Research Data Management Policy

Section 1 - Purpose

(1) This Policy outlines how the principles and responsibilities of the Macquarie University Code for the Responsible Conduct of Research (the Macquarie Research Code) must be applied to the management of research data at Macquarie University.

(2) This Policy outlines the expected approach to the collection, storage, and dissemination of research data that facilitates reuse of data by the research community, while honouring ethical or confidentiality requirements and any contractual obligations imposed on Macquarie University by its sponsors, collaborators, partners or funding bodies.

Background

(3) Research data and primary materials are a valuable product of research activity which can assist in promoting open inquiry and debate, complementing other research outputs and publications, providing research transparency, and justifying research outcomes.

(4) Following the Management of Data and Information in Research: A guide supporting the Australian Code for the Responsible Conduct of Research, issued by the National Health and Medical Research Council, the Australian Research Council and Universities Australia, 2019, this Policy guides researchers to comply with their responsibility to "retain clear, accurate, secure and complete records of all research including research data. Where possible and appropriate, allow access and reference to these by interested parties".

Scope

(5) This Policy applies to anyone who conducts research or research support under the auspices of Macquarie University, in accordance with the Macquarie Research Code.

(6) The expected practices are described for all researchers but additional requirements may be imposed on those involved with external/overseas institutions, or where required by local legislation (for example in the case of Higher Degree Research Candidates subject to cotutelle or joint PhD agreements, or Researchers with a conjoint appointment). Researchers must comply with and disclose all relevant legal and regulatory requirements that pertain to their data or information collected, used or disclosed (refer to the Research Data Management Procedure regarding data governance documentation).

(7) Researchers may consult with a Research Data Steward for advice in relation to the implementation of this Policy.

(8) This Policy adheres to and complements the Macquarie University Cyber Security Policy (and associated information) and will be reviewed periodically, following changes to those policies, to the Australian Code for the Responsible Conduct of Research or related policies or when new data infrastructure are available to support researchers to manage their data.

Section 2 - Policy

Part A - General Principles

Research Data Management Responsibilities

(9) Researchers must adhere to the following research data management obligations:

a. data creation, collection, storage and management, and the provisions for its access from its creation or collection to its preservation must be documented early in the research process (via a Data Management Plan);
b. appropriate records of relevant approvals, protocols, authorisations and other administrative documents such as ethics and financial approvals, receipts, and consent forms must be retained;
c. data must be accurate, secure, complete, and documented with enough detail to enable verification of research results and to reflect what was communicated, decided, or done throughout the research process.
d. where possible, data and materials should be digitised, recorded in a durable and retrievable form, stored in institutionally sanctioned infrastructure (refer to the Research Data Sensitivity, Security and Storage Guideline), appropriately indexed and described by metadata, and backed up. Where appropriate, metadata should contain the location of any physical object being digitised Data should conform to disciplinary standards and protocols, comply with relevant laws and regulations, and meet publisher and funder requirements;
e. data and any digital research objects must be retained to substantiate research results and published claims. Researchers should determine what data and materials need to be retained based on legislative requirements, conditions imposed by publishers or funders conditions and best practice conditions in a discipline;
f. data security and data access arrangements for data or materials must be implemented in proportion to the sensitivity of the information to protect it from inappropriate access and use (a Data Breach). Minimum security considerations to be applied to data containing sensitive information are outlined in the Research Data Sensitivity, Security and Storage Guideline;
g. where possible and appropriate, data and materials must be collected or generated, stored, and licensed in a way which incorporates the Findable, Accessible, Interoperable, and Reusable (FAIR) data principles, with access allowed to interested parties, thereby enabling its reuse in future research;
h. plans for end-of-project publication of data outputs or archiving/accession of data to appropriate repositories must be documented and be based on the principle of 'As open as possible, as closed as necessary' (EU Horizon2020). In the absence of justifiable reasons such as respect for cultural sensitivity (see for example 9.e.) or unmanageable risks regarding sensitive or highly sensitive information (8.f)), researchers should lodge data outputs publicly within a data repository, without mediation, under an open license. Datasets that cannot be exposed without mediation should be lodged in a repository supporting mediated access with clear instructions describing access conditions. Valid reasons must be articulated explicitly for mediating or restricting data access or for more restrictive licensing;
i. data management for all projects involving human research must comply with the National Statement on Ethical Conduct in Human Research 2023. In the case of identifiable (or plausibly re-identifiable) personal data or data from human research participants, the consent obtained must be adhered to. Such consent must clearly state proposed data retention, confidentiality, access and reuse conditions;
j. data management for all projects involving animal research must adhere to protocols approved by an Animal Ethics Committee as per the Animal Research Act 1985 and the principles of the Australian Code for the care and use of animals for scientific purposes including the 3Rs (Replacement, Reduction and Refinement) by making data re-usable where possible;
k. project-specific protocols must be articulated and followed when they require measures that are beyond those outlined in this document or associated procedure or guideline, relevant University Policies, discipline-specific practices, or relevant laws, regulations, and guidelines;

l. inappropriate actual or potential use of, access to, or loss of sensitive data must be reported in accordance with the Cyber Security Policy. Reportable incidents could include incidents that have implications on personal security, commercial-in-confidence data, work health and safety, privacy or security matters involving trade, defence or the environment. Privacy and data breaches must be reported to the Chief Information and Digital Officer via the IT Service Desk reporting system and to the appropriate Human Research Ethics Committee (or other committee if applicable) for appropriate management; and

m. while all researchers are responsible for data and materials management, the Principal (lead) Investigator of a research project is ultimately responsible for ensuring that data is managed in accordance with this policy, or for ensuring their assigned Research Data Custodian has the authority and the relevant training and experience to do so.

Data Ownership

(10) This clause outlines data governance and ownership standards and must be read in conjunction with the University's Intellectual Property Policy:

a. generally, Macquarie University will own all Data and the Intellectual Property created during the collection or use of the Data, subject to the Intellectual Property Policy and any third-party agreement;

b. Macquarie University may assert ownership of Data and the Intellectual Property created during the collection or use of the Data for Higher Degree Research (as defined in the Intellectual Property Policy), in accordance with the Intellectual Property Policy;

c. Higher Degree Research Students or Macquarie University staff may request to retain a copy of the Data collected or created during their research. The request must be made to a Research Data Steward;

d. where a research project involves multiple institutions, an agreement must be reached at the outset of the project that covers the management, custodianship, storage, retention and disposal of the Data at each institution (following the Macquarie University Collaborative Research Standards); and

e. with respect to the ownership of the data used in or generated by research involving Aboriginal and Torres Strait Islander peoples and communities, data governance should reflect the considerations of the AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research. In general, Macquarie University or its researchers may hold data or information but should not make decisions about the access to or reuse of this data or information without proper consultation with its Indigenous owners, if any.

Supervising Research Trainees

(11) Research mentors and supervisors have a responsibility to help researchers under their supervision develop the necessary skills for appropriate data management, understand their obligations, and meet data management requirements.

Part B - Roles and Responsibilities

(12) The key roles and responsibilities of those involved in implementing this Policy are documented in this part.

Researchers, Staff and Students

(13) Researchers, Staff and Students must:

a. ensure their conduct reflects the principles and responsibilities in the Macquarie University Code for the Responsible Conduct of Research and the practices described in this Policy;

b. implement the requirements of the Research Data Management Procedure;

c. report potential breaches of the Macquarie University Code for the Responsible Conduct of Research or data breaches in a timely manner; and

d. undertake training and/or seek advice as required to implement this Policy and FAIR data principles.

Research Data Custodian

(14) The Research Data Custodian of a given project is an authorised person (either the Principal Investigator or another investigator nominated by the Principal Investigator) responsible for the collection, storage, or transmittal of data, ensuring the implementation of this Policy, including developing and maintaining a Data Management Plan (DMP).

(15) The Research Data Custodian will be the administrative point of contact for queries regarding the data, including for access to data which is classed as Mediated (restricted or special access) for a given research project. This will often, but not always, be the Principal Investigator on a project. Research Data Custodians may change over time.

Research Data Steward

(16) The Research Data Stewards are designated by the Deputy Vice-Chancellor (Research) or their nominee, to help ensure that data are classified appropriately and managed in accordance with their sensitivity. They also help provide appropriate training to users who interact with data. To accomplish these tasks Research Data Stewards will collaborate with IT, Library, Legal, Research Ethics and Integrity, and Risk Management teams.

(17) Research Data Stewards may:

a. review and provide advice on Data Management Plans regarding compliance with policy or audit selected projects for compliance;
b. assist the University in the event of an external audit, including granting access to data as required;
c. inherit custodianship if all associated research investigators sever relationships with Macquarie University;
d. review requests for or assist with bespoke data management or curation arrangements; and
e. provide support to researchers with data management, curation, or access queries.

Authorised User

(18) An Authorised User is a person who is authorised by the Research Data Custodian to access and use the data as specified in the applicable Data Management Plan (DMP).

(19) Authorised Users must:

a. use the data only for the purpose specified in the DMP; and
b. comply with controls established by the Research Data Custodian, outlined in the DMP.

Principal Investigator

(20) The Principal Investigator of a project is chiefly responsible for:

a. ensuring that data and materials management for any project under their lead is undertaken in accordance with the Macquarie University Code for the Responsible Conduct of Research;
b. clearly articulating governance, ownership, custodianship and intended use of any data collected, generated or collated as part of their research at the start of a research project;
c. creating a research data management plan, nominating Research Data Custodians, ensuring relevant members of the project team are granted appropriate access to the active data (at appropriate times) as Authorised Users, and that all Authorised Users are equipped with the knowledge and skills to implement this Policy; and
d. ensuring that adequate controls are in place to ensure the accuracy, authenticity, and integrity of the data.

(21) The Principal Investigator of a project may assign some of the above responsibilities to a Research Data Custodian.

Research Integrity Advisor

(22) Research Integrity Advisors have a good knowledge of the Macquarie University Code for the Responsible Conduct of Research, accompanying procedures, policies and processes and provide advice to those with concerns or complaints about potential breaches of the Macquarie University Code for the Responsible Conduct of Research.

Section 3 - Procedures

(23) Refer to the Research Data Management Procedure.

Section 4 - Guidelines

(24) Refer to the Research Data Sensitivity, Security and Storage Guideline.

Section 5 - Definitions

(25) The following definitions apply for the purpose of this Policy. These definitions have been adapted and modified from the Australian Code for the Responsible Conduct of Research, 2018, the Management of Data and Information in Research: A guide supporting the Australian Code for the Responsible Conduct of Research, the National Statement on Ethical Conduct in Human Research 2023, and from other resources noted in the associated information section of this Policy (including the Australian Data Archive):

a. Breach means a failure to meet the principles and responsibilities of the Macquarie Research Code (including failing to meet the standards or Policies accompanying the Macquarie Research Code). Breaches occur on a spectrum from minor to more serious breaches. A serious breach of the Macquarie Research Code which is also intentional or reckless or negligent constitutes research misconduct.
b. 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, e.g., Excellence in Research for Australia (ERA).
c. Data Breach is the accidental or deliberate access or exposure of University information (including research data) to unauthorised parties. Potential or actual data breaches must be reported appropriately and follow the Cyber Security Policy.
d. Data Lifecycle refers to the various stages that data goes through from initial creation or capture to eventual archiving, publishing and/or re-using. Data have a longer lifespan than the project that creates them and there is a need for all stages of the data lifecycle to be planned.
e. Data Management Plan describes the management of data through all stages of its lifecycle and includes documentation of how data will be created, collected, stored, and managed, and the provisions for access to data from its creation or collection to its preservation (refer to the Research Data Management Procedure).
f. Data Output refers to any output that communicates, disseminates, or makes available the research data to the

public or to other researchers. Data outputs may include digital data assets, techniques, algorithms, and software. Data outputs should be deposited in a data repository with appropriate Data Access Conditions.

g. Data Access Conditions may include specific restrictions on how people can gain access to the data output and should be determined by the sensitivity of the information.

h. Data Repository is information infrastructure, which may also be known as a data library or data archive. A data repository is used to store datasets for the long term, and often supports data sharing or publication (open or mediated access) and data reuse reporting. Data outputs should be deposited in a domain-specific, domain-general, or institutional data repository. The submission of a data output into a data repository is often a requirement for publication or funding. Data repositories differ in their ability to offer the users the ability to set mediated access conditions. Research data containing sensitive information which is being deposited into an online repository should have appropriate mediated access conditions assigned to it (refer to 'Mediated Access' or 'Specialised Access' in the Research Data Sensitivity, Security and Storage Guideline. A restrictive license is also applied that limits redistribution of the data.

i. Data Security refers to the process of protecting data from unauthorised access and data corruption throughout all stages of its lifecycle. This may include practices such as: data encryption, two factor authentication, backup and other key management practices that protect data across all applications and platforms. Refer to the Research Data Sensitivity, Security and Storage Guideline.

j. Data Sensitivity Categories ensure that research data at Macquarie University are grouped according to the level of sensitivity of the information in that data. The categories are: General, Sensitive, and Highly Sensitive. The level of sensitivity determines the minimum security and safety measures that should be applied to the data throughout its lifecycle and helps determine the access conditions that can be applied to any data output being generated and disseminated. Data outputs being generated from data classed as General, can usually have Unmediated or Open access and be liberally licensed. Data outputs being generated from data classed as Sensitive or Highly Sensitive would generally have Mediated Access (Restricted or Specialised) conditions applied - with varying degrees of mediated access and more restrictive licensing. Refer to the Research Data Sensitivity, Security and Storage Guideline for more details.

k. De-identified or Anonymised Data: These terms are often used interchangeably, but for the purposes of these documents:

    i. De-identified data refers to data which has had any direct and indirect identifying details (eg personal or locational) removed or transformed to protect privacy or confidentiality. However, identifying information may be able to be re-associated with the data later or re-identification may be possible in association with other publicly available data.

    ii. Anonymised data refers to data which has been collected or processed in a manner that makes it permanently impossible to identify individuals or locations from it. This usually involves stripping all identifiable elements from the data, for example by aggregating or summarising the data to such a general level that individuals or locations cannot be identified, and the data cannot be re-engineered to identify individuals or locations.

l. FAIR Data principles are implemented to enhance and improve the findability, accessibility, interoperability, and reusability of data and relate to the machine-actionability of the data (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention). They should be considered and reflected in the management of data.

m. Intellectual Property (IP) includes all copyright and all rights in relation to inventions (including patent rights), registered and unregistered trademarks (including service marks), registered and unregistered designs, confidential information, and circuit layouts and all other intellectual property rights resulting from intellectual activity in the academic, industrial, scientific, literary, and artistic fields recognised in domestic law anywhere in the world.

n. License offers researchers and institutions a standardised way of sharing data outputs with others and understanding their rights to use a data output generated by other researchers without infringing copyright.

When a data output is being produced at Macquarie University, a licence (such as one that is defined by the Creative Commons, see further resources) should be assigned which sets out the uses that may lawfully be made of the data output, and specifying the conditions under which its future use must comply.

o. Macquarie Research Code (Macquarie University Code for the Responsible Conduct of Research) sets out the principles that underpin an honest, ethical, and conscientious research culture and the expectations for the conduct of research under the auspices of Macquarie University.

p. Macquarie Research Code Procedure (Macquarie University Research Code Complaints, Breaches and Investigation Procedure) outlines the process for managing complaints, concerns or allegations, regarding the conduct of research and describes how potential or actual departures from the principles and responsibilities outlined in the Macquarie University Code for the Responsible Conduct of Research, should be reported, assessed, investigated and managed.

q. Mediated Access data has been deposited in a repository but requires a request from any would-be user, followed by the granting of permission for the use of the data either from a representative of the repository (e.g., 'restricted access' data in the Macquarie Research Data Repository) or from the researcher(s) who produced the data, typically the designated Data Custodian (e.g., 'special access' data in the Macquarie Research Data Repository).

r. Metadata refers to documentation that is created throughout the data lifecycle. Metadata is essential information describing data, providing the context for a data output, and making the data discoverable and reusable.

s. Research can be defined as the creation of new knowledge and/or the use of existing knowledge in a new and creative way to generate new concepts, methodologies, inventions and understandings. This could include synthesis and analysis of previous research to the extent that it is new and creative.

t. Researcher is any person (or persons) who conducts or assists with the conduct of research under the auspices of Macquarie University - may include staff members (academic and professional), visiting students, visiting fellows, volunteers, honorary and adjunct title holders, Emerita/us Professors, occupational trainees, and any student in any course at the University who conducts or assists with the conduct of research at or on behalf of the University.

u. Research Output refers to any output that communicates, disseminates, or makes available the products of research. A research output may include any form (hardcopy, electronic, creative work or other) of academic or public communication of the research from any stage of the research process (e.g., a professional blog, web-based publications, books, performances, book chapters, conference papers, journal articles or data outputs).

v. Sensitive or Highly Sensitive Information refers to information of a personal or sensitive nature which must be protected against unwarranted disclosure. Refer to the Research Data Sensitivity, Security and Storage Guideline for guidance in assessing the sensitivity of data and appropriate security measures to be applied. Sensitive information may include, but is not limited to, health-related data; personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs; financial information; genetic data or biometric data processed solely to identify a human being. Sensitive information may also relate to heritage or cultural information or information which may pose a risk to the environment or animals (such as the location of endangered species). Sensitive information must be safeguarded with appropriate data security practices following the Research Data Sensitivity, Security and Storage Guideline. Protection of sensitive information may be required for legal or ethical reasons, for issues pertaining to personal privacy, for the protection of animals or the environment or for proprietary considerations.

Status and Details

Status	Current
Effective Date	6th September 2021
Review Date	7th September 2024
Approval Authority	Deputy Vice-Chancellor (Research)
Approval Date	11th December 2022
Expiry Date	Not Applicable
Responsible Executive	Sakkie Pretorius Deputy Vice-Chancellor (Research) +61 2 9850 8645
Responsible Officer	Shawn Ross Director, Digitally Enabled Research +61 2 9850 7010
Enquiries Contact	Shawn Ross Director, Digitally Enabled Research +61 2 9850 7010