Management of Data and Information in Research

A guide supporting the *Australian Code for the Responsible Conduct of Research*

Contact us

To obtain information regarding NHMRC publications or submit a copyright request, contact:

E: nhmrc.publications@nhmrc.gov.au
P: (02) 6217 9000

Contents

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

iii

1.   Introduction

This guide supports the implementation of the *Australian Code for the Responsible Conduct of Research* (the Code), which articulates the broad principles and responsibilities that underpin the responsible conduct of Australian research.

In particular, this guide is intended to assist institutions and researchers to adhere to relevant principles of the Code, including:

• Principle 2, 'Rigour in the development, undertaking and reporting of research', which requires that research be characterised by attention to detail and robust methodology and that researchers avoid or acknowledge biases.

• Principle 3, 'Transparency in declaring interests and reporting research methodology, data and findings', which requires researchers to share and communicate research methodology, data and findings openly, responsibly and accurately.

• Principle 7, 'Accountability for the development, undertaking and reporting of research' so as to comply with relevant legislation, policies and guidelines and ensure good stewardship of public resources used to conduct research.

The responsible conduct of research includes within its scope the appropriate generation, collection, access, use, analysis, disclosure, storage, retention, disposal, sharing and re-use of data and information. Institutional policies that are developed to govern the conduct of research require proper attention to ethics guidelines, privacy legislation and guidelines, other relevant laws, regulations and guidelines, as well as research discipline-specific practices and standards and models for best practice.

The Code and this guide apply to all research conducted under the auspices of Australian institutions. These institutions vary in size, maturity, experience and organisational structure. They range from large and complex universities, to small privately funded institutes. Accordingly, it is acknowledged that different institutional policies and processes are capable of fulfilling the aim of this guide and attempts have been made to ensure that there are appropriate options for flexibility in its application.

2.   Responsibilities of institutions

Research institutions have a responsibility to develop and implement policies and provide facilities and processes for the safe and secure storage and management of research data and primary materials in order to:

• allow for the justification and verification of the outcomes of research

• maximise the potential for future research

• minimise waste of resources of value to researchers and the wider community.[1]

> **Under the Code, institutions have the responsibility to:**
>
> R3   *Develop and maintain the currency and ready availability of a suite of policies and procedures which ensure that institutional practices are consistent with the principles and responsibilities of the Code.*

---

1   This objective is also consistent with accepted standards requiring that data and information be findable, accessible, interoperable and re-usable. See https://www.ands.org.au/working-with-data/fairdata and https://www.fair-access.net.au/fair-statement

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

1

Institutional policy 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.

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

2.1 Provision of training for researchers

Institutions must provide ongoing training and education that promotes and supports responsible research conduct and assists all researchers and those in other relevant roles to follow the institution's data management policies and other relevant disciplinary-specific policies.

Under the Code, institutions have responsibilities to:

R4 *Provide ongoing training and education that promotes and supports responsible research conduct for all researchers and those in other relevant roles.*

R5 *Ensure supervisors of research trainees have the appropriate skills, qualifications and resources.*

2.2 Ownership, stewardship and control of research data and primary materials

Institutional policies should provide guidance about ownership, stewardship and control of research data and primary materials, both during a research project and after the project has been finalised.

Ownership of research data and primary materials can be difficult to determine, especially when the research involves multiple researchers or externally sourced data or information, and may need to take account of the character or status of the primary materials. An institution may choose to assert ownership of research data, or to grant ownership to researchers. Institutional policies should clarify the criteria that will be used to determine the status of research data and primary materials in these circumstances.

With respect to the ownership of data and information used in or generated by research involving Aboriginal and Torres Strait Islander peoples and communities, institutions or researchers may hold data or information; however, they should not make decisions about the access to or reuse of this data or information without proper consultation with its Indigenous owners, if any.

Institutional policies should cover cases where researchers move between institutions or employers and where research data are held outside of Australia. Agreements covering ownership, stewardship and control of this data or information should be reviewed whenever there is movement or departure of research staff.

As a general rule, the most satisfactory arrangement will be that the materials and data retained at the end of a project are the property of the institution that hosted the project, another institution with an interest in the research, or a central repository. Institutional policies should clarify any exceptions to these arrangements.

Institutional policies on these matters should not unnecessarily impede the normal use of research data and primary materials by researchers for research and scholarly purposes, including their sharing and communication.

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

2

Upon the determination of ownership of research data, a relevant licence should be considered for the purpose of clarifying the status of the data for re-use by third parties.

In some instances, research may be conducted partly or entirely using data or materials that are owned by another party. In such cases, neither the institution nor the researcher can assert ownership. Therefore, to meet the requirements outlined in this guide, both the institution and the researcher should document the source of the data or materials and describe the access arrangements. Such arrangements should be in place to allow justification and verification of the outcomes of research.

2.3   Storage, retention and disposal

Institutional policies should address the storage, retention and disposal of all research data, whether held in an institutional repository or externally.

Retaining the research data is important because it may be all that remains of the research work at the end of the project.

The storage, retention and disposal of research data should:

- be consistent with any copyright or licensing arrangements that are in place
- be in accord with research discipline-specific practices and standards
- comply with relevant privacy, ethical and publication requirements
- comply with other relevant laws, regulations and guidelines.

The period for which data should be retained should be determined by prevailing standards for the specific type of research and any applicable state, territory or national legislation. In general, the minimum period for retention of research data is 5 years from the date of publication. However, for any particular case, the period for which the data should be retained should be determined by the specific type of research, subject to any applicable state, territory or national legislation. For example:

- for short-term research projects that are for assessment purposes only, such as research projects completed by students, retaining research data for 12 months after the completion of the project may be sufficient
- for most clinical trials, retaining research data for 15 years or more may be necessary
- for areas such as gene therapy, research data must be retained permanently (e.g. data in the form of patient records)
- if the work has community, cultural or historical value, research data should be kept permanently, preferably within a national collection.

Institutional policies should clarify the requirements for short and long-term storage of research data and how any disposal of data and information is to be undertaken and recorded.

Good archival practice includes scheduled review of items in long-term storage.

2.4   Safety, security and confidentiality

Institutions must have policies on the ownership of, and access to, databases and archives that is consistent with confidentiality obligations, legislation, privacy principles and other guidelines.

These policies must require that:

- researchers are informed of relevant confidentiality agreements and restrictions on the use of research data
- computing systems are secure
- information technology personnel understand their responsibilities for network security and access control
- those holding primary material, including electronic material, understand their responsibilities for security and access.

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

3

2.5   Access by interested parties

Institutional policies should describe how to make research data, including the outputs of research, available to interested parties both within and outside of the institution, giving particular consideration to licensing and access arrangements. Institutional policies should address options for sharing data via open access and via mediated access (i.e. access to data or information with the assistance of a data custodian or other authorised person).

> **Under the Code, institutions will:**
>
> R8   *Provide access to facilities for the safe and secure storage and management of research data, records and primary materials and, where possible and appropriate, allow access and reference.*

Licensing provides a standardised way for researchers and institutions to share research data with others and to govern subsequent use of that data. When considering licensing for this purpose, the least restrictive option, such as a Creative Commons Attribution licence[2], is encouraged.

Where the sharing of research data has been requested and access has been refused, the reasons for not sharing the data should be transparent and justifiable.

For guidance on intellectual property matters, institutions and researchers should consult institutional policies and/or seek further advice from appropriate professionals. For research involving Aboriginal and Torres Strait Islander peoples, consult the relevant guidelines.[3]

2.6   Facilities

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.

Research institutions offering facilities for the storage of personal or sensitive data or information should have a policy describing their responsibilities as data custodians for the security of and access to the data and information.

The policy should also address control or ownership of data facilities and archives in which data or information is stored.

Research data may be published in international, national, or discipline-based repositories, such as international databanks, in addition to institution-based storage or archiving. Institutions should consider maintaining a record of the research data generated by their researchers and where it is stored.

3.   Responsibilities of researchers

In accordance with the Code, researchers must adhere to their institution's policies related to management of data and information, relevant laws, regulations and guidelines, and research discipline-specific practices and standards.

The responsibilities of researchers with respect to management of data and information in research should be clear from the beginning of a research project. The development of a data management plan for this purpose is strongly encouraged.

---

2   https://creativecommons.org.au

3   https://www.nhmrc.gov.au/research-policy/ethics/ethical-guidelines-research-aboriginal-and-torres-strait-islander-peoples

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

4

A data management plan should be developed as early as possible in the research process and should include, but not be limited to, details regarding:

- 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, and
- what information from the data management plan, if any, needs to be communicated to potential participants.[4]

While it may not be practical to keep all the primary material (such as ore, biological material, questionnaires or recordings), durable records derived from them (such as assays, test results, transcripts, and laboratory and field notes) must be retained and accessible.

Planning for the management of data, particularly for retention of research data, should include considerations of practicality and cost.

The guidance provided in this section should be read in conjunction with section 2 (Responsibilities of institutions) of this document.

> ***Under the Code, researchers will:***
>
> *R22 Retain clear, accurate, secure and complete records of all research including research data and primary materials. Where possible and appropriate, allow access and reference to these by interested parties.*

If not otherwise clarified in institutional policy, researchers should:

- retain clear, accurate, secure and complete records of all research data and primary materials
- retain and be able to produce on request all relevant approvals, authorisations and other administrative documents, such as ethics and financial approvals, receipts and consent forms
- where possible and appropriate, allow access to research data and primary materials, in particular, to enable to facilitate the sharing of research data. This access should be facilitated by the use of indexes or catalogues of data and information generated, accessed and used during the research
- respect any project-specific conditions of consent or confidentiality obligations
- adhere to project-specific protocols that require measures beyond those required by institutional policy or relevant laws, regulations and guidelines, or research discipline-specific practices and standards
- report any inappropriate use of or access to or loss of data, in accordance with applicable institutional policies and, where relevant, other reporting schemes such as the Notifiable Data Breaches scheme[5]
- ensure that agreements are in place to govern circumstances in which researchers leave the project or move from one institution to another during the course of the project
- ensure that agreements are in place between institutions for managing responsibilities set out in this guide for data and information in multicentre or collaborative research projects.

---

4   Guidance on data management plans is provided in the *National Statement on Ethical Conduct in Human Research* at Chapter 3.1, Element 4.

5   https://www.oaic.gov.au/privacy-law/privacy-act/notifiable-data-breaches-scheme

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

5

In order to optimise project efficiency and avoid information loss and duplication, researchers should employ good management practices. These practices vary across disciplines, but the essential elements include:

- stable storage formats and regular backup to a source external to an individual computer
- version control and other relevant mechanisms for datasets, algorithms, models and software configuration management
- workflow documentation with provenance information for instruments (use and calibration) and software used
- adherence to appropriate national and international standards for scientific terminology and information encoding.

Research data can be the subject of Freedom of Information requests, and in such circumstances, there is an expectation that any information that is delivered will be provided in an understandable format and state.

3.1   Retention and publication

The central aims of retention of data and information are to enable the justification of outcomes of the research and the facilitation of sharing of research data.

Researchers have primary responsibility for deciding which research data and primary materials are candidates for long-term retention and wider accessibility.

In addition to legal requirements and the requirements of funders, government bodies and publishers, the following criteria should be considered in deciding which research data and primary materials should be retained:

- uniqueness and non-replicability
- reliability, integrity, and usability
- relevance to a known research initiative or collection
- community, cultural or historical value
- economic benefit.

Researchers should consider appropriate approaches to maximising the benefits of valuable data and information in the context of any required or reasonable restrictions on sharing the data or information, such as the character of the consent provided,[6] permissions granted by any data custodians and the requirements of privacy laws. Researchers should be sensitive to the tension created by these competing imperatives.

In addition to standard publication requirements, options for researchers include publishing or making their research data available through data centres, national and international collections, or through online repositories maintained by institutions and research communities.

Researchers should adhere to established national and international standards for data description and structuring to facilitate tracking of references. These standards include using Digital Object Identifiers for datasets, ORCID IDs for researchers, and standard terminology for scientific concepts.

Published research data generally require some kind of online description (i.e. metadata) and should be findable, accessible, interoperable, and re-usable, both manually and with automated tools. This requires researchers to include appropriate context (descriptive, technical, methodological, access, and provenance information) either within the data structure or in separate metadata records for the research data.

Researchers should consider the options for licensing of research data in order to provide clear parameters around the use and re-use of this data. When considering licensing for this purpose, the least restrictive option, such as a Creative Commons Attribution licence, is encouraged. Researchers should be prepared to justify the use of more restrictive controls.

---

6   For information on consent and privacy considerations in research see the *National Statement on Ethical Conduct in Human Research*.

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

6

3.2   Managing confidential and other sensitive information

Researchers must exercise care in handling confidential or other sensitive information used in or arising from a research project. Research data and information to which obligations of confidentiality or other sensitivities may apply commonly fall into one of the following categories:

- data or information that is commercial-in-confidence or that is inherently confidential and which has been provided in confidence (e.g. secret and sacred religious or cultural practices, or information on the location of vulnerable species)
- sensitive data or information subject to privacy legislation (e.g. identifiable human medical/health and personal data or information)
- data or information subject to classification regimes and other controls (e.g. national security information, police records or information and primary materials subject to export controls).

Researchers must ensure that the security and privacy measures[7] that are used for research data and primary materials are proportional to the risks associated with the confidentiality or sensitivities of these data and materials. These measures relate to storage, access and sharing of the data and information and should be recorded in a data management plan.

Sensitive research data may be appropriately shared through mediated access arrangements and the application of a risk assessment framework.

3.3   Acknowledging the use of others' data

The Code and *Authorship: A guide supporting the Australian Code for the Responsible Conduct of Research* require that the work of others is appropriately referenced and cited in the presentation, publication or sharing of research. This principle applies to all data and information used as an input to a research project. In referencing and citing the work of others, researchers should follow accepted norms and standards for scholarly literature and can reasonably expect that their work is acknowledged by others.

3.4   Engagement with relevant training

Researchers should engage with relevant training and education provided by or through their institution.

4.   Breaches of the Code

Institutions should manage and investigate concerns or complaints about potential breaches of the Code in accordance with the *Guide to Managing and Investigating Potential Breaches of the Australian Code for the Responsible Conduct of Research* (the Investigation Guide).

Breaches of the Code that are related to management of data and information in research (see also Section 2.1 of the Investigation Guide) include, but are not limited to:

- falsification of research data or primary materials
- fabrication of research data or primary materials
- failure to notify the institution and relevant authorities in a timely manner of a data breach or instance of inappropriate access to data held by the researcher
- failure to retain clear, accurate, secure and complete records of all research including research data and primary materials
- failure to adhere to the conditions of any institutional policy or project-specific approvals that relate to the retention, sharing or destruction of research data or primary materials

---

7   Guidance on managing and sharing confidential data and information is provided in the *National Statement on Ethical Conduct in Human Research*. See Additional Resources for more information on recommended security and privacy measures.

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

7

- selective retention of research data or primary materials so as to hinder the verifiability of a research output or access request
- failure to apply appropriate security controls to research data or primary materials
- failure to obtain necessary approvals or acting inconsistently with a condition of any approval granted in relation to the management of research data or primary materials.

In assessing the seriousness of a breach, any confidentiality obligations or other sensitivities that apply to the research data or information should be considered (Section 3.2, above), in addition to other relevant factors (see Investigation Guide, Section 2.2).

Additional Resources

Resources referred to in this guide, or supporting the principles or responsibilities outlined in the guide include:

Australian Government Research Ethics and Research Integrity Guidelines

https://nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2007 [*Australian Code for the Responsible Conduct of Research, 2007*]

https://nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018 [*Australian Code for the Responsible Conduct of Research, 2018*]

https://nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018 [*National Statement on Ethical Conduct in Human Research, 2007* (Updated 2018)

https://nhmrc.gov.au/research-policy/ethics/ethical-guidelines-research-aboriginal-and-torres-strait-islander-peoples [Ethical guidelines for research with Aboriginal and Torres Strait Islander Peoples]

https://aiatsis.gov.au/research/ethical-research/guidelines-ethical-research-australian-indigenous-studies [*Guidelines for Ethical Research in Australian Indigenous Studies*]

Australian Research Data Commons (ARDC, formerly Australian National Data Service (ANDS))

https://www.ands.org.au/working-with-data/fairdata

http://www.ands.org.au/guides/sensitivedata

Creative Commons

https://creativecommons.org.au

European Union General Data Protection Regulation (GDRP)

https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en#abouttheregulationanddataprotection

Guidelines on Personal data breach notification under Regulation 2016/679

F.A.I.R.

https://www.fair-access.net.au/fair-statement

Five Safes

Managing the Risk of Disclosure: the Five Safes Framework [ABS]

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

8

National Archives of Australia

http://www.naa.gov.au/information-management/records-authorities/types-of-records-authorities/GRA/gra37/index.aspx

Office of the Australian Information Commissioner

https://www.oaic.gov.au/privacy-law/privacy-act/

https://www.oaic.gov.au/privacy-law/privacy-act/notifiable-data-breaches-scheme

https://www.oaic.gov.au/agencies-and-organisations/guides/guide-to-securing-personal-information

Management of Data and Information in Research:
A guide supporting the *Australian Code for the Responsible Conduct of Research*

9