Australian Research Data Commons

Research Data Management Framework for Institutions

An output of the
Australian Research Data Commons Institutional Underpinnings Program
February 2023

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative

The ARDC is enabled by NCRIS

CC BY
Text & infographics only.

DOI: 10.5281/zenodo.6392340

Suggested citation:

Australian Research Data Commons. (2023) Research Data Management Framework for Institutions.

Viewed online at:
http://doi.org/10.5281/zenodo.6392340

FOREWORD

Data is the life-blood of the modern research ecosystem. For universities, developing the infrastructure and frameworks to create a data-rich institutional approach to research is vital to every part of their operation. Harnessing this data to tackle the world's greatest problems and build research capacity is essential for our nation.

The World Economic Forum has identified data as one of the world's biggest challenges.[1] They propose that "Data can help us tackle our largest societal challenges, including climate change, inequality, global health and economic resilience". Often, however, data is inaccessible, poorly described and poorly managed.

Now is a time of great experimentation and learning. The skills, policies and integrated landscape of systems required to support a modern university are still in development. The Australian Research Data Commons Institutional Underpinnings program offers universities a unique opportunity for sharing knowledge and creating a framework that can be adapted innovatively to address this challenge. This emerging framework supports a new approach to ethics and responsibility for curating collections — an approach that is transdisciplinary and visionary.

Bringing together 25 universities has been no small achievement. Tackling the breadth of data issues for research has also taken a major commitment, focused on reshaping and integrating data services from all areas in our universities. I congratulate all involved in this world-leading program and encourage all researchers and those providing supporting services to carefully read this report and take action.

Roxanne Missingham OAM
University Librarian
Australian National University

Cover Image — ben-leung - tz_Jr5QcxF8 / unsplash.com

The ARDC is enabled by NCRIS

MESSAGE OF THANKS FROM THE ARDC'S CEO

Good research data management helps to drive research and innovation. When data is managed well, researchers can get more out of the data that they collect, and can find new value in existing data. Universities, and other research institutions, play a large and vital role in the management of Australia's research data assets. They provide the infrastructure and support that enables researchers to manage their research data effectively, and underpin the creation and longevity of the national-scale data collections that enable new discoveries.

The Institutional Underpinnings program has demonstrated the enthusiasm of universities to come together and collaboratively address the challenges of supporting research data management in a varied, fast-changing and ever-expanding research data landscape. The 25 universities who developed this framework represent a major part of Australia's research sector, with an enormous amount of experience in supporting research data management. I thank them for their participation in this program. Their commitment of time, effort and expertise, and their willingness to share with one another, has resulted in a valuable resource for all research institutions. By working together in this way, we move towards a more joined-up research landscape that supports collaboration and excellence in Australian research.

Rosie Hicks
Chief Executive Officer
Australian Research Data Commons

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative
The ARDC is enabled by NCRIS

ABBREVIATIONS

AIATSIS
Australian Institute of Aboriginal and Torres Strait Islander Studies

ARC
Australian Research Council

ARDC
Australian Research Data Commons

ADKAR
Awareness, desire, knowledge, ability and reinforcement

ATSILIRN
Aboriginal and Torres Strait Islander Library, Information and Resource Network

CARE
Collective Benefit, Authority to Control, Responsibility, and Ethics

CODATA
Committee on Data of the International Science Council

COPE
Committee on Publication Ethics

DCC
Digital Curation Centre

DMP
Data management plan

DP
Digital preservation

DOI
Digital object identifier

DReSA
Digital Research Skills Australasia

DVC
Deputy Vice-Chancellor

DVCR
Deputy Vice-Chancellor – Research

FAIR
Findable, accessible, interoperable and reusable

FAQ
Frequently asked question

GDPR
General Data Protection Regulation

HASS
Humanities, arts and social sciences

HDR
Higher degree by research

HPC
High performance computing

ICMJE
International Committee of Medical Journal Editors

IP
Intellectual property

NCRIS
National Collaborative Research Infrastructure Strategy

NHMRC
National Health and Medical Research Council

ORCID
Open researcher and contributor identifier

PDCA
Plan-Do-Check-Act

RACI
Responsible, accountable, consulted, informed

RAiD
Research activity identifier

RDA
Research Data Alliance

RDM
Research data management

RISE
Research Infrastructure Self-Evaluation

ROR
Research Organization Registry

SEO
Socio-economic objectives

SWOT
Strengths, weaknesses, opportunities, and threats

TEQSA
Tertiary Education Quality and Standards Agency

TOP
Transparency and Openness Promotion

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

CONTENTS

CONTENTS

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

EXECUTIVE SUMMARY

The Institutional Underpinnings program is led by the Australian Research Data Commons (ARDC). In this program, 25 Australian universities collaborated over 18 months during 2021 and 2022 to develop a national framework for institutional research data management (RDM). The framework is intended to inform institutions' design of policy, procedures, infrastructure and services, and improve coordination of RDM within and between institutions.

The participating universities identified 19 elements essential for RDM in a university. Of these, they identified 9 elements as opportunities for immediate collaborative action. Expert working groups developed recommendations and calls to action for these 9 elements. The universities tested the effectiveness of this guidance in local projects. The findings from these projects were combined with feedback from national infrastructure providers and other Australian research institutions to produce the final version of the framework presented here.

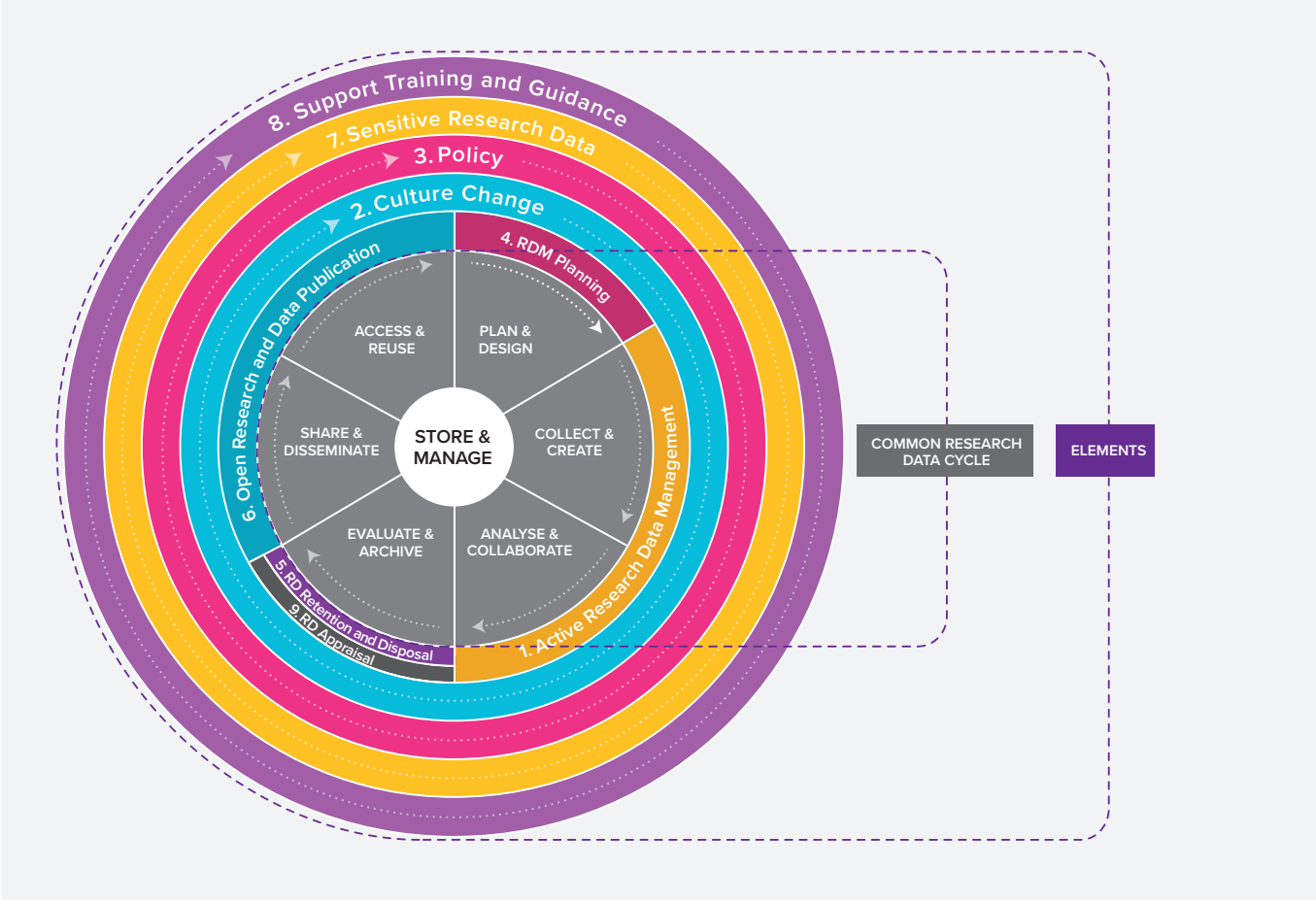

Figure 1. *Schematic of a common research data life cycle, and how the framework elements map to it. (Adapted from the Research Data Lifecyle by LMA Research Data Management Working Group under a Creative Commons Attribution-NonCommercial 4.0 International License.)*

The ARDC is enabled by NCRIS

EXECUTIVE SUMMARY

The Institutional Underpinnings program is led by the Australian Research Data Commons (ARDC). In this program, 25 Australian universities collaborated over 18 months during 2021 and 2022 to develop a national framework for institutional research data management (RDM). The framework is intended to inform institutions' design of policy, procedures, infrastructure and services, and improve coordination of RDM within and between institutions.

The participating universities identified 19 elements essential for RDM in a university. Of these, they identified 9 elements as opportunities for immediate collaborative action. Expert working groups developed recommendations and calls to action for these 9 elements. The universities tested the effectiveness of this guidance in local projects. The findings from these projects were combined with feedback from national infrastructure providers and other Australian research institutions to produce the final version of the framework presented here.

The 9 fully developed elements are as follows:

- **Element 1: Active Research Data Management** addresses institutional approaches to providing the infrastructure for managing research data during the life of the research project for the purposes of conducting that project. Institutions have a responsibility under the *Australian code for the responsible conduct of research*[2] to give researchers access to infrastructure to use throughout active RDM. This part of the framework deals specifically with selecting the active RDM infrastructure by an institution. Because active RDM takes place during the life of the research project, it has a large impact on a researcher's ability to conduct their research. Effective active RDM solutions reduce the burden on researchers' workload and prevent them from turning to non-endorsed solutions that reduce institutional oversight and expose both institution and researcher to risk.
- **Element 2: Culture Change** is the shifting of RDM practices within an institution towards more effective RDM. This part of the framework includes both approaches to changing institutional staff attitudes and practices, and changing the institutional processes, guidance and incentive structures that motivate and support these attitudes and practices. Culture change must be actively considered when improving institutional RDM to ensure that the goals identified by the institution are translated into practice.
- **Element 3: Policy** lays out the principles that govern the institution's approach to RDM. Effective RDM policy gives an institution a structured approach to meeting its regulatory requirements and ensuring that the required roles, responsibilities, processes and procedures are in place for effective RDM.
- **Element 4: Research Data Management Planning** addresses institutional considerations for planning the management of data emerging from a research project. This planning is typically done by the researcher with the support of the institution, which may supply tools to guide the researcher through the planning process. RDM planning ensures that researchers carefully consider the management of their research data, leading to better RDM practice within an institution. As they take the researcher through the planning process, institutions can introduce researchers to institutional infrastructure and processes. RDM planning may be documented by the researcher in a data management plan, which should help with institutional oversight of research data, and can be used to inform the provision of RDM infrastructure and services.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

EXECUTIVE SUMMARY

- **Element 5: Research Data Retention And Disposal** refers to the decisions about what to do with data at the end of a research project. Institutions are responsible for large quantities of research data, and retaining this data to meet their regulatory requirements, to back up the integrity of their research, and as a valuable asset in its own right. However, not all data can be retained indefinitely, so institutions must have systems in place to allow for the appropriate disposal of data. Good policy, systems and procedures for retaining and disposing of research data become increasingly important for institutions as the cost of storing ever-increasing quantities of data rises.
- **Element 6: Open Research And Data Publication** can improve the visibility and impact of research. The research sector is moving towards a more open model, where the data underlying research is made available in aid of research integrity, reproducibility, collaboration and innovation. Research funders and publishers are increasingly requiring that research data be made open. By enabling open research and data publication, institutions help researchers to meet these requirements.
- **Element 7: Sensitive Research Data** is data that presents a risk to persons, groups, the environment or society at large if it is disclosed or mishandled. Special protections are required when managing sensitive research data within an institution. Institutions must therefore account for data sensitivity when putting in place RDM infrastructure and procedures. Ensuring that sensitive data is managed appropriately protects institutions from legal and reputational risk.
- **Element 8: Support, Training And Guidance** addresses institutional approaches to providing researchers with the essential knowledge to manage data effectively. Effective support, training and guidance enables researchers to effectively use the services and infrastructure provided by an institution, provides research efficiencies, and improves RDM practice across the institution.
- **Element 9: Research Data Appraisal** is the process of reviewing data collections to assess their ongoing value and determine retention requirements. Effective appraisal requires sufficient information about the data to be appraised, well-defined triggers for appraisal decisions, and clarity within the institution about who has the responsibility to appraise data.

A further 10 elements identified by the universities were seen as priority areas but were not an immediate focus of the working groups. However, considerations for these elements raised by the program group are summarised in this framework:

- **Element 10: Data Sharing And Access**
- **Element 11: Cybersecurity**
- **Element 12: Data Ownership**
- **Element 13: Digital Preservation**
- **Element 14: Funding And Sustainability**
- **Element 15: Governance**
- **Element 16: Identifiers And Metadata**
- **Element 17: Non-Digital Material**
- **Element 18: Standards And Guidelines**
- **Element 19: Indigenous Data Management.**

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

INTRODUCTION

Background

The ARDC is a National Collaborative Research Infrastructure Strategy[3] (NCRIS) facility, providing Australian researchers with a competitive advantage through data. Through strategic partnerships between the ARDC and key stakeholders, projects in this initiative will develop national-scale data assets that support leading-edge research. National data assets are created and maintained through cross-institutional collaboration and require long-term custodians. The Institutional Underpinnings program focuses on universities as major producers and custodians of Australian research data. This framework was therefore developed by universities with their needs in mind.

Other research institutions may benefit from the guidance provided by the framework, but, given the differences in structure and drivers between universities and these other research institutions, not all of the advice may be directly applicable. The program aims to foster a more cohesive, collaborative approach to RDM across Australia's universities and research institutions more generally, and to support an overall uplift in their RDM capability in line with the responsibilities outlined under the *Australian code for the responsible conduct of research*[2].

RDM is a complex challenge for institutions, requiring the coordinated effort of multiple stakeholders within the institution. This challenge is becoming more pressing as the quantity of research data generated within institutions increases and the sector moves towards increasingly open research practices, with a focus on data availability for both integrity assurance and reuse. Institutions are also facing financial pressure to improve the efficiency of their RDM, because storage costs are no longer decreasing fast enough to allow all research data to be stored indefinitely. The Institutional Underpinnings program aims to help institutions to face these challenges by developing a coordinated approach so that they can pool their efforts and learn from one another's experiences, rather than each attempting to develop their own solutions in parallel.

Australia's universities are seizing the emerging opportunity for data management at the enterprise level to support research efficiency, integrity, excellence, and innovation, and align to the Findable, Accessible, Interoperable and Reusable (FAIR) principles[4] and the Collective Benefit, Authority to Control, Responsibility and Ethics (CARE) principles[5] for data. The Institutional Underpinnings program allows the institutions to pool resources and define their own priorities and roles in contemporary data-enabled research.

The Framework

The universities participating in Institutional Underpinnings have produced a framework that outlines those elements required for effective institutional RDM, including policy, procedures, services and infrastructure. The Research Data Management Framework for Institutions is intended to inform institutions' design of policy, procedures, infrastructure and services, and improve coordination of RDM within and between institutions.

The framework is intended specifically for use by Australian universities, but may be useful to other research institutions. It provides an institutional-level perspective, and is aimed at supporting decision-makers and service providers.

Researchers produce and manage research data within institutions and rely on institutions for support and services. While the framework is intended to benefit researchers, it is not intended to provide advice or guidance directly to the researcher, but rather to assist institutions in supporting their researchers.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Our Approach

The Institutional Underpinnings program presented an unprecedented opportunity to approach the challenges of RDM with a large group of Australia's research institutions. We took a co-design approach, in which the group collectively determined the structure and content of the framework based on their shared priorities. This co-design process was iterative, with multiple opportunities for feedback and revision by key stakeholders across all participating universities, as well as input from the broader sector.

In our approach, we acknowledged that RDM is a large and complex challenge to which no one institution has the perfect solution. Given the limits of this project, it would be impossible to produce a comprehensive and complete RDM framework that perfectly served all of Australia's institutions. Additionally, RDM needs to constantly change with advances in technology, regulations and agreed best practice. For these reasons, we focused on how we could produce the most shared value for institutions in the face of their current challenges.

To make the most of this opportunity for collaboration, we focused on those areas where most institutions could benefit from drawing lessons from those institutions who are leading in the sector, and also areas where collective work is required for progress. The latter might occur either because a problem is so challenging that combined effort is needed, or because the solution requires agreement between institutions (for instance, to facilitate cross-institutional research collaboration). We also aimed to highlight those places where agreement is difficult and where shared solutions are needed but do not exist.

Institutions have different resources and needs, based on their scale and the maturity of their RDM capability. For this reason, it was important that the framework be technology-agnostic rather than prescriptive. Our approach was therefore to provide the principles that underlie successful solutions to problems, with examples that show how these are applied in practice.

The Process

All 43 of Australia's universities were directly approached to participate in the program. Of these, 34 registered their interest and took part in a pre-program workshop in late 2020 where their desired outcomes for the framework and their participation in the program were identified. Subsequently, 25 of these universities applied to take part in the program, all of whom were successful.

The participating universities were:

Australian National University
Bond University
Charles Darwin University
Curtin University
Edith Cowan University
Federation University Australia
Griffith University
Macquarie University
Monash University
Queensland University of Technology
Swinburne University of Technology
The University of Adelaide
The University of Melbourne
The University of Queensland
The University of Sydney
University of Canberra
University of New England
University of New South Wales
University of Southern Queensland
University of Tasmania
University of Technology Sydney
University of Western Australia
University of Wollongong
Victoria University
Western Sydney University.

The ARDC is enabled by NCRIS

INTRODUCTION

Because RDM is complex and involves multiple parts of a university, participation was at the institutional level, with support from key stakeholders across the relevant business units at each university. During the co-design phase, each university put forward a key representative to take part in discussions and decision-making; they were also required to have a mechanism for internal consultation to ensure that all relevant views from their institution were represented.

The editorial committee was composed of nominees from the participating universities. Editorial committee members were selected to ensure that the committee represented a range of different RDM-related roles, areas of expertise, and experience with different university scales and needs.

Work on the framework began in 2021. We began by collectively identifying concepts and issues in RDM that were important, challenging, and/or the subject of current work and development. The editorial committee arranged these into broad topic areas (the "elements" of the framework), and the group collectively described these elements and identified the key considerations and challenges faced for each element. This framework outline was then reviewed by internal stakeholders across the participating universities, and was refined to produce an initial list of 16 elements.

Of these 16 elements, we collectively identified 8 elements for immediate development within the program:

- Element 1: Active research data management
- Element 2: Culture change
- Element 3: Policy
- Element 4: Research data management planning
- Element 5: Research data retention and disposal
- Element 6: Open research and data publication
- Element 7: Sensitive research data
- Element 8: Support, training and guidance.

These 8 elements were identified as those with the highest current importance to the participating universities. The universities were then asked to nominate experts to develop the elements, and, where possible, test and validate them in the next phase of the program. Later in the process, the group identified the need for an additional element — Appraisal.

For these 9 elements, the universities nominated 99 experts to participate in working groups. The working groups produced recommendations and guidance for each element.

Each of the participating universities designed a project to implement, test and validate a section of the draft framework in their local environment. These projects gave us feedback on the draft guidance, additional insights, and reusable outputs that are openly available for other universities.

We also consulted with other universities, research institutes and national infrastructure providers. The expert working groups and editorial committee incorporated the feedback from this broader sector consultation, and from the university projects, to produce the final version of the framework. During this process, 2 additional elements were identified for future development, giving a final set of 19 elements.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

The 19 Elements Of The Framework

The 19 elements of the framework offer a high-level overview of the key considerations for providing institutional RDM.

The 9 elements that were developed by expert working groups include an in-depth analysis of each of these topics. This includes the commonalities and differences in approach between universities, and the factors that drive those differences; the common challenges; and the principles that underlie successful solutions across universities.

Each element includes highlighted recommendations for universities seeking to implement or improve the particular aspect of institutional RDM.

The elements also identify gaps in knowledge or resources, and places where collective action is needed for the sector to advance. These are highlighted in 2 ways:

- 'Needs for shared information' highlight where participating universities identified a need for better understanding of approaches between institutions. This sharing could be achieved either through projects which actively collect information from universities on these topics, or through increased sharing in national forums such as the Australian Research Data Management Community[6] or other relevant communities of practice.[7]
- 'Calls to action' are highlighted to identify gaps or needs that can be addressed only by collaborative action between universities. These calls for collective action may form the basis of future collaborative projects.

Links to additional materials provided by the participating universities (such as case studies and example documentation) are in the digital version of this framework.[8]

It is important to note that the content in the working group outputs may not be comprehensive — working groups focused on the most important parts of an element or on the parts where they were best able to provide collective guidance.

Resources listed at the end of each element section are correct at the time of publication.

Potential Future Directions

As well as the calls to action identified within each element, the participating universities and sector feedback also identified more general next steps and future directions to build on the work represented in this framework. These next steps could be pursued by the Institutional Underpinnings program, or by projects conducted by others in the sector.

Shared terminology: Early discussions identified that it would be too time-consuming to develop a single shared set of definitions for key terms during the program, given the large number of partners. Each working group was therefore encouraged to be clear about how terms are used within specific parts of their guidance. However, there was strong support within the program for a longer-term effort to develop shared definitions of key terms (including "research data" itself) to clarify and simplify future discussions and guidance.

Extending the framework guidance: There was strong support for guidance to be developed for elements 10–19. Although the participating universities selected 9 of the elements for immediate development by working groups, all 19 elements were recognised as essential for university RDM, and collaboratively developed guidance on the remaining elements would be beneficial. Similarly, there was support for further developing some of elements 1–9; some of these contained a large number of potential sub-topics, and working groups were required to narrow their focus to those where they could immediately provide the greatest benefit. Revisiting these elements would allow further important guidance to be provided.

Implementation and maturity models: The framework recognises that universities have differences in scale and resources, and that their RDM may be at different levels of maturity. Care was therefore taken to avoid being prescriptive about how RDM services should be provided. Each university may need to approach the framework guidance differently depending on their needs and priorities. However, universities might benefit from assistance in setting their priorities and creating an implementation strategy for RDM improvements. One approach would be to build a maturity model based on the framework guidance, to help universities identify what they can realistically aim for at different levels of maturity. Alternatively, universities who are at an early stage of their RDM development might benefit from the description of a baseline set of tools or services that allows them to provide effective RDM. However, careful consideration is needed to avoid setting a standard that is not relevant to the needs of all universities.

Measuring impact: The framework guidance is intended to help institutions to review and improve their RDM services. The participating universities identified a need to evaluate the impact of these efforts, so that benefits can be assessed against costs and ongoing institutional support can be encouraged. This evaluation could be strengthened by collaboratively identifying key metrics (quantitative and qualitative) for measuring impact.

ELEMENT 1: ACTIVE RESEARCH DATA MANAGEMENT

Summary

This element provides institutions with guidance to ensure that research practice is efficient and impactful, and that research data is managed according to requirements (such as those outlined in the Australian code for the responsible conduct of research (the Code)[2]) and aligns to the FAIR data and CARE principles[4,5]. This element outlines some of the infrastructure considerations required to enable effective RDM and highlights the importance of user-focused design in ensuring adoption by researchers. Integrating RDM platforms with existing research administration systems is recommended. The importance of 'soft' infrastructure elements, such as governance and communication, is also discussed.

Why Active Data Management Is Important

Providing infrastructure, systems and services that facilitate good active RDM is an important responsibility of the institution. The Code states that:

> *Research data controlled by the institution and/or its researchers should be stored in facilities provided by or approved by the institution. These facilities, including information technology, must comply with privacy requirements and other relevant laws, regulations and guidelines, and research discipline-specific practices and standards related to safe and secure storage of data and information.*

Although the Code refers specifically to storage here, active data management is bigger than this. Active data management refers to the management of research data at any stage during the life of a research project, including the selection of what to collect/acquire, collection/acquisition, storage, analysis, visualisation and collaboration. It ends when the data is either disposed of or moved to long-term storage after project reporting is complete, and does not include sharing data for reuse after the life of the original project.

Figure 2. *The life cycle of research data. (Adapted from the Research Data Lifecyle by LMA Resesarch Data Management Working Group under a Creative Commons Attribution-NonCommercial 4.0 Internatinal License.)*

Effective active data management provision involves a combination of the technical infrastructure that supports these functions, the policy and procedure that shapes its use, and the people who maintain and run that infrastructure and assist researchers in interacting with it. This element focuses primarily on the provision of technical infrastructure; other considerations that contribute to effective active data management can be found in the policy, culture change, and support, training and guidance elements.

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

It should be noted that the Code mentions facilities provided by or approved by the institution. An institution may meet the infrastructure needs of its researchers in several ways. The infrastructure may be provided and maintained from within the institution or obtained from a commercial provider, and may include approved external services and platforms.

Institutions have several drivers that can shape decision-making on infrastructure for active data management. They need to ensure that the services they provide can be maintained and supported in a cost-effective and sustainable manner. They also require a good understanding of the quantity and value of data that they hold, sufficient information to make decisions about its retention and disposal, and sufficient oversight over that data to ensure compliance with legal and regulatory requirements.

While it is the institution's responsibility to provide (or endorse) the infrastructure for good active data management, it is generally the researcher who will be responsible for managing the data. Researchers are obliged to manage their data appropriately according to their responsibilities under the Code as well as other regulatory, legislative and ethical requirements.

Researchers' decisions about their use of active data management services are driven by many factors, including the need to complete their research efficiently, their knowledge and understanding of the options available to them, and their ability to use those options effectively and receive support when they have difficulty doing so.

A common active data management challenge for institutions is "shadow IT" — the use of non-institutionally endorsed services to capture, store and manage data. The use of shadow IT can open both institutions and researchers to risk, and makes it difficult for institutions to keep track of the data that they are responsible for. Researchers may turn to shadow IT when they are not provided with the services that they need to manage their data, when they do not understand how to access these services effectively, or when they find them overly burdensome. These difficulties can be worsened if the institution does not clearly articulate how research data should be managed.

However, even when researchers use institutionally supported systems for active data management, institutions can still have difficulty maintaining oversight over their data holdings, particularly where institution-supported systems have originally been provided for corporate or teaching and learning purposes rather than with the research use case in mind. A lack of governance and oversight by the institution can result in competing advice from different areas of the institution, multiple hurdles with the same information having to be provided to various areas of the institution and, in the worst cases, a shift of responsibility for all data governance decisions to the researcher.

Differences In Approach And Need

When it comes to providing infrastructure for managing active research data, there is no one-size-fits-all solution. Before deciding where and how to store and manage research data, institutions should understand the nature of the data their researchers are working with in terms of data types, size, and sensitivity. These considerations apply whether institutions are designing their own systems or purchasing them commercially.

Depending on the type of research being conducted at an institution, the profile of active data management needs may be quite different. Factors influencing this may include:

- Very large datasets
- Compute-intensive research
- Research requiring extensive pre-processing or specialised (including proprietary) software to handle data
- Data that requires particular technical or procedural protections due to its sensitivity
- Physical (non-digital) data assets such as biospecimens, signed consent forms or soil samples.

The profile of the researchers will also affect the needs of the institution. For instance, institutions with large numbers of adjunct positions or higher degree research (HDR) students will need to consider how to manage research conducted on devices that are not owned by the institution.

Institution size and resources can influence the selection of active data management infrastructure. It is important to note that size is distinct from maturity. An institution that deals with a smaller volume of research data may not need the level of automation that is required to manage much larger volumes of data, and so may develop quite different systems and infrastructure as compared to a larger institution. Resourcing will also play a role in decision-making, as institutions must ensure that they can sustainably support the infrastructure that they put in place within their budgets.

Finally, institutions differ in their internal structure and history of RDM solutions. These factors will lead to researchers and service providers having existing cultures, relationships, expectations and workflows that must be taken into account when making changes to how data is managed — a solution that works well for one institution may be difficult to implement at another.

Recommendations And Advice

Properties Of Successful Approaches

Clear Data Governance

Effective active RDM relies on a clear understanding of who has the responsibility to make decisions about data and its management, and on the implementation of systems to record and facilitate the upholding of those decisions. This governance spans from decisions about high-level, data-related policy to decisions about access to particular datasets. If it is not clear who “owns” important functions, they may fall through the gaps.

Poor data governance can affect the ability to provide effective data management services. For example, difficulties in integrating systems for active data management are worsened when it is not clear who has the overall responsibility for developing those integrations and ensuring that business processes are aligned.

Poor data governance can also affect researchers' interactions with institution-provided data management systems. Researchers may receive conflicting advice from different parts of the institution, or may be unable to find anyone who is able to advise them. If researchers' trust in the ability of the institution to manage their data appropriately is damaged, they may be further incentivised to turn to shadow IT.

One step towards improving research data governance is to map out the decisions and responsibilities related to RDM; for an example, see the *Monash Health data governance framework*[9].

Recommendation 1: Clearly articulate who is responsible within the institution for making decisions about the governance of active research data management.

User-Focused Design

For an institution's active data management solution to be effective, it must be used by researchers. It is therefore vital that their needs are understood. Researchers are likely to become frustrated with a system that does not have the functionality they require, is cumbersome to use, or appears to create unnecessary administrative load. A system that meets the technical and compliance requirements of the institution but does not consider user experience is unlikely to be successful, because researchers will be incentivised to seek shadow IT solutions that better meet their needs.

Researchers must therefore be involved at all stages of the planning process, both when existing systems are being updated, and when new ones are being developed/implemented. Local service providers need to understand the profile of needs at their institution, and identify where particular systems might place unexpected burdens on researchers.

Recommendation 2: Involve researchers at all stages of the Active Data Management solution planning process.

It is also important to recognise where needs vary between different research domains. For example, different domains have different needs for access control. HASS research may involve the contributions of citizen researchers who need to be able to access and add to a dataset; Indigenous data may require controlled access that allows communities to access their own data; and health research often requires that data can be accessed only by authorised personnel in a controlled environment.

Ideally, institutionally provided (or institutionally approved) systems should make research easier so that they are more attractive to researchers. One way that this can occur is through 'compliance by design', so that researchers are less burdened by the need to track and meet their compliance requirements.

An example of the co-design process is described in the case study: *Developing and implementing Monash Secure eResearch Platform*[10].

Recommendation 3: Make sure that investment in research data infrastructure is driven by the 3 principles of **researcher co-design, security by design** and, where applicable, **privacy by design**.

The ARDC is enabled by NCRIS

Clarity Of Communication

To be effective, an institution's active data management system must not only be used by researchers — it must be used correctly. It is therefore important that researchers understand what is expected of them when using these systems.

Institutions should clearly communicate to researchers (and other key stakeholders) what they can and cannot do with a given platform or service, so that obligations around security and privacy, for example, can be met. We also recommend providing information about why those recommendations/restrictions are in place, so that researchers understand them better. For an example, see the University of New South Wales (UNSW) guide, *Storing your research data*[11], which outlines the features of key supported storage platforms and highlights issues with common unsupported platforms.

Recommendation 4: Clearly outline the appropriate use cases and use terms for each platform and service, and disseminate information about endorsed platforms and services through multiple channels.

Different stakeholders may use the same terms quite differently, which can create problems for communicating about data management requirements. For example, a service provider may have redundancy in a system that allows retrieval of data in case of a catastrophic loss. If a researcher is told that the data is "backed up", they may interpret this to mean that they can retrieve their data if they accidentally delete it. This could lead users to unintentionally misuse services or select inappropriate services for their needs. We recommend clearly defining terms during stakeholder consultation and when presenting information about the correct use of institutional systems.

Integration For Seamless User Experience

To provide the best user experience, the active data management system should be viewed as a whole, rather than as a set of disparate services. Researchers will have to interact with the system at many different points, and a common complaint is that they are asked to provide the same information over and over again, creating an unnecessary administrative burden. Where possible, information should be passed between parts of the system to reduce double-entry and to allow the automation of processes.

As well as the technical challenge of creating integration between platforms, there are associated procedural challenges. A well-integrated system requires the identification of key metadata to support good management and automation, determination of the points where this metadata will be collected, and a common source of truth for this metadata (for more information see Element 4: Research Data Management Planning and Element 16: Identifiers and Metadata). This can be challenging when different parts of the system are managed by different business units within the institution with different goals and drivers. The metadata required for retention and disposal decisions should also be considered in these discussions.

The ARDC is enabled by NCRIS

Ongoing Investment Is Essential

Active data management systems should not be “set and forget” — they require active development, support and investment. Any IT infrastructure will require ongoing maintenance. User needs will shift over time as research practices advance, and may change dramatically in a short time if a change of strategy alters the research profile of the institution. The regulatory environment is not static, and there are also constant advancements in technology and associated best practices. Commercial providers can also make changes to agreements and technical specifications, which may necessitate a quick response to maintain the effectiveness and compliance of active data management systems. It is therefore very important that institutions recognise and support the need to factor ongoing investment into their budgeting around active data management.

Consideration Of Institutional And Individual Legacy

Any active data management solution will involve not just technology but also people and processes. The legacy of past RDM solutions will affect each of these and must be considered carefully. Institutional factors go beyond the migration of data and metadata into the new system — existing business processes may also need to be adjusted.

Similarly, researchers will have existing workflows based on the systems that they are used to (often using shadow IT), which may be seriously impacted by a shift to a new active data management system. We should also remember that for some researchers their unapproved solution for RDM represents a considerable personal investment of time and effort, and there is an emotional cost to moving away from this solution (particularly if the new solution does not have the ease of use or customisation of the old). Careful change management is therefore important.

The incentive structures for researchers are also important to consider. Under current incentive structures, time-poor researchers may invest less in RDM activities, which offer less reward than, for example, publication.

ELEMENT 1: ACTIVE RESEARCH DATA MANAGEMENT

Platform Selection

The platforms selected to underpin active data management will vary according to institutional needs. The following list of needs is not exhaustive; however, we outline some key considerations below:

- Disaster recovery
- Ability to support data discovery
- Interface with instruments that collect data
- Ease of local collaboration
- Ease of external collaboration (particularly the management of authorisation and authentication)
- Ease of use with required research tools, analysis environments and/or analysis workflows
- Ability to maintain central control/oversight
- Location of storage
- Ability to manage access, and audit actions performed on the data
- Security and privacy, with the availability of tiered access for sensitive data
- Capacity for version control
- Ability for remote access and syncing across devices
- Cost.

A need for shared information 1: Institutions are encouraged to share with one another information about the active data management platforms they are using, including their benefits and limitations.

Designing active data management systems requires an understanding of the types of data that must be managed and the requirements for those data types. We recommend the creation or adoption of a clear research data classification scheme, co-designed with researchers, to help with this task. Platforms can then be assessed against these requirements, and procedures developed to ensure that platforms are used appropriately.

Institutions often begin by designing an active data management solution for straightforward use cases, and then treating other data types as an exception. We recommend that the management of data with complex requirements (for example, sensitive data, big data) be considered from the beginning, as it is almost inevitable that the solution will need to handle these data types.

A need for shared information 2: Institutions are encouraged to share with one another examples of the standards they are using for classifying their data.

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

Setting Expectations

It is unrealistic to expect an institution to provide or support every platform that its researchers might request. A careful balance should be struck between supporting diversity of research and maintaining a sustainable service. Institutions should instead aim to understand what most use cases require, and to provide or approve a minimum set of platforms and systems that meet these needs. Research support capability[12] should be available to assist researchers who have needs that go beyond the standard offering. Ideally, the platforms and systems required to support edge-case needs can be put through a risk assessment to ensure that they are appropriate for the requested use, although the ability of an institution to assess these risks depends on resourcing and the volume of such requests.

Resources

- Monash University case study: Developing and implementing Monash Secure eResearch Platform (SeRP). 2021. Available from: https://ardc.edu.au/wp-content/uploads/2022/03/02z-Active-Data-Management-attachment-Case-Study-Monash-SeRP.pdf
- UNSW guide: Storing your research data. August 2020. Available from: https://research.unsw.edu.au/data-storage-and-tools

The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Summary

Culture change is required for researchers to adopt institutional changes and best practice in RDM, which should protect them and the institution from acknowledged risks such as litigation, fines and reputational damage. We recommend institutions take stock of the current state of RDM and develop a strategic roadmap towards a desired future state. We outline important aspects of successful change, regardless of desired future state, and which therefore can be used in conjunction with the recommendations and approaches for other elements in this framework. We offer detailed advice on how to track and monitor the culture change and any risks associated with this, and on the resourcing and communication required for successful cultural change, including staff training and incentives for behaviour change. Recommendations for institutions are highlighted throughout the section as are the needs for greater sharing of information across institutions on approaches to culture change.

Why Culture Change Is Important

> Behavior change is hard. Whatever its faults, the status quo is familiar and the words are known. The status quo is also easy to maintain. Just do nothing, inertia takes care of everything. We even have a tendency to defend the status quo. We'd rather believe that the way it is, is the way it should be.
>
> **Brian Nosek**
> *Center for Open Science, 2019* [13]

Many universities find themselves in a position where the RDM status quo is less than optimal. Current practice might not comply with legal or ethical requirements, thus exposing institutions to legal, financial or reputational risk. It may not conform with the expectations of publishers or funders, limiting the ability of researchers to compete in those arenas and endangering the institution's standing. Or researchers may rely on RDM approaches that are inefficient or error-prone, reducing the power, integrity, scalability and impact of research. Modern, good-practice approaches to RDM are a necessary (if not sufficient) condition for fostering trust in research, whether by addressing the reproducibility crisis evident in some disciplines or by simply improving the transparency of research, opening it to scrutiny and verification.

Good RDM practices are a priority for some researchers, but many will find RDM priorities a challenge when placed among the many other professional and institutional obligations and demands. Without intervention, the status quo at an institution is unlikely to change. Since people are part of social and cultural systems, it is useful to consider this widespread change in practice or behaviour as 'culture change'.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Cultural systems set norms, provide incentives, and impose requirements, giving organisations tools for changing behaviour.[13] Culture change and supporting areas such as risk assessment and change management are mature domains, with literature and examples that universities can use to think about and effect change.

Academic institutions are experiencing a difficult financial climate resulting in reduced personnel and roles being merged, so, without additional staff, capacity is limited for dedicated culture change work to be done. Important and necessary changes need appropriate funding and resourcing to ensure that they can achieve intended outcomes. In terms of culture change around RDM, institutions are exposed to real and potentially unnecessary risks which include litigation, fines and reputation damage. Also a risk is the more subtle reputational damage of obsolete practices and infrastructure, which may undermine an external partner's trust in the institution.

Here, we introduce approaches to culture change at universities in the context of RDM — approaches that universities may be able to adapt to their local context, recognising that research cultures, maturity levels, resourcing, degree of centralisation, and other institutional characteristics vary greatly among Australian universities. Subjects covered include broad approaches to culture change, narrower supporting approaches such as risk assessment and change management, and more specific advice on practical topics such as research engagement, community building, and communication planning. We provide a brief review of key literature and examples, and discuss individual ability and motivation, and alignment of institutional culture towards the desired change.

A note on language: Terminology about change includes culture change, organisational change, behavioural change and change management. Often, the first 3 focus on strategic elements of change, while change management is tactical. Here, we mainly use the term 'culture change'.

Culture Change Topics Identified By The Program

The Institutional Underpinnings program identified several topics of discussion during the early participant workshops, and from feedback and review by participating institutions:

1. Required culture change for researchers and research teams – aspects such as awareness; better project budgeting, planning and implementing of RDM; records and documentation; compliance checks; developing protocols; and leadership, especially in open research
2. Required culture change for universities – aspects such as governance and a more policy-driven approach; clearer institutional roles; advocacy; clearer communication, for example at induction, grant touch points, at promotion; compliance tracking with systems and procedures; providing central consultancy; planning and explicit commitment to foster open research; and changing 'business as usual'
3. Embedding culture changes through training and onboarding – offering training programs provided to all staff and students; communities of practice; and champions
4. Networks and communities – including local and national; research champions; institutions sharing best practice, exemplars and evaluation measures
5. Culture change for funding bodies – they are in a strong position to demand greater commitment to and compliance; clear direction on their requirements for RDM; advocacy for grants to cover RDM.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Our focus in developing this element has been on unpacking topics 1 and 2 — required culture change for researchers and research teams; and required culture change for universities. Topics 3 and 4 are better incorporated as practical advice on how to achieve culture change. Topic 5 was not a focus because of uncertainty about the level of impact that might be achieved in the program timeframe, and because funder concerns can be complex.

Key Aspects Of Culture Change

Culture Change And Change Management

When embarking on any new RDM initiative, large or small, careful consideration for the level of culture change required is essential. It is important to understand that facilitating change is often a complex matter requiring the development and execution of a carefully planned culture change strategy. That strategy needs to consider a range of key factors including the current state, the desired target state, key stakeholders, factors likely to impede culture change, and the application of an appropriate change management methodology.

Change management can at times be confused and/or conflated with culture change. However, the distinction is important. Change management can assist with driving culture change, but is generally part of the broader culture change strategy.

Generating Leadership Buy-In

Since successful culture change in large organisations will often, if not always, be driven from the top, generating 'upward' buy-in is necessary. The Prosci change management approach, in particular, calls out the importance of executive-level sponsorship[14]. Buy-in from institutional leadership helps ensure that the required change is understood to be important, and that the nature of the change is clear across the entire institution. In the context of RDM, the Deputy Vice-Chancellor – Research (DVCR) (and, to some degree, the Vice-Chancellor) often drive change, and need to actively and visibly advocate for the desired changes.

As such, leaders of RDM change initiatives should engage early with the DVCR to ensure that key institution leaders are aware of the initiative, support it, and understand the risks of failure. Institutional RDM leaders should also be certain that the DVCR clearly understands the importance of supporting, communicating and promoting the changes that are needed across the institution.

Varying levels of engagement from institutional leadership, as well as differences in effective strategies for engaging leaders, have been reported. Language around risk, competitiveness, and the 'future-proofing' of research is likely to be important, but different institutional contexts and leadership personalities mean the combination of these or other factors that work best will vary from one institution to another. See the sub-section on Risk management and quality assurance for a discussion on risk. Since this area is both important and difficult, we have flagged it as a particular area of need for information-sharing across institutions.

The ARDC is enabled by NCRIS

Defining Success

To be able to say that a culture change strategy has been successful, the desired future state needs to be articulated by the institution and measures of success clearly defined. This means that processes to track changes in culture (for example, behaviours, attitudes and outcomes) over time need to be put into place at the beginning of the change process to collect meaningful data, and measurable final-state goals need to be defined. Tracking changes by collecting data along the way, then analysing and interpreting trends in the data, will help leaders determine if the changes are leading the institution in the desired direction. If not, course corrections need to be made to address and potentially redirect efforts to complement new learnings and behaviour. See our advice below on Understanding Current and Target States of Practice, Behaviour and Culture and Developing a Change Strategy.

Different Cultural Approaches For Different Demographics

A successful cultural change plan needs to take into account the varying demographics and, thus, motivations of cohorts.

One large factor that will affect behaviour is the researcher's career stage. Institute directors will have varying motivations, quite different to HDR candidates, for example. Institute directors will focus on areas such as managing the risk of data breaches, recording audit information, and predicting data infrastructure needs. On the other hand, HDR students will focus more on learning what research resources are available to them to meet the needs of their projects, and on future-proofing their careers, and might appreciate a more structured or guided data management plan, compared to a senior researcher. Also, open discussions between HDR supervisors and their students may at times surface shared concerns, allow for targeted advice, potentially facilitate peer group support, and can be an opportunity to raise RDM practice with senior researchers in a familiar context.

Some disciplines will have differing needs. Research by Akers and Doty[15] found significant variability in RDM practices, attitudes, needs, and interest in support services within and between disciplines. Health and medical researchers, for example, will have many questions about handling sensitive data and will appreciate guidance on secure storage and sharing. Creative art researchers will be more focused on openly publishing their work, but with proper acknowledgement of copyright and intellectual property.

Institutions may encounter challenges with reduced staff capacity or resourcing constraints. For further discussion, see our advice on Resourcing and support and Leveraging existing skills and staff. In such environments change efforts may need to strongly align with institutional priorities related to research. A phased approach might be more achievable, with gradual and achievable targets, rather than a plan to move many facets of RDM to a state of best practice. Another strategy may be to work with specific discipline cohorts in areas of strategic priority to progress change. Alternatively, governance approved documents such as university policies, statements or risk assessments can represent official positions and may be the most important foundations for change, rather than eliciting priorities from engagements with researchers or assessing researcher practices.

The Culture of Institutional Support and Professional Staff Networks

To effectively implement culture change it is vital to have support staff and professional staff engage with the process. Professional staff can provide the support, guidance and training necessary to engage and enable researchers in their RDM.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Relevant professional staff should be involved or enlisted in collaboration and capacity-building activities in an early phase of culture change. Consider that their involvement in collaboration and capacity building may need to be enabled with skills development in RDM or culture change. Investment in ongoing education and skills development for professional staff can be essential to grow and uplift their knowledge and expertise, to provide support and to drive change.

Engaging professional staff can have long-lasting effects and can deepen institutional change and adoption beyond the timelines of implementation initiatives or projects.

A need for shared information: Institutions are encouraged to share with one another case studies and examples of culture change. When documenting a case study, consider the following questions:

- What models are you using to inform RDM culture change at your institution?
- How did your institution assess the current state and target state? Did you have experiences with Research Infrastructure Self-Evaluation (RISE), generic analyses such as Strengths, Weaknesses, Opportunities, and Threats (SWOT), or other approaches? How effective were they?
- How did your institution approach the assessment, articulation, and management of risk related to institutional RDM? Who did you consult when assessing and developing strategies for managing risk? What risks and potential harms related to RDM did you identify?
- How did you generate buy-in from your institution's leaders? How did this buy-in help your culture change initiative? Alternatively, did you need to work with an apparent lack of buy-in from leaders and what was your approach?
- How did you manage the resourcing of your culture change initiative?
- What rewards and incentives have you implemented?
- Did you find it helpful to have working definitions of particular terms at your institution? Terms and concepts related to change at institutions can be confusing to newcomers. What terms and concepts should be better defined in this document?
- Does your institution benchmark against (inter)national good practice? If so, how?
- How does your institution monitor the progress of cultural change? Does it differ from your initial stocktake or is it an extension of those activities?
- How has your institution solicited feedback? What combination of surveys, interviews, and other approaches worked for you? What other information or metrics have you used to assess the progress of change?
- How has your institution acted on feedback during a change process? How did you communicate those actions?

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Recommendations And Advice

It is important to articulate:

- the current state of practice, behaviour and culture
- the target state
- a change strategy mapping a pathway from current state to target state.

Defining the target state, and prioritising aspects of it in the pathway or change plan, will also benefit from an articulation of risks and benefits, including the institution's sensitivity to risk.

Understanding Current And Target States Of Practice, Behaviour And Culture

Undertaking organisational change benefits from a clear roadmap. It allows an institution to plan the path of how to move from the former state to the target state. Without clear articulation of either the start or end point, it will be more difficult to plan change or to convince stakeholders that change is needed.

The Current State

It is important that an institution takes stock of the current state to understand what challenges they will face when trying to change the culture around RDM. This stocktake should aim to understand the current cultural practices around RDM, as well as the infrastructure, processes and services that support best practices in RDM. An institution's appetite for change will be influenced by the fall-out of past events (for example, recent change processes which may have led to change fatigue), current pressures that might reduce or limit workload capacity associated with change (for example, changes to staffing, or workload models), and, inevitably, the funds available.

Several approaches and tools can be used during an institutional stocktake to assess the current state. Approaches aimed at institutional RDM include the Digital Curation Centre's Research Infrastructure Self-Evaluation (RISE) framework[16] (others are cited and compared on the RISE website). RISE offers a way for organisations to evaluate 'readiness levels' for various aspects of institutional RDM. General-purpose approaches to assessing the current state, such as SWOT analysis, are also commonly used in institutional settings. Institutions may need to use some combination of RDM-specific and general-purpose approaches to characterise various aspects of their current state.

The information needed to define the current state can be gathered through surveys, interviews or focus groups, sometimes used in combination. Feedback indicates that interviews might be more useful than surveys, as problems or strengths can emerge more fully through a long-form interactive discussion. When interviewing or surveying staff, it can be useful to include a range of stakeholders, including institution leadership and research support staff, in addition to researchers. Some existing metrics, such as the number of datasets in an institutional data repository, might also be useful. Metrics can provide baseline data before any change process begins, may provide comparative data during the change process, and may highlight gaps or inconsistency in service provision, resources, infrastructure, documentation, communication or governance structures. Potential measures of success to monitor can be found below, in the sub-section on Continual Assessment.

The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

RDM-specific and generic approaches to assessing the current state will also likely point towards desired target states (for example, 'we should eliminate this threat', 'we should pursue this opportunity', or 'we should raise this maturity/readiness level'). A target state can also be informed by problems and opportunities identified from surveys, interviews, focus groups or metrics. Discussions with institution leaders are likely to be important for defining the target state and promoting buy-in to that vision of the future.

Note that, as with other types of technology-related business analysis, users or clients (researchers, in this case) tend to articulate solutions before clearly defining problems (for example, 'I need a larger network-attached storage device' rather than 'I have 10 terabytes of data and no place to store it'). Experience suggests directing survey respondents, focus group participants, and interviewees towards problem definition before discussing solutions.

With the current state documented, the gap between the desired target state and the current state can be better articulated, and a clear path can be mapped between them. Leaders must take the opportunity to communicate the vision, strategic goals, and motivation (for example, risks and benefits) to ensure buy-in from the beginning, but with a sound understanding of the current state.

Recommendation 1: Understand the current institutional RDM state.

The Target State

Once the current state is understood, a culture change strategy should take into account the current culture and the methods to bring positive change into the culture space. Aiming for an initial target state as part of a staged approach has merit, with your 'low-hanging fruit' first in line. A state that can be achieved by leveraging existing infrastructure can provide early successes that contribute to sustaining momentum towards the objective of cultural change. Additionally, a staged approach can assist with limited resourcing. An important consideration is an evidenced acknowledgement of the resourcing required for an approach to be successful. In the current climate, where workloads are already heavy and resources tight, there is a real risk of increased workloads with new tasks relating to RDM for researchers and professional staff. Current workloads need to be adequately addressed or strategies will risk partial implementation and loss of momentum. It is important to highlight implementation as distinct from adoption, with adoption being the end goal. Proper communication planning and cultural adaptation need to happen to get the push from 'implementation' to 'adoption'.

For researchers, the positives or incentives of moving towards the objective of change need to be clear and well communicated, as does the need for risk management and compliance. Policy agendas and/or the rationale for change objectives should be communicated. With threats of data breaches, auditing needs and future technical resource budgets all weighing on universities, there is merit in top-down support driven by leaders who understand the effects of these external pressures. So, finding benefits and reward mechanisms for researchers, and highlighting the value of RDM planning and aligning to the FAIR data and CARE principles[4,5], are essential. There can be value in engaging with compliance; for example, asking researchers to estimate how much storage they will need is also useful for estimating technical budgets for future years.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

One of the most valuable tools for identifying a desired target state is likely to be benchmarking against other institutions, within Australia or overseas, and seeking other external sources for good practice in institutional RDM. Your institution is likely to find another institution with a solution to problems you are facing. Much can be learned by reviewing other institutions' research technology web pages, and meetings can often be arranged with research technology or research data staff. Attending events and reading outputs from national and international organisations such as the ARDC, Research Data Alliance (RDA), the Center for Open Science, the Committee on Data for Science and Technology (CODATA), the Digital Curation Centre (DCC) and others is essential. Familiarity with statements or manifestos about RDM can also inform development of target states, especially if they are endorsed by funders, publishers or other relevant organisations. Finally, when defining the target state, convincing stakeholders that the changes are necessary requires articulating current-state risks and target-state benefits. Experience at many Australian universities has shown that articulating risk, in particular, has been crucial to convincing stakeholders of the need to change and securing the resources necessary to drive change. As such, we discuss risk management and quality assurance below.

An Approach To Getting Started

One approach to identifying high-level aspects of current and future state is to analyse key stakeholders' values, aims and experiences with respect to different aspects of RDM. This could include early engagement with both researchers and institutional stakeholders (support staff or senior leaders). Engagements could take the form of discussions, workshops, interviews or surveys. Structuring the conversation around values, aims and experiences can reveal insights:

- Values – What is important to researchers (or the institution)? Culturally, these values are what we see value in spending time on, or value in expending effort to adopt better practice. (indicates a willingness to adopt change)
- Aims – What are researchers (or the institution) trying to achieve? (indicates ideal practices, workflows)
- Experiences – What have researchers (or the institution) tried so far? How did this affect us or how well did things work? (indicates current behaviour, gaps or evidence of practice).

Discussions and questions around values, aims and experiences can be focused on either researchers or the institution, but should also be focused on elements of RDM to provide more constructive analysis. For example: What are researchers' aims in data management planning? Why do researchers find it valuable spending time on data management planning? What have been their experiences? (Do they have a data management plan? Do they benefit from it?)

In general, insights on aims can indicate opportunities for future state, insights on experiences can indicate current state, and insights on values may indicate 'low-hanging fruit', incentives, or areas that require more significant culture change.

Developing A Change Strategy

Once an organisation articulates current and target states, a pathway can be mapped that guides how the organisation and its staff will move from the current state to the target state. Producing this roadmap is perhaps the core activity in developing a behaviour change or culture change strategy. Culture change is a mature field; this brief introduction seeks to provide an orientation to it.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Broadly speaking, culture change or behavioural change has 2 components — individual change and institutional change. Most culture change literature focuses on the individual's ability and willingness to adopt new behaviours or practices (see Appendix 1: Comments on Culture Change Models). Nosek[13], writing about culture change in a research context, however, takes for granted that researchers want to, and can, 'do the right thing', and instead focuses on institutional culture, including norms, incentives and requirements. At most universities, changing RDM practice is likely to require both researcher change and institutional change. A researcher must be convinced to adopt better practice and be trained to implement it. An institution needs to ensure that better practice is not only as easy as possible, but also normative, rewarding and, at least in some cases, mandated.

Change leaders will determine the philosophy behind the change approach. Will it focus on the institution, and assume that staff are ready and willing to change (for example, Nosek[13]), or will it assume that staff need to be persuaded and educated (most other change models, such as the Kübler-Ross Change Curve®[17])? Will it use evidence to persuade people to choose change (as in 'nudge theory'[18])? Will it focus on supporting staff on their emotional journey through the change process, as with the Satir change model[19] or the Kübler-Ross change curve? Or will it be driven from the top down, perhaps with required training and penalties (an approach criticised by Dobbin and Kalev[20])?

Many guides to common change management models can be found online; two concise places to start are Mulholland[18] and Lucid Content Team[21]. Also, Appendix 1: Comments on Culture Change Models provides an overview of several models.

Recommendation 2: Understand the desired future institutional RDM state, develop a strategy to achieve it, and regularly monitor progress towards it.

The Applied Advice sub-section below includes discussion on key activities and engagements that are often used to drive culture change relating to RDM in institutions. More details on key activities and engagements can be found in other framework elements, especially Element 4: Research Data Management Planning, Element 8: Support, Training and Guidance and Element 3: Policy. The change models introduced above deal with the more generic process of change, but strategies for change will vary according to local circumstances.

Be aware that assessing progress during the change process comes with the risk of "feedback fatigue", which may reduce engagement with the process. While determining relevant ongoing assessment, consider what metrics are readily available, assessments that require less direct contributions from researchers, or restrict direct assessment to a few specific points in the process.

Continual Assessment

To ensure success, ongoing monitoring of long-term changes in behaviour is needed to assess if adopted approaches are successful. Monitoring progress ensures that the interventions around infrastructure, policies, processes, outreach, training, support, and other areas are producing the desired results. Regularly soliciting feedback, acquiring relevant metrics or other indicators (for example, service uptake, enquiries, training attendance), and evaluating progress during a change process reveals the course of change, providing early indications about whether or not the organisation is progressing towards the desired target state.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

This information can indicate when change-related activities are not working effectively, so that mid-course corrections can be made. Taken early, such adjustments to the change process can be modest and incremental, limiting disruption and building confidence among stakeholders, who will appreciate that change activities are being monitored and that any adjustments are informed and deliberate (this is especially important considering the stress inherent to times of change). As well as identifying flaws in the change strategy and informing the adjustment of interventions, soliciting feedback and evaluating progress can reveal unforeseen barriers to change, again allowing informed and deliberate adjustments to the change process. Many of the change models presented in Appendix 1: Comments On Culture Change Models incorporate or assume a degree of feedback from stakeholders, perhaps most explicitly in the 'Check' phase of the Plan-Do-Check-Act (PDCA) model[22]. While monitoring progress, however, it is important to bear in mind that changes in productivity or outcomes will not likely rise along a smooth upward trajectory. Some change models (for example, Satir and Kübler-Ross) recognise that productivity may initially decline as time, energy and resources are directed away from usual activities towards the process of change itself (and its ancillary emotional and organisational effects).

Assessment measures may include:

- Training attendance
  - Attendance numbers, if training is non-mandatory; a measure of HDR students versus academics, and/or by discipline, can also highlight changing attitudes in particular cohorts; percentage of researchers who have completed online relevant RDM training
  - Teachers/trainers can indicate attitudes, number and nature of questions, and participation levels, for guidance (could be more relevant for mandatory workshops)
  - Relevant workshops include those on research data, sensitive data or storage.
- Storage requests
  - Number of storage requests and size of storage allocations
  - File activity and movement between storage and RDM systems; for example, inactive files on active research data management systems, or lack of movement of files into research data repositories
  - Requests may provide evidence of changes in data management practice; for example, storage requests for individuals versus requests for project or collaborative groups. Patterns of unusual requests may identify unmet needs; for example, cybersecurity concerns.
- Data management plans
  - Proportion of researchers and research groups that have submitted a plan
  - Updating of plans after initial submission
  - Comparison of successful grants against plan submissions
  - Number of non-mandatory plans submitted
  - Number of plans archived with retention data or records.
- IT processes in place
  - Viewer counts and clicked links to RDM-related webpages
  - Number of helpdesk requests recorded for RDM consultations.

- Publications
  - Number of datasets published
  - Number of successful grants, completed or nearing completion, which have associated metadata in a data repository
  - Number of researchers using ORCID identifiers associated with their research outputs
  - Proportion of publications in open research journals
  - Citation metrics on reuse for published datasets.
- Surveys
  - Longitudinal surveys and interviews to gauge attitudes and practices.
- Ethics and integrity
  - Decreases over time in retractions or misconduct related to RDM
  - Number of requests from ethics committees for clarification or better explanations relating to RDM
  - Comparison of successful ethics applications against data management plan submissions.

Soliciting, Evaluating, And Acting On Feedback

The process of soliciting feedback from researchers, research support staff and other stakeholders, and then evaluating and acting on that feedback, is a valuable part of the change process. Sincere investigation of the impact of change-related activities, and reactions to it, is crucial to generate stakeholder buy-in. If leaders fail to listen to or address the concerns of people affected by the culture changes, it may lead to resentment and reluctance to adopt new processes and behaviours (see, for example, Dobbin and Kalev[20]).

Acquiring meaningful feedback from researchers and other stakeholders requires an intentional feedback approach. Vehicles for feedback may resemble those used in your initial stocktake, including surveys, focus groups, and longer-form interviews (with similar caveats as to their relative effectiveness). The Carpentries use surveys to assess the short- and long-term impact of their Software Carpentry and Data Carpentry training workshops[23], for example, which may provide inspiration for the sorts of questions to ask researchers after activities are undertaken. It is likely that consolidation of feedback from multiple sources will be necessary to accurately monitor the change process. Considering the challenge represented by change processes, and that resistance to change is inevitable, feedback should be assessed carefully, with some expectation for negativity even when the process is going well. Comparing feedback from different sources — for example, from research support staff in ethics and integrity, research services, IT or the library — may offer perspectives to contextualise feedback from researchers (for example, integrity officers may see improvements in ethics applications and a reduction in risk, or grant officers may see improvements in funding applications and higher success rates, even as researchers complain about new requirements). It may be valuable to cross-check against other metrics related to compliance (for example, via audits of data practice or tracking the number of data-related breaches) or improved practice (for example, tracking the number of published datasets or articles published in journals with stringent requirements related to FAIR data or the Transparency and Openness Promotion (TOP) guidelines). As was the case when a target state was being defined, benchmarking against other institutions is also likely to be helpful (for example, where gaps are being closed). Combining information from multiple sources is most likely to increase understanding.

The ARDC is enabled by NCRIS

Feedback must also be acted upon — and be seen to be acted upon. Actions based on feedback may include modifying the change strategy or process itself, adding new infrastructure, modifying policy, or providing additional training and support. When such changes are made, they should be communicated to stakeholders as a demonstration that the feedback and evaluation process is being undertaken in good faith.

Finally, as part of assessing change, it is important to recognise when the target state has been reached. Change is fatiguing and, in the short term, reduces productivity. Concluding a change process by acknowledging — even celebrating — the fact that a goal has been met, a transition completed, and a new (improved) status quo reached is an important final step in the process.

Risk Management And Quality Assurance

At Australian universities it has been observed that, although priorities vary between universities, risk management and mitigation have often been major drivers of organisational change around RDM. Risk management is a mature field, with its own conceptual frameworks[24]. As such, a working knowledge of core risk concepts and terminology can help RDM proponents make a case to institutional leaders for change and explain the necessity for change to researchers and other stakeholders.

Generally speaking, risk management begins with an assessment process identifying hazards, and assessing likelihood, potential impact and severity. Once risks have been assessed and prioritised, they can be managed to reduce exposure to legal, financial or reputational damage. The goal of risk management is not necessarily to eliminate all risk, but rather to understand risks and potential impacts, minimise risks to the extent possible, decide what risk an organisation is willing to bear, and develop approaches to manage remaining risks. Finally, since not all risks can be known beforehand, building organisational robustness or “antifragility”[25] may also be part of a risk management approach.

Risk management intersects with RDM culture change in two ways. First, risks should be identified and evaluated as part of articulating the current state. And second, a culture that maturely assesses and manages risk, with purpose and self-awareness, should be part of the institutional RDM target state. and risk planning should inform the direction of that culture change. Articulating risks may help convince stakeholders and institution leaders of the need for change and resources to drive the change. Developing a visual display of the risks, such as an institutional risk matrix, may be useful in engaging with senior stakeholders and institutional leaders.

The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Determining risks, risk tolerance, and risk mitigation measures requires conversations with staff from the various organisational units that support RDM. Such units may include:

- Ethics and integrity, who can articulate the risk, risk management approaches, and risk tolerance related to research ethics and integrity compliance generally, including human and animal ethics compliance
- IT security, for risks related to cybersecurity
- The library, for risks related to copyright and reuse
- Research services, for risks related to funder and publisher compliance, as well as compliance with other governance, such as classified or sensitive information
- Office of General Counsel, for legal risks related to, for example, privacy acts, record acts, data breaches
- Risk management, for overall institution risk strategy, and sometimes specialised areas like international compliance (for example, EU General Data Protection Regulation)
- DVCR for general guidance on risk appetite in relation to research data.

To begin conversations with these stakeholders, it is useful to have some familiarity with the language of risk management, and an idea of the types of risks associated with RDM at universities. Common risks and associated consequences include:

- loss of high value or irreplaceable research data
- non-compliance with privacy laws, which can lead to large fines, reputational damage to researchers and to the institution in the event of a breach, and subsequent risks of failing to attract outstanding researchers and HDR candidates
- non-compliance with record retention law, which can result in fines, and reputational damage in the event of data loss
- non-compliance with human ethics or research integrity requirements, which can stop research, render researchers or institutions ineligible to participate in future research in the event of a breach, and lead to reputational damage
- inability to comply, or non-compliance, with funder RDM requirements, which may prevent researchers from applying to certain grant schemes, make their applications uncompetitive or, if a grant is won but requirements are not met, sanctions from funders regarding future eligibility, with downstream effects on research income, ranking and reputation
- inability to comply with publisher RDM requirements, which may make it impossible for researchers to publish in certain (often high impact) outlets, with downstream effects on research income, ranking and reputation
- misalignment of institutional policy and RDM practice, which can reflect a current state that encourages non-compliance, or an inability to comply, and brings with it the above associated risks
- an ineffective institutional RDM approach, particularly with respect to infrastructure, training and support, may also hamper recruitment and retention of research staff.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

To conclude, identified risks should be included in the current-state review of the change strategy. Good practice leading to reduced risk, and better management of residual risk, should be part of the target state. The assessed severity of risks can inform prioritisation of activities during the change process. Ideally, a purposeful approach to risk itself can also be incorporated into the target state.

Applied Advice

Resourcing And Support

Resourcing must be acknowledged for culture change to be successful. In the current climate, workloads are already heavy and resources tight, so any changes, increased workloads, or new tasks relating to RDM will be inadequate alone. The real risk of inadequate resourcing is partial implementation, loss of momentum and, consequently, inadequate institutional RDM.

Minimum resourcing recommendations are difficult to estimate, but dedicated time is required to develop the strategy, governance and coordination, and to meaningfully engage all participants in the change. A staged approach may assist where resources are limited. Opportunities to dedicate existing advocates to change activities may arise.

Resourcing Advice

- Employ people with experience in and knowledge of managing culture change.
- Work to secure or develop desirable skills and resources that should be available to your culture change initiative, such as experience in: designing policy, resources and tools; seeking out appropriate platforms to endorse, promote and support; designing training; and delivering digital literacy training.

Leveraging Existing Skills And Staff

Skilled research support staff can be found in a variety of roles across most institutions, and some consider RDM a part of their expertise or role. This can range from providing RDM services to advising on specific domain platforms or situations such as ethics or contracts.

At a minimum, staff providing all levels of support (online advice, help and enquiries, expert consultation) might also recommend that researchers use particular services or engage in a particular initiative. In culture change initiatives, informing research support staff can increase general awareness and broaden uptake of RDM benefits.

Broader support staff can be leveraged to support culture change initiatives, either as formally seconded resources, or by informally managing their involvement. This comes at a cost, potentially in salary, management overhead, or the effort required to gain support from their line management. But in circumstances where centrally managed culture change is insufficiently resourced to engage many researchers across many disciplines, enlisting broader support staff can make a big difference.

The ARDC is enabled by NCRIS

Support staff can also provide deeper access to their existing client bases, which can further embed the required culture change across organisational units, beyond structural changes, and/or beyond the end of funded initiatives/projects. Models for this include:

- Building an informal culture change working team across organisational units, usually led by a dedicated role or initiative.
- Leveraging experts on a per project or activity basis.
- Involving staff in running communities of practice.
- Doing a skills audit to identify strengths and areas for development. This can be a useful way to involve a broad cross-section of staff, can lead to skills development in RDM or culture change, and can identify and develop existing change resources and/or champions.
- Working towards a one-stop RDM FAQ or virtual help service. This can better leverage any past or future collaborative change efforts by better communicating services, resources, infrastructure and/or governance. It may also lead to staff sharing resources across organisational units to maintain change efforts.

Engaging Stakeholders

Stakeholder engagement, broadly speaking, is the process by which an organisation involves people who may be affected by the decisions it makes. Methods range on a continuum from communication through to collaboration and participatory co-design. Ensuring inclusivity and diversity in engagement enlists managers/leaders in solving the problem, exposes them to different groups, and encourages visibility and discussion.[20] One way to conceptualise building cultural change in RDM, applicable to any element of the framework, would be to consider lenses of engagement:

- 'Top-down' – Leadership, strategy, policy, funding, rules and frameworks play a role in setting the incentives that shape research culture.
- 'Bottom-up' – Researchers have the potential to make behavioural and attitudinal change at the level of research groups, centres and institutions, which may form new social norms.
- 'Horizontal' – Professionals from across organisational units who support researchers and data infrastructure may be able to help with communication, communities, training and making change easier.

When people are forced to change through rules, compulsory training, and penalties, there is usually resistance and the opposite behaviour from that which is intended occurs. This has been demonstrated in the failures of corporations to change their behaviour towards inclusivity and diversity in recruitment and promotion.[20] Allowing people to feel that the choice is theirs leads to better outcomes.

The ARDC is enabled by NCRIS

Advice To Ensure Meaningful Engagement

- Communicate any policy agenda carefully; for example, the research culture agenda has evolved, the importance of engaging researchers and disciplines, the collective efforts to engender community.
- Gather and reflect insights from conferences, discussion meetings, and previous communities of practice.
- Survey the user-experience of key stakeholders to understand gaps, needs and what the community has in common.
- Invest in building relationships within the relevant areas of professional services and between professional services and the academic community impacted by the changes.
- Encourage researchers to follow and join broader communities of practice, such as those relevant communities facilitated by the ARDC.
- Create an RDM working group and recruit champions in academic areas.
- Build RDM into the existing research professional development framework, if one exists, with links to research training competencies.
- Consider ways for service centres to have open discussion with researchers (for example, build it into existing events and activities) and open communication with library, IT, learning and teaching, and research services. This can encourage the breaking down of silo boundaries.
- Avoid highlighting or presenting technology as the solution to all issues or concerns. Human intervention is key to understanding, articulating, supporting and designing for researchers' needs.

The next 3 sections explore interconnected properties of each aspect of engagement.

Engaging Researchers And Disciplines

In the context of RDM, researchers are the key stakeholders. Engaging with researchers is necessary to help achieve buy-in and is a prerequisite for success of all the aspects of culture change. Engaging with researchers about the risks and benefits of RDM to them, and to the institution, enables them to understand the rationale and justification for change. This leads to them better appreciating what is required and facilitates researchers partnering to implement and adopt the change. Implementation relies on adoption which relies on buy-in.

Engaging with researchers is also a 2-way process, providing the institution with the ability to acquire a researcher-centric view of service and solution design. Custom RDM solutions that are aligned with researcher perspectives and provide immediate value to researchers can be designed and implemented, resulting in less resistance and more uptake.

Understanding how different academic disciplines use and engage with RDM will inform and influence how new RDM initiatives are communicated. For example, creative arts researchers are likely to have different needs to genomics researchers, so understanding discipline-specific RDM needs can help ensure that those needs are met.

The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Engaging with researchers also assists with the collation of a strong evidence base that can be used to help convince institutional leaders, academic leaders, and influencers across the institution to endorse, support and champion RDM culture change.

Practical Deeper Advice[26]

- Inform the community. An ingredient for successful engagement is to carefully articulate and communicate the rationale for the objective.
- Foster "champions". These should be researchers who promote best practice and who will ideally need workload allocation and recognition for this role. Authors of *The four building blocks of change* recommend role modelling as one of the 4 key actions that influence employees' mindsets and behaviour.[27] Champions and case studies could serve as effective role models for best-practice RDM behaviour.
- Target the discipline-specific resources, examples and case studies. Of the 5 levels of intervention in the Center for Open Science's strategy for culture change, 2 of them are to "make it easy" and "make it normative"[13]. Discipline-specific resources can help to make best practice easy for researchers, and champions can help to make it normative.
- Consult the community. Ensure broad engagement with researchers and/or research support staff from different academic divisions or representing different RDM workflows. This approach will better identify practices, risks, pain points, and appropriate channels for communication and downstream engagement.
- Involve and collaborate with academics in the planning and implementation process. The degree of academic involvement can vary, ranging from an advisory capacity to co-development of solutions and culture change initiatives. The spectrum of participation proposed by the International Association of Public Participation[28] (IAP2) offers an interesting insight.
- Adopt sensible lead-time expectations. Any requests for a researcher's time need to be framed with an explicit awareness that they need sufficient lead time. Even researchers leading RDM practice have commitments and higher priorities that may preclude short-timeframe expectations.
- Correlate benefits gained from good RDM with increased research impact, where possible. Wade Kelly provides an emerging, holistic view of what could constitute research impact.[29]

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Building Communities

Communities of practice, arguably, are critical to cultural change. The recent shift to remote working has diminished the physical signifiers that more readily supported the scaffolding of communities.

Several factors make it hard to build a community:

- The turnover of doctoral researchers and research staff alike presents challenges for establishing a lasting community culture.
- Variety in the size and structure of institutions means that solutions need to be relevant to the context.
- Building community is both everybody's and nobody's responsibility. Everyone has some sense of duty to contribute to a safe and inclusive community culture, but without delivery being in a named individual's portfolio, efforts risk being fragmented, duplicated or missed entirely.

Some institutions may need a different approach if they are without an identified lead position or a named responsible individual. They might adopt a collegial or informal approach between different units, and change might focus on collaboration or areas where interests overlap.

Recommendation 3: Clearly define ownership for building the RDM community at the institution.

With increasing pressure on professionals supporting researchers to frame their work in terms of longer-term and structural actions that promote a safe research culture, the community-building activities that were once a nice-to-have are now fundamental. This is not to say that individual community-focused interventions will, on their own, transform research culture, but the sum of different kinds of support can have a transformative effect on how researchers experience that culture.

Planning Communication

A well designed and executed communication plan is critical to the success of any organisational change program or initiative. Communication needs to be designed to flow downward, upward and sideways within the organisation. The communication plan is a centre piece to all widely used change frameworks and exists to ensure that all participants in a given activity understand the reasons for the proposed change, and what is expected of them in terms of learning and complying with new procedures or systems. Participants in change also include support professionals, so articulating what messages or changes are expected of them is also important.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENT 2: CULTURE CHANGE

Recommendation 4: Take a coherent approach and be consistent with messaging.

- Carefully communicate policy agendas and/or the rationale for the objectives.
- Develop terms of reference for a working group to be created and seek endorsement from the top.
- Agree on a communications plan to execute and the key people needed to execute it.
- Leverage multiple communication methods and technologies including both "push" methods (bulk communications, such as email, staff news) and "pull" methods (those requiring opt-in recipient discretion, such as web resources, blogs, subscription lists). A variety of text and multimedia modes will also help ensure that messages are effectively communicated.
- Share and articulate the objectives, strategy and plan for raising awareness.
- Encourage research leaders and mentors to lead RDM practices and related discussions among the research team.
- Encourage research leaders and mentors to lead both education and compliance within their discipline and/or schools.
- Provide a marketing toolkit to aid in any communication. This could include reusable collateral such as slides with embedded consistent messages. Messages should succinctly outline the most important calls to action, the reasons, and potentially the risks or benefits that resonate.
- Build change communication into inductions and orientation of new staff, both researchers and research professionals.
- Incorporate meaningful "carrots" into the communication plan. For example, at Edith Cowan University and Bond University, upon completion of a data management plan, data storage in an instance of SharePoint is autogenerated for researchers.
- Leverage existing communication events and outreach opportunities, such as researcher forums and department research meetings. Change fatigue and workload constraints can mean researchers choose what they will get the most value from, so, ideally, bring the communication to them.
- Consider post-initiative/post-project engagement strategies in communication plans, which may involve broader staff to sustain and continue the culture change. Consider communicating and setting these expectations with staff, especially articulating what messaging or communication changes are expected of them in the long term.

(Much of this section references presentations from Vitae Connections Week 2021[30].)

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Providing Appropriate Training

As indicated, providing the 'what/why/how' for RDM is essential for change. The responsibility of the institution to provide guidance in the form of training is widely understood and is accepted as a foundational activity to provide those important change messages. Having executive support to instigate mandatory training can further set the cultural attitude of the institution around the importance of RDM.

General high-level advice for planning training:

- Ensure training resources and materials are tightly linked with consistent messages from any communication plans.
- Consider career-stage-specific training or professional development needs (HDR students vs postdocs), which could include offering training as part of induction or as onboarding recommendations.
- Make training available so that it is timely to research or the HDR project life cycle. Early intervention is better than retrofitting, but retrofitting is better than no change.
- Ensure that pathways to next steps post-training are clear; for example, further training and guidance.
- Ensure that training reinforces where to find up-to-date support documentation, services and personnel.

Making Change Easy And Possible

If change is not easy, only highly motivated people will make the effort to learn new systems, new processes and enact these new behaviours to become habits. Remove barriers to the new processes and behaviours, and people should only need convincing and upskilling.

General high-level advice for making change easy:

- When planning for initiatives, improvement works and projects, prioritise the ease of use of systems, integration and processes that will, ideally, step researchers through decisions and considerations. Ease of use can be prioritised in activity principles, goals, user-experience resourcing, and specific activities such as user-engaged co-design and testing.
- When planning initiatives, improvement works and projects, prioritise and test for systems and processes that work together to funnel behaviour in the desired direction.
- Make options clear and easy to decide — provide frameworks for decision-making.
- Try to have options for most use cases; for example, storage types, sizes, tools for data capture, sensitive data. Without sufficient infrastructure and support options, change may not be possible.

The ARDC is enabled by NCRIS

- Draw on data from other enterprise systems or consider not collecting information that can be obtained from other enterprise systems. Much less information may be required to refer to specific information in another system (for example codes or identifiers for grants, ethics, data management plans), but test that it is reasonable for researchers to have this minimum information (for example, specific identifiers). Reducing the need for researchers to provide the same information to multiple systems can facilitate processes, reduce errors, and reduce the amount of information that needs to be kept up to date (for example, in allocating storage space or registering datasets).
- Use journey mapping[31] to ensure that the process researchers are being asked to follow is as simple as possible, that it makes sense, and that appropriate scaffolding is built in at the right places to help researchers with different levels of experience through the required steps.

Rewards And Incentives

As of 2021, most academic institutions in Australia do not reward research data practices, including open-research and FAIR data practices. Academic career advancement is dependent on the number of high impact publications and the number of citations. This "publish or perish" culture has resulted in selective reporting of research outputs towards positive trends and "tidy" stories to enable publication. For researcher behaviour to change towards open research and FAIR data, the culture and the system supporting researchers needs to change to enable and reward FAIR data and open-research behaviours. Rewards are one way to incentivise change. By acknowledging changes in researcher behaviours and processes, and rewarding those that align with the goals of the desired future state, people can be motivated to keep going in the desired direction. Providing incentives and rewards for desired behaviours creates a positive feedback loop.

The following suggestions for meaningful rewards and incentives can be used to recognise the desired changes in behaviours relating to RDM:

- **Workload models** could appropriately acknowledge the necessary training time for researchers to upskill (for example, to become digitally literate, to learn how to use appropriate platforms, to learn how to publish data with discipline-specific metadata, to transfer/deposit data from active storage to archive/repository, or to appropriately document and annotate data) and the additional time it will take to make data FAIR and publish open research.
- **Promotion criteria** could acknowledge some best practices in RDM in academic promotions processes (for example, data publications; service roles, such as a data champion; publications in journals that require FAIR data and open research to be considered or ranked highly for transparency and reproducibility; funding secured from organisations that require FAIR data and open research to be considered or ranked highly).
- **Annual professional development reviews** could include a section where staff can plan their incorporation of RDM best practices, and review their progress and training needs. A dedicated RDM section would actively encourage researchers (by showing it as an expectation of managers) to develop skills and practice in this area over the following year (for example, training in RDM, upskilling in digital literacy, publishing in a journal that supports open research, or publishing a dataset in a discipline-specific research data repository conforming to discipline-specific metadata standards).

The ARDC is enabled by NCRIS

- **Position descriptions / selection criteria** could acknowledge skills and experiences in best practices in RDM at all levels.
- **Institutional research grants and funds** could require best practices in RDM or, at a minimum, a data management plan. Or even further, research data could be recognised as an institutional grant category (either specifically or as part of a grant) to encourage some applicants to include a focus on establishing significant research data or databases or data publishing. A data focus can be valuable for grants involving infrastructure or for Australian Research Council (ARC) Linkage Infrastructure, Equipment and Facilities scheme near misses. As a further incentive, the institution might co-contribute to successful applicants an RDM specialist and infrastructure support for the project.
- **Career pathways** could be specified for those who display best practices in RDM (for example, for someone who is a data champion in their department and has mentored/advised/trained other researchers in best practices, pathways to data steward/architect roles, and more academic type positions could be included).
- **Annual awards** that acknowledge excellence in research data and best practices in RDM could be introduced (for example, Vice-Chancellor awards, such as in learning and teaching, could be expanded to include research data and RDM).
- **Institutional researcher profiles** could incorporate information about research datasets and published data outputs (for example, in Pure, Symplectic Elements). Not only would the datasets be more discoverable, but the publishing of datasets at the institution would be more obvious (datasets could be the result of FAIR data initiatives, open-research funding, journals conforming to the Transparency and Openness Promotion guidelines[32]).

Resources

Please see Appendix 1: Comments on Culture Change Models for a list of models related to culture change, with key content notes.

The ARDC is enabled by NCRIS

ELEMENT 3: POLICY

Summary

This element provides institutions with guidance to ensure that their research data is managed according to legal, statutory, ethical, publisher and funding body requirements, such as those outlined in the *Australian code for the responsible conduct of research*[2] (the Code). Here, we offer guidance on the steps for developing an institutional policy for RDM, including, for example, establishing the required roles and responsibilities and cross-institutional consultation and communication. We recommend content for an effective institutional policy and include considerations for supportive guidelines to underlie policy. Finally, we outline recommended approaches for successfully implementing an institutional RDM policy.

Why A Research Data Policy Is Important

How institutions and researchers manage their data is increasingly important as universities look to increase the integrity, quality, impact and openness of research. Good research data also underpins public trust and confidence in research and is an integral part of responsible stewardship of research funds.

A research data policy is vital for institutions to ensure that research data is managed according to legal, statutory, ethical, publisher and funding body requirements. Institutional responsibilities include the provision of access to facilities for safely and securely storing and managing research data, records and primary materials and, where possible and appropriate, to allow access and reference. The policy also provides safeguards against misconduct, such as data fabrication and data falsification.

For researchers, the policy is a framework for ensuring data created from research is accurate, authentic and verifiable; it also ensures they are able to meet their responsibilities in relation to obligations under the Code and to manage their data in a holistic way.

A good research data policy should support the institution's research strategy. This may relate, for example, to maximising the impact of research by publishing or sharing data; supporting open research and the FAIR and CARE principles[4,5]; or creating a vibrant and collaborative research culture.

The policy, relevant infrastructure and tools will differ between institutions, and should be tailored to support both the nature of the research and the level of RDM maturity achieved at the university. The RDM policy for the institution should be complemented by a suite of related policies and procedures about managing data, contributing to an effective research environment, and supporting collaborative, applied and translational research.

A Policy Is Not Always Named As A Policy

In some institutions, policy about RDM is embedded within another policy or policies, rather than being a standalone policy. In other institutions, this information is set out as a guideline or principles, rather than a policy. The form of the policy depends on processes and choices made at the university. Such a document still meets institutional requirements under the Code and the Australian Government Tertiary Education Quality and Standards Agency (TEQSA) guidance note[33], which specify that an institution should have a suite of policies and procedures.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Research Data Versus Enterprise Data

Institutions are responsible for managing and governing both research data and enterprise data. Some principles apply to both domains; for example, compliance with legislation and cybersecurity. However, there are significant issues that are relevant only to research data; for example, data sharing and intellectual property ownership. The nature of these differences is such that a specific research data policy is important to provide institution support for effective RDM. There may need to be alignment or references to the policies for all data at the institution.[34]

Drivers For Policy

Drivers for an RDM policy may vary between institutions, and individual needs should be considered; however, the overarching drivers are likely to include:

- complying with legislation and regulatory obligations
- meeting funding requirements
- meeting project-specific requirements
- enabling and facilitating collaboration
- supporting the effective operation of the institution in managing risk
- supporting the research strategy, including through optimising engagement and impact, including through increased citations
- increasing research impact through faster and easier access to research data
- increasing research efficiency by sharing results
- supporting research integrity through transparency and reproducibility
- building researcher capacity
- safeguarding against misconduct such as data fabrication and data falsification
- clarifying ownership, copyright and licensing provisions
- ensuring secure and responsible handling of sensitive data
- ensuring safe and secure storage, including protection of data and privacy
- supporting ethical conduct and appropriate use, including of culturally sensitive and Indigenous data, including the CARE principles[5]
- managing the cost of storage and services and ensuring the sustainability of these services
- identifying, nurturing and supporting best practice
- maximising the benefits of open research to the researcher and the institution, including through adoption of the FAIR data principles[4].

A policy framework is an efficient means to make compliance expectations from various sources explicit and clear in one place. It should include related policies, procedures and guidance on available services, infrastructure, storage and training. Best practice beyond meeting compliance requirements may also be considered.

Meeting Regulatory, Legal, Funder And Publisher Requirements

Institutions and their researchers have regulatory, legislative and contractual obligations that they are required to meet. These are important matters for RDM and indicate the complexity of the landscape for policy and practice.

All research institutions are required to have some form of RDM policy, although the content may differ. The RDM policy should assist researchers and the institution to comply with their obligations and requirements. The RDM policy should sit within a framework of policy and other documents, including intellectual property, employment agreements, and documents implementing legislative requirements. The nature of an institution's decision on an RDM policy requires consideration of these related documents in order to establish a policy that best meets their individual needs.

Regulatory Requirements

All institutions and their researchers are expected to comply with the Code, which "outlines the expectations for the conduct of research in Australia or research conducted under the auspices of Australian institutions".

Australian institutions are also "strongly encouraged to follow the advice in the Guides" that support the Code. The Code is clear that institutions are expected to have a suite of policies and procedures related to RDM; however, the current version is more concise than previous versions, so, generally, more nuanced guidance is required at the institutional level.

The guide to *Management of data and information in research*[49], released with the Code, details how to comply with the Code. It states:

"Institutional policies should include guidance for managing research data and primary materials that addresses the following:

- *ownership, stewardship and control*
- *storage, retention and disposal*
- *safety, security and confidentiality*
- *access by interested parties."*

As for the scope of the policy, the guide also states:

"Policies should apply to all research conducted under the auspices of the institution and may be influenced by the funding arrangements for the project."

It is important for each institution to provide clear definitions on what constitutes data; whether it is produced, used or held by the researcher; and who the policy applies to, including consideration of research students, adjuncts, honorary staff, emeritus staff and visitors.

The ARDC is enabled by NCRIS

TEQSA has requirements for appropriate corporate governance that are particularly significant for research data created by HDR students. The TEQSA *Guidance note on research and research training* also requires universities to establish and implement "an institutional research training policy framework". In this context, the policy framework is required to address:

"[...] specific matters that institutions need to achieve in relation to:

- *the rights and responsibilities of research students and supervisors*
- *monitoring the progress of research students*
- *communication of research outputs by students*
- *the resolution of disputes."*

Legislative Requirements

The handling of research data is governed by a range of state/territory and commonwealth legislation. A major consideration is privacy legislation. Legislation such as the *Privacy Act* (Cth), *Privacy and Data Protection Act 2014* (Vic), *Privacy and Personal Information Protection Act 1998* (NSW), *Archives Act 1983* (Cth), state archives acts, and the *Health Records Act 2001* (Vic) require that researchers handle personal information in accordance with the relevant privacy principles. The Office of the Australian Information Commissioner provides a list of privacy laws in each state and territory.[35]

Other legal requirements, such as compliance with record-keeping standards and intellectual property, including patent applications or the commercialisation of research, may also be relevant.

Clinical trials (drug and device trials) are governed by the Therapeutic Goods Administration. Section R6E2 of the *Guideline for good clinical practice*[36] requires full and verifiable capture of data and decades-long retention of data for investigational medicinal products. Likewise, ISO 14155: 201137 requires the same level of record-keeping and long-term data retention for investigational medical devices.

The examples above are not definitive and there will be a range of legislation relevant to the research being undertaken at the university; also, requirements may differ between jurisdictions. Such legislation should be referred to in the policy.

Call to action: Institutions are encouraged to collectively describe details of relevant legislation that affects aspects of research data in their state or territory and other relevant jurisdictions.

Funders' And Publishers' Requirements

Funders may have their own requirements for RDM. The ARC and the National Health and Medical Research Council (NHMRC) have integrity policies that require universities to develop and implement policies on RDM. Other funders may also have requirements that relate to how data is collected, stored, shared, owned, published or disposed of.

Since February 2014, the ARC has required researchers to outline how they plan to manage research data arising from ARC-funded research. The requirement includes the proper management of research data and primary materials by researchers, along with university policies addressing data ownership, storage, retention and "appropriate access [...] by the research community".[38]

The ARDC is enabled by NCRIS

Major funders also generally require that outputs, often including data, produced through publicly funded research be made openly accessible through an institutional or discipline repository.

Many projects also have specific requirements related to data size, versioning, security, sharing, retention, backup and research data management plans.

Publishers and other relevant bodies may also have requirements that should be considered or referenced in an RDM policy. The Committee on Publication Ethics (COPE) and the International Committee of Medical Journal Editors (ICMJE) both require retention and long-term storage of the data underpinning publications. Specifically, the ICMJE "believes investigators have a duty to maintain the primary data and analytic procedures underpinning the published results for at least 10 years".[39]

Other Stakeholders' Expectations

Different groups or organisations may have their own expectations on appropriate management and sharing of data that fall outside of legal requirements. Explicit understanding of these is crucial to build trust, enable future collaboration, and meet the individual's and the university's accountability.

The ARDC is enabled by NCRIS

Establishing and Updating the Policy

Most institutions already have documented processes in place to draft, agree and sign off a new policy.[40] Table 1 summarises the steps in the process and related considerations for each step.

Table 1. *Establishing a policy for RDM – 14 Steps*

STEP	RELATED CONSIDERATIONS
1. Identify the need for the policy.	Check the policy suite for what already exists, conduct a gap analysis and identify related policies. Triggers to changing the policy can be: ■ changes to legislation and guidelines ■ something going wrong ■ a request from researchers ■ the normal review cycle.
2. Set up a roles and responsibilities (RACI[41]) matrix of stakeholders and their roles in the policy/process.	Identify the policy owners and stakeholders. This will vary for each institution and may ultimately sit with the DVCR. This step is about identifying responsibilities and roles, not just for the owner of the policy but also the roles in policy development and associated activities within the group responsible for developing the policy. Each institution will be operating under a different governance framework. This will determine who has the authority to make decisions related to or affecting research data at the enterprise level and who holds primary responsibility for RDM. An early step in policy development or revision is to refer to and understand the current delegations of authority, consider if these arrangements are still appropriate or propose changes. The roles identified for responsibilities will normally include: ■ data executive ■ data owner ■ data custodian ■ data steward ■ data stakeholder. It is important to consult with relevant parties such as the research division, data governance, privacy, conduct and integrity, copyright, data management and eResearch, training administrators, library, IT, and the Office of General Counsel. If there is an RDM or eResearch committee, it can be useful to have them review the policy. Identify key components of the policy and the policy development. This can involve people who are responsible for data ownership, active data systems, archive/repository, training, draft policy development, communication and documentation record-keeping.
3. Gather information.	In gathering information it can be useful to perform a gap analysis, or conduct user surveys, identifying, for example, what was missing in the old policy.

The ARDC is enabled by NCRIS

ELEMENT 3: POLICY

STEP	RELATED CONSIDERATIONS
4. Draft the policy. This should be done by a person(s)/ department(s) with expertise in the area.	Benchmark or identify the key objectives of the policy. Ensure there is the capacity and support to draft the policy. Consider ongoing resourcing implications and capacity for the research office, eResearch and research integrity, for example. It is worth thinking about future expectations and developments and planning for these in the policy. The policy will require senior level buy-in so it is important to take that into account while drafting the policy, procedures and guidelines.
5a. Ensure processes and systems align with policy.	Policy and procedures can be made mandatory only once the required services and infrastructure are in place. It can be useful to have a champion to ensure actions are completed.
5b. Map out the impact of the policy on other departments or other relevant units.	Consult other departments or other relevant units to find out whether operational capacity is available and what the implications of the policy will be.
6. Develop a communication plan for the process of establishing the policy.	Communicating, raising awareness and managing expectations during the consultation process is important, to take researchers and executives on the journey. It is especially important when expected changes are significant. It can be helpful to draft answers to common major concerns.
7. Consult stakeholders.	The draft policy should be reviewed by all relevant stakeholders. It is worth sending it to existing committees, or if they do not exist, consider arranging meetings. Stakeholders include users/researchers of the policy.
8. Consider external consultation.	Examples of external organisations that could be consulted are: the ARDC, leaders in policy and open research, international experts, or perhaps an external consultant.
9. Consult with required departments or relevant units (for example, Legal, Risk Management).	The departments/units that must be consulted for *any* new policy can vary by institution. There is usually a standard list of departments that need to agree on all policies.
10. Release the draft policy for institutional-wide comment.	Sufficient time should be allocated to this step to ensure inclusive and effective feedback can be gathered.
11. Incorporate university-wide feedback.	Depending on the feedback received, consider whether the relevant stakeholders need to be consulted again.
12. Seek approval through standing committees.	This will vary by university and may include research and research training committees or the academic senate. This will be dictated by the university governance structure.
13. Approve the policy. Publish the policy.	Approval is by the policy owner. Plan training for staff involved in the roll out and application of policy. Final approval to publish a policy often requires an implementation and communication plan. The roll out can be led by a department other than that of the policy owner.
14. Compile FAQs for the policy.	Compile common questions encountered during consultation processes and draft answers. Have clear reporting and escalation lines for emerging or other issues.

Recommendation 1: Follow the 14-step process to establish a policy for RDM.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Considerations For A Good Research Data Policy

A good policy is framed in a way that relates to its key audience, which, for a research data policy, is primarily researchers. For most universities, the policy document hierarchy is as follows:

1. Legislation
2. Policy
3. Procedures and standards
4. Guidelines
5. Local (university-specific) protocols.

Policy documents are principle-based, with operational detail incorporated in procedures and guidelines. They are written and presented clearly and succinctly, taking a holistic perspective.

There is also an ARDC guide describing the content of an RDM policy.[42]

Components Of A Research Data Policy

Institutional policy documents usually contain key headings or components, as described below. Universities seeking to develop a new research data policy, or revise their existing policy, may find these headings a useful starting point. Additionally, Swinburne University of Technology has shared a template research data policy which universities may reuse.[43]

Definition Of Terms

A policy normally contains a section that defines the terms used in the policy. The definitions should include a definition for research data. Neither the Code nor associated guides contain a definition for research data, and definitions set by universities vary.

Some institutions' policies apply only to data in a digital format, and others include primary and physical materials. Software and the code used to generate or analyse the data is included in some instances. Universities will also differ in whether non-traditional research outputs are included in the definition of research data.

Purpose Of The Policy

Most policies contain a short description of the purpose of the policy. Here is an example from the University of Wollongong:

- *"To provide guidance regarding the management of research data and primary materials throughout the research data lifecycle*
- *To outline the requirements articulated in The Australian Code for the Responsible Conduct of Research regarding the responsibilities for the management, storage, access, retention and disposal of research data and primary materials"*

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Institutional Aims And Values

Policies can contain references to the strategic aims and values of the university, such as supporting the reproducibility of research, verifying research findings, sharing information with the public, conducting research ethically, respecting Indigenous data sovereignty, or providing transparency to the research conducted at the university.

These aims and values can be supported by referring to existing external principles and guidelines such as the FAIR data and CARE principles[4,5]. For an example of how this can be done, see the references to CARE and Indigenous data sovereignty in the template RDM policy provided by Swinburne University of Technology[42].

Roles And Responsibilities

This section of the policy describes at a very high level the positions and roles outlined in the policy or associated procedures, and their responsibilities.

It should identify in detail who is responsible for the policy and who has overarching responsibility for implementing and maintaining the policy; this information, including definitions of delegations and authority, usually resides in procedures and guidelines.

The responsibilities of the researcher, supervisor and the institution as defined in the Definition of Terms section of the policy (above) should also be included, along with relevant delegations.

Where roles in the policy include positions such as data custodian, data owner, or responsible manager, these roles should be defined in the Definition of Terms because arrangements will vary between institutions, departments and research groups.

Responsibility for identifying project-specific legislation should also be considered.

It is also important to describe the responsibilities of the university under this policy. Here is an example:

The institution is responsible for:

- providing infrastructure necessary and appropriate for the safe and secure storage of research data and primary materials and records
- relevant training in best practice RDM, including training in the effective use of research data infrastructure.

It can be useful to appoint a responsible officer for implementing and exercising the responsibilities in the policy, and to outline the responsibilities of specific parts of the institution for providing services, such as training, storage and ethics guidance.

The ARDC is enabled by NCRIS

The variations in role definitions across institutions and the different considerations are worth highlighting. Here are some examples:

- There is some variation between universities in the definition of a researcher. The University of New England defines a researcher as a Researcher Person (or persons) who conducts, or assists with the conduct of, research. Such a broad definition will include PhD students, masters by research students and honours students. Research assistants may also be included. Universities should consider whether their definition should also apply to supernumeraries (adjuncts, honoraries, emeritus staff) and visitors.
- Macquarie University's research data management policy[44] identifies several roles, including research data custodian, research data steward, authorised user, principal investigator and research integrity advisor. The role of research data custodian over a dataset may be transferred over time.
- Victoria University requires that "Responsible Owners/Managers/Custodians of the data and materials" be identified on a research data and materials plan when accessing storage. These can be data managers, principal or chief investigators, or heads of school; however, at least one of them should be based at the university.

'Data owner' can be a complex term and will vary within the institution by arrangement. Broadly speaking, data owners hold legal rights to data, and have responsibilities for that data. In some cases the data owner will be the researcher, in others it may be the contractual parties or the institution. In some cases it may be a team or a board that is responsible for the data collection. Ownership may also be held by or shared with other parties, such as patients or Indigenous communities. The rights and responsibilities associated with data ownership may be distributed between parties in different ways depending on the arrangement. For this reason it is important to clearly specify what is meant by 'data owner' where this term is used.

Research Data Governance

Research data governance is not always a separate section of the policy. However, it is important to consider who can make decisions about the research data collection and who will have access to the collection. This does not have to be an individual and can be covered by a committee.

Documentation And Metadata

Throughout the research process, metadata and documentation is collected and should be stored alongside the research data. This allows the responsible persons and the institution to make decisions about the data throughout the research process and when sharing the data. It can also help inform decision-making on retention and disposal of the data. It is important to clarify expectations placed on researchers in this area.

The ARDC is enabled by NCRIS

ELEMENT 3: POLICY

The University of Wollongong's research data management policy[45] provides an example for wording about collecting and storing metadata:

> *"Metadata is the information describing the attributes of an item or resource that enables it to be identified, retrieved and managed over time.*
>
> - *Administrative metadata may consist of dates, file size, type, creator details, location, licensing information and retention periods to assist with the management of the dataset.*
> - *Descriptive metadata provides information for ease of discovery and retrieval such as an abstract, title, keywords, categories, versions and unique identifiers. Supplementary contextual information should be captured to enable the reuse and interpretation of the data.*
> - *Structural metadata explains how the data is organised and how it relates to other collections, for example, indexes, chapters, database fields, page numbers and XML schemas.*
>
> *Metadata must accompany the research data for as long as it is retained, and should be updated as necessary.*
>
> *Any process, software program or code used to generate or analyse the research data should be documented and recorded in the metadata."*

RDM Planning

It is useful to have a principle in the policy regarding RDM planning, whether it is mandated for specific types of data and research, or for all research. Some funders and institutions require data management plans to be developed and lodged as part of the RDM planning process, although the detail is usually included in a related procedure.

Storage

This section of the policy describes what research data storage the institution provides and supports, including multiple-site or cloud backup, where appropriate. It should state expectations about the use of storage options available to researchers, which may include institutional, project, discipline, national or personal storage facilities, and the situations in which these are, or are not, appropriate.

Access Management Controls And Cybersecurity

This section of the policy describes what the institution expects of those providing access to data. It includes considerations for providing access to sensitive data, including culturally sensitive data, human data, or data that is commercial in confidence.

It can be useful to provide a classification of data types in a procedure, with associated storage and management solutions.[46]

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Use Of Pre-Existing Data

This section of the policy describes what the institution expects of researchers who are using pre-existing or third-party data. This may include data obtained from databases and collections, provided by other organisations via contractual agreements or research collaborations, or derived from publicly accessible sources such as social media. It may also include licensing and copyright considerations, documentation of and adherence to the terms and conditions of data access, and appropriate acknowledgement of pre-existing data.

Sharing Of Data For Reuse

This section of the policy describes what the institution expects of researchers in the sharing and reuse of their data. It may elaborate on what data may (and may not) be shared, when, how, and with whom. It may include recommendations on copyright, publishing, the use of specific licences, and procedures for drafting data-sharing agreements, when appropriate.[47]

Retention And Disposal Of Data

This section of the policy describes the minimal retention periods and the institutions expectations for retention and disposal of research data and records. It may also outline expectations about longer retention beyond the minimum for specific collections or under specific circumstances. It should include considerations and arrangements for the permanent disposal of data.

Related Legislative Requirements

The policy should include references to related legislative requirements, which in some cases differ depending on state/territory. References include:

- the relevant state or territory privacy legislation[48] or the national Privacy Act
- any relevant data-breach reporting obligations
- Commonwealth/state/territory legislation relating to retention and disposal of records
- funder requirements, such as the Code and accompanying *Guide for management of data and information*[49]
- ethics guidelines such as the *National statement on ethical conduct in human research*[50] and the Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS) *Code of ethics for Aboriginal and Torres Strait Islander research*[51]
- TEQSA's *Higher education standards framework* (threshold standards)[52] (sections 4.1 together with 5.2 and 6)
- relevant project-specific legislation, including international legislation to be met at the project level.

Handling Data Breaches

The policy (or associated procedure) should provide some instruction for how to handle breaches of appropriate RDM standards; for instance, by referring to existing protocols for reporting data breaches at the institution.

Relationship To Other Policies And Procedures

The policy should provide a clear framework that points to other related policies and more detailed procedures. These procedures should contain more detail and be adjusted as infrastructure or processes change at the institution. Consider where each of these procedures fits into the life cycle of the research process and the life cycle of the data.

Example related policies are listed in Table 2, arranged from the more frequently mentioned to the less frequently mentioned by participating universities. The naming and scope of these policies will differ across institutions.

Table 2. *Examples of related policies (in descending order of frequency of being mentioned by participating universities)*

- Intellectual Property Policy and Procedure
- Privacy Policy
- Open Access Policy
- Records and Information Management Policy
- ICT Information Management and Security Policy
- Cybersecurity Policy
- Research Integrity Policy
- Ethics Policies
- Authorship Policy
- Copyright Policy
- Data Governance Policy
- Collaborative Research Standard
- Research Code Complaints, Breaches and Investigation
- Research Misconduct Policy
- Enterprise Architecture Policy
- Administrative Access Scheme Policy
- Right to Information Policy
- Data Storage Policy
- Publication and Dissemination Standard
- Information Classification and Handling Policy
- Authorship Policy.

The ARDC is enabled by NCRIS

ELEMENT 3: POLICY

An RDM policy should reference the relevant policies, each of which should reference the RDM policy.

Institutions may also have related procedures and guidelines that set out in detail what processes are in place and how a researcher can meet the expectations. Table 3 lists examples, again arranged from the more frequently mentioned to the less frequently mentioned. The naming and scope of these policies will differ across institutions.

Table 3. *Examples of related procedures and guidelines*

- Records and information management procedures
- Authorship procedures
- Human research ethics procedures
- Research data management guidelines
- Research data management procedures
- Research data sensitivity, security and storage guidelines
- Data management guiding principles
- Research integrity and misconduct procedures
- Right to Information procedures
- Lab notebook guidelines
- Engagement of cloud computing services procedures
- Information infrastructure and use regulations
- Administrative access scheme procedures
- Animal Ethics Committee procedures
- Biosafety procedures
- Information privacy procedures
- Information asset and security classification procedure
- Cloud computing procedures
- Intellectual property guidelines
- Data governance framework
- Higher degree research confirmation.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Recommendations And Advice

Procedures To Support Policy

An RDM policy may be accompanied by procedures. A procedure can be used to detail how a researcher is expected to meet the requirements of a policy and can help establish some of the mechanisms by which an institution can track compliance. Good procedures may contain best-practice examples to support researchers and paint a clearer picture of what is expected of them in practice. Standards and procedures are generally easier to approve and change than a policy, so processes that are subject to regular change are better formed as a procedure rather than being contained in the policy. These procedures, and other resource sheets, may be referenced from the policy.

Procedures Related To Sensitive Data

Procedures related to the management of sensitive data are valuable documents that often contain a data-classification schema. This is very useful for researchers when making decisions on appropriate places to store their data and mechanisms to provide access to these data.

Sensitive data must be handled in a manner appropriate to the nature of its sensitivity. While often focused on human subject data, it may also consider data related to, for example, archaeological sites, corporate entities or endangered species.

Procedures should include the nature of permission and consent. That includes the importance of capturing metadata and managing information so that consent is recorded and approvals are clear, including respecting the desires of participants and communities.

Procedures should also address consent for the use of third-party data where those whose data has been collected cannot necessarily directly consent to its use; for example, data collected from Facebook and Twitter.

Sensitive-data procedures may also include privacy guidelines and outline the difference between de-identified, re-identified, re-identifiable and anonymised data, as well as the potential needs of third parties. The principles related to sensitive data may be set out in the RDM policy, or in a separate policy. Procedures relating to sensitive-data classifications should be referenced from this section of the policy.

Implementing the Policy

Most institutions have documented processes to implement a new policy. Some change and communication support may be required to support and encourage uptake. Monitoring and review of the policy should also be factored in.

The ARDC is enabled by NCRIS

Table 4 lists steps to implement a policy and related considerations.

Table 4. *Implementing the RDM policy – 9 steps*

STEP	RELATED CONSIDERATIONS
1. Consult broadly during policy development.	Relevant parties for consultation include those parts of the university that deal with research administration, data governance, ethics, privacy, legal advice, conduct and integrity, data management, eResearch services, and training and guidance. Use a consultation log to record engagement and input critical to the development of the communication plan.
2. Approve the new policy.	When the new policy is approved, target communication to the areas/people who are required to implement new processes (in accordance with the communication plan).
3. Assign responsibility for the communication plan.	The policy sponsor is responsible for monitoring the implementation process and ensuring staff are provided with appropriate information and training.
4. Complete and agree on the implementation plan.	The implementation plan is usually prepared before the new policy is announced. There may be a standard template for policy implementation that can be used. This usually covers elements such as: ■ stakeholders ■ support, advice, guidance or training ■ awareness / internal communications plan ■ timing of expected changes; for example, prospective policy ■ requirements and phased progressive release of procedural requirements ■ responsibilities for actions ■ general sensitivities or barriers to communication ■ implementation risks and mitigations.
5. Complete supplementary materials and integrate with existing systems and orkflows.	Materials should be constantly updated and evaluated and communicated as issues emerge.
6. Consider the level of culture change required. Develop and implement a strategy.	Consider measures to incentivise compliance, such as provisioning data storage, which is usually more effective than penalising non-compliance. Where training plays a role in culture change, ensure it is delivered in context. Policy is an important element for culture change activities, which may form a separate initiative.
7. Audit / check compliance.	Any audit should be agreed by the policy sponsor and would normally be held either before the implementation of the policy or as part of the post-implementation audit. The audit may identify organisational maturity or evaluate compliance with the policy. Timing of the audit and checking against the objective of the policy is important. Having control over that makes it much more useful. Performing an audit is quite resource intensive. When the university is undertaking a TEQSA re-accreditation, an assessment of the effectiveness of the policy may be relevant.

STEP	RELATED CONSIDERATIONS
8. Conduct a post-implementation review.	The policy will be subject to review according to your institution's policy review cycles. Use the nominal 'review by' date set by your institution. The purpose of the review should be to understand whether the policy is being implemented and is achieving its intended aims/purpose. The review needs to be carefully planned and should be the responsibility of the policy sponsor. It can be useful to use the risk and mitigation initiatives identified in the policy implementation plan.
9. Consult stakeholders / survey users of the policy.	This is generally part of the post-implementation review.

Recommendation 2: Follow the 9-step process for implementing an RDM policy.

A need for shared information: Institutions are encouraged to share with one another examples of their post-implementation review processes.

Resources

- List of all available university data policies: https://sites.google.com/ardc.edu.au/rdmcommunity/resources/institutional-policy
- ARDC Guide on Institutional research data management policies and procedures. ARDC; 2021. Available from: https://docs.google.com/document/d/1E2wQbBlbQ48EJZH_JAJ01giUreK0Alo9/edit
- UNSW's research data management guidance for researchers: https://research.unsw.edu.au/research-data-management

ELEMENT 4: RESEARCH DATA MANAGEMENT PLANNING

Summary

This element provides institutions with guidance to achieving good practice in planning research data management which includes standardisation, tooling and selection of applications. We offer advice on implementing the RDM planning infrastructure, the core elements of a data management plan (DMP), and how RDM planning outputs may be used. We describe detailed steps for RDM planning and offer perspectives on how to engage different stakeholders.

Why Planning For Research Data Management Is Important

RDM planning is the forward planning of the management of data generated from a research project or projects. It can include identifying the types of data that will be generated, how that data needs to be stored and handled, who can access it, how long it must be stored, among many other factors (the essential ingredients of a DMP are described below in more detail).

Researchers are typically the ones responsible for most RDM planning, as they have the best understanding of their data. Institutions seek to encourage and support good RDM planning, as it results in better RDM. One approach taken by institutions is to require or encourage the researcher to write a DMP, which formally documents their planning decisions. However, RDM planning is not restricted to the writing of a specific plan, and may include other processes that encourage researchers to consider and plan for the management of their data. RDM planning should not be a once-off process, but rather continue through the project to ensure that the plan reflects the most up-to-date understanding of the data.

Some institutions implement DMP tools that guide researchers through the production of a DMP, especially if an institution is large and requires consistency. For other institutions, support of RDM planning may be ensuring that RDM procedures, such as an agreed/optimised workflow, are followed. The framework considers the spectrum across institutions and does not mandate the existence of a DMP tool (such as a structured template). The *National statement on ethical conduct in human research* (3.1.45)[50] requires that a DMP is created prior to requesting ethics approval for all data requiring ethics approval.

The RDM planning explanations, calls for action and recommendations in this framework should be read and applied in concert and aligned with the *Australian code for the responsible conduct of research*[2] (the Code).

Our goal is to advise institutions about how to set up the systems, tools and processes related to RDM planning across the institution. This is not a guide that tells researchers how to plan RDM for an individual project.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

There are some institutional challenges associated with RDM planning:

- Resourcing to develop and implement an institutional approach to supporting RDM planning, which varies greatly across institutions
- Culture change, which is required within institutions and is likely to be slow
- Gaining institutional recognition of the importance of an institutional approach to RDM planning
- Incentivising and standardising RDM planning within the institution
- The different units within institutions that are responsible for implementing an RDM planning approach – where should the responsibility lie?

The following factors contribute to the variation in institutional approaches:

- Institutions are at different stages of maturity in their planning approaches.
- Culture change is required across institutions and change is typically slow.
- Institutional recognition of the importance of RDM planning varies across institutions.
- Incentivising and standardising RDM planning needs to occur both within institutions and nationally (to meet ARC requirements, for example).

Purpose Of RDM Planning

All researchers are connected by their use of data. However, accompanying the benefits are risks and responsibilities for creating, storing and sharing data. The purpose of RDM planning is to plan ahead to ensure institutions and researchers maximise benefits, minimise risks and meet their responsibilities such as adhering to FAIR data and CARE principles. Consequently, RDM planning facilitates the availability of data and enhances the impact and quality of research outputs.

Successful RDM planning approaches have many components. Institutions are tasked with developing governance and policy, strategic planning, resourcing, defining roles and responsibilities, staffing, training, and infrastructure. To be successful, RDM planning approaches need to coordinate between many levels of governance (from institution to individual researcher) and in many areas of an institution (across research project teams, research support and central services teams). Engaging with these key stakeholders is critical for the uptake of good RDM by researchers.

Engagement between institutions' key stakeholders and researchers occurs at different levels and even across institutions. For example, policies developed by committees or in research administration offices affect the roles and responsibilities of data custodians. Policies and resources can be used to engage with data custodians to give them important context and background for effective RDM planning. RDM planning tools accompanied by training and support can assist researchers achieve their research goals while increasing the uptake and proper use of such tools.

Ultimately, institutions and researchers benefit from a shift in focus from compliance and DMPs to engagement with the process and practice of RDM planning. Researchers realise the benefits of RDM planning and fulfil their RDM responsibilities when it is part of their research practice. Achieving this level of RDM practice requires support, training and education in RDM planning, and awareness of the variability of research contexts. Thus, understanding stakeholders' needs and the barriers to uptake is critical for ensuring risks and responsibilities of RDM are adequately addressed at all levels within an institution.

Benefits Of RDM Planning

The benefits of RDM planning can be understood by considering the perspectives of the different stakeholders.

Benefits to the institution may include:

- increased commercial/community trust and confidence
- greater understanding of the research landscape at the institution through analysis of research and research data holdings
- identification of gaps in the research process and where researchers might require support within the research life cycle
- increased ability for institutions to effectively manage storage capacity which, in turn, informs future investments; knowing early in the research life cycle that researchers will be requiring storage means the institution can plan to ensure enough storage capacity is available
- retention of institutional knowledge through the tracking and reporting of data
- identification, capture and reuse of more research outputs
- compliance with government and funding agency requirements
- greater business intelligence capabilities around RDM
- mitigation of risks of reputational damage
- improved readiness for audits and changes in funding agency strategy towards open access, and demonstrated commitment to be an institution of influence
- improved RDM practice and culture if RDM planning occurs through the research life cycle.

Benefits to the institution are enabled when:

- the institution can provide support to their researchers that translates to an increase in good RDM practices
- reporting on research is underpinned by planned data management that applies FAIR data practices
- RDM roles are defined, which aids in the use of supplied resources such as research data storage.

ELEMENT 4: RESEARCH DATA MANAGEMENT PLANNING

RDM planning benefits researchers in multiple ways, such as:

- improved data publication quality and more effective data sharing with project collaborators
- rapid provisioning of IT resources decreasing unproductive time at the start of a research project while ensuring compliance with research data policy and legislation
- enhanced researcher profiles and potentially new-found audiences and collaborators through dissemination, citation and reuse of data
- data transparency around access, consent and licensing
- less lag-time to data analysis and publication
- easier access to raw and processed data needed for research
- time savings when data is better organised and easier to find
- reduced risk of data being stolen, lost or misused
- understanding of roles and responsibilities
- supervisors kept up to date with evolving RDM planning (if managing research students)
- having an evidence trail facilitating compliance with research integrity; for example, provenance for investigations and disputes.

Benefits for researchers are enabled when:

- RDM processes and systems are linked throughout the life of a project, from RDM planning at the start to data publishing at the end
- the process of RDM planning is easy to use and understand, making it more likely to get buy-in for engaging in the practice
- DMPs are ongoing documentation of the process facilitated by systems that help ensure researchers remain engaged with the practice
- support for completing and reviewing the DMP originates from research supervisors and lead investigators on research projects
- researchers understand that RDM planning is more than compliance for obtaining access to resources such as data storage facilities
- RDM planning provides a ‘chain of custody’ for project data, from primary materials through to reduced data. This strengthens project data output and scientific conclusions based on that data.

Benefits to the research community are enabled when:

- a robust approach to the practice of RDM planning is agreed across institutions; this supports the research community and promotes good RDM practice nationally, even when researchers move across institutions
- the community has knowledge of FAIR and CARE principled data management, which aids the practice of RDM planning and documentation through to publication and reporting to funding agencies and communities.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Planning Considerations

The following is a discussion of common considerations that should be helpful for institutions embedding RDM planning practice.

Overarching Institutional Approach And Understanding

- Keep the focus on the benefits of RDM planning for researchers and institutions.
- Recognise that some research is multidisciplinary and requires RDM planning that is synthesised and harmonised, though not necessarily standardised.
- Ensure that the RDM planning approach or framework references the Code.
- Create a working group or community of practice to support the development of any RDM planning strategies, tools, support, documents.
- Bear in mind that RDM planning is often driven by the requirements of funders and can be influenced by the research life cycle.

Integrating With Primary Materials Management

It is advisable to integrate RDM planning with primary materials management as this supports the research life cycle. Primary materials are possibly the largest type of research data generated. RDM planning must include primary materials management, or it will exclude a significant number of key research areas. RDM planning cannot be confined to digital data only, given that much digital data is derived from primary materials. Data integrity begins with primary materials integrity.

Publishing DMPs

- DMPs have been published in Europe and America.
- Publishing DMPs may be stipulated as a requirement of funding.
- Publishing documented DMPs may help support knowledge exchange and understanding of what happened during the project, highlighting areas related to data context, licensing and IP, and management and handling of the data.
- Documenting RDM planning facilitates publication of a DMP, if required.
- Documenting RDM planning can also support the publishing of a dataset, so being able to integrate the two processes could be beneficial and support data publishing.
- Machine-actionable RDM should be considered — it may enhance the ability to integrate tools, while supporting automation and validation of DMPs.

The ARDC is enabled by NCRIS

Making RDM Planning Mandatory

- Some funding bodies (for example, ARC) have a clause requiring that a DMP be created.
- At many universities, the completion of a DMP is mandatory for HDR students and is seen as something that needs to be done to enable them to commence with their studies. Students are encouraged to see that the DMP is a "living document" and that it supports good RDM during their research project.
- RDM planning is best practice and should be an integral part of high-quality research training. Specifically, RDM training can be included in researcher inductions and periodic research refresher training.
- *The National statement on ethical conduct in human research* (3.1.45)[50] requires that a DMP is created for all data requiring ethics approval.
- It is generally accepted that RDM planning and documentation in a DMP is good practice but it is not widely undertaken by researchers.

Connecting RDM Planning To Other RDM Processes

- The information captured in a DMP can influence other decisions related to RDM, such as storage allocation and data-capture tool selection.
- Decisions could be related to the sensitivity of the data being described and its impact on licence choice. Highly sensitive data may require a more restrictive licence to the data when publishing.
- Decisions could be related to the ethics process and impact of any ethics clearance required for the research.
- Information captured in RDM planning can also influence decisions on retention/disposal which in turn assists with not only storage allocation but also the overall storage management, for example capacity planning.

Planning For Training / Culture Change

RDM planning can be used as a training / culture change tool, as a survey tool, or as a way to follow up on the actual use of services (which is more useful than just compliance).

Understanding and incorporating the development of an RDM planning process into other processes helps with gaining buy-in and engagement, and facilitates suitable integration.

Through the practice of RDM planning, students and academics can be trained to understand the requirements for managing, publishing and preserving data, and FAIR and CARE principled data management in its entirety.

When planning for training and culture change, consider stakeholders' perspectives:

Infrastructure Managers

Having research infrastructure managers involved in RDM planning aids them (and the institution) in capacity planning, tool integration and support that covers requirements across the research life cycle.

Research Ethics And Integrity

- Integration with the ethics application process supports RDM requirements and streamlines activities for researchers. Making the DMP accessible to the ethics committee will facilitate the ethics approval process for the research project.
- Data management and integrity breaches such as fabrication, falsification, use without consent, and re-identification are significant areas of focus that can be reduced or avoided with an effective RDM planning approach.

Library

- Library staff should be involved in various aspects of RDM planning, aiming to help researchers early with publishing data and metadata, archiving and preserving data, as well as managing collections to increase impact and citation.
- They play a key role in all phases of RDM planning, knowledge/assistance for access to data, data licences, copyrights and use of third-party property.
- At some Institutions, the library is the hub of RDM, with staff conducting the training in essential/foundational RDM or providing consultation services for RDM.

Graduate Research School

- The graduate research school should be made aware of the RDM planning as they are the point of contact for HDRs (especially for new HDRs).
- At UNSW, the Graduate Research School manages HDR project information and requires a DMP from students to provide support and manage information on the student, course and supervision, which are all tied to the DMP execution.

Researchers

- Regardless of institutional size, a better balance of incentives and benefits is needed to increase motivation, not just an emphasis on compliance (administrative work). For example, the University of Wollongong uses data storage as a "carrot" which can be accessed only after completing the DMP.
- Researchers may be motivated only if they "have to"; for example, for compliance with grant conditions.
- RDM planning training included in researcher induction and periodic refresher training will advance their data management capability.

The ARDC is enabled by NCRIS

Research Office

- At small institutions, the research office has a broad scope of activities across the institution. The research office is involved in RDM planning as it covers a broad spectrum of activities and is ultimately responsible for good reporting to internal and external stakeholders. DMPs offer insights into the institution's research assets, capabilities and strengths. As an example, at Bond University the need for high performance computing (HPC) identified from a DMP led to the purchase of HPC service access.
- The idea of research 'data stewards' could be operationalised as a position or a role, as dictated by the institution budget and/or research strategy. They would be a significant stakeholder and, perhaps, an example of best-practice support for effective RDM planning.

Research Community

Having the research community involved in planning should result in:

- greater transparency of their research
- greater confidence in scientific research outcomes and recommendations that drive public policy.

Selecting Planning Tools

The tooling component of RDM planning relates to how RDM planning is documented, managed, published and referenced.

Here, we discuss tooling considerations for best-practice RDM planning, depending on the stage of digital integration an institution is at. Each recommendation may be applied independently.

Tooling for RDM planning can and should be integrated with other systems at the institution (for example, ethics management, storage allocation and grants management systems) so that DMP information can be shared and decisions automated accordingly. These integrations help to build the bigger picture of RDM and the components that need to be considered when planning a project because they reduce administrative burden, both for researchers and administrators, and increase uptake. The FAIR data and CARE principles should also be considered when it comes to integrating with other processes. These principles can inform the adoption of strategies to support RDM. For examples, see the ARDC's recent guidelines for making project data outputs FAIR[53]. Another process to consider is that of making RDM planning and its documentation machine-actionable.

The ARDC is enabled by NCRIS

When selecting a tool, consider stakeholders' perspectives:

Researchers

If the RDM planning tool can be integrated with other systems at the institution, the completion of fields in some forms may be able to be automated, reducing the effort of the researchers in completing forms.

Institutions

- The institution will benefit from this development of a national framework for RDM planning tools as this will provide guidance and a structured approach to the development of the RDM planning tools.
- Standardising RDM planning tools across institutions can help interoperability with other systems.
- A solution that works for several institutions leverages the potential benefits of a community forming around this solution and the community's influence on solution development.
- Having common approaches to the development of RDM planning tools across institutions enhances conversations about RDM planning and the support institutions can provide for each other.
- Integration with other research administration systems — such as ethics and human-resources feeds — is particularly important to ensure data quality and consistency in the tool used.
- Using an RDM planning tool enables better reporting and auditing. This is good for many reasons. It also assists in research integrity investigations — the research integrity assessment officer (or appropriate authority) can conduct discreet investigations before engaging a respondent.

Institutional DMP Tools Community

- The development of a portal for RDM planning tools can support the community in understanding the tools required, their capabilities, processes, applications etc. It also allows enhanced interactions between institutions, enabling sharing of thoughts and issues.
- Similarly, the development of a collaborative community of practice can enhance interactions between institutions.

Library Staff

Library staff will need to work with the researchers to help them to understand the RDM planning tools.

Research Office

The research office should be consulted as they support various processes related to RDM planning tools.

ELEMENT 4: RESEARCH DATA MANAGEMENT PLANNING

Call to action 1: Institutions are encouraged to reach agreement on the following considerations:

- What goes in the ethics system and what goes in the DMP, and in what order?
- Can RDM planning be linked with the institution's (sensitive) data classification framework?
- Can standard lab-operating procedures be used in RDM planning?
- How might we align tooling for RDM planning with FAIR data and CARE principles?
- How might we define a minimum standard with RDM planning tools?
- How might the framework better support our institutions in meeting the minimum standard?
- How might we share and standardise DMPs and/or planning documentation across collaborators and institutions?
- How can we use the scalable nature of the framework across institutions that are at very different stages (for example, taking DMPs from being an optional Word document to a technical, integrated, machine-readable and actionable plan), especially when many of the changes rely on funding to implement developments or solutions?

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Recommendations And Advice

Based on the information presented and our review of the considerations above, we recommend the following steps for planning RDM and for planning tooling for RDM.

Recommendation 1: Follow these steps for planning data management.

Institutions and RDM communities:

1. Create an understanding of the purpose and benefits of RDM planning.
2. If implementing DMPs using, for example, a tool or standardised documentation, reach agreement internally on the business drivers/purpose.
3. Make a conscious decision about the timing and level of rigour required for RDM planning and its documentation, given that RDM planning is essential for all research.
4. Consider methods for interrogating and analysing the content of DMPs to determine their usefulness and to find any gaps in the research life cycle at the institution.
5. Create an online presence to allow for support and training activities in relation to RDM planning.

Researchers and RDM communities:

1. Clearly define and describe what constitutes RDM planning to help researchers understand why they are engaging in the process and for the institution to understand how it can resource, support and leverage the RDM planning process.
2. Further investigate the idea of publishing DMPs to understand the benefits it could bring to the publishing of datasets and the exchange of knowledge.
3. Evaluate responsibilities and accountabilities using a matrix such as RACI[54] to understand who is responsible and accountable as well as who needs to be consulted and informed in relation to RDM planning.
4. Consider integrating RDM planning throughout the research project life cycle to keep it current and relevant.
5. Embed the Code, and FAIR data and CARE principles in the RDM planning processes and activities so that researchers, institutions and the research community can maximise the benefits from research.

Recommendation 2: Follow these steps for planning RDM tooling.

Institutions:

1. Prioritise the identification of RDM planning tools in use and seek to achieve interoperation and coordination.
2. Enable RDM planning to be an ongoing process through integration of RDM tooling with relevant systems.
3. Create a roadmap for tooling to show best practice, no matter what level of digital integration the institution is at. The roadmap could be as simple as a Word document or as comprehensive as a fully integrated system.

Institutions, researchers and communities:

1. Seek to align RDM planning tools with FAIR data and CARE principles.

Minimum Standards For RDM Planning

RDM planning practices vary widely across institutions due to differences in resourcing, research specialisations and capacity. RDM planning and its documentation will inherently differ across fields of research as it is dependent on the information and data being captured within research projects. However, there are essential elements which all RDM planning should include.

First, the institution needs to understand the requirements of planning and the initial application of minimum standards of DMPs in line with the FAIR and CARE guiding principles for scientific RDM and stewardship. Institutions should review existing systems and processes to benchmark against recommendations.

As discussed earlier, there are significant benefits for both institutions and researchers to invest in best-practice RDM planning. The Code states that researchers must retain clear, accurate, secure and complete records of all research, including research data and primary materials. RDM planning should therefore be seen as essential for best-practice research and be embedded in the research culture of institutions.

At a granular level, RDM planning should be seen as a practice that supports and strengthens research quality and output, and should be recognised as an ongoing process (ideally, documented), with the core foundations of being user-friendly, flexible, trackable and reportable. Institutions' ability to provide a clear curation of project RDM planning and associated data and publications will strengthen Australia's global position in the open access and reproducible research environment.

The ARDC is enabled by NCRIS

Recommendation 3: Strive for best-practice RDM planning.

Institutions:

- Support the curation and development of RDM planning by appointing or identifying persons responsible for this support, such as data stewards.
- Establish RDM teams specifically to support RDM, including primary materials management.
- Conduct RDM planning development, review and updating within an online platform for ease of tracking and for version control. This could be an integrated laboratory information management system or electronic notebook system.
- Identify and action potential RDM planning review triggers, such as changes to legislation, research platforms or tools; funding body requirements; or institutional structure.
- Link the triggering of RDM planning reviews and updates to an existing governance process(es) such as ethics, grants management reporting, or risk management systems. Set dates for automatic triggering of the review of RDM planning, and automate the notification of the documentation author of updates to RDM planning.
- Where appropriate, and in accordance with confidentiality restrictions, use RDM planning key indicators for reporting to inform strategic planning of supporting infrastructure and/or resources.
- Where possible, link in the RDM planning to the allocation and generation of digital data storage. Greater engagement of researchers with RDM planning and its documentation will offer insights into the need for and future planning of infrastructure requirements, such as data storage.

Researchers:

- Use equipment identifiers such as asset numbers in the RDM planning documentation to identify the equipment used to generate data.
- Link to or reference RDM planning in the laboratory information management system and electronic lab notebooks to provide contiguous data associations with a project.
- For primary materials, identify the storage location and the retention-and-disposal schedule for those primary materials.
- Where appropriate, and in accordance with confidentiality restrictions, make RDM planning documentation available as open access resources, for example, for collaboration.

The ARDC is enabled by NCRIS

ELEMENT 4: RESEARCH DATA MANAGEMENT PLANNING

Recommendation 4: Meet the minimum standards for RDM planning.

Institutions:

- Define persistent identifiers (for example, DOI, Research Activity Identifier (RAiD), Research Organization Registry (ROR), ORCID and/or institution-defined identifiers) for RDM planning documentation, which clearly link the RDM planning with project data and publications referencing that data.
- Include management of primary materials as part of RDM planning.
- Provide support for the curation of data, provision of training, and assistance with RDM planning practice and documentation. Clearly identify the responsible group within the institution for RDM support and development to ensure a "single source of truth".
- Try to give researchers clarity on the documentation they need to create for RDM planning.
- Provide secure digital storage for RDM planning information.

Researchers:

- Know that RDM planning is required for all research projects and will be reviewed regularly. Consider RDM planning as part of high-quality research training for HDR students as well as researcher induction and periodic refreshers.
- Consider how to effectively review/assess/self-assess your RDM planning processes once in place.
- Be flexible with your RDM planning practice and its documentation to accommodate and adapt to the needs of different disciplines and align with the Code, and FAIR data and CARE principles.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Essential Ingredients For Data Management Plans

Project-specific DMPs will differ between institutions and fields of research depending on the information and data that needs to be captured within each research project. However, there are essential elements which all DMPs should include. *Management of data and information in research*[49], one of the supporting guides to the Code, recommends the following inclusions for DMPs:

- Physical, network, system security and any other technological security measures
- Policies and procedures
- Contractual and licensing arrangements and confidentiality agreements
- Training for members of the project team and others, as appropriate
- The form in which the data or information will be stored
- The purposes for which the data or information will be used and/or disclosed
- The conditions under which access to the data or information may be granted to others
- The DMP information, if any, that needs to be communicated to potential participants in the research project.

In addition to these essential inclusions identified by the Code, and to align with the minimum standards identified in this framework, institutions should include the following information:

- Project details
- Persistent data identifier
- Data governance
- IP and copyright
- Data storage
- Primary materials storage plan and appropriate retention schedule
- Review schedule for the DMP.

To further support DMP development, Appendix 3 identifies components of DMPs from the institutions who participated in the development of this framework.

Resources

The following examples are included for institutions to consider:

- ARDC. Data management plans: https://ardc.edu.au/resource/data-management-plans/
- DMPTool – a tool to support planning: https://dmptool.org/about_us
- DMPonline – a tool for creating a plan: https://dmponline.dcc.ac.uk/
- Research data management planning checklist: https://libguides.federation.edu.au/rdm/checklist
- ARDC documentation on RDM considerations: https://ardc.edu.au/resources/aboutdata/

The ARDC is enabled by NCRIS

ELEMENT 5: RESEARCH DATA RETENTION AND DISPOSAL

Summary

This element provides institutions with guidance to ensure the mandated retention and disposal of ever-increasing volumes of research data on appropriate systems in line with legislative and funding body requirements. We highlight the importance of embedding the metadata and persistent identifiers for research data to enable integration across research administration and management systems. We recommend the development of a decision-making model that triggers the appropriate retention and disposal of research data. The need for greater sharing of information across institutions is also highlighted.

Why Data Retention And Disposal Is Important

Research data is an immense, diverse and valuable resource that requires specific management by researchers and institutions over extended periods of time. This is particularly evident once a research project's active phase has been completed, when long-term plans for data storage, retention and disposal need to be implemented, and the responsibility for ongoing management of data is passed to different stakeholders and administrators.

Storage, retention and disposal of research data require specific management actions that take place often many years after the research project has been completed, so processes need to be developed that ensure that the information needed to drive future management decisions and processes is carried forward. These processes are integral to reining in burgeoning storage costs while enabling FAIR data and CARE approaches to research data, as they ensure long-term management provides access to data for reuse in subsequent research.

Storage

The rate of data generation in research is growing exponentially, presenting institutions with a 'grand challenge' to provide adequate and appropriate storage infrastructure at an affordable cost, now and into the future. It is important to recognise that not all research data is of the same 'value' to institutions and researchers, and there may be substantial advantages in classifying data to determine appropriate storage requirements. For example, research data can be moved to less costly 'low tier' storage options once a research project has been completed, with storage costs decreasing for less frequently accessed data.

It is apparent that mechanisms and processes need to be developed that allow institutions to evaluate the status and management requirements of stored data so it can be managed optimally, and that these mechanisms move data to appropriate long-term retention and curation platforms according to a predetermined timeline.

Retention

Management of data beyond the end of a research project's active phase is complex. Both researchers and their institutions are required to store research data on appropriate systems for mandated retention periods, specified through legislation and funding body requirements. Retaining accurate records of research processes, datasets and outputs is integral to defending the integrity of all research findings, allowing published research to be reproduced and verified.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Research data is required to be retained for a minimum of 5 years, and sometimes this is extended to much longer periods (including permanent retention), so access to and management of long-term curation platforms is essential. Effective long-term management of data at scale can be achieved only by implementing automated processes that rely on adequate metadata and persistent identifiers being provided for each dataset.

Disposal

After mandatory retention periods have elapsed, data needs to be consigned to various disposal processes (should the data be reused, destroyed, or retained indefinitely?). At predefined trigger points, data needs to move to an 'end state' (disposal) based on decisions/requirements that have been made by various stakeholders, to fulfil:

- obligations to funding bodies, commercial partners, publishers
- expectations of research communities and reviewers
- regulatory requirements.

Typically, these actions occur many years after the research project has concluded. We need to ensure that all relevant data management information (plans, persistent identifiers, metadata) is carried forward with a project's research data so that the appropriate disposal directions can be applied. Note that disposal processes will differ between institutions, and that advice from institutional cybersecurity experts will need to be sought.

A need for shared information 1: Institutions are encouraged to share with one another examples of predetermined timelines and trigger points appropriate for managing retention and disposal of research data.

Differences In Approach And Need

While all research institutions are required to comply with research data retention legislation and funders' requirements, most have adopted different approaches to managing data and investing in storage infrastructure, and show different stages of 'maturity' in implementation of RDM processes.

Funding of infrastructure (including provision of cloud storage) plays a fundamental role in management of research data. Generally, research-intensive institutions have larger amounts of data to store and manage, but they also attract greater funding to help finance infrastructure, whereas smaller institutions need to provision suitable data storage and retention platforms within a tighter budget. Accordingly, each research institution has its own flavour of storage and retention platforms and its own approaches to moving and managing data across systems. This lack of standardisation makes sharing and reuse of data difficult, which hinders research outcomes and impedes movement of researchers and data between institutions.

The ARDC is enabled by NCRIS

Recommendations And Advice

We propose the adoption of a standardised approach to the storage, retention and disposal of data. The following recommendations are institution-agnostic and provide guidance towards the consistent management of research data within and across institutions.

The key to standardisation in data environments is to embed the use of metadata and persistent identifiers for datasets across systems and research projects. This enables development of connected systems and drives machine-actionable data movements throughout the research data life cycle, and is particularly valuable for long-term data management.

Each institution should adopt a decision-making model that will be used to define and implement the processes (automated and manual) that govern management of data within research projects. The processes and governance should determine the key metadata and identifiers (associated people and project context) that are required.

Identifiers include DOIs for publications and data collection, ORCIDs, RAiDs, field of research (FOR) codes and ROR codes. Use of data standards and standardised identifiers is beneficial to researchers and to administrators, and allows research projects, data and personnel to be managed across multiple systems.

Recommendation 1: Adopt data standards and identifiers that link researchers and research projects.

Internal integration enables communication and alignment of processes and systems, which is vital to the development of automated data and research management functions such as grant administration, research metrics, publications, records management and research data storage.

Note: Using common identifiers and metadata is critical to effective systems integration.

Recommendation 2: Design integration tools for administrative and research management systems.

A need for shared information 2: Institutions are encouraged to share with one another examples of metadata and identifiers for system integration.

Call to action 1: Institutions are encouraged to include data retention requirements in their data classification model, preferably standardised or aligned across institutions.

The ARDC is enabled by NCRIS

ELEMENT 5: RESEARCH DATA RETENTION AND DISPOSAL

Call to action 2: Institutions are encouraged to collaborate to develop a common decision-making tool that can accommodate local (institution-specific) policy.

Recommendation 3: Develop a broadly applicable decision-making tool (matrix, decision tree, rubric) that can accommodate local (institution-specific) policy.

Characterisation of use cases informs local requirements for decision-tool development. In many cases, well-validated use cases can be found in outliers, as they are small in number and known for their specific requirements; for example, large capacity storage uses, or requiring strict access control policies.

A need for shared information 3: Institutions are encouraged to make available to one another use cases that show approaches for making decisions about data retention and disposal.

The capture and propagation of metadata promotes accuracy and utility, and enables appropriate decisions to be made about retention and disposal of data. Combined with unique researcher and project identifiers (recommendation 1), specific metadata should be added to all research data, as early as possible, through multiple touchpoints (for example, grant application, human research ethics application, research clearance, resource provisioning). The metadata should be a single point of capture, and enable machine-actionable processes that drive auto-population across systems and minimise the burden on users (avoiding duplication of inputs). These metadata also facilitate movement of data through storage, retention and archiving platforms.

Recommendation 4: Ensure contemporary capture and propagation of metadata for research projects.

This framework can be extended to the management of physical research objects. Best practice indicates that physical research objects should have digital metadata records created that contain information about the provenance and management requirements of the object, including its physical storage and retention requirements. This metadata record should be stored as part of the research data holding for a project, and be included in decisions about retention and disposal of data.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Resources

- Braxton S et al. Should we keep everything forever (reprised)? Preservation review of research data in repository as an art and a science. Poster. Illinois Library. 2018. https://hdl.handle.net/2142/103221
- Briney KA, Coates H, Goben A. Foundational practices of research data management. Research Ideas and Outcomes. 2020; 6:e56508. https://doi.org/10.3897/rio.6.e56508
- Whyte A, Wilson A. How to appraise & select research data for curation. Digital Curation Centre, Australian National Data Service. 2010. https://www.dcc.ac.uk/sites/default/files/documents/How%20to%20Appraise%20and%20Select%20Research%20Data.pdf
- Beagrie N. What to keep: a Jisc research data study. Jisc. 2019. https://repository.jisc.ac.uk/7262/1/JR0100_WHAT_RESEARCH_DATA_TO_KEEP_FEB2019_v5_WEB.pdf
- Dorey J, Hurley G, Knazook B. Appraisal guidance for the preservation of research data. Zenodo; 2022. https://doi.org/10.5281/zenodo.5942236
- Kung JYC, Campbell S. What not to keep: not all data has future research value. Journal of the Canadian Health Libraries Association / Journal De l'Association Des bibliothèques De La Santé Du Canada. 2016;37(2). https://doi.org/10.5596/c16-013
- Natural and Environmental Research Council (NERC) Data Value checklist. 19 August 2022. https://www.ukri.org/publications/nerc-data-value-checklist/
- Shabou BM et al. Algorithmic methods to explore the automation of the appraisal of structured and unstructured digital data. Journal of Records Management. 2020;30(2). http://dx.doi.org/10.1108/RMJ-09-2019-0049
- University of Cambridge PrePARe project team. Selecting what to keep and what to bin (Checklist). Cambridge University Library. 2012. https://www.repository.cam.ac.uk/handle/1810/243754
- University of Reading. Data selection and appraisal checklist. 2016. https://www.reading.ac.uk/research-services/-/media/project/functions/research-and-enterprise-services/documents/rda-selectionappraisalchecklist.pdf?la=en&hash=0C286C721876278DC7DC00DA8EAC4F5A

The ARDC is enabled by NCRIS

ELEMENT 6: OPEN RESEARCH AND DATA PUBLICATION

Summary

This element provides institutions with guidance to improve research data discoverability and impact, in line with increasing demands from funders and governments for greater openness and data reuse, and for making research data FAIR[4] to both humans and machines to achieve a greater return on investment. We acknowledge that barriers exist to adopting open research, data publication and FAIR approaches, and provide recommendations on open research statements; incentives for making data FAIR, if not fully open; and metrics for tracking research data impact. We explore infrastructure considerations, outlining repository features that best enable data publication, discoverability and reproducibility and we provide practical advice on future-proofing RDM to enable institutions to respond to the movement towards increasingly open and FAIR research data.

Why Open Research And Data Publication Are Important

Funders and major research bodies worldwide are increasingly supporting an open research model (Plan S[55], UNESCO Recommendation on Open Science[56], for example). Even previously intransigent publishers are beginning to negotiate transformative agreements that will make a lot of new research more openly available. Therefore, any approach to RDM should be future-proofed by providing pathways for institutions to develop an open research culture and practice that supports FAIR data principles.

The higher concept of open research comprises not only access to publications and data sharing, but also the documentation of data and research processes to allow research results to be assessed, tested and reproduced. Reproducibility is key.

- Open research metadata includes record-, dataset-, and project-level metadata; administrative metadata; and data provenance, including version history.
- Open research processes include ‘paradata’, such as data-capture and processing techniques and methods, software, notebooks and samples.

Both metadata and paradata are required to enable sharing that is ‘as open as possible, [only] as closed as necessary’. Open research more broadly includes the research process itself and the use of open tools, including open software, to improve access, understanding and reproducibility.

Publication of data and data sharing, and of FAIR data, while increasingly commonplace, does not necessitate that those data are fully ‘open’. For example, there may be no clear licensing of data being shared which means reuse would be problematic. Shared data may also lack the metadata to allow the data to be understood, reused or usefully interrogated. There is also the risk that online publication venues are vulnerable to a range of cybersecurity threats. In some cases, university repositories are taking responsibility for the creation and maintenance of the data record which includes the requisite “wrappers” (metadata) that enable data sharing and reuse.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

The focus of Research Data Management Framework for Institutions discussions has been on publication of research data and data repositories. We acknowledge that guidance about data-sharing policies, practice and interpretation at institutional and national levels is also required, as is coverage of what institutional and national infrastructure exists to facilitate research and data sharing, and what might still be needed. Our recommendations point towards ideal solutions, to focus institutions towards emerging and future practice.

Common Drivers Of Open Research For Researchers

- The Code, particularly the guide for the *Management of data and information in research*[49] and the NHMRC's *National statement on ethical conduct in human research*[50], increasingly expect research-data metadata to be made openly available within a specific timeframe. Researchers who do not comply may find themselves disadvantaged in future grant rounds.
- Publishers increasingly expect underpinning data to have a persistent identifier and to be made available (either directly, or via a link to the repository where the data is housed) when manuscripts are submitted. Some publishers will not accept papers where data cannot be provided at lodgement to support conclusions. Researchers who are unable to meet these requirements will increasingly have submissions rejected by major publishers.
- Funders and governments are increasingly demanding greater openness and data reuse to ensure a greater return on their investment and to enable impact measures.

Common Drivers Of Open Research For Institutions

- To drive cultural change and improve research discoverability and impact, some universities are creating and endorsing open-research statements (Griffith University[57], the University of Melbourne Library[58]).
- Universities have strict responsibilities under the Code for the management of data assets for which they have received grant funding. If only for risk-management purposes, institutions need to establish robust and appropriate data-stewardship standards for the management of those assets in partnership with the researchers who created them.
- Under the Code, research data controlled by the institution and/or its researchers should be stored in facilities provided by or approved by the institution. Research data may be published in international, national, or discipline-based repositories such as international data banks, in addition to institution-based storage or archiving. Institutions are also advised to maintain a record of the research data generated by their researchers and where it is stored.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Key Challenges To Open Research And Making Data Fair

The barriers to open publication, data sharing, open research and to making data FAIR are significant. They include:

- lack of guidance pointing to optimal approaches at both the national and institutional level on pathways/models/goals for open research
- no clear licensing for a vast range of existing research data, and uncertainties about data ownership
- lack of metadata, rendering available data less useful and less reusable
- lack of institutional support, advice and training on 'best practice' open research or FAIR approaches
- lack of suitable publication venues for sharing and publishing data
- restrictive ethics conditions precluding data sharing and open research in many disciplines
- data sensitivity precluding data sharing and open research in many disciplines
- lack of awareness of existing national infrastructure facilitating open research and data sharing
- lack of incentives for managing data well, without which meeting FAIR or open-research requirements is difficult
- lack of incentives to practise open research, including a lack of recognition of the effort by universities and funders
- commercial constraints, such as patents and intellectual property (IP) ownership by industry funders
- inconsistent IP policies in individual institutions
- lack of clarity on who drives this for institutions; for example, who provides academic leadership, who leads institutional change, who provides day-to-day support
- lack of researcher skills and knowledge in open-source alternatives to proprietary software, which results in approaches not conducive to open research practice
- third-party data providers having explicit licensing, use, storage, and destruction requirements that preclude data sharing and open research
- legal regulation requirements in different jurisdictions; for example, issues that prevent data linkage at a national scale because of different state-based privacy laws and legislation, such as the *Defence Trade Controls Act 2012*
- lack of participant consent to data sharing in human research
- lack of knowledge about appropriate data management planning practices
- lack of incentives to publish or share data
- lack of awareness of existing institutional infrastructure facilitating open research and data sharing.

The ARDC is enabled by NCRIS

Differences In Approach And Need

A one-size-fits-all approach to open research and publication is not appropriate as institutions differ markedly in size, scale and funding. The ability of an institution to develop a future-focused data-sharing, FAIR and open research culture depends on:

- adequacy of local advice, training and support within the institution
- skills and knowledge of researchers and research support staff
- researchers' awareness of local systems and processes that facilitate sharing, including local data management systems
- the existence of an institutional repository for sharing data
- flexibility and adaptability of the local research culture
- existing data and research policies and frameworks
- adequacy of supporting infrastructure for data sharing, such as metadata capture capabilities and access controls
- local incentives, which do not need to be financial.

Recommendations And Advice

Open-Research Statement

Recommendation 1: Adopt an open-research or FAIR data statement as a driver for cultural change on sharing data/research, and enrol influential academics as champions of change.

The Council of Australian University Librarians has a useful Statement on Open Scholarship[59] which includes calls to action on advocacy, infrastructure and training.

Best Practice and Incentives

Recommendation 2: Adopt best practices for data management, with open research, FAIR data and data sharing as end aspirations, where appropriate.

The ARDC is enabled by NCRIS

Ideas for incentives for FAIR and open data include:

- an internal grants round open only to those practising open research or making their data FAIR
- institutional support for making data open or FAIR – this might involve storage provision; consulting/advisory support; advice with data clean-up/de-identification/other requirements; help with metadata creation; repository selection advice and assistance with deposit
- institutional support for open-source alternatives to proprietary software as this will help support open research
- recognition of open research or FAIR data practices in university promotion rounds
- public relations promotion of academics practising open research to help demonstrate the societal impact of their research
- tenure prioritisation for academics who practise open research
- financial support for article processing charges for researchers practising open research or making data FAIR
- development of ORCID profile metadata feeds for publications and datasets
- development of standardised metrics for open data publishing
- recognition of open research and FAIR data practices in rankings, such as Excellence in Research Australia rankings, to encourage institutions to support/encourage open research and sharing practice.

Metrics for Research Data

Metrics can be set at different levels:

- Journal level – the impact of a journal within a field
- Article level – the comparative impact of a specific article / data artefact within a field
- Author level – the impact of an author's publications within a field.

There needs to be more consideration and development of uniform metrics that are unrelated to journal article publication and that can be applied consistently across disciplines.

To be useful to open research and open data purposes, research metrics need to be expanded to include metrics for data and data reuse, citations for physical samples and research software, metrics around the adoption of FAIR principles, and also the 'real world' impact – for example, the release of an important climate or health dataset and how that release both drives and underpins future research. Currently, vendor subscription-based tools that manage data metrics (citations) are expensive third-party subscription products that are niche- and discipline-dependent, and are therefore not appropriate for all researchers.

The ARDC is enabled by NCRIS

Additional data metrics could include:

- number of research data downloads
- number of publications citing the data
- citations of the data outside of traditional publications
- citations that count the use of the data in teaching, as this is an area that is rapidly growing and that needs recognition
- adoption/use of the data in policy development.

Recommendation 3: Create and adopt uniform research data metrics as a mechanism for measuring the impact of sharing data and research.

A need for shared information: Institutions are encouraged to share with one another examples of metrics for data, for physical samples, for research software, metrics around adoption of FAIR principles, and metrics for the 'real world' impact of research.

Data citation practices — whether for data or the physical samples used in research, or the research software used — are still very immature, and are most certainly not standard procedure for most researchers when writing up their research. References to data or samples, when they appear at all, are covered in the text rather than in the list of references accompanying the work. Data or samples that do happen to have their own persistent identifiers and are listed in the references can be found only by searches or API calls, which means that a lot of potentially reusable data is being missed.

Wish list for implementing data metrics

- An automated mechanism to use standard identifiers such as DOIs / persistent identifiers to enable tracking of data citations.
- A call to action for publishers and downstream services such as CrossRef to surface any occurrence of a persistent identifier. This could at first be limited to DOIs, but should be expanded to include a growing list of globally acknowledged unique persistent identifiers such as RORs and RAiDs.
- A means for services to more easily generate COUNTER[60]-compliant usage metrics for research data, such as the IRUS UK service61. COUNTER seeks to ensure consistent and credible reporting of research data usage, which is important for understanding how data is being reused.
- Metrics relating to alignment of repository datasets to the FAIR data and CARE principles, made visible to curators of the data (for asset management purposes) but also to researchers who are looking for data to reuse.
- Links between dataset publication metrics and researcher profiles, just like any other researcher outputs. Currently, DOIs for publications typically come from CrossRef; whereas DOIs for data typically come from DataCite. Perhaps a means to link CrossRef and DataCite could be developed.

The ARDC is enabled by NCRIS

Solutions For Publishing Data

Greater sharing and openness on a widespread scale can only occur when processes for researchers are clearly stated and supported by local policy, and where researchers can adequately access advice, support and training to enable open-data, data-sharing and FAIR practices. This necessitates the provision of adequate, well-managed and suitable data storage by the institution. In some petascale data cases — for example, in astronomy, climate and geophysics — data cannot be stored at a single institution, so a caveat addressing these large data challenges needs to be factored into the framework.

The optimal solution would be storage in a data repository that provides adequate descriptive, structural and technical metadata, clear licensing and reuse conditions, and a permanent identifier such as a DOI to foster data sharing. The repository would need strong high-level commitment with associated adequate and sustained funding to guarantee long-term storage for data.

Repositories have clear advantages over other methods for sharing data as they are institutionally backed and governed, provide the kind of structured metadata required for efficient reuse of data, often manage data curation/preservation, and have support staff to assist researchers with usage.

Selecting A Repository

Institutions should encourage the publication of data via appropriate repositories. Where there is no local institutional repository to provide this service, robust disciplinary repositories such as the Australian Data Archive[62] could be used. International recommendations are to use discipline-based repositories first, before considering institutional, particularly where there is a data volume issue, for example with petascale datasets. Disciplinary repositories also offer discipline-based quality assurance / quality control on the content of the data, which very few institutional repositories can offer.

Where researchers habitually deposit in external repositories rather than in an institutional one, institutions should investigate how to harvest at least the metadata for any datasets or dataset records published to these repositories (for example, Zenodo, Mendeley, Dryad[63]) to provide a local record of the research. Institutions could also offer guidance to researchers about choosing a suitable data repository so that researchers find the optimal solution for their own needs.

However, as data volumes continue to grow exponentially, the issue of long-term, discipline-focused storage infrastructure needs to be addressed at a national level.

The following is a wish list for data-sharing platforms (the ARDC Data Quality Interest Group[64] provides advice in this area):

- Data quality reviews of published data and published metadata should occur, if possible. (At the level of the institution, discipline-specific quality reviews will be hard to resource, but could involve departments/colleges participating in peer review of data, the way they already do for papers and grant applications).
- High quality curation in data repositories should be preferred (a minimum standard needs to be agreed).

The ARDC is enabled by NCRIS

- Minting and/or publication of all persistent identifiers (such as DOIs, International Geo Sample Numbers (IGSN), ORCIDs, RAiDs, ROR codes, handles) and licences should be accommodated. All items within compound objects should be identified, possibly with individual identifiers.

Ideally, repository features should include:

- a statement that covers data curation, preservation and repository required compliance around data deposited (for example, appropriate sharing must be supported by research ethics)
- minimum community-agreed profiles of metadata standards to foster data discovery; for example, the ARDC Registry Interchange Format – Collections and Services Schema (RIF-CS) or the Australian Government agreed profile of the Geographic Information – Metadata Standard (ISO-19115)
- standardised data citation
- citation counts that meet a standard for reporting; for example, COUNTER[60]-compliant
- globally unique persistent resolvable identifiers
- continuity/viability of service
- cybersecurity of service
- data integrity checks i.e. checksum
- storage size / file size limits
- physical location, for example Australia, offshore
- licence options
- documented storage procedures such as storage provisioning and operations in place, leveraging of existing storage operations, recovery procedures in place, technology and hardware migration planning
- deposit agreement (i.e. correct permissions)
- term of deposit/retention periods
- funding source; for example, for data affected by ARC/NHMRC policy
- a range of access controls, including the availability of mediated access, which would also include levels of access; for example, access to summary metadata only
- documented methods of versioning data
- functionality for open peer review to help improve quality of data
- CoreTrustSeal certification[65].

Recommendation 4: Implement an institutional data repository with metadata capture capabilities.

Solutions To Discourage

Publishing Data On A Website

Publishing data on an institutional web page is not recommended. Institutions frequently go through reorganisation and restructures, which means the data's home may be ephemeral. The data will also lack metadata, and a website location is far less discoverable than a more structured service such as a repository.

A personal website is even less preferred as a publication venue. Not only is it likely to be ephemeral, lack metadata, and have the same difficulties of discovery as an institutional website, but it will be much more prone to cyber-insecurity, given its location in a commercial space.

Data publication on a website will not meet most journal submission requirements for data that supports a publication to be stored in a repository and linked via a unique persistent identifier.

Depositing Data With The Publisher As Part Of Publication

While this option may be superior to a website as a publication and sharing venue, the risk is also great that the data could be removed or lost. If data does need to be deposited with a publisher, the researcher should have other arrangements to store data long term and securely (publisher-agreement dependent). Also, increasingly, many journals will no longer accept data as supplements.

Depositing Data In A Preprint Archive Along With The Preprint

While this option may be superior to a website as a venue, the risk is also great that the data could be removed or lost. If data does need to be deposited along with a preprint, the researcher should have other arrangements to store it long term and securely.

Future-Proofing For Open Data

Here we provide practical actions that an institution can take now to enable them to better keep up with the movement towards increasingly open data.

Recommendation 5: Future-proof for open data.

- Embed data-sharing / FAIR data practices in institutional RDM policy.
- Embed sharing practice in ethics applications to foster future data sharing:
  - Encourage non-specific and extended consent with ethics committees so that data reuse can encompass wider research use than the original research idea. Ethics applications need to be future-proofed to enable data reuse. Currently, there is little understanding about the type of consent sought and what implications that has for sharing data later. Researchers need to better understand the pathways they create for themselves from the outset.
  - Store/capture ethics approvals within dataset metadata.
- Teach greater awareness of data licensing and the different licences. *The Open Data Handbook*[66] by the Open Knowledge Foundation provides a list of conformant licences[67.] The ARDC *Research data rights management guide*[68] is also comprehensive.
- To deter misuse, publish all data with a licence and a recommended citation. This includes the licensing of commercial data-compilation mechanisms, such as commercially controlled psychological tests and data visualisation software.
- As each licence, agreement or contract is potentially unique, it might be beneficial for the researchers to see a visible snapshot of what they can and cannot do with their data on, for example, a dashboard or even included in the DMP. Make it possible for MOUs, data-sharing agreements, and other agreements and contracts to be saved together alongside the research data.
- Establish a desirable baseline of requirements for data-sharing platforms.
- Automate a plain-English, decision-tree-style workflow to help researchers identify an appropriate licence for their data.
- Negotiate data-sharing agreements early in the project as these are critical to resolving ownership, IP and the rights to publish and reuse.
- Take appropriate measures to protect data from inappropriate loss, theft and misuse of access.
  - Accidental publication of sensitive data
    - Mitigation: Store the data classification as part of the metadata. Consider processes to check for sensitive data before publication (for example, automation and through APIs).
  - Misuse of data
    - Mitigation: Publish data with a licence and a recommended citation to deter misuse.
  - Theft of data
    - Mitigation: Publish data with a licence and a recommended citation to deter theft and provide proof of ownership and priority. Use existing systems to report offenders; for example, notifying publishers and funders of the theft.
- Develop guidance on how to confirm whether data is open/FAIR to encourage researchers to use open research data.

Resources

- ARDC Citation and identifiers: https://ardc.edu.au/resources/working-with-data/citation-identifiers/
- ARDC DataCite DOI service: https://ardc.edu.au/services/identifier/doi/
- ARDC FAIR data guidelines for project outputs: https://ardc.edu.au/resource/fair-data-guidelines-for-project-data-outputs/
- ARDC FAIR Self assessment tool: https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/
- ARDC IGSN minting service for physical (material) samples: https://documentation.ardc.edu.au/display/DOC/IGSN+Service
- ARDC Research data rights management guide: https://doi.org/10.5281/zenodo.5091579
- Australian code for the responsible conduct of research, 2018; in particular, the guide: Management of data and information in research: https://www.nhmrc.gov.au/sites/default/files/documents/attachments/Management-of-Data-and-Information-in-Research.pdf
- Beijing Declaration on Research Data: https://www.codata.org/uploads/Beijing%20Declaration-19-11-07-FINAL.pdf
- CoreTrustSeal Trustworthy Data Repositories Requirements: Glossary 2020–2022: https://zenodo.org/record/3632563
- CSIRO 5-star data rating tool: https://data.csiro.au/collections/collection/CIcsiro:27133v5/DItrue
- IRUS-ANZ: https://irus.jisc.ac.uk/irus-anz/portal/
- Joint FORCE11 & COPE Research Data Publishing Ethics Working Group Recommendations: https://zenodo.org/record/5391293
- Klump J, Wyborn L, Wu M, Martin J, Downs RR, Asmi A. Versioning data is about more than revisions: a conceptual framework and proposed principles. Data Science Journal, 2021;20(1), p.12. DOI: http://doi.org/10.5334/dsj-2021-012
- Lowenberg D, Chodacki J, Fenner M, Kemp J, Jones MB. Open data metrics: lighting the fire (Version 1). Zenodo. 2019: http://doi.org/10.5281/zenodo.3525349
- Mozersky J, Walsh H, Parsons M, McIntosh T, Baldwin K, DuBois JM. Are we ready to share qualitative research data? Knowledge and preparedness among qualitative researchers, IRB Members, and data repository curators. IASSIST Q, 2020;43(4). https://doi.org/10.29173/iq952
- Nasir JA, McCrae JP. iLOD: InterPlanetary File System based Linked Open Data Cloud. Paper presented at the MEPDaW@ ISWC. 2020. Retrieved from: http://john.mccr.ae/papers/nasir2020ilod.pdf
- Open Data Handbook Glossary: http://opendatahandbook.org/glossary/en/
- Parsons MA, Fox PA. Is data publication the right metaphor? Data Science Journal. 2013;12, WDS32–WDS46. https://doi.org/10.2481/dsj.WDS-042

The ARDC is enabled by NCRIS

- Pasquetto IV, Borgman CL, Wofford MF. Uses and reuses of scientific data: the data creators' advantage. Harvard Data Science Review. 2019;1(2). https://doi.org/10.1162/99608f92.fc14bf2d
- Research Data Alliance. Data usage metrics WG recommendations. 8 March 2021. https://www.rd-alliance.org/group/data-usage-metrics-wg/outcomes/rda-data-usage-metrics-wg-recommendations
- Research Data Alliance interest groups: https://rd-alliance.org/groups/interest-groups
- Sorbonne Declaration on Research Data Rights: https://sorbonnedatadeclaration.eu/
- The Open Research Toolkit: http://bit.ly/open-research-toolkit

The ARDC is enabled by NCRIS

ELEMENT 7: SENSITIVE RESEARCH DATA

Summary

This element provides institutions with guidance to ensure compliance with ethical and regulatory requirements, and manage risks associated with sensitive data — such as risk to individuals, risk to environments, risk to corporate interests, and legal and reputational risk to the institution. We recommend the adoption of a "risk and consequence" sensitive-data classification scheme aligned to infrastructure platforms, with appropriate protections for the different data-classification levels. We provide an initial step towards consistency of data classification between institutions. We also touch on the requirement for considering ethics, and the desirability of aligned sensitive-data classifications for the sector.

Why Sensitive Research Data Is Important

Sensitive data can be difficult to define, making it hard to have meaningful discussions about how to manage sensitive research data. We can identify classes of data that are definitely sensitive – for instance, individual patient records, information about the nesting location of an endangered species of bird, or test results that reveal the most effective way to produce a microorganism that could be used as a bioweapon. However, it is harder to agree upon the properties that cause all of these types of data to be grouped into the "sensitive" category. To sensibly address this area, it is helpful to consider our aims when discussing RDM for sensitive data.

For the purposes of institutional RDM, all of the types of sensitive data discussed above are united by the need for special protections and controls to be put in place around their management, including the collection, storage, access and sharing of those data. Institutions are motivated to address the sensitivity of data to ensure compliance with ethical and regulatory requirements and manage risk. Data can pose many kinds of risk if mishandled or exposed (including risk to individuals, risk to environments, risk to corporate interests, risk to national security, and legal and reputational risk to the institution).

Depending on the kind and severity of the risk posed, different levels of protection and control are appropriate. For this reason, it can be useful to think of sensitivity as a dimensional quality of a dataset rather than a category. There are costs associated with placing protections on data, both financial (more secure infrastructure can be more expensive to buy and support) and in terms of ease of access and use, so it is not sensible to apply the same protections to all data. More-sensitive data requires very high levels of protection that are not required for less-sensitive data. The degree of sensitivity reflects the severity of the risk posed by the mishandling or exposure of those data.

Viewing sensitivity as a dimensional quality of research data is also helpful when sharing sensitive data for reuse. By its nature, the dissemination of sensitive data must be controlled. But at the same time, the Code, the NHMRC *National statement on ethical conduct in human research*[49] publishers, and funders all encourage data sharing, particularly in line with the FAIR data and CARE principles[4,5]. The level of risk posed by the exposure of a particular dataset can be weighed against the benefits of its reuse when determining whether it should be shared and under what conditions.

The classification of data sensitivity has therefore been identified as a vital component of institutional RDM. Data sensitivity classification schemes set out different levels of data sensitivity, defined according to some set of characteristics of the data. The sensitivity of data according to the classification can be recorded in its metadata. Different protections and handling procedures can be defined for each level.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Sensitivity classification schemes have a variety of uses within an institution. They can be used to indicate to researchers what platforms are appropriate for storing and analysing particular kinds of data. They may be usefully referred to in policy and procedure which lays out the governance and handling requirements for sensitive data. They may be used by ethics committees to ensure that appropriate controls are in place to manage risks present in the data. They may also be used by IT staff to check that storage systems do not contain data that is too sensitive for those systems, and to understand what controls need to be in place when undertaking operations such as migrating data.

A long-term goal is for universities to achieve a degree of consistency in classification levels, and lay out the protections required and protections currently in practice at those classification levels. A degree of consistency could simplify decision-making around RDM infrastructure and platforms, simplify the use of infrastructure for collaboration, and aid in the movement, sharing and transfer of research data between institutions. Recognition of specific protection levels could further advance trust in the public research sector generally.

Our guidance focuses on the classification of data sensitivity, but this is only one aspect of the management of sensitive data. Other topics include:

- the provision of handling guidelines for researchers (such as information on how to reduce the identifiability of data)
- governance arrangements for sensitive data (including assessing data requests, for instance by using the Five Safes model[72])
- methods for sharing sensitive data safely with external parties
- guidance for working with Indigenous data (such as the use of Traditional Knowledge labels).

Guidance on these topics would be a valuable addition to any future versions of this framework.

Differences In Approach And Need

The Benefits Of Classifying Sensitive Data

Some universities have no formal data-sensitivity classification scheme. Instead, they may provide guidance to researchers about the appropriate protections (and therefore platform choices) through other means, such as within RDM planning tools. Alternatively, an institution may simply use their enterprise information management regulations to assess the protections required for research data.

Choosing not to have a data-sensitivity classification scheme avoids the effort required to develop, introduce and maintain such a scheme. However, the benefits to having a classification scheme should be weighed against this reduced effort. For instance, a classification scheme allows the sensitivity of a dataset to be clearly recorded within its metadata, making it easier to ensure that the data is appropriately managed and protected. A classification scheme also allows universities to provide broad advice about the protections that are required and the platforms that are supported for data of a given level of sensitivity. This is especially useful for universities operating at a scale where one-on-one advice for every research project is not feasible. Further, it would be easier to assess and work towards increasing consistency in the management of sensitive data across universities if they were to use comparable classification schemes.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

For these reasons we encourage the development or adoption of a data-sensitivity classification scheme (following the principles outlined below).

We also recommend that this scheme be developed with consideration of the specific needs of research data (although it may be beneficial to align it with an existing enterprise information management classification scheme). Research data is less uniform and has a different risk profile and user requirements than enterprise data. A classification scheme that was developed specifically for enterprise data is unlikely to be fit for purpose when applied as is to research data.

Recommendation 1: Adopt a classification scheme specific to sensitive data.

Factors Determining The Protection Required

Universities vary in the specific protections required for data of a given level of sensitivity. These differences are driven by several factors, including:

Location: The regulatory environment surrounding the management of sensitive data varies from state to state. For example, some states require that particular forms of sensitive data be stored onshore, while others require that the same data be stored within the state.

Available platforms: Universities may be reluctant to impose requirements on researchers which they know cannot be met by the current suite of supported platforms. Additionally, depending on the current infrastructure at an institution, a given level of protection may be considerably more costly in terms of both resourcing for the institution and burden of effort or restricted access on the researcher. These differences will affect the balancing of risk mitigation against cost, and can result in different decisions about the appropriateness of particular protections between universities.

Advice of legal and cybersecurity departments: An institution's understanding of both the regulatory environment that constrains the management of sensitive data and the current best practice for protecting those data depends on expert advice and interpretation. Universities have legal departments and cybersecurity teams to provide this advice and interpretation, and the advice of these groups may differ between universities. How an institution responds to this advice can also depend upon its risk appetite.

Recommendation 2: Define protections required for recognised classification levels of sensitive data.

Recommendations and Advice

Developing Or Adopting A Classification Scheme

Gain Clear Institutional Recognition And Support

The adoption of research data classification will be a significant change to institutional policy or procedures. Element 2: Culture Change of this framework recommends that those leading the change should engage early with research leaders (for example, the DVCR, academic division leaders) to ensure that they are fully aware of and support the objectives.

Academic institutions can have limited capacity for dedicated work to be done without additional staff and funding. Recognition may be required of the appropriate resourcing and, potentially, funding for research data infrastructure if they are not already in place.

Classify Only Data Sensitivity

Universities may require that their research data be classified on many dimensions. For instance, it may be important to record information about the quality of data to inform reuse, or the value of the data to inform retention and disposal decisions. However, it is important to classify these different dimensions separately, as combining too many types of information into a given classification scheme results in classifications that are more complex and less useful than they need to be. We recommend that the data-sensitivity classification scheme reflects only the sensitivity of data, and that other, orthogonal aspects of the data be classified separately.

Classify Based On Risk/Consequences

A potential pitfall is to attempt to define classification levels in terms of the types of information contained within the data (for example, “data containing personally identifying information”). The problem with this type of classification scheme is that data can be sensitive in many ways, and the sensitivity of a particular type of information may vary depending on contextual factors and over time. This can lead to very complex definitions of the classification levels, which need continuous updating. We therefore recommend that the levels instead be defined by the severity of risk/consequences of mishandling or exposure of the data.

Recommendation 3: Define sensitive-data classification levels by the severity of risk/consequences of mishandling or exposure of the data.

The ARDC is enabled by NCRIS

Have A Small But Comprehensive Set Of Levels

To ensure broad engagement and applicability, we recommend the following:

- Use an ordinal structure, from the lowest to highest risk/consequences/protections.
- Keep it simple, with relatively few levels (many examples have between 3 and 5). This will make the communications and underlying infrastructure planning as simple as possible. There is a trade-off between granularity and simplicity – a very complex scheme will allow for fine-tuning of requirements, but if researchers and infrastructure providers find it to be too complex then uptake may be affected.
- Consider how all levels of sensitivity (from open to extreme risk) will be approached. Ideally, the full range of levels from open to extreme protection should be covered within the scheme, so that all data can be given a classification. However, many universities are unable to meet the protection requirements of the most high-risk data. Some universities include an "extreme" classification for these data and acknowledge that no local systems are available to provide appropriate protections; others do not include this level, and instead treat data that cannot be managed locally as an edge case.
- Avoid labelling or defining the levels using language that matches exactly or clashes with other classifications to avoid confusion. A university adopting a scheme from another institution should adjust the wording to avoid clashes that could occur when the scheme is transposed to their local context. Some institutions have adopted colour-coded levels to avoid difficulties with verbal labels.
- Consider and test audience comprehension when defining classification levels. The same scheme may be used to classify many different types of data across an institution, particularly if the same scheme is used for both administrative and research data. It is important to ensure that the definitions can be interpreted for all types of data that must be classified under the scheme. For this reason, we discourage the use of examples in high-level descriptions of the scheme, as they are likely to cause confusion (for instance, researchers may find it hard to understand how the scheme applies to research data when reading examples of the classification of administrative data). Examples can be provided in guidance tailored to, and clearly marked for, specific audiences.
- To ensure the scheme is comprehensive and useful, we recommend aligning the classification levels to research and business requirements, and assessing these against existing research services and infrastructure.

Consider Mapping To The Institution's Information Classification Scheme

Universities will generally have an existing information classification scheme for administrative data. The relationship between classification schemes for administrative data and research data should be carefully considered. It may be helpful to develop the classification scheme for research data from that used for administrative data, or to develop a single scheme that can be applied to both contexts. The benefits of closely aligned or shared classification schemes include increased priority with leadership, access to a broader pool of resources for developing the classification scheme, and improved understanding between business units. If a combined scheme is used, special attention should be paid to ensuring that researchers understand how to apply the scheme to research data.

The ARDC is enabled by NCRIS

Universities have found it useful to have a simple explanation of the relationship between data classification schemes and any overarching institutional information governance, policy or procedure (and potentially with important external information governance, such as state government classification schemes). This could outline any agreed relationship by design; relationships between responsibilities; levels; information types; activities; or risks outlined in each. This may be most useful for engaging internal stakeholders such as central records, information management or governance groups, IT or cybersecurity.

Supporting Use Of The Classification Scheme

Work With / Provide Guidance To Researchers To Classify Risk

Classifying the sensitivity of data requires the involvement of the researchers who understand the data. However, researchers are likely to need assistance to properly interpret the classification scheme and correctly classify their data.

Both under- and over-classification of data sensitivity can be detrimental. Under-classified data will not be given the appropriate protections, increasing the risk of exposure. Over-classification can lead to data being locked down to a point that access and reuse are obstructed. Researchers (and particularly HDR students) may be over-cautious in classifying their data early in a project, when they do not yet have a fully developed understanding of that data. Less-sensitive derivatives of a dataset may also be developed over the course of a project. It is therefore important that there is sufficient flexibility in the system for classifications to be revised if the researcher later deems it appropriate.

Universities may provide assistance with classification to researchers in different ways. Where there is sufficient resourcing, it may be possible to provide one-on-one assistance to the researcher. If the institution operates at a scale where this is not possible, assistance may come in the form of written guidance or interactive classification tools such as surveys or decision trees.

Element 8, Support, Training and Guidance, has more information about approaches.

Because data can be sensitive in so many ways, it is not possible to determine based on a checklist approach that a dataset is not sensitive. However, a checklist (or more sophisticated tool) may help researchers to understand the aspects of the data that could lead to increased sensitivity. Additionally, some properties almost guarantee that a dataset has a high level of sensitivity and these are worth identifying, depending on the institution's research profile; for instance, data that contains health information about identifiable individuals. Identifying these properties helps to ensure that the majority of sensitive datasets are correctly classified, though it is important to be clear that such a list cannot be comprehensive, and that the researcher should consider other potential risks and consequences of exposure of their data.

Identify The Required Protections For Each Sensitivity Level And Appropriate Platforms To Provide These Protections

Once the sensitivity levels are identified, the university should determine the minimum protections that should be applied for data at each level. These protections should ensure that regulatory requirements are met and should be proportionate to the risk present in the data. Unnecessarily high protections are likely to be both more expensive to run and harder to access, impeding research work.

ELEMENT 7: SENSITIVE RESEARCH DATA

There should be clear guidance that points researchers to the appropriate institution-supplied (or approved external) platforms that meet the protection requirements for each sensitivity level. Researchers cannot be expected to be cybersecurity experts, and may not be able to assess the appropriateness of a given platform against the requirements without expert assistance. However, we do recommend that the protections required at each sensitivity level are clearly stated. This will help more technically able researchers to contribute to the design of appropriate solutions for edge-case projects, and may also assist with compliance by explaining to researchers why they are restricted to certain platforms for certain types of data.

Recommendation 4: Provide guidance for researchers on the protections afforded through endorsed infrastructure solutions for sensitive data.

Put Procedures In Place For Managing Edge-Case Requests

It is not feasible for universities to provide platforms that cover all use cases for all levels of data sensitivity (see Element 1: Active Research Data Management). It is helpful to have procedures in place for assessing alternative solutions for edge cases, and to understand whose responsibility it is to approve these solutions. Edge cases can be data with unusual requirements, or platforms that are not within the suite provided or approved by the institution.

If the sensitivity of the data has been classified (and any additional requirements outlined), then an expert team such as cybersecurity can assess new platforms against the required protections. However, this task takes time and effort, and if the platform is to be used for an extended period of time this assessment may need to be repeated to check for changes that would affect the appropriateness of the platform (for instance, a commercial provider may relocate their data storage from Australia to offshore). For this reason, there needs to be a good case for why existing solutions are not sufficient.

In some cases, data may require specific protections that the institution cannot provide in a cost-effective way. In this case, the cost of providing the appropriate platform may need to be recovered from the research project. Alternatively, project partners may be able to provide the necessary platform; for instance, if the Department of Defence places certain requirements on the storage of their data, then it may be more appropriate for researchers to access that data from within the department's own environment.

Assist Researchers In Securing The End-Points

A significant challenge in risk management related to sensitive data (and all cybersecurity) is "securing the end-points". Many data breaches are a result of human error or preventable activity; for example, theft of laptops or other devices offsite. For this reason, a more comprehensive risk mitigation strategy should be complemented by training and upskilling, guiding researcher practice and instilling good practices in research students. It may be helpful to seek general advice from the IT team that manages desktop environments. Some institutions already provide advice on how to better secure personal devices not managed by the institution. If not, consider the potential opportunity for existing IT support to develop and maintain guidance that is more research-focused.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Provide Guidance For Ethics Committees

Not all projects dealing with sensitive data go through an ethics committee. However, for those that do, the ethics committee will want to know that the data is being handled in a way that minimises risk. A clear data-sensitivity classification scheme will help them assess whether the planned RDM protocols are appropriate for the level of sensitivity.

For edge cases, it may be harder for ethics committees to assess whether the planned RDM solutions will be appropriate. For this reason, ethics committees should have access to expert guidance; for instance, by including an RDM expert on the committee, or by giving the committee the option to request an assessment by institutional IT or cybersecurity.

Recommendation 5: Provide guidance for ethics committees on appropriate approaches for managing sensitive research data.

Building Towards Alignment In The Sector

Align Classification Schemes

Aligning classification schemes across the sector will make it easier for universities to share and collaborate on sensitive datasets. If universities follow the design guidelines above, then it should be possible to find the equivalent levels of sensitivity across different schemes. The draft Classification Crosswalk[69] between different classification schemes demonstrates this. An initial step towards the goal of cross-institutional alignment, the classification crosswalk is an analysis of the typical structure of risk- and protection-based classification levels across existing examples. For those institutions embarking on classifying sensitive data, it could provide some quick insights.

One approach to building towards alignment is to adopt an existing classification scheme from another institution. When taking this approach, an institution may need to make adjustments to fit the local context; for instance, to change language to bring it into line with other institutional guidance.

Call to action 1: Research institutions should continue adding to and developing the draft sensitive-data classification-level crosswalk.

Call to action 2: A shared data-classification scheme that can be implemented across institutions, and allow automation of data management processes, is an eventual goal. Institutions should consider how this classification scheme could be developed.

The ARDC is enabled by NCRIS

Align Protections

Less straightforward is sector alignment of the protections required for different levels of sensitive data. Earlier, we listed reasons why universities may differ in the particular protections that they require. Also, because cybersecurity is a fast-moving area and best practice will change over time, gaining agreement between universities on this topic is difficult. We therefore recommend that a community of practice be formed to discuss what protections different universities (and other research institutions dealing with similar data) currently require for sensitive data; changes in best practice; and the extent to which a more aligned "standard" can be developed together. The ARDC's Australian Sensitive Data Interest Group[70] is one such forum through which these common interests could be explored.

Call to action 3: Institutions should come together to discuss cybersecurity protections for sensitive data.

Setting Expectations

No matter how carefully designed and sophisticated the classification scheme and risk-assessment tools provided by the institution, the researcher will always share the responsibility for making sure that the sensitivity of their data is well understood and managed appropriately. The institution is responsible for supporting the researcher in mitigating the risks associated with their data, but cannot be expected to be able to assess and manage that risk without the expert input of the researcher.

Like other aspects of RDM, the management of sensitive data requires ongoing investment. Universities need to monitor changes to regulatory requirements and stay on top of new cybersecurity threats, changes to best practice in the management of sensitive data, and the implications of updates to agreements with commercial providers. Platforms for managing sensitive data also require ongoing support.

It is unreasonable to expect that an institution will be able to sustainably support the platforms required for all sensitive-data use cases as standard offerings. Some data may require unusually high levels of protection that is too expensive for most universities to implement and maintain. Instead, we recommend identifying and catering for use cases that represent the bulk of research taking place at the institution (for example, 80 percent of cases) and expect to manage the edge cases separately. This will enable a focus on meeting a smaller number of requirements in a sustainable manner.

It is important to be clear about which requirements are being met by the institution's classification scheme and standard platforms for sensitive RDM. The classification scheme may address the institution's baseline responsibilities but is not guaranteed to cover additional requirements that may apply to specific datasets (for instance, specific access requirements placed on a dataset by an external provider). Researchers should be given sufficient information about the protections provided by the institution's standard offerings to be able to determine if they require special arrangements to meet any additional requirements.

The ARDC is enabled by NCRIS

Applied Advice

Template Project Plan for an Institutional Classification Scheme

Using some of the advice above, we have prepared the following example project structure for implementing an institutional data classification scheme. The structure is neither definitive nor required, but the high-level activities may aid in designing your own structure, or in communicating the level of activity and commitment required. Depending on institutional context, this may strengthen an institution's business case for resources.

Pre-Development Phase

- Typical objectives: to build the case for the change and investment
- Likely activities:
  - Conduct institutional risk assessment, and general assessment of risk to researchers.
  - Develop and consult on objectives.
  - Engage leadership, existing governance, obtain recognition.

Development Phase

- Typical objectives: to engage broadly and achieve buy-in, and to manage expectations with a phased approach, as the extent of implementation may be unknown to begin with
- Likely activities:
  - Co-design with researchers and develop classification levels.
  - Develop a guided tool, guidance (for researchers, ethics) and governance models.
  - Formally test the classification levels and guidance, evaluating usability, inclusivity, and alignment with objectives.
  - Engage with cybersecurity experts to identify the required protections for each classification level.
  - Assess existing research services and infrastructure.
  - Recommend changes (including governance, procedures and service changes).

Implementation Phase

Activities vary greatly with existing environments, objectives, and development outcomes, but may involve implementing:

- service and support changes
- infrastructure changes
- training
- pathways to appropriate platforms
- ways to manage edge cases
- ethics processes
- ongoing governance.

Resources

Sensitive Data – Classification Crosswalk spreadsheet (an initial analysis of alignment of classification levels across a number of schemes, including some government schemes and some universities).
https://ardc.edu.au/wp-content/uploads/2022/03/08z-Sensitive-Data-attachment-Classification-Matrix.pdf

The ARDC is enabled by NCRIS

ELEMENT 8: SUPPORT, TRAINING AND GUIDANCE

Summary

This element provides institutions with guidance on how researchers and research students can gain knowledge essential for managing their research data effectively. We outline the challenges associated with educating researchers in RDM so that institutions can address these when developing strategies for providing support, training and guidance. We discuss successful approaches, and provide information on essential components of the approach to education.

Why Support, Training And Guidance Are Important

RDM is complex, institutionally specific and multifaceted. To ensure all researchers and research students have the knowledge essential for managing their data effectively, institutions should provide support, training, and guidance.

Support encompasses the available training and educational resources provided to staff and students about RDM. Training includes a structured event or experience that has clear learning objectives and requires participants to be able to demonstrate the extent of how the objectives were met. Guidance includes online tools or resources to guide practice and behaviour and includes mandatory checklists and policies or procedures.

Advice includes avenues for help, one-on-one consultations and/or direct advice from the relevant staff across the institution about RDM – including in response to bespoke arrangements or RDM requirements for collaborative projects.

Together, support, training and guidance are a primary vector by which all other elements of this framework can be delivered to researchers or enabled for researchers. For example, knowledge gained through education, combined with skills acquired through training, is a key component of culture change; researchers performing RDM planning need guidance on how to go through the process and what must be done with the plan; appropriate cybersecurity approaches must be communicated and supported.

Institutions can ensure researchers are equipped with the knowledge and skills to comply with national and institutional RDM requirements by providing them with foundational education and skills in RDM through the provision of support, training and guidance. This enables researchers to conduct compliant research effectively and ethically.

To adequately provide support, training and guidance, institutions need to overcome several challenges. In most cases, these challenges are common, but they may vary depending on the size and priorities of the institution. It is important for institutions to address these challenges when planning their strategy for providing support, training and guidance.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

There are 5 major common challenges:

- Financial and human resources are limited.
- Time-poor researchers generally do not see RDM training as a priority.
- Within an institution, it is not always clear who should offer training, or support and guidance on RDM.
- There are discipline-specific approaches to working with data and not all subject-matter experts are experts in training.
- Expertise and systems required for good RDM are sometimes distributed across different organisational units of an institution. So, a coordinated approach is needed to align potentially different priorities/expectations/requirements across these areas.

Differences In Approach And Need

It is not feasible to have a one-size-fits-all approach for all institutions. Different organisational units within each institution will likely need to collaborate and specialise in the delivery of support, training and guidance for RDM. At a whole-of-institution level, the aim is to bring together a suite of training, information and resources with the appropriate level of support, ranging from general to specialised.

Institutions are at different points of their RDM journey. RDM maturity is likely to predict both the skills and needs of researchers and the availability of existing training materials. For institutions beginning to establish RDM processes, there may be little in the way of existing materials. For these institutions, priorities in providing training advice and guidance will generally include developing an initial range of documentation and training packages. This can be resource-intensive and take time to develop, so the first training aims should be to educate researchers in the foundations of RDM.

On the other hand, institutions that are much further along their RDM journey should already have a support, training and guidance package. In this case, the main aim in providing training might be to refine/mature the current RDM practices and culture within their institutions, embed RDM into quality research practices and research training, and/or provide more advanced data skills training.

Given the different aims and priorities, institutions should consider the minimum required competencies their researchers will need. For some competencies, a conceptual, theoretical, or abstract level of understanding may be sufficient and may be sufficiently addressed by shared or common training and materials. Other competencies may require the local development and delivery of practical, applied or concrete skills and guidance.

The approach adopted by each institution is heavily influenced by the level of resourcing available and the infrastructure underpinning support, training and guidance and its useability. In terms of resourcing, this will include a range of contributing factors such as availability of sufficient funds, staff and the effort required in developing support, training and guidance for the institution. The availability of funds will affect an institution's decision on whether it leverages existing infrastructure to provide support, training and guidance for its researchers or if it invests in upgrading its infrastructure. The availability of financial resources gives an institution the option of procuring external vendors to provide the training and/or infrastructure upgrades. If finances are limited, institutions may turn to in-house solutions.

The ARDC is enabled by NCRIS

Setting Expectations

Institutional contexts for support, training and guidance vary greatly between institutions, so setting common baseline features is difficult.

However, all institutions should expect that:

- RDM will, in most cases, be a secondary concern for researchers, so review the recommendations on approval, support and engagement (see Recommendations 1 and 2)
- resourcing for support, training and guidance will be limited (see Recommendations 3 and 4)
- support, training and guidance may need to be developed in-house and localised training resources maintained on an ongoing basis (see Recommendations 3 and 4).

Recommendations And Advice

Properties of Successful Approaches

Different institutions may have different approaches to providing support, training and guidance, depending on their needs. However, successful approaches generally exhibit 2 common features:

- The approach has approval and committed support from senior institution staff, approximately at the level of Deputy Vice-Chancellor (DVC) (refer to Element 2: Culture Change).
- The institution engages in genuine and broad stakeholder engagement and dialogue, involving:
  - researchers
  - research support staff
  - technical support staff
  - external stakeholders – funding bodies and technical partners.

Recommendation 1: Secure approval and committed support from senior institution staff — approximately at DVC level — for the approach to providing support, training and guidance.

Recommendation 2: When designing the approach to providing support, training and guidance for RDM, engage in genuine dialogue with researchers about their needs.

The ARDC is enabled by NCRIS

Essential Considerations When Planning Support, Training And Guidance

- Is the institution seeking to meet minimum compliance requirements or go beyond the minimum requirements and find further benefits of increased support, training and guidance?
- To what degree and in what ways will support, training and guidance be outsourced?
- Within the institution:
  a) Who has the relevant expertise and knowledge relating to the information/content to provide support, training and guidance?
  b) Who has skills in providing training?
  c) Are a) and b) the same people?
  d) Are the providers appropriately supported in providing support, training and guidance? For example, given time and resourcing, are they supported in their own skills development, and are they aided in scheduling, promoting, and managing participants?
  e) Are providers able to maximise the reach of training? If not, how can they be supported? For example, visible senior sponsorship, mandatory training, promotion.
  f) The relative benefits of online vs in-person training depends on institutional circumstances. For some universities, online is more efficient for time-poor researchers and trainers; and for some online training is harder to implement with current resources.
  g) Use language/terminology that aligns with other guidance and existing platforms at the university.

Institutions should consult their research ethics committees when developing training materials. This ensures that the training package content will address the requirements of each committee when it comes to management of sensitive data. Human-research ethics committees, in particular, are an excellent source of knowledge for researchers on the minimum requirements for storing sensitive and/or medical data that must be kept private and secure at all times. These committees also have insight into requirements for managing Indigenous data, as instructed by the AIATSIS *Code of ethics for Aboriginal and Torres Strait Islander research*[51].

Roles and responsibilities

Institutions should seek formal recognition (by senior staff or internal agreement) for a defined and coordinated set of roles and responsibilities for support, training and guidance. Include an outline of who is responsible for determining the level of resourcing allocated; what is considered within the scope of the support, training and guidance; and how support, training and guidance are to be delivered. At a minimum, define roles and responsibilities for the following units within an institution:

- Research office or equivalent
- IT department or equivalent
- Record-keeping department or equivalent.

The ARDC is enabled by NCRIS

Depending on local organisational structure, other units may also be included, such as:

- Library or equivalent
- Research ethics office or equivalent
- Research integrity office or equivalent
- Graduate research office or equivalent
- eResearch office.

Recommendation 3: Seek formal recognition, by senior staff or internal agreement, for a defined and coordinated set of roles and responsibilities for providing support, training and guidance. At a minimum, define roles and responsibilities for the research office, the IT department and the record-keeping department or equivalent.

Technical Considerations

Institutions should consider whether the following platforms/systems are required:

- An online platform to allow users to access training materials and self-help materials (publicly accessible or restricted to institutional users)
- An online meeting platform (such as Zoom, Webex, MS Teams) to support remote training
- A method of allocating queries, or a customer relationship management system, to allow support and guidance queries to be systematically tracked and allocated.

Training Outcomes/Competencies

When developing training packages, we recommend that institutions consider the outcomes listed below as the minimum required competencies researchers will need to gain when undertaking institutional training for RDM.

For some of these competencies, a conceptual/theoretical/abstract level of understanding is sufficient and is identified as such; these competencies can often be adequately addressed by shared/common materials. Other competencies require the development of practical/applied/concrete skills and are often more suited to localised training approaches.

Minimum required competencies for researchers:

- Identifying what research data is, including sensitive data
- Knowing the policy framework, including the Code2, legislation, funder requirements, journal requirements
- Applying relevant institutional policy and procedures
- Knowing the institutional RDM support and systems
- Planning for active RDM in an institutional context
- Creating and maintaining a DMP through the research life cycle

The ARDC is enabled by NCRIS

- Considering security, access, backup, obsolescence, sensitivity; having an understanding of why and how to make data secure, and how these requirements affect data storage and access
- Understanding how key RDM practices are aligned with the research project life cycle (i.e. from research proposal to completion of project)
- Being aware of security implications that come with movement of data and ensuring that data is moved across platforms/resources as securely as possible
- Knowledge of managing data in temporary resources
- Knowledge of how to archive data properly and for how long when a project is completed or when leaving the institution
- Knowing general data publication benefits
- Being aware of the benefits of open data and open research
- Being aware of FAIR and CARE principles
- Being aware of copyright and licensing issues
- Being aware of ethics issues in data management.

Training Package Topics

Depending on the priorities of the institution, consider developing training packages that allow researchers to develop their skills and knowledge in the following areas:

- Understanding the concepts of research reproducibility
- Skills to implement the FAIR data principles
- Skills to implement the CARE principles
- Knowledge of discipline-specific research data repositories
- Enacting best practices in data documentation and data organisation (version control, file-naming conventions and metadata), sharing data (including licensing), data publishing, reuse (both reusing data and enabling the reuse of data)
- Knowledge of data storage/publication options relevant to their data type (for example, images, models, code, literature reviews, video, creative works)
- Knowledge of setting up effective/relevant backup policies
- Knowledge of file types/documentation and how they affect preservation/reuse
- Knowledge of the systems that allow tracking of data impact/reuse/citation
- Knowledge of licensing options for data
- Knowledge of how to publish data and follow relevant data citation practices.

Applied Advice

When getting started, institutions should consider including support, training, and guidance in the consultation and planning process before implementing any RDM solutions.

ELEMENT 8: SUPPORT, TRAINING AND GUIDANCE

In some successful case studies, institutions found that setting up a single point of contact for researchers was critical, as researchers usually have trouble knowing which part of the institution is best placed to give advice on a particular topic. This can be a source of great frustration in many institutions.

These successful case studies also recommend that institutions consider using or adapting existing resources rather than developing new resources from scratch. Institutions should also consider training their support staff and actively encourage/support them to join national or international communities of practice. This will build a body of knowledge that can be shared within the institution. Laying these good foundations will lead to future-proofing the successful delivery of support, training and guidance.

While staff delivering training will develop competencies over time, institutions will need to consider how to provide upskilling opportunities for future trainers, such as train-the-trainer approaches. Training trainers has a positive effect on the effectiveness of the delivery. Setting up communities of practice may be viable.

Making refresher training available is another option. These could be annual but, as requirements may change every few years, timeframes will vary between institutions.

Institutions are also recommended to consult relevant ethics bodies when developing support, training and guidance. All RDM training should be aligned with relevant ethics requirements since many projects will require ethics approvals.

HDR students may require more specialised training packages. Consider if their training can be linked with their HDR milestones.

Be aware of any training being conducted in each unit/department, and consider whether it can contribute core material to an effective RDM training package. In many cases, library personnel, research ethics personnel and graduate research office personnel may already be delivering training on RDM relevant to their area. The institution should tap into these resources when developing training, to prevent re-inventing the wheel and to ensure that consistent messages are delivered.

> **Recommendation 4:** Consider training approaches that leverage existing staff resources to build an institutional body of knowledge.

Areas that should be consulted about RDM support, training and guidance (not an exhaustive list):

- Library or equivalent
- Research ethics office or equivalent
- Research integrity office or equivalent
- Graduate research office or equivalent.

It is also recommended that institutions establish a network of RDM champions who can be available locally to provide advice on why RDM is important and how to do it, and to encourage and promote use of the systems in place.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Recommendation 5: Consider reusing shared and nationally/internationally standard materials and resources as much as possible, adapting them for local relevance and applicability to all stakeholders within the institution.

Sharing training materials on open-access platforms such as Digital Research Skills Australasia (DReSA) is also recommended, so that other institutions may access and reuse them. Such practices will eventually allow for aligned RDM training packages across institutions that have common features/characteristics and language/terminology.

Resources

- 23 Data Things Program: https://au-research.github.io/ARDC-23-things/
- AARNet resources: https://www.aarnet.edu.au/aarnet-train-the-trainer-workshops, https://conference.eresearch.edu.au/events/its-not-yoga-but-sounds-like-it-flexibility-training-and-stretch-goals-in-2021/
- ARDC FAIR Data 101 self-guided. 2021: https://zenodo.org/record/5094034#.YQyLbY4zaUk
- Clare C et al. Engaging researchers with data management: the cookbook. Cambridge, UK: Open Book Publishers, 2019: https://doi.org/10.11647/OBP.0185
- Data organisation for spreadsheets for ecologists (Data Carpentry): https://datacarpentry.org/spreadsheet-ecology-lesson/
- Delineating the successful features of RDM training: a systematic review: https://doi.org/10.1080/1360144X.2021.1898399
- Digital Research Skills Australasia (DReSA): https://dresa.org.au/
- Dryad. Good data practices for creating reusable data: https://datadryad.org/stash/best_practices
- FAIR Handbook for training: D7.4 How to be FAIR with your data. A teaching and training handbook for higher education institutions: https://zenodo.org/record/5837500#.Y13Tq3ZBxnI
- FOSTER Consortium. Managing and sharing research data. Zenodo. 2018, November 26. https://doi.org/10.5281/zenodo.2630562
- FOSTER Consortium. The open science training handbook: https://book.fosteropenscience.eu/
- How to name files, by Jenny Bryan for Reproducible Science Workshop: https://speakerdeck.com/jennybc/how-to-name-files
- Interview worksheet for faculty (Purdue University Libraries): https://docs.lib.purdue.edu/dilsymposium/2013/interviewinstruments/1/ http://dx.doi.org/10.5703/1288284315510
- Library Carpentry: https://librarycarpentry.org/
- Mantra RDM training: https://mantra.ed.ac.uk/ (adapt it to provide local context)

- Principles Aligned Institutionally-Contextualised (PAI-C) RDM Training: https://dresa.org.au/materials/principles-aligned-institutionally-contextualised-pai-c-rdm-training
- Research Data Management and Sharing – MOOC that introduces the concepts and issues around research data management: https://www.digitalresearchservices.ed.ac.uk/training/research-data-management-and-sharing-mooc
- Research Data Management Toolkit, updated by Jisc 2021: https://www.jisc.ac.uk/guides/rdm-toolkit
- TIER Protocol: https://www.projecttier.org/tier-protocol/
- Yu F, Deuble R, Morgan H. Designing research data management services based on the research lifecycle – a consultative leadership approach. Journal of the Australian Library and Information Association. 2017;66(3):287–298. DOI: 10.1080/2475158.2017.1364835. https://www.tandfonline.com/doi/full/10.1080/24750158.2017.1364835

The ARDC is enabled by NCRIS

ELEMENT 9: RESEARCH DATA APPRAISAL

Summary

This element provides an overview of the importance of appraising research data, and the challenges and barriers for effectively appraising the growing research data assets of Australian universities. The key challenges identified include identifying and collecting the information required to support appraisal; lack of clarity around governance arrangements; and understanding the trigger points for appraisal actions. Our calls to action highlight opportunities for addressing the challenges and barriers.

Why Appraising Research Data Is Important

From an archival context, appraisal is the process of reviewing records to determine value. For research data collections, appraisal may be undertaken to determine:

- retention requirements
- ongoing value
- reusability
- conformance with legal/ethical requirements.

Appraisal is not a one-off process — research data collections need to be appraised over time.

Appraisal systems supporting well-managed research data collections typically include some combination of infrastructure, software, policy and expertise. Effective appraisal systems support the management and use of research data collections, and provide confidence that the data is being managed appropriately. Beyond meeting legal requirements, appraisal practices identify data which has ongoing value in terms of reuse/importance, reflecting a university's research priorities and values.

Often several stakeholders play a role in appraising research data. The creator of the data, the data custodian, support services such as library and archives staff, and records management staff may all be involved in appraisal or in supporting appraisal processes. Providers of IT infrastructure for research should also be informed about retention and disposal decisions made by those responsible.

Efficient appraisal requires access to contextual information about the research data collection. Information about what the data is, why it was collected, and what the retention requirements are is necessary for facilitating its appraisal, particularly when the original creator or custodian is no longer available. Recording sufficient information about research data can also enable automated actions, where appropriate.

Knowing when to undertake appraisal actions, and what data to actually appraise, can be challenging. Logical decision points such as commencement of a research project, completion of a research project, allocation of data storage, publication of research output/s, in practice may not be as clear cut as they first seem.

We explore these challenges in further detail below, making recommendations and calls to action to address them, where appropriate.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Challenges In Appraising Research Data

Despite a growing awareness of the need for appraising research data, there are common barriers to effective appraisal practices for research data at Australian universities:

- Contextual information required to support efficient appraisal of research data collections may be missing or difficult to locate.
- There can be a lack of clearly defined decision points for undertaking appraisal actions.
- Responsibilities for the management of research data, and data governance more broadly, can be unclear.

In any institution, one or more barriers may prevent effective/efficient appraisal. However, to some extent barriers may also depend on the size of the institution and/or the size of the research data collection. We highlight this in the following sub-sections, where relevant.

Identifying Information to Support Appraisal

Collecting contextual information/metadata about a research data collection is necessary to enable its appraisal over time. Where such information does not exist, it can be difficult and time-consuming to retrospectively collect and, therefore, presents a significant barrier to appraisal. This is particularly the case where custodianship has been transferred (either explicitly or by default) to the university, and the original custodian is no longer available. Orphaned data collections — those that no longer have active data custodianship — are a growing challenge across the sector, a challenge which is often compounded by a lack of contextual information about the data.

Data custodians should be encouraged to capture information about a data collection at any time as part of their broader data management practice. However, there are some points in time where capturing or reviewing this information is particularly timely:

- At the commencement of a research project and/or when a storage allocation is provided
- At the conclusion of a research project
- When an appraisal task has been undertaken.

Data management plans may provide an opportunity for capturing such information, noting that research data collections and the research project defined in a data management plan often do not neatly align. Another opportunity for data capture is during provision of research infrastructure, particularly research data storage.

Capturing information about research data collections is useful to universities of any size. For smaller institutions, it can facilitate more efficient appraisal practices; at larger institutions there may be opportunities to use such information to automate processes, such as identifying collections that may be eligible to be moved to archival storage.

The ARDC is enabled by NCRIS

Articulating Data Governance And Data Management Requirements

Obligations for managing and retaining research data are mandated by university policies and state and territory archives through retention and disposal authorities. At this higher level, requirements and responsibilities for meeting data retention obligations are relatively well defined.

However, although data governance structures often exist for corporate data at Australian universities, they typically do not cover research data. At many universities, this results in a lack of coordination and consistency in appraisal systems for research data.

In many instances, responsibility for the implementation of data management policies and requirements including appraisal fall on the data custodian. Depending on the context, data custodians may have limited or no access to expertise and support to appraise data. Where support does exist, appraisal expertise may be provided only on an ad hoc advisory basis, as support staff such as archivists and record managers often lack resources too. University archives have not generally been tasked with playing an active role in appraising research data collections, unless such collections have been identified as cultural collections.

The size of the institution and the number/volume of data collections often influence approaches to appraisal. In larger institutions with significant volumes of research data, appraisal is more likely to be left to researchers by default, due to the scale of the challenge. Providing widespread access to appraisal expertise at research-intensive universities is not feasible in practice. By contrast, smaller institutions with lower volumes of data may have more scope for collaborative appraisal whereby support staff (such as archivists, record-keepers) work with data custodians. However, scalability is also becoming an increasing issue for smaller institutions as data volumes continue to grow.

The lack of coordination and consistency in appraisal systems for research data combined with limited resourcing and access to expertise means that appraisal often does not happen. Where it is done, it may not be done consistently, or with any overarching guiding principles to reflect the university's broader aims.

Understanding Appraisal Trigger Points

Although appraisal is an ongoing process, it can be helpful to identify and articulate common trigger points. However, this is more challenging than it may initially seem. Trigger points may include: the end of a research project, a change in custodianship, the end of the minimum retention period, or when data is being moved or migrated.

When applied to a specific context, however, other factors also require consideration. Research data collections and research projects often do not neatly align, so identifying the data produced from a completed research project, and therefore what data to appraise, may not be a simple process. Longitudinal research data, for example, commonly relates to numerous research projects over time.

Likewise, transfer of custodianship may occur by default, without any specific handover process taking place to trigger an appraisal process. This is particularly the case with orphaned research data which no longer has an active custodian. The widespread challenge of appraising and managing orphaned data collections is compounded when custodianship is unclear, making it hard to access the data and gain permission to appraise it.

The ARDC is enabled by NCRIS

Smaller universities may be able to take a more proactive approach to providing resources to data custodians to facilitate appraisal, enabling a more active collaboration between researchers and research support staff. At larger universities, the scale and growth of data and data collections often means such collaboration is not widely feasible.

Recommendations And Advice

There is no single solution for facilitating more effective appraisal practices for research data at Australian universities. However, we can identify pragmatic approaches for addressing the key barriers. Our recommendations and advice should be considered in conjunction with other framework elements, particularly retention and disposal and sensitive data.

Sufficient Contextual Information About Research Data Collections Should Be Captured To Enable Efficient Appraisal Practices

Irrespective of the size of the institution, the definition of a common set of information/metadata to facilitate appraisal of research data at Australian universities would be helpful. Working towards standardisation of the capture and storage of this information could facilitate efficiencies in appraisal processes and provide opportunities for automation. Examples of automation include generating notifications for appraisal to be done, or identifying data collections that may be ready for archiving.

Broad categories of information to capture and store include:

- Descriptive information, such as format, technical requirements, size
- Retention requirements
- Relationships to other entities: people, projects, data, funders, ethics applications
- Access restrictions / rights / data classification.

Institutions should, as much as possible, streamline requirements and opportunities for capturing contextual information to avoid duplication of effort. Similar information may be required as part of an application for research infrastructure provision (for example, access to data storage or compute facilities) and for data-management planning processes.

Call to action 1: Further to the recommendations of the retention and disposal element of the framework, universities should collaborate on a common set of information/metadata about research data collections to facilitate efficient data appraisal practices.

Data Governance, Including Stakeholder Roles And Responsibilities, Should Be Clearly Defined And Articulated

Many Australian universities already have data-governance frameworks covering administrative data but these frameworks are often not applied to research data, for many reasons, depending on context. Research data is often complex, and is not homogenous, while ownership and custodianship are generally not well understood. The scale of the data, both in terms of number of collections and size, can introduce further complexities.

The ARDC is enabled by NCRIS

To improve coordination and implementation of good-practice data management and appraisal approaches, universities would benefit from implementing data governance frameworks that specifically address research data.

To define what an appropriate research data governance framework might look like, further consultation and investigation would be required before more specific recommendations could be made on suitable approaches. It seems likely, however, that existing data governance structures could be adapted with some modification to apply to research data.

Call to action 2: Universities are encouraged to collaboratively agree on common elements for good-practice Australian research data governance, making use of existing governance approaches, where appropriate.

As part of a research data governance framework, a university should also consider explicitly aligning the appraisal of research data with its overarching values and goals, which will vary depending on the university's focus.

Recommendation 1: Consider developing guidance on appraisal practices that reflect the institution's values and priorities.

Appraisal Trigger Points Should be Clearly Defined and Well Understood

Lacking clearly defined and understood appraisal trigger points has been identified as a key barrier to effective appraisal. A trigger point could be any point in time where an action could be taken, whether that action is automated or through human intervention. The action could be an appraisal-related process and/or the capture of contextual information to facilitate management and appraisal. Examples of trigger points include:

- The end of a research project
- When custodianship is transferred from one stakeholder to another (for example, from researcher to university)
- The end of the minimum retention period
- When a storage allocation is provided for storing the data
- When the data is moved to a new storage location (for example moved to cold storage).

There are significant challenges in defining such trigger points; for example, identification of end dates is complicated by the continuation of research beyond official project dates or longitudinal research data.

Further investigation is required to better understand how to provide advice to data custodians on when to undertake appraisal processes, as they also need to be considered within the specific local context.

ELEMENT 9: RESEARCH DATA APPRAISAL

Recommendation 2: Consider opportunities for identifying and supporting common appraisal trigger points, minimising manual interventions where possible.

Call to action 3: The significant challenges, lack of clarity, and desire for common practice requires collaborative action to identify and define practical appraisal trigger points and guidance on how to integrate them into an institution's local systems and workflows.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

ELEMENTS 10 TO 19

The following 10 elements of the framework were identified by participating universities in early workshops and feedback, but were not further developed by the expert working groups.

While all elements were considered priority areas for universities, the expert working groups focused on an initial set of 9 elements based on an assessment of the best opportunities for collective advancement, current institutional priorities, and availability of expert contributors.

Captured here are explanations of the remaining 10 elements (numbered 10–19), their importance to Australian research institutions, and the related challenges and opportunities. Also captured are key questions resulting from the participant discussions, which institutions may find useful to consider when addressing these areas of RDM. These questions will also help to guide future collaborative work in the sector.

Element 10: Data Sharing And Access

While several early framework discussions focused on the combined themes of data sharing, data publication and open research, the expert working group discussions focused more on the current movement towards open research and data publication. It is important to acknowledge that sharing is commonplace, it occurs during and as part of research projects, it is often concerned with ease of collaboration, and it does not necessarily involve 'open' or 'publishing'. Because sharing data is crucial to active research, institutions often provide data-sharing platforms and infrastructure. Researchers and custodians need an understanding of data ownership, the risks of sharing, and the implications of sharing agreements and data licensing.

Questions To Consider

- What does it mean to licence a dataset and what is the value of doing this?
- Do institutions need to provide training, or rather tools, that guide researchers through licensing?
- How do we bring clarity to topics such as data ownership, owner rights, portability of data, and delegation of rights?
- How do we address complex datasets that combine data from different institutions or many different sources?
- How are institutions addressing cross-institutional research projects and data management agreements?
- How do we facilitate data handover between institutions, industry bodies, external parties – for example, for commercialisation purposes?
- How are data-sharing requests managed and should institutions support processes for mediated access?
- How do we appropriately identify risks in sharing or aid researchers in identifying the risks?
- What are the recommended requirements for data-sharing platforms, and should these platforms also take into account data publishing?

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

- As data sharing agreements, contracts and licences can be unique, how do institutions make these visible or aid researchers in keeping track of data and licences?
- How do aspects such as sensitive data, participant consent, and external partners affect the sharing of data, and how should institutions guide towards best practice?

Element 11: Cybersecurity

Appropriate cybersecurity is the cornerstone for keeping research data safe and secure. It is an important aspect of good RDM.

Many institutions have IT departments that have dedicated cybersecurity expertise to oversee their institution, but need input from research data specialists to understand and support the particular needs of research data.

Researchers are required to ensure that their data has been stored securely, for the long term, and sometimes with appropriate access controls in place to comply with legal, ethical, and funder requirements. Institutions, therefore, need to provide guidance on the level of security attached to various institutional data services, and how to appropriately choose and correctly use them. A key consideration is providing researchers and institutions with an overview of cybersecurity elements.

Element 7: Sensitive Research Data makes a call to action encouraging institutions to come together to discuss cybersecurity protections for sensitive data.

Questions to Consider

- How are we building cybersecurity into infrastructure in support of and on behalf of researchers?
- What cybersecurity actions are required and what are institutions doing to secure research data (for example, encryption, enabling multi-factor authentication)?
- How do we monitor the security of and access to data?
- What constitutes a notifiable research data breach and how are institutions managing, identifying and being made aware of breaches?
- Is there a best practice model?
- What is the indicative list of (research) legislation relevant for cybersecurity?
- What are the institutional responsibilities — including systems and provision of training — and who needs to be involved (for example, IT, ethics, research services)?
- What are the responsibilities of and requirements on researchers?
- How are researchers made aware of what is supported?
- What cybersecurity training is available at institutions?

The ARDC is enabled by NCRIS

Element 12: Data Ownership

A national institutional RDM framework could provide guidance on data ownership and usage rights, seeking harmonisation with institutional policies on data ownership with the aim of achieving a national approach.

Questions to Consider

- How clearly is this stated in institutional policies at this point?
- Is clarity required on data ownership, retention rights, usage rights and publishing rights and limitations when working with complex projects, commercial entities, on industry-funded projects, with governments, with international partners?
- How do institutions detect and address gaps when intellectual property and RDM policies overlap or interact?
- How do we address data ownership for citizen science projects, especially as researchers move across institutions?
- How might we better manage ownership and access when researchers and students change institutions or move on?
- How do we manage ownership and access when data sources retain ownership, or impose usage restrictions, or include complex agreements from multiple providers?
- How do we better manage the audit trail of custodianship?
- How do we record, define and manage compliance with Indigenous data ownership, community ownership and data sovereignty?

Element 13: Digital Preservation

Digital preservation (DP) involves the processes of assessing, collecting, organising, maintaining, preserving, and providing continued access to content determined to have long-term value, in ways that ensure authenticity, reliability, integrity, and usability (UNESCO and the International Council on Archives, *Universal Declaration on Archives*, 2011[71]).

The goal of DP is to ensure that digital assets are managed well and cared for so that they remain usable and understandable over long periods of time – through both technology and organisational changes. This involves institutional elements such as policies, strategies, and institutional commitment to maintain and provide access to digital content.

Like other elements of this framework, institutional DP intersects with other framework elements. But as an institutional element of an RDM framework, DP goes beyond providing training in cultural awareness and preservation techniques to academics.

DP also goes beyond infrastructure practices to ensure the safety and accessibility of data, such as in long-term retention (for example, bit-level preservation, including, but not limited to, checksum monitoring).

The ARDC is enabled by NCRIS

DP includes content preservation and technological dependencies (software and hardware), which incorporates developing and maintaining risk profiles of digital collections in an institution's custody, technology watch processes, through to preservation actions (for example, identification of existing file formats; migration of proprietary and/or legacy file formats to open and/or recommended preservation file formats).

It also includes establishing safeguards against accidental and/or malicious deletions or corruptions, agreed-upon actions for rectifying issues discovered, and complex standardised metadata support (for example, descriptive, technical, and preservation metadata to aid in providing context, ongoing maintenance and improved discoverability).

While retention processes may partially facilitate ongoing access to digital assets, a more comprehensive institutional DP approach is necessary to ensure long-term reuse, integrity and reproducibility.

Element 5: Research Data Retention and Disposal highlights the need for long-term curation platforms, ensuring adequate metadata and persistent identifiers, and the need for institutions to develop decision-making tools.

Questions to Consider

- How do institutions select research data assets of significance and enduring value while balancing with the best use of capacity?
- How can institutions select assets while taking into account their institution's unique profile, their custodial roles and key stakeholders?
- When does selection occur?
- How can identification be built into all stages of the data life cycle?
- How are institutions committing to long-term research-data preservation in terms of workforce, ongoing functions and infrastructure?
- How do we foster a research- and RDM-aware workforce skilled in DP as the importance of this institutional responsibility increases?
- What funding models work for ongoing and long-term preservation of assets, given the different timelines and needs of active data management and IT support?
- Can a more holistic perspective be achieved in the interests of maintaining assets across segmented approaches to resourcing?

Element 14: Funding And Sustainability

Research institutions, researchers and research support staff need advice on funding and seeking funding for RDM. Guidance may also be needed for funding agencies mandating RDM. Clarity is required on components and proportion of research funding and institutional resourcing that can be used for RDM. Specific guidance is required on how to fund RDM sustainably.

As part of research funding by researchers and institutions, RDM should be planned and factored to be sustainable and cost-effective. RDM will scale to varying sizes and with varying priorities of research institutions.

The ARDC is enabled by NCRIS

Questions To Consider

- What should funders mandate with regard to RDM policies or RDM planning?
- What can researchers ask funders for in terms of funding for RDM services – rules and expectations?
- Resourcing models for RDM services will differ, but what percentage should be expected to come from institutional block grants?
- How can institutions aggregate different sources of funding in the most efficient way, promoting recognition of the role of co-funding, and helping researchers maintain a sense of ownership?
- How can institutions best address resourcing challenges — funding for both people and infrastructure? This should be viewed as a responsibility for the enterprise.
- Is it possible for institutions to support all RDM for all research activities?
- Are costs to smaller institutions disproportionately large and can larger collaborative use help?
- Are institutions being driven to engage external providers and at higher costs?
- What is the role of the ARDC and the National Collaborative Research Infrastructure Strategy in funding and sustaining RDM?
- How can non-traditional research outputs be sustained, curated and managed, given they bring further challenges?

Element 15: Governance

There are several levels of governance to be defined and addressed by the framework. Governance cascades through several levels, from institutional governance of the RDM process down to individual responsibilities for mediated access to restricted data. Levels can also include governance over multiple projects, programs of long-running studies, or governance specific to projects.

Institutions have responsibilities for aligning with the Code and its guidelines on RDM.[2] Repository governance is important, with many institutions managing institutional repositories, some hosting research data, and with some researchers using disciplinary repositories to both store and share data.

Some datasets generated by research require ongoing and mediated access, for example:

- via an expression-of-interest process for usage and access that is assessed on merit and as appropriate
- Australian Bureau of Statistics restricted datasets
- locally curated datasets
- population health characteristics data to be shared only with registered users.

Such a process can also require subsequent publications or resulting research be properly attributed and tracked.

The ARDC is enabled by NCRIS

Furthermore, collaboration with other institutions can involve joint responsibility for shared data which may require multi-institutional governance.

Questions To Consider

- What is optimal for institutional data governance, what does it need to address, and what models exist?
- Is a separate governance structure required for RDM or is it covered by other research governance structures, such as in ethics?
- Do existing data governance structures include risk management assessment and planning around data security, privacy, IP and research reproducibility?
- Are there data and information management policies already in place that include research data and how are such policies enforced?
- How are governance structures communicated and how are academics best engaged?
- What principles should governance approaches consider adopting; for example, The Five Safes[72] for effective decision-making and risk management, FAIR and CARE principles[4,5], and Indigenous guidelines (see Element 19: Indigenous Data Management)?
- What are the key roles for institutional governance of research data?
- Is there a need to define multiple roles; for example, data owner, data steward, data custodian, data manager?
- Will certain roles be crucial points of contact for stakeholders involved in planning, retention, disposal or decision-making?
- Are there examples or guidance on how to map such roles to researchers, research support and administrative staff?
- How are roles in external policies mapped to locally defined roles and responsibilities; for example, the broad definition of “researcher” within the Code?
- Does governance play a role in ensuring researchers’ needs are adequately met with infrastructure and appropriate types of storage?
- How do we ensure the “long tail” of smaller data is adequately governed as well as big data?
- How does governance address the transfer of custodianship when researchers leave, or the ongoing management of data once a project is complete?
- How are institutional repositories governed and what are the related policies around data deposit, data curation, data preservation, and data access?
- Could open research be a driver to kickstart discussions around inadequate areas of research data governance?
- Is mediation of access to research data an inappropriate use of researchers’ time, what models exist to manage this task, and how should institutions support this?
- For data that has multi-institutional requirements, where does governance for this kind of data sit, how is this agreed (for example, by prior agreement, contract), what effect do legal jurisdictions have, and how are issues resolved?

Element 16: Identifiers And Metadata

Having a range of different identifiers is important. Feedback indicates a need for all stakeholders — including institutions, platforms, service providers, and funders — to collaborate towards common approaches in adopting and implementing identifiers and collecting metadata. This could facilitate many aspects of institutional RDM.

Institutions need guidance on best-practice metadata management and on the responsibility within institutions for managing metadata.

The metadata goal is to enable data to be identified, retrieved and managed over time by capturing information including data contents, quality, format, collection/creation methodology, storage location, usage requirements, management requirements, and people involved with collection.

The identifier goal is to unambiguously reference resources (digital and physical) within a given context, typically by means of a string or number conforming to a formal identification system (such as Dublin Core Metadata Schema).

Guidance is required on how to practically apply and manage the various identifiers in use; for example, DOI, ORCID, RAiD, ROR. This includes how to use identifiers to link systems; when to use local or public, persistent or temporary identifiers; and which identifiers to use for consistency between institutions.

Questions To Consider

- With the proliferation of identifiers and their use, how do we minimise/standardise?
- What identifiers are required by funding bodies and how are they used?
- How are identifiers linked to impact and researcher incentives?
- What do we mean by metadata, what are the standards or vocabularies or data dictionaries, and what is discipline-specific metadata?
- What are the responsibilities of researchers, institutions and support staff in capturing and maintaining metadata and identifiers?
- How can machine-readable metadata and identifiers play a role, especially in aligning data to the FAIR data and CARE principles?
- How can metadata be practically applied to datasets across multiple platforms and research environments, especially when data is moved, copied or transferred?

Element 17: Non-Digital Material

Institutions and research fields need guidance on the objectives and management of non-digital research material to align with the guidance related to digital data, records and other digital objects. It is recognised that such guidance will involve the digital records about the non-digital materials, including the metadata required for the management, retention, disposal and related decision-making.

Questions To Consider

- How do we advise and engage on planning for, management of, and digital integration of non-digital or non-digitisable primary material?
- How does this supplement, support or work together with guidance on digital materials?
- Is there a need for a national approach (like, for example, biobanking) or identifier strategy for non-digital and/or non-digitisable primary material?
- How should field-specific standards and identifiers relating to non-digital primary materials interact across fields of research?
- Should shared vocabularies relating to these primary materials be promoted?
- Should RDM policies encompass non-digital/non-digitisable primary material?
- Should digital and non-digital material be managed separately or integrated?
- How should the decentralised management and cross-field nature of non-digital materials be addressed?

Element 18: Standards And Guidelines

Standards here are defined as comprising the mandatory legislation or other types of regulation with which institutions must comply, whereas guidelines are not necessarily enforceable.

Standards include the Code, international standards such as General Data Protection Regulation (GDPR), and national and state legislation and regulations around data, archiving, retention and disposal.

Guidelines may comprise local institutional recommendations, processes, 'how-to' guides, and best-practice advice from institutions active or leading in RDM and related topics.

Standards and guidelines are often attached to institutional policies. A national framework might facilitate the standardisation and reuse of locally developed guidelines, and raise awareness of existing best-practice advice.

Questions To Consider

- Standards and guidelines are important to this framework but do they need to be addressed separately?
- What is the definitive list of standards that a national framework will need to address, including international standards (for example, GDPR), funder requirements, national regulatory standards (for example, information standards), state legislation and standards regarding data retention and disposal, other more local policies and standards, and standards required for specific kinds of data (for example, personal information, sensitive data)?
- How are conflicts between standards addressed, especially when research crosses jurisdictional boundaries?
- What standards apply to specific disciplines and do such standards differ from the normal risk mitigation and management applied to research in general?

- Where does responsibility within the institution sit for keeping up with amendments and additions to standards, legislation and regulations?
- What are researchers' responsibilities and expectations for general awareness of standards, legislation and regulations?
- Where should expertise and detailed knowledge on these issues sit?
- Should there be a single institutional entry point for advice and guidance on these issues?
- How can researchers know when they have done enough to meet a standard and who might check for compliance with standards?

Element 19: Indigenous Data Management

Having received significant input from partners and external feedback, we recognise the common institutional need to capably manage Indigenous research data, including the data of the Aboriginal and Torres Strait Islander communities and data from other First Nations peoples around the world. This need may also go beyond data about individual people and communities to include culturally significant places, flora and fauna.

Indigenous data is an embedded and crucial part of everyday RDM. How Indigenous RDM is best represented with this framework is yet to be determined. Discussions with the participating universities have indicated a desire for good Indigenous RDM to be embedded in institutional systems and practices, rather than viewing it as an isolated activity. There are existing Indigenous initiatives working on RDM, such as the Improving Indigenous Research Capability project[73], which might be usefully supported or contributed to by collaborative effort from universities.

Resources

- Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS): https://aiatsis.gov.au/about/who-we-are/vision-mission-and-functions
- The AIATSIS Code of ethics for Aboriginal and Torres Strait Islander research, related research and committee resources: https://aiatsis.gov.au/research/ethical-research
- AIATSIS guides and resources, include policy resources: https://aiatsis.gov.au/research/guides-and-resources
- The Aboriginal and Torres Strait Islander Library, Information and Resource Network (ATSILIRN), including the ATSILIRN protocols for libraries, archives and information services: https://atsilirn.aiatsis.gov.au/about.php
- The CARE Principles for Indigenous Data Governance: https://www.gida-global.org/care
- Traditional knowledge identifiers / Traditional Knowledge Labels: https://localcontexts.org/labels/traditional-knowledge-labels/

The ARDC is enabled by NCRIS

RECOMMENDATIONS

Key recommendations for institutions seeking to improve RDM are collected here for easy reference. Each element provides context and additional considerations for these recommendations.

Active Data Management

Recommendation 1: Clearly articulate who is responsible within the institution for making decisions about the governance of active research data management.

Recommendation 2: Involve researchers at all stages of the Active Data Management solution planning process.

Recommendation 3: Make sure that investment in research data infrastructure is driven by the 3 principles of researcher co-design, security by design and, where applicable, privacy by design.

Recommendation 4: Clearly outline the appropriate use cases and use terms for each platform and service, and disseminate information about endorsed platforms and services through multiple channels.

Culture Change

Recommendation 1: Understand the current institutional RDM state.

Recommendation 2: Understand the desired future institutional RDM state, develop a strategy to achieve it, and regularly monitor progress towards it.

Recommendation 3: Clearly define ownership for building the RDM community at the institution.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

RECOMMENDATIONS

Recommendation 4: Take a coherent approach and be consistent with messaging.

- Carefully communicate policy agendas and/or the rationale for the objectives.
- Develop terms of reference for a working group to be created and seek endorsement from the top.
- Agree on a communications plan to execute and the key people needed to execute it.
- Leverage multiple communication methods and technologies including both “push” methods (bulk communications, such as email, staff news) and “pull” methods (those requiring opt-in recipient discretion, such as web resources, blogs, subscription lists). A variety of text and multimedia modes will also help ensure that messages are effectively communicated.
- Share and articulate the objectives, strategy and plan for raising awareness.
- Encourage research leaders and mentors to lead RDM practices and related discussions among the research team.
- Encourage research leaders and mentors to lead both education and compliance within their discipline and/or schools.
- Provide a marketing toolkit to aid in any communication. This could include reusable collateral such as slides with embedded consistent messages. Messages should succinctly outline the most important calls to action, the reasons, and potentially the risks or benefits that resonate.
- Build change communication into inductions and orientation of new staff, both researchers and research professionals.
- Incorporate meaningful “carrots” into the communication plan. For example, at Edith Cowan University and Bond University, upon completion of a data management plan, data storage in an instance of SharePoint is autogenerated for researchers.
- Leverage existing communication events and outreach opportunities, such as researcher forums and department research meetings. Change fatigue and workload constraints can mean researchers choose what they will get the most value from, so, ideally, bring the communication to them.
- Consider post-initiative/post-project engagement strategies in communication plans, which may involve broader staff to sustain and continue the culture change. Consider communicating and setting these expectations with staff, especially articulating what messaging or communication changes are expected of them in the long term.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Policy

Recommendation 1: Follow the 14-step process to establish a policy for RDM.

Recommendation 2: Follow the 9-step process for implementing an RDM policy.

RDM Planning

Recommendation 1: Follow these steps for planning data management.

Institutions and RDM communities:

1. Create an understanding of the purpose and benefits of RDM planning.
2. If implementing DMPs using, for example, a tool or standardised documentation, reach agreement internally on the business drivers/purpose.
3. Make a conscious decision about the timing and level of rigour required for RDM planning and its documentation, given that RDM planning is essential for all research.
4. Consider methods for interrogating and analysing the content of DMPs to determine their usefulness and to find any gaps in the research life cycle at the institution.
5. Create an online presence to allow for support and training activities in relation to RDM planning.

Researchers and RDM communities:

1. Clearly define and describe what constitutes RDM planning to help researchers understand why they are engaging in the process and for the institution to understand how it can resource, support and leverage the RDM planning process.
2. Further investigate the idea of publishing DMPs to understand the benefits it could bring to the publishing of datasets and knowledge exchange.
3. Evaluate responsibilities and accountabilities using a matrix such as RACI[54] to understand who is responsible and accountable as well as who needs to be consulted and informed in relation to RDM planning.
4. Consider integrating RDM planning throughout the research project life cycle to keep it current and relevant.
5. Embed the Code, and FAIR data and CARE principles in the RDM planning processes and activities so that researchers, institutions and the research community can maximise the benefits from research.

The ARDC is enabled by NCRIS

RECOMMENDATIONS

Recommendation 2: Follow these steps for planning RDM tooling:

Institutions:

1. Prioritise the identification of RDM planning tools in use and seek to achieve interoperation and coordination.
2. Enable RDM planning to be an ongoing process through integration of RDM tooling with relevant systems.
3. Create a roadmap for tooling to show best practice, no matter what level of digital integration the institution is at. The roadmap could be as simple as a Word document or as comprehensive as a fully integrated system.

Institutions, researchers and communities:

1. Seek to align RDM planning tools with FAIR data and CARE principles.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

RECOMMENDATIONS

Recommendation 3: Strive for best-practice RDM planning.

Institutions:

- Support the curation and development of RDM planning by appointing or identifying persons responsible for this support, such as data stewards.
- Establish RDM teams specifically to support RDM, including primary materials management.
- Conduct RDM planning development, review and updating within an online platform for ease of tracking and for version control. This could be an integrated laboratory information management system or electronic notebook system.
- Identify and action potential RDM planning review triggers, such as changes to legislation, research platforms or tools; funding body requirements; or institutional structure.
- Link the triggering of RDM planning reviews and updates to an existing governance process(es) such as ethics, grants management reporting, or risk management systems. Set dates for automatic triggering of the review of RDM planning, and automate the notification of the documentation author of updates to RDM planning.
- Where appropriate, and in accordance with confidentiality restrictions, use RDM planning key indicators for reporting to inform strategic planning of supporting infrastructure and/or resources.
- Where possible, link the RDM planning to the allocation and generation of digital data storage. Greater engagement of researchers with RDM planning and its documentation will offer insights into the need for and future planning of infrastructure requirements, such as data storage.

Researchers:

- Use equipment identifiers such as asset numbers in the RDM planning documentation to identify the equipment used to generate data.
- Link to or reference RDM planning in the laboratory information management system and electronic lab notebooks to provide contiguous data associations with a project.
- For primary materials, identify the storage location and the retention-and-disposal schedule for those primary materials.
- Where appropriate, and in accordance with confidentiality restrictions, make RDM planning documentation available as open access resources, for example, for collaboration.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

RECOMMENDATIONS

Recommendation 4: Meet the minimum standards for RDM planning.

Institutions:

- Define persistent identifiers (for example, DOI, Research Activity Identifier (RAiD), Research Organization Registry (ROR), ORCID and/or institution-defined identifiers) for RDM planning documentation, which clearly link the RDM planning with project data and publications referencing that data.
- Include management of primary materials as part of RDM planning.
- Provide support for the curation of data, provision of training, and assistance with RDM planning practice and documentation. Clearly identify the responsible group within the institution for RDM support and development to ensure a 'single source of truth'.
- Provide secure digital storage for RDM planning information.

Researchers:

- Know that RDM planning is required for all research projects and will be reviewed regularly. Consider RDM planning as part of high-quality research training for HDR students as well as researcher induction and periodic refreshers.
- Consider how to effectively review/assess/self-assess your RDM planning processes once in place.
- Be flexible with your RDM planning practice and its documentation to accommodate and adapt to the needs of different disciplines and align with the Code, and FAIR data and CARE principles.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Retention And Disposal

Recommendation 1: Adopt data standards and identifiers that link researchers and research projects.

Recommendation 2: Design integration tools for administrative and research management systems.

Recommendation 3: Develop a broadly applicable decision-making tool (matrix, decision tree, rubric) that can accommodate local (institution-specific) policy.

Recommendation 4: Ensure contemporary capture and propagation of metadata for research projects.

Open Research And Data Publication

Recommendation 1: Adopt an open-research or FAIR data statement as a driver for cultural change on sharing data/research, and enrol influential academics as champions of change.

Recommendation 2: Adopt best practices for data management, with open research, FAIR data and data sharing as end aspirations, where appropriate

Recommendation 3: Create and adopt uniform research data metrics as a mechanism for measuring the impact of sharing data and research.

Recommendation 4: Implement an institutional data repository with metadata capture capabilities.

The ARDC is enabled by NCRIS

RECOMMENDATIONS

Recommendation 5: Future-proof for open data.

- Embed data-sharing / FAIR data practices in institutional RDM policy.
- Embed sharing practice in ethics applications to foster future data sharing:
  - Encourage non-specific and extended consent with ethics committees so that data reuse can encompass wider research use than the original research idea. Ethics applications need to be future-proofed to enable data reuse. Currently, there is little understanding about the type of consent sought and what implications that has for sharing data later. Researchers need to better understand the pathways they create for themselves from the outset.
  - Store/capture ethics approvals within dataset metadata.
- Teach greater awareness of data licensing and the different licences. The *Open Data Handbook*[66] by the Open Knowledge Foundation provides a list of conformant licences[67]. The ARDC *Research data rights management guide*[47] is also comprehensive.
- To deter misuse, publish all data with a licence and a recommended citation. This includes the licensing of commercial data-compilation mechanisms, such as commercially controlled psychological tests and data visualisation software.
- As each licence, agreement or contract is potentially unique, it might be beneficial for the researchers to see a visible snapshot of what they can and cannot do with their data on, for example, a dashboard or even included in the DMP. Make it possible for MOUs, data-sharing agreements, and other agreements and contracts to be saved together alongside the research data.
- Establish a desirable baseline of requirements for data-sharing platforms.
- Automate a plain-English, decision-tree-style workflow to help researchers identify an appropriate licence for their data.
- Negotiate data-sharing agreements early in the project as these are critical to resolving ownership, IP and the rights to publish and reuse.
- Take appropriate measures to protect data from inappropriate loss, theft and misuse of access.
  - Accidental publication of sensitive data
    - Mitigation: Store the data classification as part of the metadata. Consider processes to
  - Misuse of data
    - Mitigation: Publish data with a licence and a recommended citation to deter misuse.
  - Theft of data
    - Mitigation: Publish data with a licence and a recommended citation to deter theft and provide proof of ownership and priority. Use existing systems to report offenders; for example, notifying publishers and funders of the theft. check for sensitive data before publication (for example, automation and through APIs).
- Develop guidance on how to confirm whether data is open/FAIR to encourage researchers to use open research data.

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

RECOMMENDATIONS

Sensitive Data

Recommendation 1: Adopt a classification scheme specific to sensitive data.

Recommendation 2: Define protections required for recognised classification levels of sensitive data.

Recommendation 3: Define sensitive-data classification levels by the severity of risk/consequences of mishandling or exposure of the data.

Recommendation 4: Provide guidance for researchers on the protections afforded through endorsed infrastructure solutions for sensitive data.

Recommendation 5: Provide guidance for ethics committees on appropriate approaches for managing sensitive research data.

NCRIS National Research Infrastructure for Australia
An Australian Government Initiative
The ARDC is enabled by NCRIS

Support, Training And Guidance

Recommendation 1: Secure approval and committed support from senior institution staff — approximately at DVC level — for the approach to providing support, training and guidance.

Recommendation 2: When designing the approach to providing support, training and guidance for RDM, engage in genuine dialogue with researchers about their needs.

Recommendation 3: Seek formal recognition, by senior staff or internal agreement, for a defined and coordinated set of roles and responsibilities for providing support, training and guidance. At a minimum, define roles and responsibilities for the research office, the IT department and the record-keeping department or equivalent.

Recommendation 4: Consider training approaches that leverage existing staff resources to build an institutional body of knowledge.

Recommendation 5: Consider reusing shared and nationally/internationally standard materials and resources as much as possible, adapting them for local relevance and applicability to all stakeholders within the institution.

Appraisal

Recommendation 1: Consider developing guidance on appraisal practices that reflect the institution's values and priorities.

Recommendation 2: Consider opportunities for identifying and supporting common appraisal trigger points, minimising manual interventions where possible.

The ARDC is enabled by NCRIS

CALLS TO ACTION

Calls to action identified by each working group are collected here for easy reference. Calls to action identify gaps or needs that can only be addressed by collaborative action between universities. These may form the basis of future collaborative projects.

Policy

Call to action 1: Institutions are encouraged to collectively describe details of relevant legislation that affects aspects of research data in their state or territory and other relevant jurisdictions.

RDM Planning

Call to action 1: Institutions are encouraged to reach agreement on the following considerations:

- What goes in the ethics system and what goes in the DMP, and in what order?
- Can RDM planning be linked with the institution's (sensitive) data classification framework?
- Can standard lab-operating procedures be used in RDM planning?
- How might we align tooling for RDM planning with FAIR data and CARE principles?
- How might we define a minimum standard with RDM planning tools?
- How might the framework better support our institutions in meeting the minimum standard?
- How might we share and standardise DMPs and/or planning documentation across collaborators and institutions?
- How can we use the scalable nature of the framework across institutions that are at very different stages (for example, taking DMPs from being an optional Word document to a technical, integrated, machine-readable and actionable plan), especially when many of the changes rely on funding to implement developments or solutions?

Retention and Disposal

Call to action 1: Institutions are encouraged to include data retention requirements in their data classification model, preferably standardised or aligned across institutions.

Call to action 2: Institutions are encouraged to collaborate to develop a common decision-making tool that can accommodate local (institution-specific) policy.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

CALLS TO ACTION

Sensitive Data

Call to action 1: Research institutions should continue adding to and developing the draft sensitive-data classification-level crosswalk.

Call to action 2: A shared data-classification scheme that can be implemented across institutions, and allow automation of data management processes, is an eventual goal. Institutions should consider how this classification scheme could be developed.

Call to action 3: Institutions should come together to discuss cybersecurity protections for sensitive data.

Appraisal

Call to action 1: Further to the recommendations of the retention and disposal element of the framework, universities should collaborate on a common set of information/metadata about research data collections to facilitate efficient data appraisal practices.

Call to action 2: Universities are encouraged to collaboratively agree on common elements for good-practice Australian research data governance, making use of existing governance approaches, where appropriate.

Call to action 3: The significant challenges, lack of clarity, and desire for common practice requires collaborative action to identify and define practical appraisal trigger points and guidance on how to integrate them into an institution's local systems and workflows.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

NEEDS FOR SHARED INFORMATION

Needs for shared information identified by each working group are collected here for easy reference. Needs for shared information are highlighted where participating universities identified a need for better understanding of approaches between institutions. This could be achieved either through projects which actively collect information from universities on these topics, or through increased sharing in national forums such as the Australian Research Data Management Community or other relevant communities of practice.

Active Data Management

A need for shared information 1: Institutions are encouraged to share with one another information about the active data management platforms they are using, including their benefits and limitations.

A need for shared information 2: institutions are encouraged to share with one another examples of the standards they are using for classifying their data.

Culture Change

A need for shared information: Institutions are encouraged to share with one another case studies and examples of culture change. When documenting a case study, consider the following questions:

- What models are you using to inform RDM culture change at your institution?
- How did your institution assess the current state and target state? Did you have experiences with Research Infrastructure Self-Evaluation (RISE), generic analyses such as Strengths, Weaknesses, Opportunities, and Threats (SWOT), or other approaches? How effective were they?
- How did your institution approach the assessment, articulation, and management of risk related to institutional RDM? Who did you consult when assessing and developing strategies for managing risk? What risks and potential harms related to RDM did you identify?
- How did you generate buy-in from your institution's leaders? How did this buy-in help your culture change initiative? Alternatively, did you need to work with an apparent lack of buy-in from leaders and what was your approach?
- How did you manage the resourcing of your culture change initiative?
- What rewards and incentives have you implemented?
- Did you find it helpful to have working definitions of particular terms at your institution? Terms and concepts related to change at institutions can be confusing to newcomers. What terms and concepts should be better defined in this document?

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

- Does your institution benchmark against (inter)national good practice? If so, how?
- How does your institution monitor the progress of cultural change? Does it differ from your initial stocktake or is it an extension of those activities?
- How has your institution solicited feedback? What combination of surveys, interviews, and other approaches worked for you? What other information or metrics have you used to assess the progress of change?
- How has your institution acted on feedback during a change process? How did you communicate those actions?

Policy

A need for shared information: Institutions are encouraged to share with one another examples of their post-implementation review processes.

Retention And Disposal

A need for shared information 1: institutions are encouraged to share with one another examples of predetermined timelines and trigger points appropriate for managing retention and disposal of research data.

A need for shared information 2: institutions are encouraged to share with one another examples of metadata and identifiers for system integration.

A need for shared information 3: institutions are encouraged to make available to one another use cases that show approaches for making decisions about data retention and disposal.

Open Research And Data Publication

A need for shared information: institutions are encouraged to share with one another examples of metrics for data, for physical samples, for research software, metrics around adoption of fair principles, and metrics for the 'real world' impact of research.

The ARDC is enabled by NCRIS

ACKNOWLEDGEMENTS

We would like to thank the following people for their invaluable contributions to this resource:

Editorial Committee

Adele Haythornthwaite, The University of Sydney
Anesh Nair, University of Canberra
Belinda Weaver, Griffith University
Francis Gacenga, University of Southern Queensland
Roxanne Missingham, Australian National University
Shawn Ross, Macquarie University.

ARDC Institutional Underpinnings Team

Nichola Burton
Lyle Winton
Frankie Stevens
Natasha Simons
Keith Russell

ARDC Program Sponsor

Adrian Burton

Expert Working Group Members

Adrian Chew, University of New South Wales
Ailie Smith, The University of Melbourne
Amanda Miotto, Griffith University
Antoinette Cass, Bond University
Ben McRae, Griffith University
Brock Askey, Bond University
Cameron Fong, The University of Sydney
Caroline Lobell, Bond University
Catherine Nicholls, Monash University
Catherine Tan, Griffith University
Cheryl Claridge, Federation University
Chris Stueven, The University of Melbourne
Claire Blankley, Edith Cowan University
Claire Carter, University of Wollongong
David Abramson, The University of Queensland
David Jung, University of New South Wales
David Poger, Microscopy Australia
Deanne Holmes, The University of Adelaide
Dianne Brown, Monash University
Dianne Jeffery, Curtin University
Em Johnson, Swinburne University
Erin Gallant, Australian National University
Eva Fisch, Monash University
Fiona Bradley, University of New South Wales
Franco Di Dio, Western Sydney University
Gary Allen, Griffith University
Gerry Devine, Macquarie University
Gnana Bharathy, ARDC
Harko Werkman, University of Tasmania
Helena Lynn, The University of Melbourne
Ian Welch, Edith Cowan University
Iftikhar Hayat, Charles Darwin University
Isabelle Bartkowiak-Theron, University of Tasmania
Jac Charlesworth, University of Tasmania
Jacky Cho, University of New South Wales
Jake Surman, University of New South Wales
Jane King, Swinburne University
Janice Chan, Curtin University
Jared Winton, The University of Melbourne
Jaye Weatherburn, The University of Melbourne
Jen Lyons, University of Wollongong

The ARDC is enabled by NCRIS

Expert Working Group Members (Continued)

Jen Rowland, Macquarie University

Jen Stanton, The University of Sydney

John Brown, Curtin University

Jonathan Schlossberg,
The University of Melbourne

Juan Cooper, Western Sydney University

Julie Gardner, Victoria University

Julie Toohey, Griffith University

Kandy White, Macquarie University

Kate Carruthers, University of New South Wales

Kate Croker, University of Western Australia

Kate Cumming, The University of Sydney

Kate Elder, Charles Darwin University

Katherine McClellan, Macquarie University

Katherine Shaw, University of Tasmania

Katina Toufexis, University of Western Australia

Katrina Trewin, Western Sydney University

Kay Steel, Federation University

Kaye England, University of Southern Queensland

Kim D'Costa, The University of Melbourne

Kim Robertson, Griffith University

Kyle Hemming, University of Canberra

Kylie Burgess, University of New England

Lauren Quinlan, Federation University

Leonie Sherwin,
University of Southern Queensland

Les Mitchell, University of Southern Queensland

Lesley Wyborn, ARDC

Liz Stokes, ARDC

Luke Burns, Edith Cowan University

Mandy Callow, University of Southern Queensland

Marc Fellman, Edith Cowan University

Matthew Bellgard,
Queensland University of Technology

Matthias Liffers, ARDC

Max Wilkinson, ARDC

Michael Lynch, University of Technology Sydney

Michelle Vincent, Swinburne University

Milica Symul, University of Canberra

Nick Rossow, Griffith University

Owen Griffiths, Griffith University

Owen O'Neill, Consultant

Patrick Splawa-Neyman, Monash University

Paul Hardaker, Western Sydney University

Paul Sou, Macquarie University

Penny Cross, University of Wollongong

Peta Scott, University of New England

Peter Bugeia, Macquarie University

Peter Neish, The University of Melbourne

Peter Smith, University of New South Wales

Philippa Frame,
Queensland University of Technology

Racheal Chesters, University of
Southern Queensland

Rob Watson, Federation University

Robin Burgess, University of New South Wales

Ryan McConville, Sydney University

Sandra Ennor, Monash University

Sarah Wright, University of Southern Queensland

Shannon Smith, Macquarie University

Sharron Stapleton, Griffith University

Stephanie Bradbury,
Queensland University of Technology

Stephen Dart, Monash University

Steve Love, The University of Adelaide

Susan Adams, Edith Cowan University

Susannah Tindall, Monash University

Vanessa Crosby, University of New South Wales

Vladimir Bubalo, Macquarie University

Willo Boniface, University of Wollongong.

The ARDC is enabled by NCRIS

APPENDIX 1: COMMENTS ON CULTURE CHANGE MODELS

This is a non-comprehensive list of models related to culture change, with key content notes.

Center For Open Science's Strategy

This strategy maps change to a position on Rogers's 'diffusion of innovations' curve. It envisions desired behaviours progressing from the base to the peak of a pyramid, from 'making it possible' to 'making it required'.

See: Nosek B. Strategy for culture change. Center for Open Science. Web log post 11 June 2019, viewed 26 May 2021, https://www.cos.io/blog/strategy-for-culture-change

The Four Building Blocks Of Change

The McKinsey model, The Four Building Blocks of Change, recommends reinforcing changes through formal mechanisms such as financial compensation, celebrating how changes made a difference to customer experience. The model starts with fostering an understanding and conviction with employees by clearly articulating what is being asked of people. This allows an alignment between beliefs and actions because believing in 'why' supports behaviour changes (and avoids cognitive dissonance and false consensus). They suggest that leaders develop a change story to help get the message out and serve as an effective influencing tool which is tailored to each group and describes how things will be better in the future state.

See: Basford T, Schaninger B. The four building blocks of change. McKinsey Quarterly. Web log post, 11 April 2016, viewed 2 June 2021, https://www.mckinsey.com/business-functions/organization/our-insights/the-four-building-blocks--of-change

Bridges Transition Model

The Bridges Transition Model is another people-focused model that recommends highlighting successes and tangible results to encourage and reinforce desired behaviours. This model has leaders focus on communicating the purpose and benefits of the change to stakeholders at the start.

See: https://wmbridges.com/about/what-is-transition/

Nudge Theory

Nudge theory recommends 'Clearly defining the change' and 'Using evidence as the best option' where leaders gather information about the current state and compare this to the desired future state. This allows stakeholders to see why the changes are needed and how the nominated path is the best course of action. Nudge theory presents change as a choice so that people choose to change, and paves the way for them to do this by showing them evidence that it is the best option. Then listening to and taking on board feedback whether or not the desired changes were chosen (i.e. listen, respond and be compassionate). This approach further recommends celebrating short-term goals and milestones to keep the momentum of change going.

See: Tahir U. What is nudge theory? Change Management Insight. 18 January 2020. http://changemanagementinsight.com/nudge-theory-in-change-management/

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

See: Mulholland B. 8 Critical change management models to evolve and survive, Process Street, web log post, 24 July 2017, viewed 9 June 2020, https://www.process.st/change-management-models/

Top-Down And Punitive Attempts To Impose Change

Dobbin and Kalev critique top-down and punitive attempts to impose change, using the example of diversity programs.

See: Dobbin F, Kalev A. Why diversity programs fail: And what works better. Harvard Business Review, Human Resource Management section, web log July-August 2016, viewed 14 September 2021: https://hbr.org/2016/07/why-diversity-programs-fail

Kübler-Ross Change Curve® - EKR Foundation

The Kübler-Ross change curve is a strategy that breaks down how people process change using the 5 stages of grief.

See: https://www.ekrfoundation.org/5-stages-of-grief/change-curve

The Satir Change Model

This model encourages leaders and managers to support employees going through a change process by providing the right support at the right time. The 5 stages they identify are late status quo, resistance, chaos, integration, and new status quo.

See: Mulholland B. 8 Critical change management models to evolve and survive. Process Street. Web log post, 4 July 2017, viewed 9 June 2020: https://www.process.st/change-management-models/

See: https://stevenmsmith.com/ar-satir-change-model/

KPMG's Culture And Conduct Assurance Framework

KPMG's Culture and Conduct Assurance Framework is a model that should keep an ongoing change process aligned with compliance requirements, particularly related to trigger points for evaluation. This model is instructive in that it includes continual evaluation of the process, but its details might be difficult to apply in the context of university RDM. This model should align with the assurance requirements appropriate for iteratively evaluating RDM culture change and conduct risk that would identify the need for further culture change efforts. It may also help to identify trigger points for undertaking evaluations.

See: https://assets.kpmg/content/dam/kpmg/cn/pdf/en/2020/01/kpmg-s-culsture-and-conduct-assurance-framework.pdf

‘Plan, Do, Check, Act’ Model

Deming’s ‘Plan, Do, Check, Act’ (PDCA) model (similar to Kolb’s reflective learning cycle) encourages a cyclical process to change. It starts with leaders identifying and analysing the problem or opportunity for change, developing a hypothesis, planning how to test it, and writing out the expected result. The hypothesis is then tested, and the outcomes of the test are reviewed, allowing participants to evaluate the solution and revise action as needed. If the plan did not produce the expected or desired results, then participants are directed back to the planning stage to try again.

See: Lucid Content Team. How to apply the Plan-Do-Check-Act (PDCA) model to improve your business. Lucidchart. Web log post, viewed 23 June 2021: https://www.lucidchart.com/blog/plan-do-check-act-cycle

See: Mind Tools Content Team. PDCA (Plan Do Check Act) Continually improving, in a methodical way. Mind Tools; 2020. Web log post viewed 24 August 2021: https://www.mindtools.com/pages/article/newPPM_89.htm

The Prosci Methodology

The Prosci Methodology is a widely used approach to organisational change that focuses on the ‘people’ side of change (as opposed to technology or infrastructure). It particularly draws out the importance of support from organisation leaders. One part of the Prosci Methodology is the ADKAR® (Awareness, Desire, Knowledge, Ability, Reinforcement) model, which recommends reinforcement to sustain change. This can be accomplished through positive feedback and recognition to encourage employees to keep following new processes and prevent reversion to old processes. The ADKAR model recommends leaders run practice or pilot sessions before the change is widely implemented to get feedback on how the planned changes will work and be received. It further suggests monitoring the impact of changes (for example, in terms of participant performance) immediately after the change, with participants providing constructive feedback. Progress should be tracked via goals and metrics set earlier at the start of the change process. If goals are not being met, then the process can be adjusted as needed to keep things on track.

See: https://www.prosci.com/methodology-overview and https://www.prosci.com/methodology/adkar

APPENDIX 2: INSTITUTIONAL DEFINITIONS OF RESEARCH DATA

University Of New England

Research data means data as facts, observations, computer results, measurements or experiences on which an argument, theory or test is based. Data may be numerical, descriptive or visual. Data may be raw or analysed, experimental or observational. Data include records that are necessary for the reconstruction and evaluation of reported results of research and the events and processes leading to those results, regardless of the form or the media on which they may be recorded.

Macquarie University

Data means research data, which includes primary materials or information held in any digital format or media, or anything that can be digitised, on which an argument, theory, test or hypothesis, or another research output is based. Data may also include other 'digital research objects' such as analytical code that support research outcomes. Research data may be in the form of facts, observations, images, computer program results, recordings, questionnaires/surveys, biographies, audio files, physical specimens or artefacts, measurements, experiences or various other forms. Data may be numerical, descriptive, visual or tactile and could be raw, cleaned or analysed. Data referred to in this Policy does not include the information about research performance or statistical research data which is used by Macquarie University for planning and budget purposes or that which is reported to government agencies, for example, Excellence in Research for Australia.

Swinburne University Of Technology

Research data may take many different forms, and may include primary research data, spreadsheets, questionnaires, notebooks, photographs, films, models, test responses and samples. The code used to generate or analyse the data may also be included.

University Of New South Wales

Research data are the original sources or material that have been created or collated to conduct a research project. They can be digital or non-digital. The response to a particular research question is based on the analysis of research data.

University Of Wollongong

The data, records, files or other evidence, irrespective of their content or form (for example, in print, digital, physical or other forms), that comprise research observations, findings or outcomes, including primary materials and analysed data. Research data referred to in this policy relates to data generated in research projects and is to be distinguished from the information about research performance and statistical research data which is used for planning and budgeting purposes.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

Victoria University

Research data: Facts, observations or experiences on which an argument, theory or test is based. Data may be numerical, descriptive or visual. Data may be raw or analysed, experimental or observational. Data includes: laboratory notebooks; field notebooks; primary research data (including research data in hardcopy or in computer readable form); questionnaires; audiotapes; videotapes; models; photographs; films; test responses. Research collections may include slides; artefacts; specimens; samples. Provenance information about the data might also be included: the how, when, where it was collected and with what (for example, instrument). The software code used to generate, annotate or analyse the data may also be included.

APPENDIX 3: FIELDS INCLUDED IN DATA MANAGEMENT PLANS FROM CONTRIBUTING UNIVERSITIES

Table 5. *Information to be included in a data management plan, as suggested by Bond University, Edith Cowan University, the University of Southern Queensland and the University of Tasmania*

	BOND UNIVERSITY	EDITH COWAN UNIVERSITY	UNIVERSITY OF SOUTHERN QUEENSLAND	UNIVERSITY OF TASMANIA
Project title / ID	X	X	X	X
Project description		X		X
FOR and SEO codes	X			X
Chief Investigator details	X	X	X	X
Supervisor details (if applicable)	X	X		X
Ethics approval number	X	X		X
Grant funding details (if applicable)	X	X	X	
Estimated costs of data management storage			X	
Data collection and/or project dates	X	X	X	X
Sensitive data	X	X	X	X
Data Identifiability		X		X
Restrictions on data or ownership / IP issues	X		X	X
Digital data storage location	X	X	X	X
Data processing requirements/resources	X		X	
Digital data size and format	X	X	X	X
Physical data storage and type	X	X		
Retention period and/or dates	X	X	X	
Data access (project team)	X	X		X
Data sharing or institutional repository	X	X	X	X
Future use of data	X	X	X	

X = included. Blank space = not included.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

APPENDIX 4: DEFINITIONS USED IN DISCUSSIONS ABOUT DATA AND RETENTION

The following commonly used terms were defined to allow group discussions to proceed based on a shared understanding of the topic in question.

Archiving

A curation activity that ensures that data are properly selected, stored, and can be accessed, and for which logical and physical integrity are maintained over time, including security and authenticity.

Copyright

A legal right over intellectual property belonging to the creator of the work. A copyright holder may use a licence to grant other people rights in the protected material, perhaps subject to specified restrictions.

Curation

The activity of managing and promoting the use of data from their point of creation to ensure that they are fit for contemporary purpose and available for discovery and reuse.

Data

Data may be thought of as unprocessed atomic statements of fact. It very often refers to systematic collections of numerical information in tables of numbers such as spreadsheets or databases.
Source: https://opendatahandbook.org/glossary/en/

Data Publishing

The publication of research data on the web, whether for sharing, to satisfy grant or publisher requirements for openness, or otherwise. Publication alone does not make the data 'open' — without associated metadata, documentation and software code, and clear licensing, the data may well be available but not usable. In some cases, university repositories take responsibility for the creation of a dataset record which includes the requisite 'wrappers' that enable sharing and reuse.

Data Quality

A measure of the "useableness" of data.

Intellectual Property

Intellectual property refers to creations of the mind, such as inventions; literary and artistic works; designs; and symbols, names and images used in commerce. IP is protected in law by, for example, patents, copyright and trademarks, which enable people to earn recognition or financial benefit from what they invent or create. Source: https://www.wipo.int/about-ip/en/

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

APPENDIX 4: DEFINITIONS USED IN DISCUSSIONS ABOUT DATA AND RETENTION

Licence

A legal instrument by which a copyright holder may grant rights over the protected work. Data and content is open if it is subject to an explicitly-applied licence that conforms to the Open definition[67]. A range of standard open licences are available, such as the Creative Commons CC-BY licence, which requires only attribution.

Mediated Access

Mediated access refers to research publications and data that are not open, but access can be requested of a mediator. Levels of mediated access exist, whereby, for example:

- access is managed by a researcher
- access is managed by a research team, or their delegate
- access is managed by a data repository
- access is managed by an institution.

Open Access

Open access is a set of principles and a range of practices through which research outputs are distributed online, free of cost or other access barriers.

Open Data

Short version: Structured data that are accessible, machine-readable, usable, intelligible, and freely shared.

Long version: Data is open if it can be freely accessed, used, modified and shared by anyone for any purpose — subject only, at most, to requirements to provide attribution and/or share-alike. Specifically, open data is defined by the Open definition[67] and requires that the data be:

A. Legally open: that is, available under an open (data) licence that permits anyone freely to access, reuse and redistribute

B. Technically open: that is, that the data be available for no more than the cost of reproduction and in machine-readable and bulk form.

Open Research

Open research comprises not just open access to publications and data sharing that is 'as open as possible, [only] as closed as necessary', but also documentation of data (metadata – for example, record-, dataset-, and project-level metadata; administrative metadata; data provenance including version history) and research processes ('paradata' – for example, data capture and processing techniques and methods), sufficient to allow the assessment, testing, and reproduction of research results.

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

APPENDIX 4: DEFINITIONS USED IN DISCUSSIONS ABOUT DATA AND RETENTION

Some definitions:

- https://en.wikipedia.org/wiki/Open_research
- https://osc.cam.ac.uk/open-research
- www.exeter.ac.uk/research/openresearch/about/explained/
- https://en.wikipedia.org/wiki/Open_science

Open Science

The movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of an inquiring society, amateur or professional. 'Open research' is preferred as it is more inclusive of disciplines such as the humanities.

Open Source

Software for which the source code is available under an open licence.

Preservation

Long-term preservation of present-day datasets is more difficult to ensure owing to uncertainty about the future of file formats, computer architectures, storage media and network connectivity. Projects that put particular stress on data preservation take a variety of approaches to solving these problems. Source: https://opendatahandbook.org/glossary/en/terms/data-preservation/

Research Sharing And Reuse

Data can be reused again and again, in ways that were never envisaged when it was collected, provided only that the data-holder makes it available under an open licence to enable reuse.

NCRIS National Research Infrastructure for Australia
The ARDC is enabled by NCRIS

REFERENCES AND NOTES

1 King R. Predictions 2022: Data can help address the world's biggest challenges - 5 experts explain how. New York (US): World Economic Forum; 19 January 2022. Available from: https://www.weforum.org/agenda/2022/01/data-can-help-tackle-global-challenges-2022-predictions/

2 NHMRC. Australian code for the responsible conduct of research. [Internet]. 2018. Available from: https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018

3 National Collaborative Research Infrastructure Strategy: https://www.education.gov.au/national-collaborative-research-infrastructure-strategy-ncris

4 FAIR data principles: https://ardc.edu.au/resources/aboutdata/fair-data/

5 CARE principles for data: https://www.gida-global.org/care

6 Australian Research Data Management https://sites.google.com/ardc.edu.au/rdmcommunity/home

7 Relevant communities of practice: https://ardc.edu.au/get-involved/communities-and-groups/

8 Digital version of this report: https://doi.org/10.5281/zenodo.6392340

9 Monash University. Health data governance framework. 2020. Available from: https://doi.org/10.26180/5f3335c306a3b

10 Monash University. Developing and implementing Monash Secure eResearch Platform. 2021. Available from: https://ardc.edu.au/wp-content/uploads/2022/03/02z-Active-Data-Management-attachment-Case-Study-Monash-SeRP.pdf

11 UNSW. Storing your research data. August 2020. Available from: https://research.unsw.edu.au/data-storage-and-tools

12 ARDC Institutional Underpinnings Element: Support, Training and Guidance. 30 March 2022. Available from: https://ardc.edu.au/wp-content/uploads/2022/03/09-Support-Training-and-Guidance.pdf

13 Nosek, B. Strategy for culture change, Center for Open Science, web log post 11 June 2019, viewed 26 May 2021. Available from: https://www.cos.io/blog/strategy-for-change

14 Prosci. Primary sponsor's role and importance. No date. Viewed online. 21 December 2022. Available from: https://www.prosci.com/resources/articles/primary-sponsors-role-and-importance

15 Akers, K, Doty, J. Disciplinary differences in faculty research data management practices and perspectives. International Journal of Digital Curation. 2013;8:5–26. Available from: https://doi.org/10.2218/ijdc.v8i2.263

16 Rans, J, Whyte, A. Using RISE, the Research Infrastructure Self-Evaluation Framework. v.1.1 Edinburgh: Digital Curation Centre. 2017. Available from: https://www.dcc.ac.uk/guidance/how-guides/RISE

17 Kübler-Ross Change Curve®. Available from: https://www.ekrfoundation.org/5-stages-of-grief/change-curve/

18 Mulholland, B. 8 Critical change management models to evolve and survive, Process Street, web log post, 24 July 2017, viewed 9 June 2020. Available from: https://www.process.st/change-management-models/

19 Smith, SM. The Satir change model. 4 October 1997. Available from: https://stevenmsmith.com/ar-satir-change-model/

20 Dobbin, F, Kalev, A. Why diversity programs fail: And what works better. Harvard Business Review, Human Resource Management section, web log July-August 2016. Viewed 14 September 2021. Available from: https://hbr.org/2016/07/why-diversity-programs-fail

21 Lucid Content Team. 7 Fundamental change management models. Lucidchart. Web log post. 2018. Viewed 23 June 2021. Available from: https://www.lucidchart.com/blog/7-fundamental-change-management-models

22 Lucid Content Team. How to apply the Plan-Do-Check-Act (PDCA) model to improve your business. Lucidchart. Web log post. 2019. Viewed 23 June 2021. Available from: https://www.lucidchart.com/blog/plan-do-check-act-cycle

23 The Carpentries. Assessment and impact: measuring the impact of our workshops and trainings. Available from: https://carpentries.org/assessment/

24 See: Australian Government. Business risks. Available from: https://business.gov.au/risk-management/risk-assessment-and-planning/business-risks

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

REFERENCES AND NOTES

25 Taleb NN. Antifragile: things that gain from disorder. Random House; 2012.

26 Based on the spectrum of public participation proposed by the International Association of Public Participation (IAP2) which is available from: https://iap2.org.au/wp-content/uploads/2020/01/2018_IAP2_Spectrum.pdf

27 Basford T, Schaninger B. The four building blocks of change. McKinsey Quarterly. Web log post, 11 April 2016. Viewed 10 September 2021. Available from: https://www.mckinsey.com/business-functions/organization/our-insights/the-four-building-blocks--of-change

28 International Association of Public Participation. The spectrum of public participation. Available from: https://iap2.org.au/resources/spectrum/

29 Kelly W. The emerging impact landscape. The Research Whisperer blog. 2019. Available from: https://researchwhisperer.org/2019/05/21/the-emerging-impact-landscape/

30 Vitae Connections Week 2021: https://connectionsweek2021.vitae.ac.uk/agenda

31 Atlassian. Customer journey mapping | Atlassian team playbook. Available from: https://www.atlassian.com/team-playbook/plays/customer-journey-mapping

32 Center for Open Science. TOP Guidelines. Available from: https://www.cos.io/our-services/top-guidelines

33 Tertiary Education Quality and Standards Agency (TEQSA). Guidance Note: Research and research training. 5 July 2018. Available from: https://www.teqsa.gov.au/latest-news/publications/guidance-note-research-and-research-training

34 A good example of this is the UNSW data and information governance framework. Available from: https://www.datagovernance.unsw.edu.au/

35 Office of the Australian Information Commissioner. Privacy in your state. Available from: https://www.oaic.gov.au/privacy/privacy-in-your-state

36 Therapeutic Goods Administration. Guideline for good clinical practice. Available from: https://www.tga.gov.au/publication/note-guidance-good-clinical-practice

37 ISO 14155: 2011. Available from: https://www.iso.org/standard/45557.html

38 ARC. Research Data Management. Available from: https://www.arc.gov.au/policies-strategies/strategy/research-data-management

39 International Committee of Medical Journal Editors. Responsibilities in the submission and peer-review process. Available from: http://www.icmje.org/recommendations/browse/roles-and-responsibilities/responsibilities-in-the-submission-and-peer-peview-process.html

40 Examples can be found at: https://www.gs.unsw.edu.au/policy/framework/index.html and https://documents.uow.edu.au/content/groups/public/@web/@gov/documents/doc/uow058645.pdf

41 Eltis. RASCI matrix definition. Available from: https://www.eltis.org/glossary/rasci-matrix

42 Australian Research Data Commons. Institutional research data management policies and procedures. Zenodo; 2021. Available from: https://doi.org/10.5281/zenodo.5784765

43 Shafee T. ARDC Institutional Underpinnings policy template (0.12.1). Zenodo; 2022. Available from: https://doi.org/10.5281/zenodo.6859513

44 Macquarie University. Research data management policy. Available at: https://policies.mq.edu.au/document/view.php?id=300

45 University of Wollongong. Research data management policy. 7 February 2019. Available from: https://documents.uow.edu.au/content/groups/public/@web/@gov/documents/doc/uow116802.pdf

46 See, for example: https://www.datagovernance.unsw.edu.au/data-classification-standard

47 Australian Research Data Commons. ARDC Research data rights management guide. Zenodo; 2019. Available from: https://doi.org/10.5281/zenodo.5091580

48 Office of the Australian Information Commissioner. Privacy in your state. Available from: https://www.oaic.gov.au/privacy/privacy-in-your-state

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

REFERENCES AND NOTES

49 NHMRC. Management of data and information in research: A guide supporting the Australian Code for the Responsible Conduct of Research. NHMRC; 2019. Available from: https://www.nhmrc.gov.au/sites/default/files/documents/attachments/Management-of-Data-and-Information-in-Research.pdf

50 NHMRC. National statement on ethical conduct in human research. NHMRC; 2018. Available from: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018

51 AIATSIS. Code of ethics for Aboriginal and Torres Strait Islander research. AIATSIS; 2020. Available from: https://aiatsis.gov.au/research/ethical-research/code-ethics

52 TEQSA. Higher Education Standards Framework (Threshold Standards) 2021. TEQSA; 2021. Available from: https://www.teqsa.gov.au/higher-education-standards-framework-2021

53 ARDC. FAIR data guidelines for project outputs. ARDC; 2020. Available from: https://ardc.edu.au/about_us/policies-and-guidelines/fair-data-guidelines-for-project-data-outputs

54 Harned B. What is a RACI chart? How to use RACI to define project roles. Blog post. Teamgantt; 15 March 2022. Available from: https://www.teamgantt.com/blog/raci-chart-definition-tips-and-example

55 cOAlition S. Plan S. Available from: https://www.coalition-s.org/

56 UNESCO. Open science. Available from: https://en.unesco.org/science-sustainable-future/open-science/recommendation

57 Griffith University Open Research Statement: https://www.griffith.edu.au/research/about-griffith-research

58 Open research at Melbourne University: https://library.unimelb.edu.au/open-scholarship

59 Council of Australian University Librarians. Statement on open scholarship. 2019. Available from: https://www.caul.edu.au/programs-projects/advancing-open-scholarship-fair/statement-open-scholarship

60 Counter. COUNTER Code of Practice for Research Data Usage Metrics release 1. Available from: https://www.projectcounter.

61 IRUS: https://irus.jisc.ac.uk/r5/

62 Australian Data Archive: https://ada.edu.au/

63 AGU. Generalist repository comparison chart. Available from: https://www.agu.org/-/media/Files/Publications/Generalist-Data-Repository-Grid.pdf

64 ARDC Australian and New Zealand Data Quality Interest Group: https://sites.google.com/ardc.edu.au/australian-data-quality-ig

65 CoreTrustSeal – Core Trustworthy Data Repositories: https://www.coretrustseal.org

66 Open Knowledge Foundation. Open data handbook: http://opendatahandbook.org/guide/en/what-is-open-data/

67 Open Knowledge Foundation. Open definition: https://opendefinition.org/licenses/

68 ARDC Research data rights management guide: https://doi.org/10.5281/zenodo.5091579

69 ARDC Classification crosswalk: https://ardc.edu.au/wp-content/uploads/2022/03/08z-Sensitive-Data-attachment-Classification-Matrix.pdf

70 ARDC Australian Sensitive Data Interest Group: https://sites.google.com/ardc.edu.au/sensitive-data-cop/home

71 UNESCO and the International Council on Archives. Universal Declaration on Archives. 10 November 2011: https://www.ica.org/en/universal-declaration-archives

72 The Five Safes: http://www.fivesafes.org/

73 ARDC project: Improving Indigenous Research Capability: https://ardc.edu.au/project/improving-indigenous-research-capability/

NCRIS National Research Infrastructure for Australia An Australian Government Initiative
The ARDC is enabled by NCRIS

FEEDBACK

We welcome your feedback on this framework. Please email contact@ardc.edu.au with any comments or questions.

ABOUT THE AUSTRALIAN RESEARCH DATA COMMONS

The Australian Research Data Commons (ARDC) enables the Australian research community and industry access to nationally significant, data intensive digital research infrastructure, platforms, skills and collections of high quality data.

The ARDC is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy (NCRIS).

The ARDC is enabled by NCRIS

Australian Research Data Commons

CONTACT

ardc.edu.au
+61 3 9902 0585
contact@ardc.edu.au

FOLLOW

@ardc_au
australian-research-data-commons
subscribe to our newsletter

NCRIS
National Research Infrastructure for Australia
An Australian Government Initiative

The ARDC is enabled by NCRIS