Research Data Sensitivity, Security and Storage Guideline

Section 1 - Purpose

(1) This Guideline provides details on data sensitivity indicators and advice for assessing and classifying data as highly sensitive, sensitive or general. It also documents appropriate security measures and storage options for active data according to its sensitivity classification.

Scope

(2) This Guideline applies to anyone who conducts research or research support under the auspices of Macquarie University, as per the Macquarie University Code for the Responsible Conduct of Research.

(3) The list of data sensitivity indicators within this Guideline is not exhaustive. Where a researcher believes their data may be sensitive or has queries relating to this Guideline they are encouraged to contact a Research Data Steward.

(4) The Guideline assists researchers to apply the principles of the Macquarie University Code for the Responsible Conduct of Research to the management of research data at Macquarie University and to direct their implementation of the expected standards.

Section 2 - Policy

(5) Refer to the Research Data Management Policy.

Section 3 - Procedures

(6) Refer to the Research Data Management Procedure.

Section 4 - Guidelines

Background

(7) Research Data may contain information of a personal or sensitive nature which must be protected against unwarranted disclosure.

(8) 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 information which may pose a risk to cultural resources, the environment or animals (such as the location of endangered species or threatened archaeological remains), to potentially valuable intellectual property, or to national security.

Part A - Sensitive Information within Research Data

(9) Sensitive information must be: protected against unwarranted disclosure, monitored for potential data breaches resulting in such disclosure and amenable to audit in the event of an actual or alleged data breach.

(10) Access to sensitive information must be safeguarded with appropriate data security practices.

(11) Data security is a shared responsibility between the University and the researcher (refer to Cyber Security Policy).

(12) Protection of sensitive information may be required for legal or ethical reasons, for issues pertaining to personal privacy and welfare, for cultural or environmental factors, for proprietary considerations, or to meet regulatory requirements.

(13) Research Data at Macquarie University can be grouped into three categories depending upon the sensitivity of its information. The categories are:

a. General
b. Sensitive
c. Highly Sensitive

Data Sensitivity Indicators

(14) Data is generally considered either Sensitive or Highly Sensitive if it contains Identifiable 'personal information' or identifiable health information. This includes:

a. '[Information or an opinion] about an individual whose identity is apparent or can reasonably be ascertained from the information or opinion.' (Privacy and Personal Information Protection Act 1998 section 4.1; Health Records and Information Privacy Act 2002 section 5.1)'
b. See also: '...information or an opinion, whether true or not, and whether recorded in a material form or not, about an identified individual, or an individual who is reasonably identifiable.' (Privacy Act 1988).

(15) Data may also be deemed sensitive due to cultural considerations, environmental, or proprietary considerations.

(16) The type of 'personal information' contained in the data, or other aspects of the data, will determine if it should be classified as 'Highly Sensitive' or 'Sensitive' as follows.

(17) Research Data is considered highly sensitive when:

a. it contains the following types of 'personal information' (adapted from the Australian Privacy Principles, Privacy Act 1988, schedule 1; see APP B.138 for more information; see also the Privacy and Personal Information Protection Act 1998 No 133, Section 19.1; the Health Records and Information Privacy Act 2002, Section 6) and is identifiable or (potentially) re-identifiable based on data contained within the record itself or if combined with other publicly available data:
    i. racial or ethnic origin
    ii. political opinions
    iii. membership of a political association
    iv. religious beliefs or affiliations
    v. philosophical beliefs
    vi. membership of a professional or trade association
    vii. membership of a trade union
    viii. sexual orientation or practices

ix. criminal record
x. health information about an individual
xi. genetic information
xii. biometric information
xiii. financial information

b. it contains information that is subject to regulatory controls and is deemed highly sensitive by a Research Data Steward or by a relevant Research Management Committee (for example, data relating to controlled technology per the Defence Trade Controls Act 2012 or information which poses a risk to national security).

(18) Data is considered sensitive when it:

a. is identifiable and contains personal information / human subject data but does not concern the sensitivity indicators listed above in clause 17 a.;
b. concerns a sensitivity indicator listed above in clause 17 a., but has been deidentified and cannot plausibly be re-identified from the data itself or if combined with other, publicly available data. However, files relating reference codes or unique; identifier keys (e.g., Personal Identifier Number or PINs) to an identifiable person must be treated as highly sensitive data;
c. contains cultural heritage location information or other heritage data where community consent for release is lacking (standards and procedures vary in different countries);
d. contains ecological or environmental data concerning rare, threatened or endangered species;
e. contains data governed by IP / commercialisation agreements;
f. contains data that one or more investigators on the project do not consent to release (agreement should be reached before a project is launched and articulated in a DMP);
g. contains non-work-related contact information, location information, or other information deemed 'private', 'confidential', or 'sensitive' by any Macquarie University policy; or
h. it is associated with a project that is under regulatory control and is deemed sensitive by a Research Data Steward or by a relevant Research Management Committee (for example, an Animal Ethics Committee per the Animal Research Act 1985 and the Australian Code for the Care and Use of Animals for Scientific Purposes or by an Institutional Biosafety Committee).

(19) Data is classified general when it is:

a. publicly available third-party data;
b. open data in the public domain or carrying an explicit, permissive license (e.g., an open Creative Commons or Open Data Commons license);
c. anonymised, aggregated or other derivative datasets based on personal information where the data cannot be disaggregated or used to reconstruct the original dataset (alone or in combination with publicly available data);
d. de-identified personal information that does not concern a sensitivity indicator listed above (17 a.) and cannot plausibly be re-identified from the data itself or if combined with other, publicly available data; or
e. not otherwise sensitive or highly sensitive.

Part B - Research Data Security

(20) Examples of how research data can be classified according to its sensitivity, and how that relates to the Information Classification and Handling Procedure include (but are not limited to) the examples provided below:

Macquarie University Research Data classification	Macquarie University Cyber Security classification	Examples
General	Public	Published research data
	Internal	Unpublished research data not covered by conditions making it more sensitive. Data considered 'general intellectual property'. Anonymised, aggregated or derivative research data relating to individuals. If uncertain, consult a Research Data Steward and/or your Human Research Ethics Committee (HREC). De-identified research data relating to individuals and not associated with a sensitivity indicator listed at clause 17 a. that cannot plausibly be re-identified from the data itself or in combination with other, publicly available data (if uncertain, consult a Research Data Steward and/or your HREC).
Sensitive	Confidential	Culturally sensitive data. Environmentally sensitive data. Data with explicit IP constraints. De-identified research data relating to individuals and associated with a sensitivity indicator listed at clause 17 a. that cannot plausibly be re-identified from the data itself or in combination with other, publicly available data (if uncertain, consult a Research Data Steward and/or your HREC). Identifiable research data relating to individuals that does not include data associated with any of the sensitivity indicators listed at clause 17 a. Data which contains information that is subject to regulatory controls and is deemed sensitive by a Research Management Committee (e.g., Animal Ethics Approval: refer to the Animal Research Act 1985).
Highly sensitive	Highly sensitive	Identifiable research data relating to an individual that includes data associated with any of the sensitivity indicators listed at clause 17 a. De-identified research data relating to individual that includes data associated with any of the high-sensitivity indicators listed at clause 17 a., which could be re-identified based on the data in the record itself or in combination with other publicly available data. Data which contains information that is subject to regulatory controls and is deemed highly sensitive by a Research Management Committee (for example if it poses a risk to national security: refer to Defence Trade Controls Act 2012).

Data Collection

(21) Data capture or collection practices vary from discipline to discipline and must be specified in your Data Management Plan.

(22) Researchers must use approved platforms for the collection, capture, or collation of sensitive or highly sensitive data where such platforms are available. The approved platforms are listed in Table 1: Data Collection, Storage, Archiving, and/or Publication Platforms.

(23) If no platform exists for your research discipline consult a Research Data Steward regarding the process for proposing use of an unapproved data platform.

Active Data Storage

(24) The Macquarie University approved storage options for data can be found in Table 1: Data Collection, Storage, Archiving, and/or Publication Platforms (appropriate security measures as per clauses 28-30 must be implemented).

(25) Custom storage solutions using Australia-based commercial web services (e.g., AWS, Azure, Google Cloud) or peak facilities (e.g., NCI, Pawsey) may also be acceptable but will require approval by a Research Data Steward via a Data Management Plan in FoRA.

(26) Bespoke on-site storage solutions may be possible and will require approval by a Research Data Steward via a

Data Management Plan in FoRA. If no platform exists for your research discipline consult a Research Data Steward regarding the process for proposing use of an unapproved data platform.

Table 1: Data Collection, Storage, Archiving, and/or Publication Platforms

(27) Table 1: Data Collection, Storage, Archiving, and/or Publication Platforms outlines the storage options endorsed by the University (staff access only).

Active Data Security

(28) Security practices must be applied to all active data to prevent unauthorised access or accidental loss. The required security controls are summarised in Table 2: Security Controls according to Data Sensitivity Classification.

(29) The sensitivity level of the data determines the security practices that must be applied during data management.

(30) Researchers are expected to obtain assistance from a Research Data Steward or IT (if needed) to meet the following requirements:

a. The standard security practices that should be applied to research data which is not classed as either sensitive or highly sensitive are:
    i. If using a personal device to store or access data it must be properly maintained (e.g., regarding antivirus software).
    ii. You must back up locally-stored data. Backups of locally-stored data should be automated, either via continuous synchronisation (as with OneDrive or Cloudstor sync clients), or should be frequent and regular (e.g., with a daily incremental backup via a shell script).
    iii. You must ensure unique, strong passwords for all services related to the data.
    iv. You should encrypt all personal or work devices on which data is stored and from which the data will be accessed.
b. If your data is classed as sensitive, the following security practices are expected:
    i. You must ensure unique, strong passwords for all services related to the data.
    ii. You must encrypt all personal or work devices on which data is stored and from which the data will be accessed.
    iii. You must back up locally-stored data. Backups of locally-stored data should be automated, either via continuous synchronisation (as with OneDrive or Cloudstor sync clients), or should be frequent and regular (e.g., with a daily incremental backup via a shell script).
    iv. If using a personal device to store or access data it must be encrypted and properly maintained (e.g., antivirus software).
c. If your data is highly sensitive, the following security measures must be applied:
    i. You should avoid storing Highly Sensitive data locally (i.e., access, edit, and analyse it in its online location without downloading it).
    ii. If you cannot avoid storing Highly Sensitive data locally, then you must only use Macquarie-issued, encrypted devices and back up locally-stored data (e.g., to SharePoint). Backups of locally-stored data should be automated, regular, and frequent. Consult a Research Data Steward for further advice.
    iii. You must ensure unique, strong passwords for all services related to the data.
d. Additional security controls may be required for defence-related research projects (or in projects deemed high-risk by the Research Risk Review Committee).

Data Storage and Access for Archiving and Publication

(31) The Macquarie University approved archiving and publication platform options can be found in Table 1: Data

Collection, Storage, Archiving, and/or Publication Platforms.

(32) Security practices must be applied to all archived data to prevent unauthorized access.

(33) The sensitivity level of the data determines the security and access practices that must be applied when data is archived and published.

a. The standard storage and access practices that should be applied to research data which is not classed as either sensitive or highly sensitive are:
    i. You must archive data in an appropriate repository (a discipline-specific research data repository, the Macquarie Research Data Repository, or a general-purpose research data repository).
    ii. You must explicitly license your data.
    iii. You should license your data with a CC-0 or CC-BY License unless you justify a different license.
b. If your data remains sensitive or highly sensitive at the time of archiving and publication, the following storage and access practices should be applied:
    i. You must archive data in an approved repository that supports mediated access and implements appropriate access controls (e.g., the Australian Data Archive, another approved research data repository, or the Macquarie Research Data Repository). Contact a Research Data Steward for advice if required.
    ii. You must select a mediated access regime for data placed in the Macquarie Research Data Repository (in consultation with a Data Custodian if necessary):
        - Restricted Access: Macquarie manages data access on your behalf, logging access to the data and ensuring users consent to stipulated Terms and Conditions of Use.
        - Special Access: You will be informed by Macquarie of each request to use the data, to give or withhold permission.
    iii. If you use the Australian Data Archive or another approved repository, you must select an appropriate mediated access regime from the options available from that repository (in consultation with a Research Data Steward if necessary).
    iv. You must explicitly license your data.
    v. You must license your data with a Macquarie Standard License Terms and Conditions of Use (for the Macquarie Research Data Repository), or a similar license restricting data redistribution, such as those offered by the Australian Data Archive or other approved research data repositories.

Table 2: Security Controls according to Data Sensitivity Classification

(34) Table 2: Security Controls according to Data Sensitivity Classification outlines the security measures that are expected to be applied (staff access only).

Section 5 - Definitions

(35) Definitions specific to this Guideline are contained in the Research Data Management Policy.

Status and Details

Status	Current
Effective Date	7th September 2021
Review Date	7th September 2024
Approval Authority	Deputy Vice-Chancellor (Research)
Approval Date	3rd September 2021
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