

INSTRUCTIONS FOR THE DATA MANAGEMENT PLAN

The Rectors' Conference of Finnish
Universities of Applied Sciences Arene



ARENE

Contents

| | |
|---|----|
| Instructions for the data management plan | 3 |
| 1. General description of the data..... | 4 |
| 1.1. Description of the data: Data and materials to be collected or that already exist and their properties | 4 |
| 1.2. Ensuring the quality of the data | 4 |
| 2. Ethical principles, legislation and the processing of personal data | 5 |
| 2.1. Personal data and data protection considerations | 5 |
| 2.2 Main responsibility for the processing of personal data, i.e. controllership | 5 |
| 2.3 Privacy policies and statements | 6 |
| 2.4 Research designs in theses requiring ethical review..... | 6 |
| 2.5 How will you manage the rights to the data and materials you use, produce and share? | 7 |
| 3. Data documentation | 8 |
| 3.1. Data documentation | 8 |
| 3.2 Data order and integrity..... | 8 |
| 4. Recording and information security during the thesis process | 9 |
| 5. After the completion of your thesis: destroying, preserving, or finding further use for and opening data | 10 |
| 5.1 Destroying data after the completion of your thesis | 10 |
| 5.2 Preserving data | 10 |
| 5.3 Further use for and opening data | 11 |
| Where can the data be opened and shared? | 11 |
| 6. Duties and responsibilities | 12 |

These data management plan instructions serve as the guideline for preparing the data management plan for a bachelor's or master's thesis. Familiarise yourself with the instructions before you draw up the actual data management plan using the [following template](#). When in doubt, remember to refer to the different sections in these instructions while you are preparing your plan.

This work is licensed under a Creative Commons Attribution 4.0 International license. View the license at <https://creativecommons.org/licenses/by/4.0/>.

Instructions for the data management plan

These data management plan instructions serve as the guideline for preparing the data management plan for a bachelor's or master's thesis. Familiarise yourself with the instructions before you draw up the actual data management plan using the following template (link). When in doubt, remember to refer to the different sections in these instructions while you are preparing your plan.

The data management plan* is prepared as an appendix to your thesis plan. Both plans describe your data and materials from different perspectives:

- **The thesis plan or research plan** describes the scientific background and research methods that apply to your data and materials.
- **The data management plan** describes the processing and further use of the data and materials as well as the research ethics, data protection, information security, and rights that apply to them. It is a practical guideline for your thesis, as it describes how you will process your data during and after the thesis process.

The plan will help you:

- prepare instructions for yourself on how to process your data and materials during the thesis process
- identify risks related to data protection and information security as well as the loss or destruction of data
- anticipate and manage the details related to access rights, and consider the ways in which the data and materials could be utilised further.

Why do I need to manage my data and prepare a data management plan?

- Your data management plan describes your competence as a higher education student.
- The management of research data and materials and the preparation of a data management plan are part of good scientific practice and provide practical experience in the management of different data sets and documents, which you will need in working life.
- By preparing the plan in advance, you will be better able to comprehend the ways in which your data and materials could be further utilised in the future.

*Terms such as *thesis data* and *thesis materials* refer to the data used in your thesis and on which the thesis is based. The data and materials you collect can include e.g. surveys, interviews, measurement results, photos, videos, workshop materials, notes or field journals.

Your thesis may also utilise ready-made data and materials obtained from other sources, such as archive sources, sound recordings, YouTube videos, photos, literature as a research subject, films, websites, discussion forum posts, medical imaging, simulations, etc. The data and materials can also be research data/materials obtained from your instructor or a data archive.

1. General description of the data

1.1. Description of the data: Data and materials to be collected or that already exist and their properties

Make a list or table of the data used in your work and describe its properties. If you use pre-existing data, follow the terms of use of the data and specify its origin, authors and sources in accordance with good research practice and legislation. Pre-existing data can be e.g. samples, images, code, or survey or interview data collected/produced by someone else. You do not need to describe every source that you will refer to in your work here.

- List or tabulate the data or material(s) you have collected and, if applicable, any pre-existing data or material(s) you use in the plan. Describe their contents, types and possible file formats.
- Data and material types: e.g. text types, images, video or audio recordings, survey data, photos, measurements, statistics, physical samples, codes.
- Explain the basis on which you have the right to use the data and materials collected/produced by others (licence or access right, right granted by the client, etc.).
- If you estimate that you will need a large amount of data and materials, we recommend assessing the disk/physical space that they will require.

Tips:

- Possible file formats include .csv, .txt, .docx, .xlsx, and .tif.
- Remember to also include the online data and materials you use and the materials you collect from social media platforms.
- Remember to specify any special or unusual programs needed to view or use the data, especially if they are coded or produced during the thesis.
- The details and processes related to your data analysis and research methods are described in the thesis plan.

1.2. Ensuring the quality of the data

Briefly describe in your plan how you will ensure that the data and materials will not change inadvertently and that the original data content is preserved. Describe how the data will remain error-free throughout its life cycle, e.g. when converting or transferring data or files from systems or during analysis.

Tips:

- Save a copy of the original data before you begin analysing and editing it.
- The quality of data can be influenced by e.g. how the data is entered, how audio and video recordings are implemented technically, how the data is named and organised, and how the data's variables are named and transcribed. Consider how your data and materials will be organised during your thesis work by describing your naming practices and folder structure.
- Save a copy of the original survey data in a table format before processing it. Use checksums.

2. Ethical principles, legislation and the processing of personal data

2.1. Personal data and data protection considerations

If your data and materials contain personal data, you must comply with the provisions of the EU General Data Protection Regulation and the Finnish Data Protection Act.

Personal data means data that can be used to identify a person directly or indirectly, for example by combining different pieces of data. Personal data may be **direct** (e.g. a name, personal identity code, e-mail address, telephone number) or **indirect** (e.g. an image, sound, specific feature or even a particular hobby). For more information on personal data, see [the website of the Data Protection Ombudsman](#).

Personal data can also be **sensitive** in nature. Such data must be handled with particular care. As a rule, students should not create theses that include the collection of sensitive personal data.

Sensitive data includes different types of information that may indicate the following:

- racial or ethnic origin
- political opinion
- religious or philosophical belief
- trade union membership
- genetic or biometric data that can be used for the unambiguous identification of a person (a photo may constitute such personal data);
- health status
- sexual behaviour and orientation
- information on criminal convictions and offences.

Tips:

- Can an individual be identified on the basis of their age and gender?
- Please note that photos and audio are also considered personal data.
- Can an individual be identified on the basis of their place of residence and profession?

2.2. Main responsibility for the processing of personal data, i.e. controllership

If your data contains personal data, you must specify its controller, i.e. who has the main responsibility for the processing of personal data. The controller is responsible for ensuring that the data is processed in accordance with all laws and regulations.

Note the following principles:

- If you are completing your thesis alone, in pairs or as part of a group, the main responsibility for the processing of personal data lies with you or your group.

- If you are working on a thesis for a higher education institution as part of a project or while you are employed by said institution, the higher education institution is primarily responsible for the processing of personal data.
- If you are working on a thesis for an organisation (e.g. a company, the public sector, the third sector) while you are employed or commissioned by said organisation, the organisation is primarily responsible for the processing of personal data, unless otherwise agreed.

2.3. Privacy policies and statements

If your data or materials contain personal data and you are its controller, you must prepare a privacy policy/statement detailing what personal data you process and how, as well as the grounds for your processing activities. In almost every case, you will also need to prepare a research bulletin and consent form for your research subjects. The research bulletin informs potential research subjects of the progress and implementation of the study and the processing of personal data.

I have prepared a **privacy policy/statement**: yes/no/not applicable to my research

I have planned how I will obtain my research subjects'/participants' **consent for the processing of personal data**: yes/no/does not apply to my research

I have prepared a **research bulletin for my research subjects/participants**: yes/no/not applicable to my research

Please note that you will often be required to separately ask your participants for their **consent for participating in the research**.

2.4. Research designs in theses requiring ethical review

Determine whether you will be able to complete a thesis on the topic you have selected at your higher education institution, and whether your higher education institution has an ethics committee that can conduct an ethical review if necessary. Follow the instructions given by your institution.

An ethical review is needed in the following designs:

1. participation in the research deviates from the principle of informed consent,
2. the research interferes with the physical integrity of its subjects,
3. the research focuses on persons under the age of 15 without informing their guardian or seeking their separate consent, on the basis of which the guardian would be able to prohibit their child from participating in the research,
4. the research involves exposing subjects to exceptionally strong stimuli,
5. the research involves the risk of causing psychological harm that exceeds the boundaries of normal everyday life to the subjects or those close to them, or
6. conducting the research may pose a security threat to the subjects or researchers or those close to them. (TENK: 2019.)

You must describe the potential harms and risks of the thesis research to the ethics committee so that the committee can assess the ethical nature of the research and determine the possible risks to the research subjects in relation to the information sought by the research. The research must be carried out in a way that best minimises the potential harm and risks to the research subjects.

If your research design requires an ethical review but one is not carried out, you may be deemed to be in violation of good scientific practice, i.e. the Responsible Conduct of Research (RCR) / Research Integrity (RI). The same applies to possible situations where the researcher/thesis worker does not comply with the instructions issued to them.

Is an ethical review needed? (yes/no, justification and implementation).

2.5. How will you manage the rights to the data and materials you use, produce and share?

In this context, research data/material refers to the basic data collected by the student or the data refined from basic data on which the results of the thesis process are based. Intellectual property rights refer to all immaterial rights that can be used to protect intangible assets, such as copyrights and related rights, trademarks, design rights, rights to the mask work of an integrated chip, utility models and patents.

During the thesis process, the student may collect research data/materials that form the basis for the thesis or the results obtained in the thesis process. If a student intends to transfer any research data/materials or rights to research data/materials collected by them to a partner or a university of applied sciences, **this transfer be agreed upon separately.**

The student may grant a **parallel right of use** to the data/material, in which case the student will always have the right to use the result themselves regardless of what they decide to transfer. The transferred right of use includes, if expressly agreed, the following:

- **the right to modify the result**, which means that the recipient of the right of use may, for example, change the form in which the result is presented or otherwise modify it in terms of its content, for example by producing their own improved version and/or
- **the right to further transfer the right of use**, which means that the right of use granted by the student may be transferred freely by its recipient, for example to its own partners, while still retaining the right of use.

The student may also **transfer the ownership of the research data/material**, which in turn means that the partner owns the result after the transfer (e.g. an object resulting from 3D modelling) and the student does not have any rights to the result after this, excluding the right to be mentioned as the author of the result as required by good practice. If the result includes an invention, the rights to the result must be agreed upon separately.

3. Data documentation

You should document your data and materials throughout the thesis process so that both you and other people can understand the nature and application of the data and materials. Think about how you can help yourself recall what your data and materials consists of in the event that you need to return to your research several years after the fact.

3.1. Data documentation

Can you recall how you processed your data and materials during your research? Data documentation plays a key role in tracking the changes made to your data. Remember to maintain at least a personal record of the stages involved in your data processing practices. Include the terms, variable names, codes and abbreviations you used.

The data documentation describes how the thesis data has been produced, what it contains and how it has been processed. Without sufficiently detailed contextual information, research data can often become useless.

You can use the following tools to document your data:

- a readme file accompanying the data that describes its main elements, such as its variables, file formats, etc.
- a research journal
- a separate document in which you record the main details of your data and materials, such as the changes made,
- their analysis stages and, for example, the meaning of each variable.

3.2. Data order and integrity

It is important to look after your data and materials throughout the thesis process. How will you keep your data organised and intact and avoid any unintended changes?

For example, in a qualitative interview study, the term original data/material refers to the collection formed by all transcribed interviews. When you analyse data and materials, you produce different versions of them. For example, a thematic analysis results in a new version of the data, which must be saved as a separate entity.

Tips for preserving data integrity:

- Keep the original data and materials separate from those used during the thesis so that you can reference them whenever necessary.
- Version management: before you begin your thesis process, think about your naming scheme for the different versions of your data, and follow your version management policy systematically.
- Focus on the life cycle of your data and materials at the beginning of the thesis process and prepare for situations where your data may change without you noticing, such as during the recording, transcription, file format conversion or saving process.

4. Recording and information security during the thesis process

Make sure to store the data securely both during and after the thesis process and that you create the necessary backups. Please note that a protected network drive is more secure than a hard drive or USB stick. Do not store any sensitive personal data or confidential information on a public cloud service (such as Google Drive).

Storing and backing up data:

- Explain where the data is stored and how it will be backed up during the thesis process.

Tips:

- Your file storage location depends on various factors, such as the size and nature of the files and any related security issues.
- Ensuring the security of your data is particularly important when processing special categories of personal data or other confidential information, such as business secrets.
- We recommend using the storage services offered by your institution.

Access control:

- Describe who has access to your data, what the persons in question can do with the data, and how you will ensure that the data is transferred safely to any potential partners.
- Please note that, in the case of personal data, your privacy policy lists all of its processors, and viewing the data is also considered processing.

Tips:

- You must pay special attention to who has access to any sensitive data. If the collection of the data was based on consent, the use and transfer of the data must comply with the terms of the consent provided by the respondent to the survey or interview.
- Follow the information security and data protection guidelines of your higher education institution and use its recommended storage locations.
- When possible, use a higher education institution account instead of your private account.

5. After the completion of your thesis: destroying, preserving, or finding further use for and opening data

Once you have completed your thesis, the data and materials collected for it must be processed systematically. You can choose one of the following options:

1. Destroy the data
2. Preserve the data
3. Further use for and/or opening data.

Please note that each option is subject to specific conditions and responsibilities. Sometimes part of the data may be stored and/or used further, while other pieces must be destroyed. Follow the instructions given by your institution.

5.1. Destroying data after the completion of your thesis

You will need to destroy the data and materials used in your thesis in the following cases:

- You did not inform your research subjects about the further use of the data in advance.
- Data containing personal data will not be anonymised (see section 5.3 of these instructions).
- You have not agreed otherwise with your institution or commissioner.

Tips for destroying data and materials containing personal data:

- Paper material can be disposed of efficiently by, for example, shredding or burning it.
- Data on a USB stick can be destroyed by destroying the stick itself.
- Simply moving the data to the recycle bin on your computer or emptying the recycle bin will not result in the permanent destruction of the data.

Follow the instructions given by your institution.

5.2. Preserving data

Discuss the storage of the data and materials used in your thesis with your thesis supervisor and commissioner. Please note that if the data contains personal data that can be used to identify people (see section 2.1) or other content that could prevent its preservation (e.g. business secrets), you will not be able to preserve it in its current condition.

Follow the instructions given by your institution on the storage of thesis data and materials.

5.3. Further use for and opening data

You must inform your research subjects of the further use of the data before the data is collected. **You must inform your participants of any further use in an understandable manner.** We recommend requesting your participants' consent to opening the data and sharing it for further use when they participate in the research.

If you want to preserve your data for possible further use by yourself or your commissioner, you are responsible for ensuring that the data is stored securely and that all personal data is protected (see chapter 4).

Discuss with your thesis supervisor whether your data is worth opening and sharing in an online data archive for others to inspect and use. Is the data generally interesting, is it methodologically valid, and could it be utilised in a new study? You need to agree with your commissioner at the beginning of your work on whether the data will be opened and shared. Make an agreement with your commissioner on the ownership or parallel use of the data (see section 2.5).

Before opening and sharing the data, it must be anonymised i.e. edited in a way that it no longer includes any personal data (see section 2) and cannot be used to identify individual people through any data combinations.

Tips:

- If your commissioner wants to make use of your data, you must ensure that it no longer contains any personal data. The matter must be agreed upon separately with the commissioner and the participants in the research.
- If your commissioner wants to have access to your data, you must ensure that, at the data collection stage, your research subjects are aware of the parties that will receive the data.

Where can the data be opened and shared?

"Opening" refers to storing and sharing data in a public data archive. If you want to open and share your data, discuss with your thesis supervisor where the data can be opened and where its metadata can be published. If you want to open and share your data, contact the support services of your institution together with your thesis supervisor.

6. Duties and responsibilities

Who is responsible for data management and what kinds of resources does it require? Briefly explain here how the tasks and responsibilities described in the previous answers have been divided. Who is responsible for duties related to data management during the life cycle of the thesis? What kinds of resources (e.g. working hours, money) are needed to store and share the data?

Tips:

- Describe the duties and responsibilities that are related to, for example, data collection, data quality, data storage and backup, information security and data protection, and the possible opening and sharing of the data.
- How long will it take to save and process the data? How much disk space do you need?
- If you are creating the thesis for a commissioner, remember to also describe their possible role.

Instructions and template for a Thesis Data Management Plan - working group

The instructions and template for the thesis data management plan were drafted and implemented by a group convened by the Avoin TKIO network, whose members included:

Heini Hassel, Diakonia-ammattikorkeakoulu
Kirsi Hänninen, Kaakkois-Suomen ammattikorkeakoulu
Mira Juppi, Savonia-ammattikorkeakoulu
Elina Kirjalainen, Jyväskylän ammattikorkeakoulu
Anne Kärki, Satakunnan ammattikorkeakoulu
Joonas Koironen, Metropolia ammattikorkeakoulu
Tarja Kupiainen, Karelia-ammattikorkeakoulu
Katja Laine, Tritonia
Henni Laitala, Kaakkois-Suomen ammattikorkeakoulu
Minna Marjamaa, Laurea-ammattikorkeakoulu
Tiina Nokkala, Turun ammattikorkeakoulu
Elina Nurminen, Metropolia ammattikorkeakoulu
Johanna Ojala, Centria-ammattikorkeakoulu
Seliina Päällysaho, Seinäjoen ammattikorkeakoulu
Anne-Marie Tuikka, Turun ammattikorkeakoulu
Kaisa Varis, Karelia-ammattikorkeakoulu