Research Data Management policy
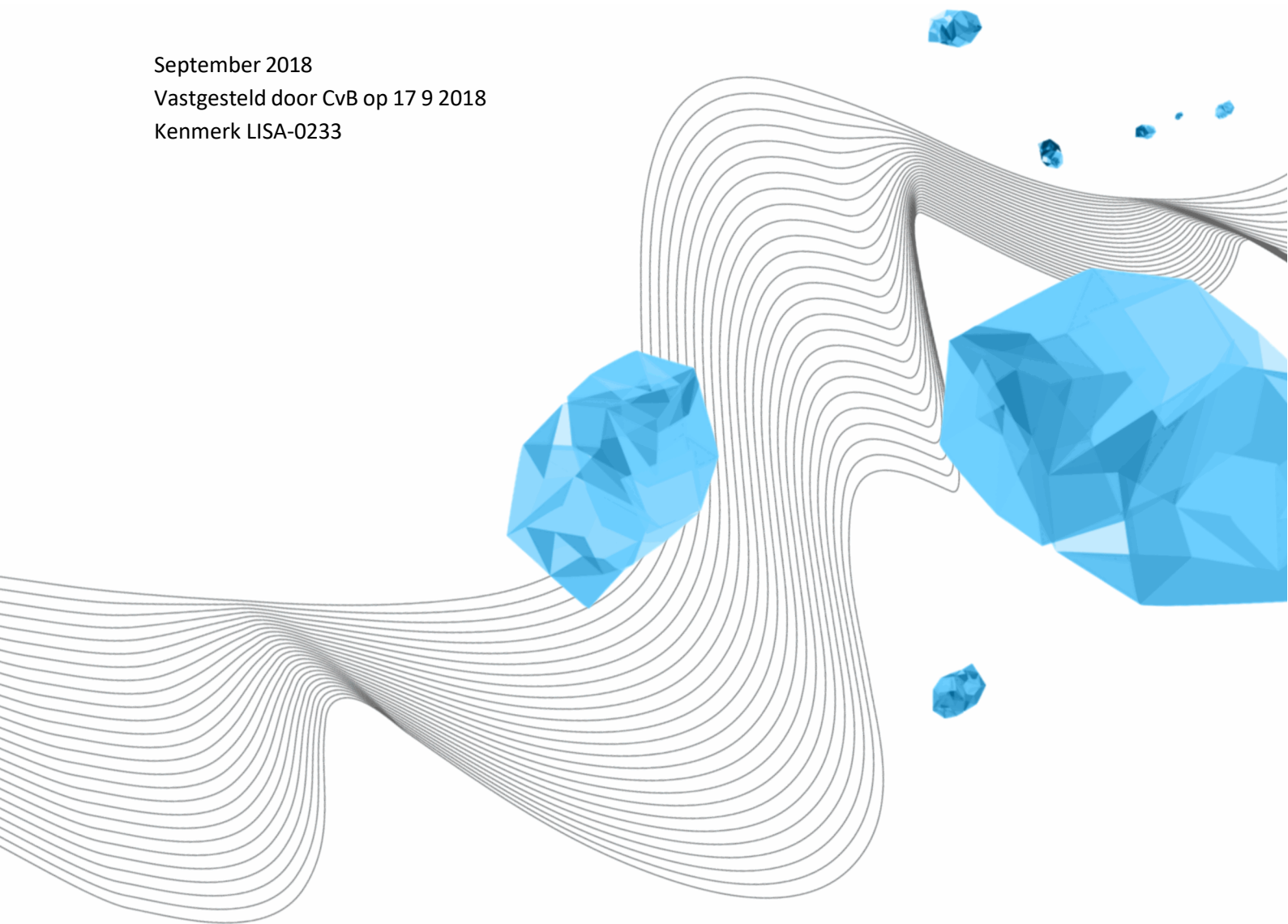# UNIVERSITY OF TWENTE

September 2018
Vastgesteld door CvB op 17 9 2018
Kenmerk LISA-0233

UNIVERSITY OF TWENTE.

CONTENTS

1 INTRODUCTION

Good scientific practice is one of the main reasons to manage ***research data[1]*** during and archive the data after the project. It is important to keep available the raw, processed and/or analyzed data, as well as the documentation necessary for understanding the data and the way it was collected, processed and analyzed. This plays an important role in accountability issues, and allows the researcher to reuse his/her own data or to return to earlier stages of the research process when needed. As a result, research will be reproducible and verifiable. Proper research data management (RDM) makes science more transparent, and improves scientific integrity and societal trust.

Open Science, FAIR data

More recently, ***Open Science*** and the wish to make optimal use of publicly-funded scientific publications and data developed into a second main reason for RDM. The results of research - publications and data- should be available for other scientists and for society. This favors both citizens and entrepreneurs, while researchers benefit in terms of funding and acknowledgement of impact in research evaluations.

As a general policy, research data underlying publications should be Findable, Accessible, Interoperable, and Re-usable (FAIR, see appendix 1). This does not mean, however, that all data should be 'open'. The basic principle is that data should be open to the public whenever possible and closed when needed. Closure, temporary closure (embargo) or limited/restricted access can be justified by, for instance, arrangements with research funders, the General Data Protection Regulation (GDPR, in Dutch: Algemene verordening gegevensbescherming, AVG), public safety, and/or ***intellectual property*** rights of third parties.

These insights are gradually translated and formalized in international, national and local policies and frameworks formulated by science institutions, research funding agencies, and policy makers (see appendix 2). In the Dutch National coalition agreement 2017, 'Open science' and 'Open access' are presented as normative.

New UT data policy

In 2015 the UT accepted the first version of a research data policy. Since that time the issue of the responsibility of universities for managing the research data they generate, obtain and (fully or partially) own has become more important. This is due to the growing attention to the proper handling of research data in scientific practice, the increasing amount of (digital) data, and the growing demand of integrity and open science. Therefore, the UT-data policy of 2015 is updated, resulting in this RDM policy, which pays more attention to the benefits of RDM for the researcher and gives more specifically defined instructions on good research data management at the UT.

The aim of this RDM policy is to stimulate the awareness of the need and benefits of good research data management (section 2) and offer clear rules for managing research data at the UT (section 3). Both sections are relevant for the main target group: the UT scientific community, e.g. researchers and research managers, who have the main responsibility of data management in research projects.

---

[1] Bold italic concepts are explained in appendix 4: Glossary

Besides, also general managers in the faculties and those who are involved in research data management support can benefit from this document, especially section 3.

This Research Data Management policy as well as practical information and general information about RDM, can be found at www.utwente.nl/researchsupport.

2 WHAT BENEFITS CAN RDM BRING RESEARCHERS?

RDM can bring researchers several benefits, which can be categorized as follows:

1. *Reputation and impact*

   - Good RDM allows for the publishing of datasets and consequently its citation. This improves the visibility of your research and therefore its citation rate, reputation, and its scientific impact.
   - When *citizen science* is used for data collection, and open data are used by societal parties, societal impact is improved.
   - More scientific and societal impact will lead to better results in research assessments and staff evaluations.

2. *Quality, efficiency, safety and integrity*

   - Good RDM improves the quality of research. Your data are always accurate, complete, authentic and reliable. Making the data early and easy available to colleagues and reviewers, provides you with an extra quality check.
   - On the basis of your data, others can quite simply verify your results. Open data reduces the risk of manipulation.
   - RDM reduces the risk of data loss by keeping your research data safe and secure.
   - After an initial investment in time, planning for your data management needs and documenting your data throughout its life cycle can save you time and resources in the long run.
   - Enabling other researchers to use your data or being able to use other researchers data prevents duplication of effort.

3. *Reuse, new results and new collaborations*

   - You can benefit from findable and reusable high quality data collected by other researchers.
   - Reuse and re-analysis of data make possible the testing of new or alternative hypotheses and methods of analysis; enables the exploration of topics not envisioned by the initial investigators; encourages diversity of analysis and opinion.
   - Combining datasets from multiple sources through meta-analysis promotes new types of research and permits the creation of new data sets, information, and knowledge.
   - Open data practices facilitate collaboration within and outside your discipline. This allows the crossing of boundaries of disciplines and institutions, and leads to co-publications.

3 RDM POLICY AT THE UT

As a general starting point the UT RDM policy implies that faculties and/or research groups

- support the implementation of the FAIR-principles: Research data must be Findable, Accessible, Interoperable, and Re-usable (see appendix 1)
- commit to the general principle that research data is classified as public data unless there are specific requirements to maintain the confidentiality of research data, such as when a researcher is bound to protect the confidential information of a collaborating company or when the data relates to human subjects
- stimulate awareness regarding RDM practices and develop a culture in which researchers are stimulated to and rewarded for sharing data
- include the provision of RDM-training and -support to researchers
- facilitate solid infrastructure and tools for storing and archiving research data both during and after the research
- monitor the development and execution of **Data Management Plans (DMP)**

Privacy regulations

When applicable, special attention must be given to handling with personal data. Specific RDM policies within a faculty must comply with the GDPR (in Dutch: AVG). For more information about privacy related issues, see:

1. UT privacy policy: www.utwente.nl/en/cyber-safety/cybersafety/legislation
2. UT privacy statement: www.utwente.nl/en/about-our-website
3. Privacy guidelines for research: www.utwente.nl/en/cyber-safety/cybersafety/privacy/guideline-for-research/

Data management plan

Good RDM starts with the writing of a data management plan . Most research funders require a DMP as part of funding application or to be handed in shortly after the start of the project. Specific RDM regulations in the faculty must describe the procedure about when and where a DMP has to be handed in and how reviewing and monitoring is organized.

- *Every research project must have a DMP. The DMP can be derived from the specific RDM* regulations and procedures *at the nearest organizational level in the faculty.*
- *Every PhD-student follows the TGS-course Research Data Management as a preparation of the writing of a DMP.*
- *The DMP has to be reviewed and monitored regularly, in line with planning and progress of the research project.*

Data storage

To prevent data loss **during a research project** *research data and related materials* must be stored properly.

- *All collected research data, including related materials (e.g. protocols, models or questionnaires), must be stored in the ISO 27001- and NEN 7510-certified facilities.* *Certified data storage*

facilities *are offered by the UT-ICT services (LISA). If applicable, terms of use of data suppliers are leading.*

- *Personal cloud services must only be used for copies and comply with legal and contractual conditions. The preferred personal cloud service is Surfdrive. This service complies to the Dutch and European privacy legislation.*
- *Storage of research data on portable devices must be avoided as much as possible, if needed should only be used for copies and must comply with legal and contractual conditions.*
- *Personal, confidential or classified research data and related materials, such as consent forms, must be stored in accordance with relevant Dutch legislation and European regulation and the VSNU Conduct code for the use of personal data in scientific research (UNL Code of conduct for using personal data in research [in Dutch]) for which UT-storage mentioned above is available.*
- *Non-digital research data and related materials, such as physical samples and lab notebooks, must be handled in accordance with clearly described procedures within the organisational unit and/or project.*

Data documentation

In line with the FAIR principles data must be well documented, during the dynamic phase of the data and especially as soon as they have become static.

- *Research data and related materials, both digital and non-digital, must be accompanied by proper metadata and documentation in such a way that it enables the verification, replication and, if possible, reuse of the data. This documentation must also contain information about property rights and terms of use.*

Data sharing

To guarantee that data can be accessed and checked during the research digital and/or non-digital research data and related materials must be shared. Implementation is dependent of intellectual property and responsibilities regarding research data, and the terms of use of data suppliers.

- *During the research project data and related materials must be shared in such a way that, apart from the researcher, it can be accessed by at least one other member of the organisational unit. Before the end of the project all research data and related materials which is needed for verification/replication and reuse, must be made available to the responsible in the organisational unit (see also RDM roles and responsibilities).*
- *Personal, confidential or classified research data and related materials, such as consent forms, must be shared in accordance with relevant Dutch legislation and European regulation and the VSNU Conduct code for the use of personal data in scientific research (UNL Code of conduct for using personal data in research [in Dutch]). The ISO 27001- and NEN 7510-certified project- and organization directory of the UT is recommended for secure sharing of personal data.*
- *In case of a Non-Disclosure Agreement with third parties, arrangements must be made about sharing data during the research.*

Data archiving

For the purpose of verification, replication and reuse research data and related materials must be archived after the research project. Implementation is dependent of intellectual property and responsibilities regarding research data, and the terms of use of data suppliers

- *Selection of research data and related materials for long-term preservation is based on what is needed for verification/replication and reuse. This must be at least the research data that form the basis of and can therefore be linked to publications. The selection to be archived can also comprise the full set of raw and/or processed data. In case of large data sets archiving costs, for both preparation and storage, should be taken into account when applying for project funding.*
- *Preferably during, but not later than 1 month after finishing the research selected data and related materials are archived in both group or faculty facilities and in a **trusted repository**, in accordance with FAIR-principles (see appendix 1) and, in compliance with legal and contractual conditions, openly accessible. The preferred archive for data from the technical and beta sciences is 4TU.ResearchData, and DANS for data from the social sciences.*
- *Selected research data and related materials must be archived at least for 10 years, unless legal or contractual regulations demand another term.*
- *Non-digital research data and related materials, such as physical samples or lab notebooks, must be archived in secure UT-provisions accompanied with clearly described access procedures.*
- *In case of a Non-Disclosure Agreement with third parties, arrangements must be made about archiving and sharing of data for verification and replication.*
- *Archived research data and related materials, both digital and non-digital, are accompanied with proper metadata for findability and good documentation for reasons of interpretation and reusability (see also data documentation and data registration).*

Data registration

Because research data is becoming a valuable asset and in the near future will be formally recognized as scientific output, it is important to know what digital and/or non-digital research data and related materials have been created or used and where these are located.

- *In addition to archiving, all digital and/or non-digital research data and related materials must be registered and described by metadata, including a link or reference to the location of the digital or non-digital objects.*
- *The preferred system for registration of digital and/or non-digital research data and related materials is Pure Research Information because of automatic ingest of metadata from other systems (such as 4TU.ResearchData). Moreover these data and related materials can be linked to the registered UT publications based on them.*

4 RDM POLICY IMPLEMENTATION

Although dealing with research data may vary a lot per research discipline, the Executive Board has formulated this RDM policy as a framework of proper and responsible handling of research data. Faculties must implement this policy at one or more organizational levels (faculty, department, group and/or other unit) and supplement it with regulations and procedures related to the specific character or circumstances of the research discipline in question.

RDM regulations and procedures in the faculty must be available on the relevant websites or intranet and communicated in such a way that every researcher can easily find these.

RDM roles and responsibilities

When implementing the UT RDM policy in the faculty it is important to identify and formalize RDM roles and responsibilities. This can comprise responsibilities like having and maintaining a RDM policy on one or more organizational levels in the faculty as well as monitoring the compliance to this policy. Also, operational RDM responsibilities must be addressed, like stimulating RDM awareness, the writing and monitoring of **Data Management Plans (DMP)**, and the implementation and maintenance of necessary RDM facilities. An overview of RDM roles and responsibilities can be found in appendix 3.

In case of third party involvement, roles and responsibilities regarding research data management must be included in agreements. Good research data management is the shared responsibility of all parties involved. In principle, intellectual property rights on research data ("database right") shall vest in UT.

In line with the university's responsibility regarding RDM, support, infrastructure and facilities are offered by service department LISA. Specific RDM support, infrastructure and facilities can also be offered by faculties or other organizational units, in some cases in cooperation with LISA.

APPENDIX 1: FAIR PRINCIPLES

Preamble

One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of, task-appropriate scientific data and their associated algorithms and workflows. Here, we describe **FAIR** - a set of guiding principles to make data **Findable, Accessible, Interoperable, and Re-usable**.

To be Findable:

F1. (meta)data are assigned a globally unique and eternally persistent identifier.
F2. data are described with rich metadata.
F3. (meta)data are registered or indexed in a searchable resource.
F4. metadata specify the data identifier.

To be Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
A1.1 the protocol is open, free, and universally implementable.
A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
A2 metadata are accessible, even when the data are no longer available.

To be Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
I2. (meta)data use vocabularies that follow FAIR principles.
I3. (meta)data include qualified references to other (meta)data.

To be Re-usable:

R1. meta(data) have a plurality of accurate and relevant attributes.
R1.1. (meta)data are released with a clear and accessible data usage license.
R1.2. (meta)data are associated with their provenance.
R1.3. (meta)data meet domain-relevant community standards.

Source and further information: https://www.force11.org/group/fairgroup/fairprinciples

APPENDIX 2: REGULATIONS, GUIDELINES, CODES, AND POLICIES RELEVANT FOR RDM

This is a list of regulations, guidelines, codes, and policies which are useful or necessary to read when writing a data policy or data management plan. This RDM policy is based or in line with these guidelines.

INTERNATIONAL

- European Code of Conduct for Research Integrity (https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf)
- EU General Data Protection Regulation (GDPR) (https://gdpr-info.eu/)
- EU guidelines for data management (http://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm)

NATIONAL

- National Plan Open Science (https://www.openscience.nl/en)
- Netherlands Code of Conduct for Research Integrity (https://www.universiteitenvannederland.nl/files/documents/Netherlands%20Code%20of%20Conduct%20for%20Research%20Integrity%202018.pdf)
- Algemene verordening gegevensbescherming (AVG, see also GDPR) (https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-nieuwe-europese-privacywetgeving, in Dutch)
- UNL Code of conduct for using personal data in research (UNL Code of conduct for using personal data in research [in Dutch])
- NWO Data management protocol (https://www.nwo.nl/en/policies/open+science/data+management)

LOCAL (University of Twente)

- Code of Conduct on ICT and Internet Use (https://www.utwente.nl/en/cyber-safety/cybersafety/legislation/gedragscode-ict-mw-en.pdf)
- Information security policy (https://www.utwente.nl/en/cyber-safety/cybersafety/legislation/informatiebeveiligingsbeleid-engels-def.pdf)
- Privacy policy UT (https://www.utwente.nl/en/cyber-safety/cybersafety/legislation/privacy-policy-university-of-twente-20161017-nieuwe-linkjes-2017.pdf)

APPENDIX 3: ROLES AND RESPONSIBILITIES

This is an overview of suggested RDM roles and responsibilities which can be implemented in the faculty.

THE RESEARCHER

- writes a data management plan (DMP) in accordance with the RDM data regulations and procedures of the nearest organizational unit in the faculty
- obtains and maintains knowledge of data management by means of RDM-course(s)
- develops and adopts appropriate procedures and processes for collecting, documenting, storing, processing, using, accessing and sharing of the collected or generated research data and for selecting and archiving the research data
- guarantees the integrity, quality, security and persistent availability of the collected or generated data
- acts in accordance with the Personal Data Protection Act (GDPR) and other legal, contractual and ethical rules and regulations
- budgets the costs and time investment for data management
- updates the DMP when necessary.

THE PROJECT COORDINATOR/SUPERVISOR(S)

- reviews the DMP written by the researcher
- checks the compliance of the DMP with the relevant RDM regulations and procedures in the faculty/university, and with legal, contractual and ethical rules and regulations
- checks that the DMP is part of the researcher's qualifier report (in case of PhD'ers)
- monitors the correct execution and updating of the DMP in accordance with the relevant RDM regulations and procedures in the faculty/university and with legal, contractual and ethical rules and regulations
- documents the agreements made on data management in the case of joint research projects or contract research where responsibility for data management rests in principle with the project coordinator
- checks the integrity, quality, security and persistent availability of the collected or generated data
- arranges the necessary resources, facilities and support for data management in the research

THE HEAD OF THE RESEARCH GROUP

- is responsible for having RDM regulations and procedures on the level of his or her own research group or disseminating information in the group about RDM regulations and procedures on other level(s) in the faculty.
- supervises and monitors the correct execution and updating of the group RDM regulations and procedures (if available)
- supervises and monitors the correct execution of data management in the research group in accordance with the group regulations and procedures or other relevant RDM regulations and procedures in the faculty and with legal, policy and ethical rules and regulations

- is responsible for the correct selection and persistent availability of data of all projects of the research group for the purpose of verification/replication and reuse
- arranges the availability of the necessary resources, facilities and support for data management in the research group
- creates and supervises awareness and keeps knowledge of the group or other relevant RDM regulations and procedures in the faculty, and data management in general, in the research group at the desired level

THE FACULTY BOARD

- is responsible for having RDM regulations and procedures at one or more organizational levels in the faculty as implementation of the UT RDM policy, and in accordance with legal, contractual and ethical rules and regulations
- supervises and monitors the correct execution and updating of the RDM regulations and procedures in the faculty in accordance with legal, policy and ethical rules and regulations
- arranges the availability of the necessary resources, facilities and support for data management in the faculty
- creates and supervises awareness and keeps knowledge of RDM regulations and procedures in the faculty, and data management in general, in the faculty at the desired level.

THE RECTOR MAGNIFICUS

- is responsible for having and maintaining a UT RDM policy which contributes to scientific integrity and societal trust
- facilitates and monitors the implementation of the UT RDM policy as a framework for good research data management on other levels in the university

RDM SUPPORT RESPONSIBILITIES

On all levels the execution of the RDM policy and operational RDM responsibilities is supported by members or organisational units within the faculties, in certain cases in cooperation with the service departments. Service departments have different and in some cases shared responsibilities for RDM support. As a basic principle, faculties determine what RDM-support they demand from the service departments.

General RDM: LISA

- Coordination of research data management support
- Advice and support on general research data management issues, during both planning and implementation
- Keeping record of and participation in national and international RDM developments

Policy issues: LISA and S&B

- Data policy or procedure preparation and implementation support on university, faculty and research group level
- Support on compliancy of data policies and procedures with relevant legislation, regulations, guidelines, etc.

Awareness and training: LISA

- Organization of courses/workshops RDM for researchers and research groups.
- Stimulation of awareness about RDM on individual and organizational level.

Funder requirements: SBD (EU-office) and LISA

- Advice and support on research data management requirements of funders.

Legal and ethical issues: General Affairs and LISA

- Advice and support on legal and ethical issues related to the collection, storage, access, sharing and archiving of research data
- Registration and investigation of data breaches

Infrastructure and facilities: LISA

- Infrastructure and facilities for collection, storage, access, sharing and archiving of research data, both internal and external UT (preferably integrated with infrastructure and facilities for data processing and analysis)
- Infrastructure and facilities for writing and monitoring data management plans and RDM procedures

APPENDIX 4: GLOSSARY

Certified storage	Storage formally approved as compliant with one or more standards. The UT network storage (M- and P-drive) is certified for ISO 27001 and NEN 7510, which means that it is suitable for storing the most highly sensitive data (e.g. medical data).
Classified data	Data classified to a certain level of sensitivity and related level of protection.
Citizen science	Scientific research conducted, in whole or in part, by amateur (or non-professional) scientists. (source: Wikipedia)
Data management plan	Formal document that outlines how data are to be handled both during a research project, and after the project is completed. It normally consists of the following sections: data collection, storage and backup, documentation, access, sharing and reuse, and preservation and archiving.
Intellectual property	Category of property that includes intangible creations of the human intellect, and primarily encompasses copyrights, patents, and trademarks. (source: Wikipedia)
Metadata	Data which provide information about a dataset, such as the creator(s), the subject of the data, the type of file, and geographic information.
Open Science	The practicing of science in a sustainable manner which gives others the opportunity to work with, contribute to and make use of the scientific process. (source: National Plan Open Science)
Personal cloud service	Storage of data in the cloud, related to a personal account. When the account is cancelled, data will be deleted automatically.
Research data (and related materials)	Data, both physical and digital, that is collected, created and/or used in the framework of a research project and (partly) intended to produce and validate research findings. In the context of research data management also data related materials, such as models, instruments, lab notebooks, protocols, questionnaires and informed consents must be taken into account.
Trusted repository	Trusted repositories are data archives for long-term preservation having a Core trust seal or Data Seal of Approval (https://www.coretrustseal.org/)