Maynooth University
National University of Ireland Maynooth

Office of the Vice President Research and Innovation an Leas-Uachtaráin Taighde agus agus Nuálaíochta

Maynooth University

Research Data Management Policy

Version: 1
Effective Date: April, 2025
Reviewed: NA
Next Review Date: January 2030

Document Location

Revision History

Version Number/Revision Number	Revision Date	Summary of Changes
0.1	19/02/2021	Initial draft
0.2	25/03/2022	Incorporating comments from VPR/University Research Committee
0.3	12/10/2022	Incorporating edits from Library Research Committee
0.4	01/10/2024	Incorporating edits from IT Services management
0.5	30/10/2024	Edits from RDO/Library, update of referenced policies
0.6	7/11/2024	Inclusion of Funder Templates and Link to exemplar DMPs
0.7	10/12/2024	Minor edits following approval at Research Committee

Consultation History

Revision Number	Consultation Date	Names of Parties in Consultation	Summary of Changes
0.7	December 2024	Fran Callaghan, Fiona Morley &Cathal McCauley	Minor edits and links updated following approval at the Research Committee
0.6	November 2024	Rachel Msetfi	Suggested Inclusion of Funder Templates and exemplar DMPs.
0.5	October 2024	Elaine Mccarthy (RDO), Fran Callaghan, Fiona Morley (MU Library)	Update of links, inclusion of additional reference policies, improved definitions in glossary
0.4	September 2024	Paddy Daly, Dearbhla O'Reilly, Cathal McCauley, Fiona Morley, Fran Callaghan	IT recommendations on policy additions and updates to glossary terms
0.3	October 2022	Library Research Committee	Link updates and edits
0.2	March 2022	VPR/University Research Committee	Incorporating feedback from VPR/University Research Committee

0.1	February 2021	Library Research Committee	Initial draft

Approval

This document requires the following approvals:

Name	Title	Date
OASAC	Draft RDM Policy	October 2024
VPR	Draft RDM Policy	November 2024
Research Committee	Draft RDM Policy	December 2024
Academic Council	Draft RDM Policy	February 2025
Governing Authority	Draft RDM Policy	March 2025

Introduction and context

Purpose of this policy: The purpose of this policy is to identify MU's position regarding the management of research data and to outline the responsibilities and requirements of the University and its Researchers.

Background: Numerous factors are now influencing the **drive for data curation**, including the desire for increased transparency in research findings, increased visibility of research projects and outputs, and increased re-use of research data. In relation to **research integrity** the correct management of research data is seen as a way to encourage good practice in the use, storage and retention of data and to guard against research misconduct[1].

National Funding Agencies are driving the need for data curation by introducing policies on open access to research data and in some cases, Data Management Plan (DMP) requirements at pre-proposal and post award stages of the funding lifecycle. DMP requirements of the national funders can be viewed at our Research Data Management Libguide. From an international perspective, all the Research Councils in the UK (RCUK) have explicit RDM policies[2] and EU funding schemes such as Horizon Europe have introduced such requirements[3]. All research undertaken by MU Researchers (whether funded or unfunded) should comply with this Research Data Management Policy.

Supporting MU policies:

This Research Data Management policy is in keeping with the principles established in the following policies:

- MU Policy on Open Access to Research: covering the long-term preservation of, and free public access to, all research publication outputs deposited in MURAL, MU's institutional repository
- MU Information and Data Security Policy covering protection, confidentiality, integrity and availability of MU information and Data assets, which includes data stored on MU IT systems, cloud services, portable devices and manual data etc.
- MU Research Integrity Policy: covering research data collection, storage, retention and disposal
- MU Research Ethics Policy: covering the use of personal data for research involving human participants
- Other applicable Research Policies, Information Security Policies and Data and Privacy Policies

[1] The *National policy statement on Ensuring Research Integrity in Ireland* (2014) and the *Research Integrity Policy (2016)* identify fabrication and falsification of data among the most serious breaches of research integrity, and identifies other forms of data-related poor practice as potentially damaging to the integrity of the research community.
[2] http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies
[3] https://www.openaire.eu/opendatapilot

If action is not taken to fully address data management needs, this poses a potentially serious risk to future research in the University. Achieving compliance with RDM principles and putting these into practice will enhance MU's researchers' competitiveness for funding applications and other collaborations. The Appendix contains a glossary comprised of definitions of key terms used in this policy.

Scope

This policy covers research data generated and managed by MU researchers.

Key Principles

1. Maynooth University recognises and is committed to the implementation of high-quality research data management as an integral part of research excellence, innovation and integrity.

2. For accountable and transparent research to take place, research data must be of high quality, correct, and retrievable for re-use. Best practice supports the principle that Research Data should be FAIR (Findable, Accessible, Interoperable, Reusable) or as FAIR as possible. Consideration of whether data should be openly available should be addressed as part of the data management plan, see point 5 below.

3. All research data should be deposited in an appropriate repository (institutional and/or disciplinary specific). Where a non-institutional repository is used, a persistent link to the location of the data should be provided via a researcher's Research Information System (RIS) profile.

4. Ownership of Research Data: Data generated by MU Researchers unless otherwise stated in the terms of the funding agency or covered by research contracts are the property of MU.

5. Research projects (and ideally all research proposals at application stage as part of good research practice) should include the preparation of a Data Management Plan. This will ensure all stages of the data life cycle are considered and planned for, which is an essential part of Good Research Practice.

6. The interests of participants generating research data must be protected, in line with the MU Research Ethics Policy and Data Protection Policy.

7. Data handling must be compliant with appropriate data retention and sharing regulations in line with MU Information and Data Security Policy, MU's Data and Privacy Policies, Research Integrity Policy and MU Research Ethics Policy.

8. Data licencing must comply with research funders requirements. If the funder does not specify a licensing requirement, Maynooth University recommends the CC-BY

licence. Other Creative Commons licences are also available and details of these can be found at https://creativecommons.org/about/cclicenses/

Roles and Responsibilities

1.1 The responsibility for effective research data management during a research project lies with the Principal Investigator. However, all researchers have a personal responsibility to effectively manage the data they create as part of good research practice.

1.2 MU will support best practice in Research Data Management by providing training, support and advisory services for researchers in the practice of drafting and reviewing data management pans (DMPs). DMPs are considered 'living documents' that should be both referred to at all points in the research lifecycle and reviewed and updated to reflect changes in the project. More information on DMP supports can be found in the library guide.

1.3 The University will endeavour to develop appropriate mechanisms and services for storage, backup, registration, deposit and retention of research data assets in support of current and future access, during and after the completion of research projects.

Review

This policy will be reviewed every 5 years commencing in 2030.

Acknowledgment
MU wish to acknowledge the use of the following documents in the preparation of this policy:

University College Cork RDM Policy (2024)
https://www.ucc.ie/en/media/research/researchatucc/researchculture/policiesproceduresandguidelines/UCCResearchDataManagementPolicy.pdf

University of Manchester RDM policy (2019)
https://documents.manchester.ac.uk/display.aspx?DocID=33802

University of Edinburgh RDM Policy (2011)

https://www.ed.ac.uk/information-services/about/policies-and-regulations/research-data-policy

Appendix 1: Glossary

Principal Investigator: The Principal Investigator is responsible for the management and integrity of the design, conduct, and reporting of the research project and for managing, monitoring, and ensuring the integrity of any collaborative relationships. Additionally, the Principal Investigator is responsible for the direction and oversight of compliance, financial, the research team, and other related aspects of the research project and for coordination with school, department, and central administration personnel to assure research in is conducted in accordance with all policies and procedures.

Research: is 'any creative and systematically performed work with the goal of furthering knowledge, including discoveries regarding people, culture and society, in addition to the use of such knowledge for new applications[4].

Researchers: Researchers refers to 'all research-active members of an institution including employees, and doctoral candidates. Persons not directly affiliated with an institution, but who, for purposes of research, make use of or are physically present at the institution, are also included in the term. Visiting researchers or collaborators will also be expected to comply with the Policy'[4]. This policy also covers emeritus staff working on MU research projects.

Research data is 'the evidence that underpins the answer to the research question and can be used to validate findings. This might be quantitative information or qualitative statements collected by researchers in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence. Research data may take the form of numbers, symbols, text, images or sounds, including computer code, annotated fieldwork observations, or a descriptive record of a physical sample'[4].

Research Data Management (RDM) refers to 'the active management of research data during the lifecycle of the project and decisions about the treatment of the data post-project'[5] and the processes applied through a research project's lifecycle that guide the collection, documentation, storage, sharing and preservation of research data.

The below definitions of FAIR Principles and the four basics of FAIR are as described by OpenAire[6]

FAIR Principles:

[4] University of Manchester RDM policy (2019) https://documents.manchester.ac.uk/display.aspx?DocID=33802

[5] UCC RDM Policy (2024) p.3. https://www.ucc.ie/en/media/research/researchatucc/researchculture/policiesproceduresandguidelines/UCCResearchDataManagementPolicy.pdf

[6] https://www.openaire.eu/how-to-make-your-data-fair

The increasing availability of online resources means that data need to be created with longevity in mind. Providing other researchers with access to your data facilitates knowledge discovery and improves research transparency.

The Four Basics of FAIR:

'Findable'	i.e. discoverable with metadata, identifiable and locatable by means of a standard identification mechanism
'Accessible'	i.e. always available and obtainable; even if the data is restricted, the metadata is open
'Interoperable'	i.e. both syntactically parseable and semantically understandable, allowing data exchange and reuse between researchers, institutions, organisations or countries; and
'Reusable'	i.e. sufficiently described and shared with the least restrictive licences, allowing the widest reuse possible and the least cumbersome integration with other data sources.

Appendix 2: RDM Plan Templates

Templates from Horizon Europe and the Health Research Board follow below. The Digital Curation Centre (DCC) maintain a list of sample plans here…
https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

DATA MANAGEMENT PLAN (Horizon Europe)

(To be filled in and uploaded as deliverable in the Portal Grant Management System, at the due date foreseen in the system (and regularly updated).

⚠ *The template is recommended but not mandatory. If you do not use it, please make however sure that you comply with the research data management requirements under Article 17 of the Grant Agreement.)*

PROJECT	
Project number:	[project number]
Project acronym:	[acronym]
Project name:	[project title]

DATA MANAGEMENT PLAN	
Date:	[dd/mm/yyyy]
Version:	[DMP version]

1.4 Data Summary

Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.

What types and formats of data will the project generate or re-use?

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

What is the expected size of the data that you intend to generate or re-use?

What is the origin/provenance of the data, either generated or re-used?

To whom might your data be useful ('data utility'), outside your project?

1.5 FAIR data

1.5.1 Making data findable, including provisions for metadata

Will data be identified by a persistent identifier?

Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

Will metadata be offered in such a way that it can be harvested and indexed?

1.5.2 Making data accessible

Repository:

Will the data be deposited in a trusted repository?

Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

Data:

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Will the data be accessible through a free and standardized access protocol?

If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

How will the identity of the person accessing the data be ascertained?

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

Metadata:

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

1.5.3 Making data interoperable

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

Will your data include qualified references7 to other data (e.g. other data from your project, or datasets from previous research)?

1.5.4 Increase data re-use

How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Will the data produced in the project be useable by third parties, in particular after the end of the project?

Will the provenance of the data be thoroughly documented using the appropriate standards?

Describe all relevant data quality assurance processes.

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

1.6 Other research outputs

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their

[7] *A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)*

projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

1.7 Allocation of resources

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

Who will be responsible for data management in your project?

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

1.8 Data security

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

Will the data be safely stored in trusted repositories for long term preservation and curation?

1.9 Ethics

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

1.10 Other issues

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

HISTORY OF CHANGES		
VERSION	PUBLICATION DATE	CHANGE
1.0	05.05.2021	Initial version (new MFF).
1.1	01.04.2022	Reformatted to align with other deliverables templates.

Research. Evidence. Action.

Health Research Board

Health Research Board Data Management Plan Template

21 July 2023

Data description and collection or re-use of existing data

How will new data be collected or produced and/or how will existing data be re-used?

Guidance:

Explain which methodologies or software will be used if new data are collected or produced and specify which community standards (if any) will be used.

State any constraints on re-use of existing data if there are any.

Explain how data provenance will be documented.

Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.

What data (for example the kind, formats, and volumes), will be collected or produced?

Guidance:

Give details on the types of data – quantitative, qualitative; generated from surveys, interviews, medical records, clinical measurements, tissue samples, genotypic data, etc.

Give details on the data format: the way in which the data is encoded for storage, often reflected by the filename extension (for example pdf, xls, doc, txt, or rdf).

Justify the use of certain formats. For example, decisions may be based on staff expertise within the host organisation, a preference for open formats, standards accepted by data repositories, widespread usage within the research community, or on the software or equipment that will be used.

Give preference to open and standard formats as they facilitate sharing and long-term re-use of data (several repositories provide lists of such ‘preferred formats’).

Give details on the volumes (they can be expressed in storage space required (bytes), and/or in numbers of objects, files, rows, and columns).

Consider and detail which data will have value to other research users and could be shared. We recognize that there are many reasons why data cannot be shared or made openly available. If you do not intend to make the data you’ve generated available to others, please provide justification for your decision.

Documentation and data quality

What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?

Guidance:

Indicate which metadata will be provided to help others identify and discover the data.

Indicate which metadata standards (for example DDI, TEI, EML, MARC, CMDI) will be used and potential community standards available.

Use community metadata standards where these are in place.

Indicate how the data will be organised during the project, mentioning for example conventions, version control, and folder structures.

Consider what other documentation is needed to enable re-use - methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, and so on.

Consider how this information will be captured and where it will be recorded, for example in a database with links to each item, a ‘readme’ text file, file headers, code books, or lab notebooks.

When describing data, please remember that file and folder names as well as variables and metadata may contain personal or sensitive data. Even if your research data contains personal data, related metadata can be published if it does not contain identifiers which could be used to identify a study subject.

What data quality control measures will be used?

Guidance:

Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeated samples or measurements, standardised data capture, data entry validation, peer review of data, or representation with controlled vocabularies.
Consider how data minimisation, pseudonymisation or anonymisation will affect data quality.

Storage and backup during the research process

1.11

How will data and metadata be stored and backed up during the research process?

Guidance:

Describe how and where the data will be stored, backed-up and managed during research activities and how often the backup will be performed. It is recommended to store data in at least two separate locations.

Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution. Storing data on laptops, stand-alone hard drives, or external storage devices such as USB sticks should be avoided. If external servers are used, please ensure that they are compliant with GDPR and any other legislation related to the data collected.

How will data security and protection of sensitive data be taken care of during the research?

Guidance:

Detail the key risks to confidentiality and security related to human participants or other sensitive data and how this information will be managed.

Explain how the data will be recovered in the event of an incident.

Explain who will have access to the data during the research and how access to data is controlled, especially in collaborative partnerships.

Explain which institutional data protection policies are in place.

Legal and ethical requirements, codes of conduct

If personal data are processed, how will compliance with legislation on personal data and on security be ensured?

Guidance:

Ensure that when dealing with personal data protection laws (for example GDPR and the Health Research Regulations) are complied with:

Ensure that the preservation and/or sharing of personal data is fully consistent with the terms of the informed consent under which the data were provided by participants.

Consider anonymisation of personal data for preservation and/or sharing (truly anonymous data are no longer considered personal data).

Consider pseudonymisation of personal data (the main difference with anonymisation is that pseudonymisation is reversible).

Consider encryption which is seen as a special case of pseudonymisation (the encryption key must be stored separately from the data, for instance by a trusted third party).

Explain whether there is a managed or governed access procedure in place for authorised users of personal data.

If you are unclear on how best to comply with legislation on personal data and security, please speak to your institutional Data Protection Officer.

How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

Guidance:

Explain who will be the owner of the data and who will have the rights to control access:

- Explain what access conditions will apply to the data? Will the data be openly accessible, or will there be access restrictions? In the latter case, which? Consider the use of data access and re-use licenses (e.g. CC-BY, CC-BY-NC, etc.)
- Make sure to cover these matters of rights to control access to data for multi-partner projects and multiple data owners, in the consortium agreement.

Indicate whether intellectual property rights (for example Database Directive, sui generis rights) are affected. If so, explain which and how they will be dealt with.

Indicate whether there are any restrictions on the re-use of third-party data.

What ethical issues and codes of conduct are there, and how will they be taken into account?

Guidance:

Consider whether ethical issues can affect how data are stored and transferred, who can see or use them, and how long they are kept. Demonstrate awareness of these aspects and respective planning.

Follow the national and international codes of conduct and institutional ethical guidelines and check if ethical review (for example by an ethics committee) is required for data collection in the research project.

Data sharing and long-term preservation

How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

Guidance:

Explain how the data will be discoverable and shared (for example by deposit in a trustworthy data repository, indexed in a catalogue, use of a secure data service, direct handling of data requests, or use of another mechanism).

Outline the plan for data preservation and give information on how long the data will be retained.

Explain when the data will be made available. Indicate the expected timely release (For data related to clinical trials - specify how long the data will be made available for). Explain whether exclusive use of the data will be claimed and if so, why and for how long. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents.

Indicate who will be able to use the data. If it is necessary to restrict access to certain communities or to apply a data sharing agreement, explain how and why. Explain what action will be taken to overcome or to minimise restrictions.

We recognize that there are many reasons why data cannot be shared or made openly available. If you do not intend to make the data you've generated available to others, please provide justification for your decision. Restrictions should be minimized as much as possible.

How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?

Guidance:

Indicate what data must be retained or destroyed for contractual, legal, or regulatory purposes.

Indicate how it will be decided what data to keep. Describe the data to be preserved long-term and consider how this data will be curated and preserved beyond the lifetime of the grant. Indicate where the data will be deposited, preferably in a trusted repository.

Explain the foreseeable research uses (and/ or users) for the data.

Indicate where the data will be deposited. If no established repository is proposed, demonstrate in the data management plan that the data can be curated effectively beyond the lifetime of the grant. It is recommended to demonstrate that the repositories policies and procedures (including any metadata standards, and costs involved) have been checked.

What methods or software tools are needed to access and use data?

Guidance:

Indicate whether potential users need specific tools to access and (re-)use the data. Consider the sustainability of software needed for accessing the data.

Indicate whether data will be shared via a repository, requests handled directly, or whether another mechanism will be used?

How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

Guidance:

Explain how the data might be re-used in other contexts. Persistent identifiers should be applied so that data can be reliably and efficiently located and referred to. Persistent identifiers also help to track citations and re-use.

Indicate whether a persistent identifier for the data will be pursued. Typically, a trustworthy, long-term repository will provide a persistent identifier.

Data management responsibilities and resources

Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

Guidance:

Outline the roles and responsibilities for data management/stewardship activities for example data capture, metadata production, data quality, storage and backup, data archiving, and data sharing. Alongside the PI, specify who is responsible for ensuring the completion of these tasks.

Specify who is responsible for the management of sensitive and confidential data as well as monitoring its implementation throughout the lifecycle of the data.

For collaborative projects, explain the co-ordination of data management responsibilities across partners.

Indicate who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised.

Consider regular updates of the DMP.

What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

Guidance:

Explain how the necessary resources (for example time) to prepare the data for sharing/preservation (data curation) have been costed in. Carefully consider and justify any resources needed to deliver the data. These may include storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges.

Indicate whether additional resources will be needed to prepare data for deposit or to meet any charges from data repositories. If yes, explain how much is needed and how such costs will be covered.