EEMCS Research Data Management Policy

Based UT RDM policy, Confirmed by CvB on 17-09-2018

Last edited / Released on 11-01-2019

Scope

This policy document applies to all activities in EEMCS in which Research Data is acquired or generated, processed, interpreted, archived, published, shared and/or distributed or deleted. This policy is based on the UT Research Data Management (RDM) policy confirmed by CvB on 17 September 2018.

Background and objectives

This policy is set to ensure proper research data management at the faculty EEMCS in order to

- Secure scientific integrity
- Stimulate reuse of the data
- Adhere to laws, policies and funder requirements

EEMCS wants to stimulate the durable storage of research data and to make it as widely accessible as possible both during and after the research. However privacy, non-disclosure agreements and data size issues may prevent the use of shared data resources and demand dedicated resources and policies.

This policy will establish how to reach these goals within the faculty EEMCS by providing clear criteria and minimum requirements for proper research data management. It follows the requirements set in the UT RDM policy.

The requirements for securing scientific integrity are also based on the S*tandard Evaluation Protocol 2015-2021* by KNAW, VSNU and NWO and *The Netherlands Code of Conduct for Scientific Practice, revision 2014* of the VSNU.

Other policies and laws to keep into account: Privacy policy UT, Personal data protection act, Ethics protocol EEMCS, Funder requirements. For links to all regulations, guidelines, codes, and policies relevant for research data management can be found in Appendix 1.

Basic principles

Good RDM starts with the writing of a data management plan (DMP). For each research project, including each PhD-project a DMP should be formulated and should be regularly updated during the project.

During the research the data should be stored in such a way to minimize risk of data loss and to maintain data integrity. Use of portable storage should be minimized and data should be accessible by at least one extra member of the research group besides the principle investigator. Storage and access should be managed in accordance with legal regulations and third party contractual requirements. During the research metadata and documentation should be kept up to date.

When finalizing the research and publicizing the results all data necessary for verification and replication should be archived in a trusted repository (with the Data seal of approval). Each data set should be described by metadata (preferably using a metadata standard format) and be accompanied by clear documentation.

The archived data should be as openly accessible as possible without restrictions on reuse, limited only by legal regulations and third party contractual requirements.

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. Therefor all archived data sets should be registered in UT Research Information.

With these basic principles EEMCS RDM policy also follows the FAIR data principles: Findability, Accessibility, Interoperability and Reusability (see also Appendix 2).

Policies

The faculty policy is derived from the UT RDM policy. Underneath the guidelines from the UT policy are copied (numbering added for convenience) for each part of proper research data management and are then complemented with the EEMCS guidelines.

Data management planning

UT guidelines

UT 1.1 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.

UT 1.2 Every PhD-student follows the TGS-course Research Data Management as a preparation of the writing of a DMP.

UT 1.3 The DMP has to be reviewed and monitored regularly, in line with planning and progress of the research project.

EEMCS guidelines

EEMCS 1.1 The DMP should be stored in such a way that they are accessible at least to the group, but preferably to the faculty. The Group Data Policy should describe where the DMP's are stored.

Data storage (during a project)

UT guidelines

*UT 2.1 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.*

*UT 2.2 **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.*

UT 2.3 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.

*UT 2.4 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 (VSNU Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek) for which UT-storage mentioned above is available.*

UT 2.5 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.

EEMCS guidelines

EEMCS 2.1 If the research group uses their own storage facility, the Group Data Policy should contain a section on management and backup policy concerning this infrastructure. The storage facility should be certified and access should be managed in accordance with legal regulations and third party contractual requirements.

Data documentation

UT guidelines

UT 3.1 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.*

EEMCS guidelines

EEMCS 3.1 For data the recommended license is Creative Commons (CC-0 or CC-by), for software, depending on third party software/libraries used, MIT or GPLv3.

Data sharing (during the project)

UT guidelines

UT 4.1 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).

UT 4.2 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 (VSNU Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek). The ISO 27001- and NEN 7510-certified project- and organization directory of the UT is recommended for secure sharing of personal data.

UT 4.3 *In case of a Non-Disclosure Agreement with third parties, arrangements must be made about sharing data during the research.*

Data archiving (at the end of the project)

UT guidelines

UT 5.1 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.

UT 5.2 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 2) 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.*

UT 5.3 Selected research data and related materials must be archived at least for 10 years, unless legal or contractual regulations demand another term.

UT 5.4 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.

UT 5.5 In case of a Non-Disclosure Agreement with third parties, arrangements must be made about archiving and sharing of data for verification and replication.

UT 5.6 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).

EEMCS guidelines

EEMCS 5.1 Preferred trusted repository for archiving research data from EEMCS is 4TU.ResearchData, for software, especially when still being developed, github together with zenodo.org can be used, as long as the version used in a publication is also stored at 4TU.ResearchData.

Data registration

UT guidelines

UT 6.1 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.

UT 6.2 The preferred system for registration of digital and/or non-digital research data and related materials is UT Research Information because of automatic ingest of metadata from other systems (such as 4TU.RD). Moreover these data and related materials can be linked to the registered UT publications based on them.

Implementation and responsibilities

EEMCS is an cluster of 3 disciplines each of them consisting of a diverse portfolio of research groups. Each research groups consists of a number of researchers. The type, class and amount of data collected, used, and generated during research is very diverse within EEMCS and differs per research group. Storage and backup facilities may also differ per research group. For this reason with respect to the implementation of this policy we distinguish 3 levels: faculty, group and individual researcher.

- On faculty level this policy document lists the criteria for proper research data management.

- In addition each group is asked to formulate their own Group Data Policy (GDP). This policy details how the faculty policy is translated to the research practices in the group.

- Finally, the individual researcher is responsible for proper execution of the research and handling of the data, in conformance with the prevailing GDP.

Roles & Responsibilities in detail

The most important responsibilities for each of the three levels are further detailed below. A full set of responsibilities copied from the UT RDM policy is available in Appendix 3.

- The Board of EEMCS is responsible for this policy and the implementation. In particular the portfolio holder research is assigned with this task. They
    - oversee the creation of the Group data policies
    - review the EEMCS data policy annually

    - support the groups in providing the necessary infrastructure

  It is also the responsibility of the faculty to arrange research data management support and expertise for the research groups. At the moment this consists of:

    - The ICT account manager for EEMCS gives advice and support on infrastructure and storage facilities.
    - The information specialist for EEMCS gives advice on research data management policies and regulations.
    - The privacy contact person for EEMCS gives advice on privacy issues and the use of personal data in research.
    - The Ethics committee of EEMCS advices on ethical issues related to research projects involving human beings.

- The chair of the research group is responsible for proper data management within all research projects performed within the group. The chair
    - oversees the writing and annually reviewing of the Group data policy. It details the type of research data collected, the roles and responsibilities within the group, and the procedures for managing research data during and after the research. In Appendix 4 the guidelines for setting up a GDP are given.
    - makes sure that everyone within the group knows about the Group data policy
    - has the final say on which data sets are archived, deleted or published openly for all research projects performed within the group (if multiple groups within EEMCS take part in a project, this responsibility escalates to the Dean of EEMCS)
    - provides the necessary infrastructure for storing and archiving the research data during and after the research

  Within each group one person is identified to be the primary contact person for research data policy. The policy should be reviewed annually.

- Every researcher is responsible for the way they deal with research data, in some cases together with the lead researcher. Appendix 5 gives an overview of data types and classes with some practical pointers for the proper handling of those data. Below an implementation checklist when preparing, executing and finalizing a research project are presented.

Research data management implementation within a research project		
Preparing	**Executing**	**Finalizing**
Follow TGS course on RDM for PhD students (UT 1.2)	Securely store your data on ISO 27001- and NEN 7510-certified facilities (UT 2.1, EEMCS 2.1)	Archive all data necessary for verification and replication in 4TU.ResearchData (UT 5.1, UT 5.2, UT 5.3, EEMCS 5.1)
Write a data management plan considering group data policies/protocols and funder requirements (UT 1.1)	Make sure your data is accessible to at least one other in the group (UT 4.1)	All archived data should be accompanied by information on property rights and terms of use (UT 3.1, EEMCS 3.1)
Store DMP at a central location within the group/faculty (EEMCS 1.3)	Store and share your data according to legal regulations and agreements with third parties, with special attention to personal, confidential or classified data. (UT	Archive non-digital data according to group procedures (UT 5.4)

	2.2, UT 2.3, UT 2.4, UT 4.2, EEMCS 2.1)	
When planning to work with human subjects or privacy sensitive data, check the UT privacy policy and the EEMCS ethics protocol.	Keep your documentation and metadata up to date (UT 3.1)	Make sure all data is accompanied by documentation and metadata such that data is findable and reusable (UT 3.1, UT 5.6)
Make agreements with third parties (UT 4.3, UT 5.5)	Regularly update your DMP (UT 1.3)	Register all your datasets in the UT Research Information system (UT 6.1, UT 6.2)
	Store non-digital data according to group procedures (UT 2.5)	

Appendix 1: 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.

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://www.eugdpr.org/)
- 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)
- The Netherlands Code of Conduct for Academic Practice (in revision). (http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/Code_wetenschapsbeoefening_2004_(2012).pdf, in Dutch)
- Algemene verordening gegevensbescherming (AVG, see also GDPR) (https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-nieuwe-europese-privacywetgeving, in Dutch)
- VSNU Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek (consultatieversie: https://www.vsnu.nl/files/documenten/Domeinen/Governance/Consultatieversie%20-%20VSNU%20Gedragscode%20voor%20gebruik%20van%20persoonsgegevens%20in%20wetenschappelijk%20onderzoek.pdf, 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)

EEMCS

- Ethics protocol EEMCS (https://www.utwente.nl/nl/ewi/onderzoek/formulieren-en-downloads/ewi142099.pdf)

Appendix 2: FAIR data 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 3: Responsibilities

This is an overview of suggested RDM roles and responsibilities which can be implemented in the faculty, copied from the UT RDM policy (DRAFT v1.5).

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: Group data policy template

General

Responsibilities

- Who are the persons in the group with specific roles and responsibilities for the data management policy and research data management of the group?

Implementation

- How this data management policy will be implemented and managed? Who is/are responsible for this?
- How do you bring and keep awareness and knowledge of data management policy and research data management in the group on the desired level?

Resources

- How human and financial resources are made and kept available in the group for research data management which cannot be paid from research project budgets?

Data management planning

- What class and type of data are usually collected, processed or analyzed in the group?
- Do data collected in the group need special attention with regard to reproducibility?
- Is there a need for specific regulations, procedures and infrastructure with regard to confidential, personal or classified data
- What data-related material should be part of the data management policy?
- What is the estimated total size of the data of the group, and what is the growth rate?
- What is the estimated number of files and the maximum file size?
- Is there a need for specific regulations and infrastructure with regard to version control?
- Is there a need for specific procedures with regard to data quality?

Data storage

- Which data and data related material are to be stored on the level of the group? (distinction raw-processed/analysed – published data)
- Who is responsible for deciding what has to be stored?
- Which medium or infrastructure managed by the group is used for storing research data?
- What is the backup strategy and how backup is managed?
- Who is responsible for managing the storage and backup of research data in the group?

Data documentation

- How is documentation of data and data related material managed on group level? Indicate for instance the use of general metadata scheme, file and folder naming convention, directory structure, etc.
- Who is responsible for proper documentation when storing data and data related material on group level?
- How the group deals with issues like copyright, IP and licences for reuse with respect to research data and data related material? Refer to codes of conduct, copyright agreements, data protection acts, data security standards, etc.
- How and by whom copyright, IP and licences of reuse of data and data related material will be managed?

Data sharing

- What are the regulations and infrastructure of data sharing and access, both internal and external?
- Who is responsible for sharing and access management of the data and data related material?

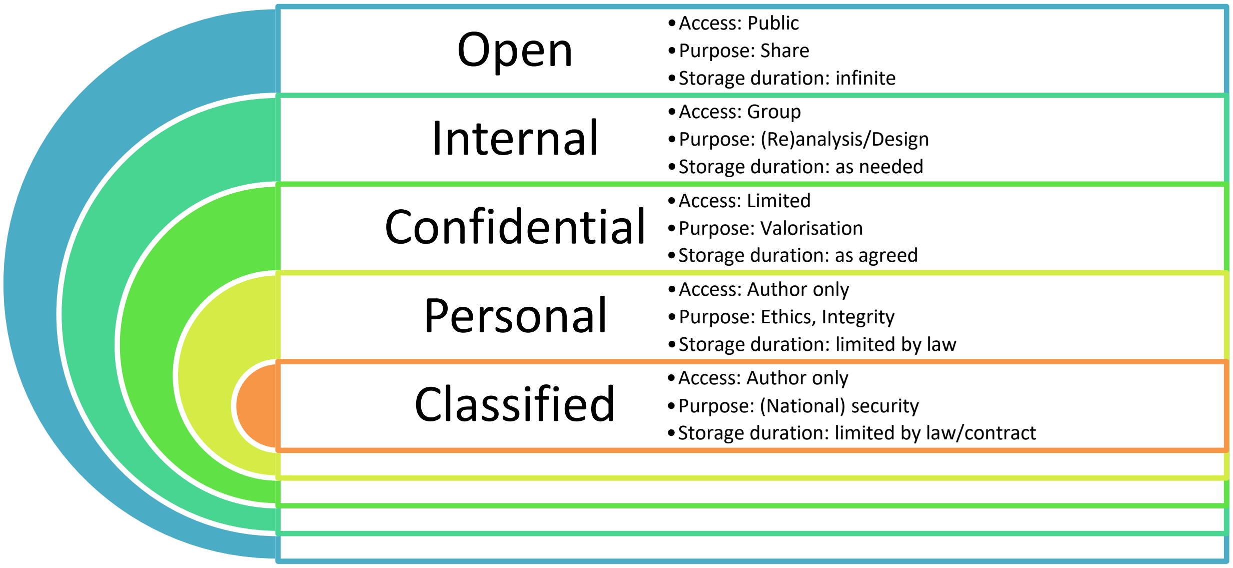
Data archiving

- Which criteria are used to decide whether data and data related material have to be archived for preservation and long-term availability or can/should be destroyed?
- Which data repository is used for preservation and long-term availability of the data and data related material?
- What is the policy of preservation and long-term availability of data linked to publications, i.e. journal articles or PhD-theses?
- Who is responsible for preservation and long-term availability of data and data related material?

Appendix 5: Research data types

Data classes

We will distinguish between 5 data classes: (1) Classified data, (2) Personal data, (3) Confidential Data, (4) Internal Data or (5) Open Data



Open
- Access: Public
- Purpose: Share
- Storage duration: infinite

Internal
- Access: Group
- Purpose: (Re)analysis/Design
- Storage duration: as needed

Confidential
- Access: Limited
- Purpose: Valorisation
- Storage duration: as agreed

Personal
- Access: Author only
- Purpose: Ethics, Integrity
- Storage duration: limited by law

Classified
- Access: Author only
- Purpose: (National) security
- Storage duration: limited by law/contract

Personal data

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/

Special data types

Furthermore there are special types of data that may need to be handled in a special way:

Software code

- When distributing open source software, the choice of license may depend on the licenses of third party open source software (e.g. libraries) that was used.
- For archiving software zenodo.org is a trusted repository that will give you a DOI for the software, and you can connect with your github account/project

Secundary data/data from others

- When using data collected or owned by others, especially confidential or personal data, there should be a *Data Processing Agreement* to stipulate under which terms the data can be used, shared, analyzed and archived.
- How to secure verifiability and replicability should also be included in the agreements
- And roles and responsibilities regarding research data management must be included in agreements.

Intellectual property

- When (Master) students are working on projects that could result in patent applications or other intellectual property, there should be agreements on ownership and possibly non-disclosure.
- When working on research that could result in patent applications, make sure to keep your research documentation very accurately

Physical data

- The group data policy should contain procedures for storing and archiving physical data
- When it is not possible to archive or preserve your physical data/samples because of costs, size or sustainability of the samples, make sure to document extensively and when possible make photos