Faculty of Science, Research Data Management Guideline

Date: *June 2023*

Document version: *1.8*[1]

Document approved by the Faculty Board on 13-06-2023

Purpose of this Document

This document is intended to serve as a guideline for researchers at the Faculty of Science for research data management. It outlines and defines a set of principles regarding research data and provides practical handlers for researchers in accordance with the regulations established in the Utrecht University policy framework for research data management (2016)[2]. The guideline can be used as a reference together with existing Institute or research-group data management protocols.

Support for Data Management

The faculty of Science has a dedicated website for Research Data Support, see Research Data Support at the Faculty of Science.

The RDM Support network from Utrecht University hosts open consultation hours every week at the University Library. For a complete event agenda, together with information on trainings and workshops see Walk-in hours & Workshops - Research Data Management Support.

Reading guides on diverse RDM topics can be found on Research Data Management Support/Guides.

RDM Support at Utrecht University developed a guide on handling personal data in scientific research. See the Data Privacy Handbook.

Contact points at the Faculty of Science:

- For research data management questions and tailored advice you can contact the faculty support team via rdm-beta@uu.nl.
- Faculty Privacy Officer: privacy-beta@uu.nl
- Ethics Review Board, Faculty of Science and Geosciences: etc-beta-geo@uu.nl. See the SG ERB website

Data Principles

1. FAIR data

In line with Utrecht University's commitment to Open Science, research data at the Faculty of Science is expected to follow the **FAIR** principles:

F - Findable: data and metadata should be findable, preferably digitally.

A - Accessible: data and metadata should be stored and maintained in such a manner that they are accessible and with clear terms of use.

I - Interoperable: data and metadata data should be stored and mantained in such a manner that the data can be easily opened, exchanged and combined with other datasets.

R - Reusable: data are licensed and described in a manner that facilitates its reuse by other users.

[1] This document updates the faculty guideline version 1.2, d.d. 2018

[2] https://www.uu.nl/sites/default/files/university_policy_framework_for_research_data_utrecht_university_-_january_2016.pdf

Additional guidance on how to make your research data FAIR is highlighted in the practical information pointers provided below. To learn more see RDM Guides/ How to Make your data FAIR

2. Data Privacy

In compliance with current regulations[3] when *personal data* is used for research purposes, researchers should incorporate organisational and technical safeguards, that guarantee the protection of personal data across the data lifecycle. Support is available at the faculty (rdm-beta@uu.nl // privacy-beta@uu.nl)

Data Management

1. Responsibilities

Research institutes can formulate their own additional specifications for data management within the university policy framework.

Principal Investigators (PI) are responsible for ensuring that students, researchers and research leaders in their group abide by the UU's data management policy, and the faculty guideline or institute protocol.

The creators of the data: researchers, students and their supervisors have primary responsibility for good stewardship of their research data. For students, it is the responsibility of the thesis supervisor to ensure that research data is stored and preserved in a sustainable manner after the project has ended.

Researchers should:

- draw up a data management plan (see below) describing proper procedures and design processes for handling data throughout the project, also considering costs related to data management and how they will be covered (institute-, subsidiary-funding or otherwise)
- draw up and record data management agreements with providers of external data or in collaboration studies;
- make arrangements for the continued operational management of the research data after the end of the research project or termination of the employment contract at the institution;
- ensure that the integrity, confidentiality and protection of the data are guaranteed;

An extended version on Research Data Management responsibilities can be found in the UU policy framework[2].

2. Data management plan (DMP)

A Data Management Plan (DMP) is a living document that specifies how research data will be collected, stored, preserved, managed, documented and used. A DMP helps to produce FAIR data, indicating actions and strategies to ensure that research data are of high-quality, secure, easy to find, understand and reuse.

2.1 Focus points from the UU policy

It is the responsibility of each individual researcher (or, in the case of a group of researchers, the Principal Investigator) to draw up a Data Management Plan (DMP) at the start of the research project and to follow up the agreements made in this plan.

2.2 Practical information

[3] Dutch Personal Data Protection Act and General Data Protection Regulation (GDPR).

- When working with personal data, DMPs are required by the Ethical Review Board of the faculty as part of the ethical approval process.
- The DMPOnline tool offers templates specifically made for researchers at Utrecht University, including templates and guidance from diverse funding agencies.
- Information on how to prepare a DMP can be found at: Research Data Support at the Faculty of Science/Plan and Design.
- Staff at the UU library (info.rdm@uu.nl) and at the Faculty of Science (rdm-beta@uu.nl) offer support to researchers for drafting data management plans.

3. Collecting and documenting research data

3.1 Focus points from the UU policy

a. Research data must be collected, registered, selected and made available for use in accordance with the applicable international standards.
b. The research data should be documented and provided with clear metadata to ensure they are findable and reusable for further research.
c. Documentation should include clear information on the party or parties holding the rights to the data and the parties entitled to use the data.

3.2 Practical information

- Documentation should include comprehensive information at project and data level, including details on file naming, version control system and folder structure to facilitate data identification within the dataset.
- Next to methodological and data-specific information, the documentation/metadata should specify sharing and access information for the data set, including contact details and licenses or restrictions placed on the data, to clarify conditions for reuse.
- Include in your metadata/documentation clear provenance information about how data was created or generated.
- To increase data findability and interoperability, use where possible (disciplinary) metadata standards and follow the scientific conventions for your discipline for taxonomies and controlled vocabularies.
- For recommended data documentation practices see Research Data Support at the Faculty of Science/Collect and Create.

4. Storing research data

4.1 Focus points from the UU policy

The data are to be stored safely, protected from unauthorized use and loss.

4.2 Practical information

- Make sure you select a storage solution suitable for your storage needs: aspects to consider are for example: data types and sizes; who owns the data and who do you need to share them with; whether you need to work in a collaborative environment; where do you need to be able to access the data from; data safety and privacy. Faculty support (rdm-beta@uu.nl) can assist you in this assessment and in using specific solutions (such as Yoda and the Beta File System)
- Use one master copy location for your data that is backed up regularly, at least also to one physically distinct location. Preferably the master copy location has automatic backup.
- Incorporate best practices for data storage, such applying safeguards to protect the integrity of the raw data and using recommended file formats to increase long-term sustainability and accessibility of your research data.
- Information on best practices for data storage and an overview of available storage solutions at the UU can be found at Research Data Support at the Faculty of Science/Store and Archive.

5. Archiving research data

5.1 Focus points from the UU policy

a. All archived data underlying publications must be available for verification for at least 10 years after publication of the results based on the data, counted from the day of publication.
b. The full set of research data connected with a research project are to be retained upon its conclusion insofar as relevant for the verifiability of the research.
c. If research data are to be destroyed, this should be done in a manner that is verifiable and is irrevocable.
d. Archived research data are to be stored in a structure that is suitable for long-term preservation and later consultation.

5.2 Practical information

- Select data to be preserved for the long-term. Archiving all research data beyond the scope of the project can be a (practical and financial) challenge for big data sets. Try to make a realistic assessment of the total amount of data that needs to be archived and contact faculty support (rdm-beta@uu.nl) about a solution.
- For data preservation and to increase reusability, data files should be ideally preserved in recommended formats (non-proprietary, unencrypted and uncompressed).
- Data may also refer to public, derived or random data sets. In that case, it may suffice to archive the data documentation and metadata that describes the location or condition(s) whereby the original data can be accessed.
- Best practices for data archiving can be found at Research Data Support at the Faculty of Science/Store and Archive.

6. Data security & privacy

6.1 Focus points from the UU policy

a. All research data are to be managed in accordance with the university's requirements governing information security, personal data protection and transparency.
b. The processing of personal data in research should be done in compliance with the *European General Data Protection Regulation (GDPR)*. For specific types of research, additional regulations might apply, such as the Medical scientific research act (*Wet Medisch Wetenschappelijk Onderzoek, WMO*).

6.2 Practical information

- Make sure you are working safely with your data. See Working safely with research data from home and Information security at the UU - Intranet
- Recommendations for handling personal data in research and an overview on available support can be found at Research Data Support at the Faculty of Science/Data Protection.
- When working with personal data, a privacy scan required by the Ethical Review Board of the faculty as part of the ethical approval process.
- For doubts about handling personal data in your research projects you can contact faculty support at rdm-beta@uu.nl or privacy-beta@uu.nl.

7. Data sharing and publishing

7.1 Focus points from the UU policy

a. Archived research data are to be made available for access and reuse at and outside Utrecht University insofar as is reasonably possible and subject to the proper precautionary measures.

b. It is not allowed to transfer exclusive rights to the data, for instance to publishers, without keeping the right to reuse and make the data openly available, unless it is required to acquire external funding.

7.2 Practical information

- It is recommended to use a suitable repository to publish research (meta)data so that data can be registered and indexed in a searchable resource. Utrecht University can assist you in using Yoda and DataVerseNL. For an overview on (disciplinary) repositories and other possibilities for publishing and sharing research data see Research Data Support at the Faculty of Science/Publish and Share.
- Make sure your (meta)data is assigned with a unique and persistent identifier, ensuring that data are findable and citable. Note that most (disciplinary) repositories will assign a persistent identifier when archiving a dataset.
- Make sure that a license (e.g Creative Commons) is provided for your data, creating clarity on data reuse conditions.
- Make metadata of your project is accessible where possible, even if data cannot be made available.
- When working with personal data, is important to consider data protection along with data FAIRness. See Data Privacy Handbook/Data Sharing for reuse for possibilities and alternatives to publishing personal data.