning.
- **Creating incentives for FAIR (findable, accessible, interoperable and reusable)² and CARE (collective benefits, authority to control, responsibility and ethics)³ compliance** through career recognition and targeted funding, to encourage researchers to publish high-quality, compliant datasets, improving overall research output.
- **Tools that empower Indigenous communities to govern their own data**, supporting autonomy over its use and access.
- **Improvements to metadata elements to apply CARE principles to legacy data held by institutions**, such as environmental samples or surveys conducted on Indigenous land, assuring ethical and accurate data stewardship.
- **Highly secure data enclaves**, particularly for sensitive human-related data, to safeguard privacy and maintain research integrity.

Software sharing and collaborative software environments

Open-source software allows broader user bases to access, improve and develop software for research. Investments in scaling and automating software tools and processes related to data collection and management could improve efficiency and access across disciplines. Specific investments under the NDRI Investment Plan could include **infrastructure to support institutional and discipline-specific repositories that enable the sharing of code** aligning with funder requirements for open research practices and **collaborative software environments that bring together tools, data, and methods in a unified "laboratory" space to promote shared problem-solving**.

Enhancing digitisation and digital curation of research collections

Research collections, including biological collections and those managed by the galleries, libraries, archives and museum (GLAM) sector, are critical to scientific research. The NDRI Investment Plan should fund initiatives that provide national coordination for the digitisation, data infrastructure and capability to curate and manage digital objects.

A nationally coordinated digitalisation program for research collections would ensure broader access and preservation and would serve as a vital resource for conservation and science. Further, targeted investments in new technologies, national standards, and digital curation would mobilise collections-based data. Underpinning investments could include:

¹ ARDC. ARDC Research Vocabularies Australia. <https://ardc.edu.au/services/research-vocabularies-australia/>.

² ARDC. FAIR Data. <https://ardc.edu.au/resource/fair-data/> (2024).

³ ARDC. CARE Principles. <https://ardc.edu.au/resource/the-care-principles/> (2024).

- **A national data infrastructure to support the secure storage, management, and accessibility of digital objects**, ensuring security, consistency and scalability.
- **Training and development programs to strengthen capability in digitisation and digital curation** to ensure institutions can effectively manage and curate digital collections.
- **A system for assigning Globally Unique Persistent Identifiers (GUPRIs) to physical objects and their digital twins**. GUPRIs allow for automatic and unique identification, including projection data, and would improve data integrity and research use.
- Transformative digital research infrastructure initiatives identified by Australia’s biological collections community, including a “**national DNA sequence library, national advanced imaging store and using AI to extend physical collection specimen discovery**” and a “**common standards-based ‘digital extended specimen’ approach to digital documentation**”.⁴

Developing a skilled workforce to underpin a digital and data-driven research ecosystem

Building capacity in digital research infrastructure requires strategic and sustained investment in education, workforce development, and resources to meet the growing demands of data-intensive research. The NDRI Investment Plan could invest in programs throughout the training pipeline to build the skilled workforce required to operate, manage and harness cutting-edge NDRI. For example:

- **Programs in tertiary institutions to integrate digital skills into teaching and research programs**, preparing a workforce to meet the evolving needs of data-driven fields. Embedding these skills across academic disciplines would foster innovation and create opportunities for early-career researchers, enhancing Australia’s future in data-intensive research.
- **Initiatives to develop, attract and retain expertise in data science**. Competitive funding, secure employment opportunities, and professional recognition are critical to fostering talent in this high-demand field. In addition, targeted investments in professionals with specialised skills, including data curation, management, software development, code optimisation, and computational support, are needed.
- **Development of school programs focused on responsible and ethical use of AI tools** to prepare future generations for the challenges of the digital economy and data-intensive research landscapes.

To discuss or clarify any aspect of this submission, please contact Mr Chris Anderson, Director Science Policy at Chris.Anderson@science.org.au.

⁴ Australia’s Biodiversity Information Partnership. *Accelerating discovery and access to Australia’s biological collections*. https://ala.org.au/app/uploads/2024/07/Accelerating_Discovery_and_Access_to_Australias_Biological_Collections_2024_July.pdf (2024).