CHECKLIST RESEARCH DATA AND SOFTWARE MANAGEMENT

EEMCS

Version 05-06-2025

This checklist will be updated regularly. For questions on research data or software, please contact the data stewards of EEMCS: Marianna Avetisyan or Shaokang Zhang.

PLANNING

1. Familiarise yourself with relevant **policies, laws and regulations (***website*)
2. If applicable, write a **data management section**. Most funders will require this as part of the funding application. (*funder requirements*)
3. Write a **data management plan (DMP)** (*DMPtool*). For PhDs, this is part of the TGS course. If necessary, determine the **costs** of your data management. (*guide*)

> **Data management plan (DMP)**
>
> A data management plan describes:
> - what data will be collected and how, as well as which software will be used for collecting, processing and analysis,
> - how will the data (and software) be saved during the research project,
> - how will the data be shared and transferred securely,
> - how will the data (and software) be made sustainably available and, if possible, published afterwards (watch this video on research data sharing),
> - how the data will be documented,
> - what legal issues are relevant, such as copyright, the right to reuse the data and the treatment of sensitive data (link for personal data: to be changed).

4. If you are processing **personal data** during your research, register your processing in the **GDPR registration too**l (accessible via DMPtool)
5. If you are developing research software as part of your project, a **software management plan (SMP)** might be necessary. *Please check the Practical Guide to Software Management Plans for guidance*

DATA DOCUMENTATION

1. Ensure all research data is stored, archived and published with **metadata** (example) **and additional documentation, including a README file** (example).
2. Add **discipline-specific** metadata or keywords (link to discipline-specific metadata schemes)
3. **Research software** should include **embedded comments** for usability.
4. Documentation for **research software** should include **user documentation and deployment documentation**, and in some cases, also **developer documentation**.

Metadata: can help others (and yourself) to find your data

Readme file: can help others (and yourself) to understand and re-use your data, and to reproduce your results

	Metadata	Readme
Descriptive information	E.g, author, contributor, title, abstract, keywords, measurement type, project ID, geomapping, time period, and subject area	E.g. software scripts, instrument settings, methodology, experimental protocol, codebook, laboratory notebook, name and version of used third-party software/tools
Structural information	E.g. Relations i.e related datasets or publications, related projects, version.	E.g. database scheme, relations between files, table of content.
Administrative information	E.g. data format, date, size, access rights, preservation period, persistent identifier (PID, to cite your data), license.	E.g. user agreements, provenance (description of the origin of the data).

Note: Make sure the research metadata is machine readable. The machine readability allows systems and software to extract, analyse, and act upon the data effectively, which is essential for tasks such as search engine optimization, data exchange between systems, automation, and ensuring interoperability. There are 3 key characteristics of machine-readable metadata: standardized format, consistency of structure and format, accessibility to machines.

DATA STORAGE

1. Store all collected data on the **UT network storage** provided by LISA. Recommended storage is Unishare. Other options can be found here: *storage decision tree*. **The stored data should be accessible by at least one other research group member.**
2. If it is difficult or even impossible to use the UT network storage, a **suitable alternative data storage** option must be decided in coordination with the **faculty ICT account manager**.
3. A **copy** of the research data can be stored in **personal cloud** services, preferably SurfDrive (OneDrive is **not** recommended).
4. Use of **portable devices** (external hard drive, USB stick or personal laptop) **must be avoided** as much as possible. Always use encryption (more info: *More info on encryption.*)
5. **Research software** should be developed using a **version control system. Open software development** is recommended, where possible.
6. **Non-digital research data** and related materials, such as physical samples, lab notebooks and informed consent forms, must be stored in accordance with **clearly described procedures** and standards within the research group and/or project and digitised where possible.
7. **For personal data:**
    a. The **loss** of personal or confidential data must be considered a **data breach**.
    b. **Consent forms** should be stored **separately** from the data.
    c. In the case of **encrypted data files**, the **key** should be stored in a **separate location** and shared with at least one other employee in the research group. Generally, this person will be the principal investigator or the chair of the research group.

DATA SHARING

1. **UniShare** is the most recommended channel to be used to share data with colleagues within the University of Twente (UT) as well as with external people. SurfDrive service can also be used, if necessary.
2. For **transferring large files** Surf file sender is recommended.
3. In case of a **Non-Disclosure Agreement** with third parties, make sure all people with access to the data fall under the agreement. Bachelor's and Master's students with access will have to sign **separate agreements** with the UT.
4. Before sharing **personal data** with external parties or receiving data from external parties, a **Data Agreement** should be drawn up, establishing the conditions under which data are shared, with the assistance of the faculty Privacy Contact Person

DATA ARCHIVING/PUBLISHING

1. Not later than one month after **publishing a scientific work** (paper, thesis or report), all data which are the basis of published results should be preserved for at least ten years and in accordance with **FAIR principles**.
2. The **recommended repository** for archiving and publishing data underlying a publication is 4TU.ResearchData.
3. **Raw data sets** that need to be archived for further use within the group can be stored in AREDA (a static, low-cost archiving solution).
4. To promote the **visibility and the sharing of your datasets**, we recommend referring to the **DOIs** of your datasets in your articles or PhD thesis. You can reserve a DOI in advance.
5. **For personal data**:
   - **Pseudonymised** personal data that does not fall under the special categories as stated in the GDPR can be archived in **4tu.Researchdata** with restricted access or encrypted in AREDA.
   - **Consent forms** and (encryption) **keys** must be archived separately from the pseudonymized data.
   - For personal data, there is a **maximum retention period** (usually 10 years) after which the data needs to be deleted or completely anonymised.
6. 4TU.ResearchData allows you to easily archive and publish your **research software** directly from your Git repo.

DATA REGISTRATION

1. All digital and/or non-digital research data and software must be **registered** and described by metadata in **UT Research Information (Pure).**
2. For each **publication** in UT Research Information, the **underlying data should be linked**, either by adding the link to the published dataset under *Electronic version(s), and related files and links,* or under *Relations.*
3. **Research software** should be registered in Pure under Data sets.