Guideline on Ethics, Privacy & Research Data Management BMS

Version 3.0 6 May 2024

1. ROLES & RESPONSIBILITIES

It is the individual researcher/first author of the project who always bears ultimate responsibility for the **proper handling** of data. Before the start of a project, researchers must always ensure that they have the appropriate expertise and capacity available to carry out the research in line with the VSNU's guidelines and Code of Conduct. This is one aspect of a professional attitude in science. Furthermore, in the case of junior researchers and PhD students, the first supervisor of the project (with a PhD) bears **joint responsibility** for proper data management, including verification of the data packages involved. The faculty will monitor events at a distance, and will perform occasional data package audits.

2. COURSES, GUIDELINES & TEMPLATES

Courses

- All **PhD candidates** are obliged to follow the TGS course Research Data Management Bootcamp (online course + interactive session + DMP). This course is organized 5 times per year, registration should be done via this link: https://www.utwente.nl/en/ctd/courses/1000227/data-management-bootcamp/.
- **Other employees** can access the content of the RDM online course on Canvas via this link: https://canvas.utwente.nl/courses/2167.
- **Students** can check the micro-lectures to get to know the topic of research data management.
- **Handling Personal Data in Research**: This Canvas course, which is open to every UT staff member, will help you understand the range of perspectives needed to build and demonstrate compliance with privacy regulation in research. Please enroll via this link https://canvas.utwente.nl/enroll/LGWB6F.

Guidelines

- Guide with examples on writing a data management section in NWO proposals
- Guide on handling personal data: Appropriate Use of Personal Data in Research According to GDPR (https://www.utwente.nl/en/bms/research/ethics/informed-consent-procedure/)

Templates

- DMP template: a DMP template with guiding information is available in the UT DMP tool, which has been accepted by funders such as NWO, ZonMw and the EU. Therefore it is recommended to write your DMP in this DMP tool.
- Informed consent template: A informed consent form template which takes data publishing/sharing after research into account is available in BMS (Dutch version, English version)

3. PRIVACY REGULATIONS

Human privacy and personal data is protected by the European General Data Protection Regulation (GDPR). Due to the nature of research in BMS, many researchers process (e.g. collect, store, analyze, etc.) personal data in research. Personal data means any information that can be traced directly (clear link) or indirectly (when it is bit hidden) to a identifiable natural person ('data subject'), for example a name, (email) address, photograph, voice/video recording, etc. Note, that indirect identifiers (age, place of birth, occupation, family composition, salary) may not be traceable as separate variables, but linked to each other or with other information, can lead to a person's identification.

Personal data should be handled according to the GDPR. UT made a **flowchart for researchers** for more detailed guidance on the appropriate use of personal data in research. Some of the limitations and requirements for processing personal data are listed here:

- **There should be a legal basis for processing personal data.** Researchers must base on at least one of the 6 legal grounds to be able to process personal data in a lawful manner.
- **Data subjects need to be transparently informed** about the fact that their personal data is being processed, for what purposes and whether their personal data will be transferred to other parties. Please check the informed consent procedure at BMS in section 4 ETHICS in this document.
- **Registration:** For every research project, to comply with GDPR, reporting any new processing which uses personal data to the Data Protection Officers (DPO) team is required at UT. If you **collect personal data for your research**, you **must record this in the data processing register** via the GDPR registration tool, with which you can also write a DMP. The Privacy Contact Person (PCP) of BMS faculty is able to support you on this. Only anonymized data is exempt from reporting in that register since it is by law no longer personal data. If the **students** process personal data in their **bachelor/master-thesis**:
    - and the student participates in an **existing project**: the student does **not** need to take any further action, assuming that the **responsible employee has already registered the research project** in the GDPR registration tool.
    - If the **research leads to a new processing**, the **student must register the processing**, where the **supervisor is recorded as the contact person**.
- **Agreements:** If you bring in someone (e.g. external research partners, app-developer) who will be processing personal data for you, this person is not allowed to use this information for his or her own purposes. You need to formalize this in a **data processing agreement** (in Dutch: Verwerkersovereenkomst).
- **Proper safeguards:** During research, personal data must be anonymized/pseudonymized as quickly as possible. Here are some basic steps of pseudonymization which are summarized by the working group of LCRDM.
- **Storage limitation:** Personal data must be deleted or rendered anonymous as soon as identification of data subjects is no longer necessary, and personal data must be stored and managed in a secure ICT system (like the UT network storage: P-drive & BMS server). More information about storing and transferring personal data can be found in 5. RESEARCH DATA MANAGEMENT section in this document.
- **The tools used for processing personal data should be GDPR compliant**, e.g. an online conference platform, software for data analysis, a survey tool (eg. Qualtrics), etc.

4. ETHICS

Ethical review

To ensure an ethically responsible research practice, it is **mandatory for staff and students from the Faculty of BMS** to submit their research project for ethical assessment **before the start of the research,** regardless of where it is conducted. This is in principle for all intended research **involving human participants** in an **indirect (i.e. file or social media research)** or **direct manner (i.e. experiments, surveys, interviews)**, and/or **using potentially sensitive data** about and/or **from individuals, groups, or organizations**.

The UT offers ethical review by one of the **4 domain-specific committees** which are facilitated by faculties, BMS runs the domain Humanities and Social Sciences (HSS). A small part of the research by our students/staff may better fit the domain CIS (e.g. cyber-related research) or NES (technology-related health research).

Currently at BMS, the BMS Ethics Web App is used for ethical review. It is recommended to submit the ethical review request **6 weeks before the start of the data collection.** Please check the ethical committee domain HSS page for more information about the ethical review procedure.

Informed consent procedure

Informed participation is an **ethical and legal requirement** for research involving human participants. **Consent for research ethics** is composed of providing information beforehand regarding study, purpose, risks, benefits, voluntary participation, as **consent as a legal basis** is used for the processing of personal data under GDPR. A '**Informed consent procedure**' consists of **an information sheet** and **an informed consent form**. There are different types of informed consent, how to choose which type to use in your research? Who are capable of giving consent? What information should be included in the information sheet and the informed consent form? Please check this page for detailed guidance on the informed consent procedure, at the bottom of this page, you can also find **example templates of informed consent forms**.

5. RESEARCH DATA MANAGEMENT

Data storage

Data storage concern all storage during the research. After a research project the term archiving applies. During the research data should be stored in such a way to minimize risk of data loss and to maintain data integrity. Research groups take measures to avoid loss of research data during the course of a research project, due to e.g. theft of laptops, fire and water damage, or a sudden leave of a researcher without the group having access to the data.

All collected research data, including related materials (e.g. protocols, models or questionnaires), must be stored in an ISO 27001- and NEN 7510-certified directory such as the Project and Organization directory (P-drive) including backups hosted by or offered through LISA, unless exceptions apply.

Where to store?
Common recommendations are listed below, complete option lists can be found in this decision tool.

Storage options	Storage type	Suitable for personal/sensitive data *	Data sharing	Available to students
P-drive	UT Network storage* (ISO 27001- and NEN 7510-certified)	Yes, by restricting access to a specific folder via ICT servicedesk	Yes: via ICT servicedesk	Yes, with invitation from supervisor via ICT servicedesk
BMS-server	UT Network storage* (ISO 27001- and NEN 7510-certified)	Yes	Yes, via BMS lab	Yes, with pre-requisite to register your project
M-drive	UT Network storage* (ISO 27001- and NEN 7510-certified)	Yes	No	No
SurfDrive	Cloud storage	Yes	Yes	No
Google drive & One drive	Cloud storage	No	Yes	Yes

* UT network storage requires VPN connection, please check the VPN setup manuals if needed.
* Personal data is not allowed to store in private devices. In case of portable storage, storage must be encrypted (see manuals here).

Data transfer & sharing
To safely transfer research data, the SurfFileSender is recommended. In case of personal data, please tick 'Encryption' before sending the data in SurfFileSender.

Data documentation
Data documentation is 'information about the research data'. Sufficient documentation information about the data and analysis scripts are essential for the data to be understandable and therefore the research verifiable. Documentation can be in various formats. A more structured way of documenting data is to create metadata for your data. Metadata is 'Data about data'. Check these pages about metadata to see what metadata need to be captured and how to add metadata to a dataset in Excel. Documentation can also be in plain text, e.g. in a README form, please check Appendix 1 for two adjustable templates of a README file.

Data archiving

Data archiving concerns data storage after a research project ends. In the light of open science and scientific integrity, sustainably archiving of static data and providing access is crucial. Data archiving aims in the first place at preventing physical data loss or destruction and securing the authenticity of data. Besides, it contributes to the quality and impact of your scientific work by enabling verification and possible reuse. For instance by allowing further analysis or follow-up research, or as a contribution to a data resource for the scientific community.

In order for the data to survive for the long term, an active preservation regime has to be applied because data automatically gets lost over time due to e.g. digital sources degrade over time ('bit rot') or file formats and software become outdated. Therefore preparations are necessary before data gets archived. One of the preparations is to convert data file formats to non-proprietary (open) and persistent formats, so data files can always be accessed (opened). Please consult the preferred file formats before you preserve your data for the long term.

It is recommended that research data is archived together with other related materials (e.g. analysis scripts, documentation materials) at the UT data archive Areda for at least 10 years. Especially personal data should not be archived in third parties unless contracts/agreements allow it and secure archive is guaranteed. However, what files need to be archived depends on the purpose of archiving, e.g. reuse or verification/reproduction, and also on the type of research, e.g. qualitative or quantitative research. Please check Appendix 2 to decide what files should be archived and what should be better destroyed.

For more information about Areda and how to prepare data for archiving, please see the UT Areda page. However, if secondary data is used in the project, please be aware that contracts and/or other written agreements between involved parties in a project may contain information about rights, limitations and licences related to these data, which sometimes may prevent secondary data to be archived.

Data publishing

To make your data and research more visible to the scientific community, in addition to archiving your data in Areda, you can also use trusted repositories to publish your data. By publishing/depositing your data set to a trusted repository, your data set gets a **persistent digital identifier** (e.g. DOI) which allows your data to be widely findable, accessible, and easily cited by others. Similar to data, your analysis syntax (e.g. R scripts/python code) may be re-used by others as well. Therefore, in addition to depositing your data to repositories, we strongly advise that you also make your code available. For making code available, you can use GitHub or upload a copy of your code to a trusted repository as you do with your data. For UT researchers, we recommend using DANS Easy repository and 4TU.ResearchData repository to make your data or analysis syntax available to others. DANS Easy focuses on humanities, life and health sciences, social and behavioral sciences, oral history and spatial sciences. A UT researcher can upload up to 50 GB of data (including code) free of charge on the DANS Easy repository. The 4TU.ResearchData repository focuses on disciplines of natural sciences, engineering and design. A UT researcher can upload up to 1 TB of data (including code) per year on the 4TU.ResearchData repository. After you decide which repository you are going to deposit your data and code to, you should also consider what license you want to have to accompany your data and/or code. A license will define what others may do or may not do with your data and/or code. Check this license selector to find the most suitable license for your data and/or code.