AUSTRALIA'S DATA-ENABLED RESEARCH FUTURE

SYNTHESIS REPORT

A collaborative project between Australia's Learned Academies, ACOLA and the ARDC





This project is the result of a partnership between the ARDC, Australia's five Learned Academies and ACOLA to ensure Australia can undertake excellent data-enabled research across all fields of research. Notably, the project sought to help build a more coherent data policy and strategic data planning environment to uplift national data infrastructure. Five domain reports were developed, and a synthesis report focused on common themes and multidisciplinary opportunities and needs. We hope that this project will transition into an ongoing national data policy and strategic planning capability.

This project received investment from the Australian Research Data Commons (ARDC). The ARDC is supported by the National Collaborative Research Infrastructure Strategy (NCRIS)

We also thank Scientell for their assistance with developing this report.

© Australian Council of Learned Academies (ACOLA) 2022

ISBN 978-0-6452054-5-9

This work is copyright. All material published or otherwise created by the Australian Council of Learned Academies (ACOLA) is licensed under a Creative Commons – Attribution – Non-Commercial 4.0 International (CC BY-NC 4.0) licence.

DATE OF PUBLICATION

June 2022

PUBLISHER

Australian Council of Learned Academies (ACOLA) 9 Gordon Street Acton ACT 2601 Australia www.acola.org

SUGGESTED CITATION

ACOLA (2022), Australia's Data-Enabled Research Future: synthesis report www.acola.org

By providing a forum that brings together great minds, broad perspectives and knowledge, ACOLA is the nexus for true interdisciplinary cooperation to develop integrated problem solving and cutting-edge thinking on key issues for the benefit of Australia.

ACOLA combines the strengths of Australia's five Learned Academies.



ACKNOWLEDGEMENT OF COUNTRY

ACOLA acknowledges all Aboriginal and Torres Strait Islander Traditional Custodians of Country and recognises their continuing connection to land, sea, culture and community. We pay our respect to the Elders both past and present.

Summary

Quality data underpin effective policy and decision making. Researchers are accessing, creating and sharing more data than ever, providing insights that can benefit Australia and the world, in fields such as healthcare, environmental management and urban development.

However, the increasing volume and variety of data bring challenges for effective and consistent planning, access, management and storage to support world-leading research. Australia's data infrastructure, assets, policies and skills must keep pace with the opportunities from the growth of data, especially for multi- and inter-disciplinary data-driven research. Data policies, infrastructure and assets vary across jurisdictions.

In 2021, the Australian Research Data Commons (ARDC) partnered with Australia's five Learned Academies and the Australian Council of Learned Academies (ACOLA) to ensure Australia can undertake excellent data-enabled research across all fields of research. Importantly, this project sought to help develop a more coherent data policy and strategic data planning environment to underpin national (system-wide) data infrastructure. We recognise that different disciplines have different requirements, although there are some common challenges and opportunities. However, it is also worth noting that interdisciplinary research has specific needs and requires coordinated support.

New forms of data and technology, such as social media and Artificial Intelligence (AI), will deeply affect research and knowledge creation; they require new ways of capturing, storing, annotating, and analysing data. There are three key actions needed to support data-enabled interdisciplinary research:

- advocating for investment in data as a public good, including beyond STEMM disciplines and industry
- 2. driving real action to implement the FAIR¹ and CARE² principles
- focusing on the shared governance and capability development for Australia's data future while maintaining a 'watching brief' on distinct discipline-specific needs.

As part of our response to the challenges, it is vital that the Academies and ACOLA engage deeply in the development of the Australian Government's proposed National Digital Research Infrastructure Strategy, under the 2021 National Research Infrastructure Roadmap, to ensure a fit-for-purpose and future-proof approach to Australian research data.

Delivering upon the challenges will lead to:

- enhanced public awareness, dialogue, social licence, and acceptability for data use
- improved data management and standards, including sharing of code and workflows
- improved data access and sharing to facilitate and maximise the use of all forms of data, (including human data), including between governments, academia and industry, and
- improved skills and training across all levels of the research system.

¹ FAIR = Findable, accessible, interoperable and reusable

² CARE = Collective benefit, authority to control, responsibility and ethics

Introduction



High quality and reliable data underpin timely, informed decision-making on critical issues of national interest, including public health, energy, food security, social cohesion, infrastructure planning and environmental management, and the expansion of our knowledge overall.

The amount and complexity of data generated across Australia's research sectors, and by the public in general, is growing. Data used in research come from observations, images, sound, measurements, samples, trials, computer programs and recordings. They can be numerical, descriptive or visual, and can be raw, cleaned, analysed or processed. Powerful sensors provide a wealth of new data, including environmental sensors and wearable devices. Social media offers unprecedented observations of societal dynamics. With this, researchers can explore topics never before accessible in ways that can benefit Australia and the world. For example, the effective use of big data could reduce Australia's annual healthcare expenditure by more than \$11 billion³. The pandemic and recent natural disasters have highlighted the importance of rapid access to data to assist with timely decision making, and subsequently the need for coordinated, integrated data infrastructure. The increasing volume and variety of data challenges effective and consistent access, management and storage. Despite the extraordinary volume of data being generated, there are still data gaps – many of which are critical for enabling evidence-based policy decisions.

Australia's research data infrastructure, assets, policies, funding and skills need ongoing investment to keep pace with the extraordinary growth of data and multi- and inter-disciplinary data-driven research. Research data collection, storage and access for many areas remain fragmented, with a complex network of policies, infrastructure and assets. However, Australia and the research sector are not alone in this challenge, and there is much optimism and action being taken to begin addressing it.

³ Centre for Big Data Research in Health (2019), Annual Report 2019

In 2021, the ARDC partnered with Australia's five Learned Academies and ACOLA to engage Australia's greatest minds, the Fellows of the Academies, to develop insights into ensuring all of Australia's research sector can undertake excellent data-enabled research. Importantly, the project sought to help build a more coherent data policy and strategic data planning environment to uplift national data infrastructure. The approach recognised that different disciplines have different challenges and starting points and considerations for interdisciplinary research.



The Australian Learned Academies' Future of Research Data project worked to understand and improve the coordination of research data priorities across disciplines. The project combines the leadership and strategic insights of the Academies and ACOLA with the national data infrastructure expertise of the ARDC.

Firstly, each Academy assessed how data assets are currently used within their discipline areas and what is needed to enable better utilisation in the future to achieve world-class research. Academies undertook extensive consultation, ran workshops, and completed scenario analysis, surveys and desktop research. These reports are available at:

- Australian Academy of Health and Medical Sciences (AAHMS) report
- Australian Academy of the Humanities
 (AAH) report
- Australian Academy of Science (AAS) report
- Academy of the Social Sciences in Australia (ASSA) report
- Australian Academy of Technology and Engineering (ATSE) report

This document identifies the key findings from across the Academy reports aimed at enabling world-leading interdisciplinary and multidisciplinary research by maximising the value of data and ensuring we have robust national data capabilities.

This work leverages and builds upon previous reports, including those by ACOLA⁴, Learned Academies⁵, the Australian Academy of Science's National Committee for Data in Science, Australian Research Data Commons⁶, OECD⁷ and the International Science Council.

⁴ Most notably Horizon Scan reports on Internet of Things and Artificial Intelligence.

⁵ Such as the Australian Academy of Science's Advancing Data-Intensive Research in Australia report and the Academy of the Humanities' Mapping International Research Infrastructures for the Humanities, Arts and Social Sciences report.

⁶ The Humanities, Social Sciences and Arts Research Data Commons report.

⁷ Notably OECD work around access to Research Data from Public Funding.

The value of data-enabled research

There is almost no end to the applications and benefits of data-enabled research. It can support personalised and targeted healthcare, underpin early action in the face of natural disasters, and provide detailed observations of how our environment and ecosystems are changing. Ultimately, more and improved data will add value to all research disciplines.



Australia, and Australians in general, face a growing number of domestic and international challenges. Each challenge is complex, affects people from diverse backgrounds differently and requires cross-sector solutions.

- We lack linked national datasets that integrate personal history, and clinical, imaging and other data that could help us tackle many health challenges, such as dementia, obesity and antimicrobial resistance.
- There are opportunities to improve and connect national datasets that integrate individual and family histories with social wellbeing indicators, such as participation in education, income and employment, housing, nutrition, and social and cultural capital. Such integration will shed light on the complex dynamics underpinning social cohesion, mobility and wellbeing, which are critical to informing Australia's evolving social policy.

- Responding to natural disasters requires early, accurate information. Assessments that leverage data and AI can more rapidly improve emergency responses, reducing impacts on communities, infrastructure and ultimately the economy.
- High-quality data can help us understand and manage our natural environment. New data streams from sensors and satellites, supported by high-performance computing, can help us develop mitigation strategies for climate change and sustainable management of water resources.
- Longitudinal studies of social, linguistic, and cultural phenomena provide deep insights into the experience and wellbeing of our population. Guidance for public policy and private investment in these areas depends on a long-term data and information base.

Many other topics benefit from collecting, synthesising, and applying large volumes and varieties of research data. These topics include agriculture, resource management, climate forecasting, wildlife conservation, infrastructure and urban planning, and financial services.

Big data, becoming bigger and complex data, becoming more complex



Advances in science and technology have created powerful, small and cost-effective sensors (leading to the emergence of what is often referred to as the Internet of Things), and other observation tools, such as wearables, CubeSats and smartphones. The data collected by these devices is richer and often real-time, which significantly enhances our ability to discover, analyse and understand changes in health, communities and the environment, potentially revealing far more about our lives. However, this is creating unprecedented volumes and velocity of data.

Big data refers to large amounts of structured and unstructured data that exceed the ability of commonly used software tools to capture, manage and process. New techniques and technologies are needed to process massive, complex datasets. Some of the most valuable research contributions come from synthesising or integrating data from different fields, for example, to understand the social and environmental factors for disease. These integrated datasets may or may not be big, but they are 'wickedly' complex from the perspective of semantics, standardisation, governance, and policy.

Data are transforming research, enabling new, complex analyses and exploration of research questions that weren't previously possible. The ability to collect, store, access, analyse, interpret, curate, and manage data is central to science and scientific collaboration.

Preserving the past; investing in the future

Universities, research organisations and funding agencies need to cover the costs of managing data across the lifecycle.

Data collections in galleries, libraries, archives, and museums are fundamental to understanding our history, culture, and identity. Many organisations are digitising content to help preserve and maximise the data utility and access to the public and researchers. However, there are limited funds for research data digitisation, and in the case of smaller, regional organisations, often a lack of expertise and infrastructure for digitisation. National conversations are needed on how best to cohesively and consistently manage these research collections, platforms and infrastructure, especially where they have ongoing national value. There may also be a need for central digitisation funding for organisations with limited capacity and where there is a national value in their holdings.

There is a need for coordinated approaches to data management, platform sustainability, research translation, and protocols for dealing with legacy datasets. Much data and associated infrastructure are funded via short-term projects. This presents challenges for archiving, re-use and sustainability.

- Large amounts of data remain in non-digital forms.
- Data are often collected for a time-limited project, complicating storage and re-use.
- New data sources can significantly enhance research insights, especially in social domains.
- We lack the capacity and capability to maximise our current data, let alone be ready for new opportunities and data needs.

New technologies, such as social media, represent significant value to the humanities and social sciences. However, traditional ways of capturing, storing, annotating, and analysing these data are increasingly unsuitable, both in their approaches and their capacity to handle the enormous volume of generated data. Significant investment in physical and human infrastructure is needed to capitalise on this explosion of new data.

Accessing and processing data

The ability to access, share and exploit data is key to preparing for and responding to societal challenges, including pandemics, biosecurity threats, resource security, natural disasters and climate change. To make new discoveries about Australia's cultures, history, environment and medical advances, and our connections globally, researchers need access to dispersed collections of qualitative and quantitative data. However, while larger volumes and varieties of research data continue to be produced, they are not always accessible to researchers.

Work by the Australian Chief Scientist and others is exploring open research access. However, this is primarily focused on publications. Open data is an extension and part of the open access agenda and requires a similar national approach.

Data from government, industry, and private organisations can be difficult to access for reasons such as privacy or security concerns, and commercial interests. Modern and responsive legal and ethical frameworks are essential to ensure responsible data use and manage risk, as well as appropriate levels of access and de-identification. Stakeholders need to unite to consider frameworks, actions and initiatives to integrate industry data with government and research data to maximise research insights. Governments in particular have an important role in planning and negotiating access to nationally significant data held or created by industry, for research purposes, which can range from environmental observations to social interactions and sentiment captured by online platforms.

- The lack of access and sharing arrangements is a significant lost economic, social and health opportunity.
- Our legal frameworks have not kept pace with the rapid advancements in data.
- FAIR and CARE principles provide a sound and agreed basis for data management and practices, but they are not mandated.

The Data Availability and Transparency Act 2022 helps enable access and use of government data for research purposes, where research is demonstrably in the public interest. The Integrated Research Infrastructure for the Social Sciences (IRISS) project⁸, which the ARDC supports, is developing and testing various platforms, standards and workflows needed to provide access to relevant government data for social research. There will be a need to determine the best infrastructure and approaches to harness these emerging opportunities. Academies can play an important role in ensuring the perspectives of their constituent disciplines are considered in future infrastructure and investment decisions, including those that facilitate cross-disciplinarity. Similar progress will need to occur in the private sector to ensure that these types of data can be used openly and transparently for public interest research.

⁸ https://ausiriss.org.au/

The FAIR (findable, accessible, interoperable and reusable) principles encourage the fair and equitable access and sharing of research data. However, these principles must be adopted consistently across Australia's research system. The principles also need to be extended to government and industry data to ensure they are discoverable and interoperable, nationally and internationally.

While beneficial examples exist, the collection and use of data have not always served everyone's interest, including First Nations peoples, and in addition to FAIR there is a need to adopt the CARE principles (collective benefit, authority to control, responsibility and ethics). First Nations knowledge is at the forefront of national discussions on Indigenous data governance and sovereignty. Aboriginal and Torres Strait Islander researchers have raised concerns over the rights of Australia's First Nations peoples to data that involves their health, communities, cultures, languages and more. Policies and principles must ensure data use supports self-determined interests. Importantly, the FAIR principles alone are insufficient to address data sharing and data use for Aboriginal and Torres Strait Islander health and medical research. Coordinated access and sharing arrangements and processes between jurisdictions, research and industry would facilitate data use. Jurisdictions, institutions and sectors often employ different data storage technologies, systems and formats. A national approach is needed to ensure both the interoperability and the longevity of data. Uniform and consistent application of metadata formatting would be an important step, but this would require national leadership to drive adoption.

A nationally coordinated, interoperable (including with international approaches⁹) and integrated ecosystem of data repositories for research disciplines would be a complex and ambitious endeavour but would provide vital data access and insights.

It is unclear how this should be structured and managed. One model, suggested by AAHMS for the health sector, could be a 'plug and play' national data spine to link data across states, territories and national assets, including new data from registries and research studies. Careful consideration and workshopping is needed, including on the appropriate architecture, interoperability and change management needed to implement it.



⁹ This can be achieved through the adoption of international data standards and participation in global networks such as the Research Data Alliance and CODATA.

Privacy, security and ethics

Ethics processes and data access requirements need to be simplified. Complex ethics applications to access datasets, inappropriate for the level of risk, can delay or prevent research. An example is the complex ethics and regulatory frameworks regarding data stored in Australia's biobanks, which are controlled at a state and territory level and are not easily integrated into national systems. There could be opportunities to work collaboratively with a range of stakeholders to progress improvements, especially the NHMRC which has a national leadership role in research ethics.

Coherent data policies, training for researchers, and provision of data expertise can reduce the burden on, and improve outcomes for, researchers and data subjects regarding data licensing, IP rights, and privacy.

However, as noted previously by ACOLA and the Australian Human Rights Commission, while technologies such as AI can improve efficiency, we need to ensure they do not perpetuate past or current social biases observed within data. For example, there is under-representation, under-reporting and poor-guality data for certain groups of people, including culturally and linguistically diverse peoples, and training of algorithms using inappropriate and incomplete data sets can perpetuate social, cultural and economic biases. Access to data must not be taken for granted. The research community and data custodians need to build a social licence and acceptability for data use in fields such as humanities, social sciences, and health and medical research by improving public awareness and trust, seeking to understand and act on public perceptions and concerns, and protecting the privacy of data collected from individuals. Several infrastructures have been, or are being, developed to improve access and management. Examples include data.gov.au for Australia's open government data, the planned data inventories by the Office of the National Data Commissioner¹⁰ and the Australian Government Department of Health's goal to create a 'National One Stop Shop' platform to streamline and coordinate the approvals process for health-related human research.

- Australia needs to further develop and build social licence and acceptability of data for research purposes.
- Australia's deep ethical expertise should be leveraged to guide the future trust and design of data infrastructure.

Social science expertise can contribute to the safe, ethical and socially acceptable design and use of data commons systems, processes, and policies. This is particularly important with applications of new technologies such as big data and Al.

Skills and training

Data is fundamentally changing the skills landscape. While there the need for enhanced and advanced data teams in institutions and national bodies such as the ARDC, we need to build the data management and use capabilities of researchers.

Notably, the changing social and research environment, including the rapid increase in video and social media, requires the adaptation of social research methods and the training of humanities, arts and social science researchers. Data scientists will also be fundamental to underpin this transition.

- Researchers need new and additional skills to utilise data.
- The training of researchers is the primary responsibility of institutions.
- However, the complexity of data analysis often requires specialist skills, requiring specific training and recognition.
- Social and video media is a new frontier requiring investment to understand and maximise the value from data.

¹⁰ Office of the National Data Commissioner (2021), New pilot program to deliver a government-wide data catalogue

The 2021 National Research Infrastructure (NRI) Roadmap states that 'People and expertise are an intrinsic and essential part of NRI. NRI staff manage, maintain, optimise and operate facilities, and a highly skilled workforce is needed to maximise the benefits of existing NRI investments, new technologies and applications.'

Australia needs to create a digitally skilled workforce and invest in data literacy training. The research sector faces shortages in data collection, analysis, interpretation and development of data-informed policy. Universities will play a key role in training graduates and researchers in core data literacy and data skills.

Data-intensive research will require specialist teams of data scientists and other data professionals with technical skills to support research, especially to create, curate, manage, share and integrate data, especially large datasets. As data infrastructures develop, there will be a need to ensure training for researchers. Data expertise is needed, particularly at the interface between discipline experts and research infrastructure facilities.

It will be important for universities and the research sector to develop pathways for career advancement in data skills alongside academic pathways. There is a need to attract and retain data scientists and professionals, especially in a globalised workforce.

While the training of researchers may be the primary responsibility of institutions, collaboration between institutions will be important to avoid duplication and re-inventing the wheel.

Improving efficiency

With the exponential increase in the volume and velocity of data, data collection will need to be more efficient. This can be achieved through techniques such as sharing code and workflows to reduce duplicate effort, and through new technologies such as machine learning, AI and other tools that can automate the management and analysis of data.

- New technologies and processes can help improve the efficiency of data-driven research, including processing.
- Investment in exascale data capacity is needed to support the scale and complexity of data.

Al can help address important multidisciplinary problems. It can analyse and identify patterns in large and complex data sets faster and more precisely. For example, Al can improve standard clinical-based assessments, particularly with image analysis, and search scientific literature for relevant studies, link information and combine different kinds of data.

Al systems are likely to become more advanced and perform a broader range of tasks with less human input. Investing in these systems could substantially enhance data-enabled research.

Social and cultural analyses are integral to future projects that develop AI systems. Exploring the connections between an AI's technical design and its cultural and social implications will be key to ensuring feasible and sustainable AI systems that benefit society.

Moving and manipulating large volumes of data is a challenge, particularly for computationally intensive disciplines such as astronomy, climate research, and bioinformatics. Distributed computing, cloud computing and 'moving code to data' are options for data integration and analysis of such large datasets for some research fields. Investment in exascale data capacity is needed to support these disciplines.

To realise the potential of Al in research, critical progress needs to be made in data planning, quality, standards, ethics as well as in the transparency of code and data.

Bringing it all together – key areas for action

While acknowledging specific discipline needs and priorities, there are three key challenges in the current state of data-enabled interdisciplinary research:

- advocating for investment in data as a public good, including beyond STEMM disciplines and industry
- 2. driving real action to implement the FAIR and CARE principles
- focusing on the shared governance and capability development for Australia's data future while maintaining a 'watching brief' on distinct discipline-specific needs.

- Embedding the principles of FAIR and CARE is critical.
- Clearer accessible documentation of data policies, metadata and assets is vital.
- Stronger investment in HASS data infrastructure will be key to interdisciplinary research that will inform good policy.
- A national data strategy, with real actions, is needed to provide a catalyst and focal point for action.



Strengthening coordination mechanisms across institutions, organisations, platforms, programs, and alliances would be an efficient and effective approach to improving data access and application. While cross-sector coordination and leadership are needed, this is a shared responsibility with a range of mechanisms that can be used. It will be important that this is documented and communicated effectively to harness and galvanise individual and collective actions.

There is increasing agreement that data is a public good and alignment is needed between research, government and industry, and jurisdictions across Australia and internationally. This includes between traditional HASS and STEMM domains to inform good policy and research translation. However, the implementation and actions need to be commensurate. Importantly, there is a need to review and potentially better address the implementation of FAIR and CARE principles, data management and sharing and data infrastructure coordination.

Documenting critical databases, data assets and infrastructure nationally will help guide data integration, prioritisation and development of a national approach. We need to ensure national coordination to avoid duplication of effort, and regulation and reporting burden. Key challenges exist in the humanities and social sciences disciplines, which have traditionally not been as data-focused or funded for data projects, and with institutional arrangements (including international standards bodies). A national and standardised approach to the development and management of infrastructure metadata and data to support interoperability is unlikely to occur without investment. The development of a national digital research infrastructure strategy, as proposed by the Australian Government in the 2021 NRI Roadmap, with research data priorities, will be instrumental in driving a coherent data policy, infrastructure and investment. It will also help ensure future research needs are met and aligned with public research and policy needs. While important, the current focus on manufacturing and other applied sciences limits the potential for research to support net economic, social and cultural growth and benefits. It will be critical that the Strategy covers the breadth of data, and ensures greater inclusion of humanities and social science priorities and the need to enable interdisciplinary research.

Nationally integrated research data assets, which can be a mix of centralised or distributed, and structured for global interoperability, for research disciplines to support inter- and multidisciplinary research are a high priority. Innovations like this would be world-leading but require high-level leadership and commitment from federal government agencies. Key will be a strong governance structure and sustainable resourcing to enable integration of datasets and provide expertise, data curation and management services, and data analysis support.

Next steps

Every day brings more data – by 2025 the World Economic Forum¹¹ estimates that 463 exabytes will be created globally each day. This presents an unprecedented opportunity for research. When harnessed, data can unlock insights into all aspects of human endeavour and our natural environment. But to access this potential, data need to be managed, accessed, processed and shared. Critical to this are soft and hard infrastructure and investment in the skills of researchers and data managers, especially to keep pace with the needs of inter-, intra- and multidisciplinary research.

The Academies' reports have set out the opportunities and challenges associated with the successful application of research data in Australia at a discipline level. ACOLA strongly encourages stakeholders to read these reports. Importantly, these reports recognise that disciplines are not equal in their starting point or priorities. However, when considered collectively, there are clear opportunities and commonalities to overcome. If addressed, through focused actions, coordination and collaboration, a research revolution can be achieved.

Coordinated and integrated national data infrastructure, assets and policies that are futureready, adaptable and sufficiently resourced, and connected with institutional and global assets, are critical to supporting scientific research and enabling transdisciplinary breakthroughs.

Academies and ACOLA have key roles, actions and responsibilities to deliver on the key challenges for data-enabled interdisciplinary research through national data infrastructure, policies and skills, which might include:

Advocating for investment in data as a public good, including beyond STEMM disciplines and industry

• Lead national discussions on the value of data, and the impact that could be realised, especially in interdisciplinary areas.

- Individually and collectively, advocate for national research data policies and strategies, with a focus on enhancing data awareness, access, skills and governance.
- Building on previous work, develop advice for government on building societal trust in data
- Support greater recognition of non-traditional research inputs and outputs and impacts that leverage data.

Driving real action to implement the FAIR and CARE principles

- Seek champions among the Fellowships to promote and encourage the adoption and implementation of the FAIR and CARE principles in exemplar interdisciplinary research areas.
- Advocate for FAIR and CARE to be a required policy for all publicly funded research and that the costs of managing research data are recognised in funding policies for universities, agencies, grants and funding councils.

Focusing on shared governance and capability development for Australia's data future while maintaining a 'watching brief' on distinct discipline-specific needs

- Work with key government research funders, especially the Australian Research Council, the National Health and Medical Research Council and Medical Research Future Fund, to determine how to drive improvements in nationally coordinated activity on data through grant arrangements.
- Commit to working with partners to assess data and digital assets within and across jurisdictions and institutions, and identify the tasks and investments needed to integrate them into a commons infrastructure.

¹¹ World Economic Forum (2019), How much data is generated each day?

- Include data issues, barriers¹² and opportunities for data-intensive research in discipline planning, including leveraging existing activities where possible such as the Humanities, Arts and Social Sciences Research Commons Infrastructure project.
- Academies and ACOLA to meet annually to identify progress and further synergies in respective data initiatives and review the data landscape.
- Promote best practice in Indigenous data governance.
- Promote and support the growth and consolidation of workforce expertise and skills, in researchers and infrastructure staff, to ensure the effective use of data.

There are opportunities and roles for the ARDC to complement and collaborate with the Academies and ACOLA, with appropriate resourcing.

- Plan and develop national digital infrastructure capability to support research priorities.
- As part of its leadership role, foster a vision and support a more coherent data policy and standards, skills and infrastructure environment, across and within domains, to accelerate data-enabled research.
- Monitor and inform key national policy settings for data-enabled research.
- Identify critical data skills for research and professional staff to inform institutional and discipline skills roadmaps.
- Support the implementation of FAIR and CARE principles through services and expertise in data structures, standards, informatics, and metadata.
- Demonstrate cross-sector and multidisciplinary data integration and governance to address grand challenges.

Finally, the Australian Government's proposed national digital research infrastructure strategy will be vital in encouraging and underpinning the delivery of nationally consistent policies, infrastructure, assets, interoperability (metadata and collection standards), governance and ethics across Australia's research sectors. ACOLA and the Learned Academies should support and provide sector leadership in its development, with a focus on leading input that aims to:

- deliver a lift in data capability and capacity across all disciplines, including multi- and interdisciplinary needs, recognising the different states and critical needs of different disciplines to improve overall maturity
- ensure coordination across jurisdictions, with consistent policies and standards
- develop protocols and processes for sharing code, data and workflows
- assist in developing positions on data preservation and retention, including agreement on what constitutes nationally significant datasets
- deliver national leadership and models for facilitating a national skills and training data agenda
- lift minimum data skills and expertise for university graduates
- promote planning and investment in the expertise required to support data infrastructure
- implement the FAIR and CARE principles for data and software across national research infrastructure
- provide insights into the benefits and applications of industry data
- build public, dialogue, awareness and social licence and acceptability for data use, and
- consideration of technologies such as AI to improve the processing and use of data.

¹² Barriers to enabling researcher access to human data held by government and industry include legislative, social licence, research ethics regulations, and secure data storage/handling infrastructure.





By providing a forum that brings together great minds, broad perspectives and knowledge, ACOLA is the nexus for true interdisciplinary cooperation to develop integrated problem solving and cutting edge thinking on key issues for the benefit of Australia. www.acola.org

Working together: the Australian Council of Learned Academies (ACOLA) combines the strengths of the Australian Learned Academies.

LEARNED ACADEMIES











Australian Academy of Health and Medical Sciences