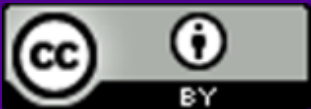


Data management for thesis supervisors

13.10.2021

Research Data Services, Tampere University Library

Saila Huuskonen & Kaisa Kylmälä



Today

- Personal information and data protection in research (Anna Rytivaara)
- Research data and data life cycle
- Data management planning
- Terms of use
- Data storage and security
- Documenting and describing research data
- Sharing research data



What is research data?

- All the materials, methods and results produced and used in the research
- Research or project data can be
 - Collected by you
 - Interviews, surveys, observation diaries, drawings, photos etc.
 - Reused data
 - Data collected by other researchers, for example surveys
 - Search datasets: [FSD](#), [Language Bank of Finland](#), [Etsin](#)
 - Data, which exist without research or a project
 - For example, [Tampereen kaupungin dataportaali](#), [European Data Portal](#)
 - New data which is created during the research
 - Data which is based on collected data and/or already existing data



Data life cycle

AFTER THE RESEARCH:

- Publishing and sharing data and metadata
 - Data archives and journals
 - Terms of reuse
- Data disposal



BEFORE RESEARCH:

- Research plan
- Data management plan
- Ethical review

DURING RESEARCH:

- Data collection
- Organizing and documenting
- Processing and analysing
- Storing and protecting
- Back-up

Data management is worthwhile

- Research data management is a part of good scientific practice
 - Secures responsible and ethical research
 - Helps to validate research results
 - Transparency, reliability, replicability and verifiability of research
- Eases the workflow and saves time
- Organising, preserving and finding data becomes easier
- Enables sharing of your data
 - Visibility and findability of the data and the researcher (data citation, merit)
 - Reduces overlapping work and makes reusing the data easier
- Tampere Higher Community and e.g. the Academy of Finland require data management plans and making your data openly available.

Start your data management process with a data management plan

1. General description of data
2. Ethical and legal compliance (e.g. informing research participants, informed consent, data protection, copyright)
3. Plan how you describe data collection and processing over the thesis project
4. Plan where you store and how you back-up your data
5. Plan how you preserve or dispose your data after the thesis is done.
6. Responsibilities and resources

- Update your plan when necessary.
- You can create your plan by [DMPTuuli-tool](#).



Terms of use

- When a student collects the data:
 - Independently → a student dictates how data is used.
 - As part of a research group → a student makes an agreement on data use with a research group
 - Data is collected for a company → a student makes an agreement on data use with a company
- When a student uses
 - supervisor's or other researcher's data → a student makes an agreement on data use with the supervisor
 - data obtained elsewhere, for example FSD, THL → complies with terms defined in a research permit and terms of use.
- Make use of [a mutual non-disclosure agreement](#), when you give confidential data to a student.
 - Ensure, that you have a right to give data to a student.

Responsible conduct of research

Ethical aspects in data management, for example:

- Informing research participants
 - Where do you inform? – face-to-face, in the beginning of a survey form, e-mail, zoom?
 - [Information sheet](#): research topic and purpose of research
 - [Consent form](#) (adults) and permission to share/archive data
 - [Privacy notice for thesis](#): how personal information is handled
- Requesting research permission from an organization where research is carried out.
- Handling sensitive information
- Protecting research participants' identity
 - Removing identifiers by anonymising or pseudonymising data

Data storage and security



Data storage at TUNI environment

OneDrive for Business

- for students and employees
- 1000 GB storage space for each user
- data can be shared also with users outside our university
- applicable for ordinary personal data

TUNI Groups

- The same as OneDrive for Business but tailored for group work
- access right management
- collaboration support tools

Personal home directory on a network drive

- All new new students and staff members get a personal home directory (P: drive).
 - The default quota for the home directory is 50 GB for staff and 10 GB for students.
- All data saved on the network drives is backed up daily.
- Read more:
<https://intra.tuni.fi/en/handbook/2677/2691/3256>

Tailored solutions

- Technical challenges
- very large file size
 - e.g. 1000 hours of video recordings
- Very sensitive data
 - e.g. patient records, personal data in GDPR category 1 R

Cloud services

- It is not essential if it is a cloud or not
- The key is the security of the selected solution
 - How it is built and maintained?
 - Probabilities for data losses and data leaks?
 - Terms of use?
- Tampere Universities have a contractual agreement with Microsoft that the O365 cloud services are e.g. GDPR compliant
 - N.B! Universities have no contract with Dropbox, Google, etc.



Online video interviews

Zoom and Teams

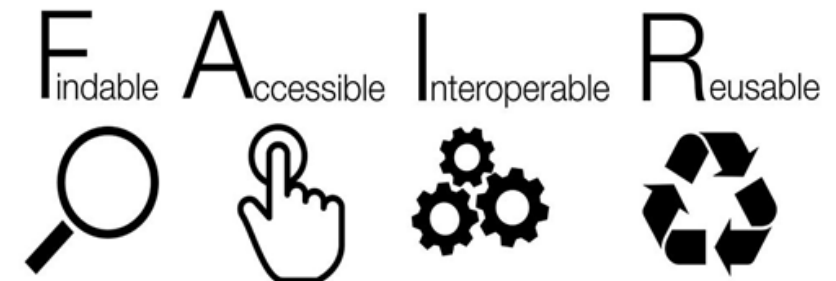
- GDPR compliant but only if **logged in with tuni account**
- Zoom: possible recording is stored at the inviewer's computer
- Teams: possible recording is stored at OneDrive

Read more:

- Security guidelines for setting up remote meetings and interviews using Teams or Zoom:
<https://intra.tuni.fi/en/handbook?page=20847>

Why should you document your data?

- Helps you and others understand your data
 - Easier to report your research findings.
- Helps other researchers understand how the data has been collected and managed
 - Makes your research more reliable
- Makes your research data easier to find and reuse
 - [FAIR-principles](#)
- Enables making your data openly available.

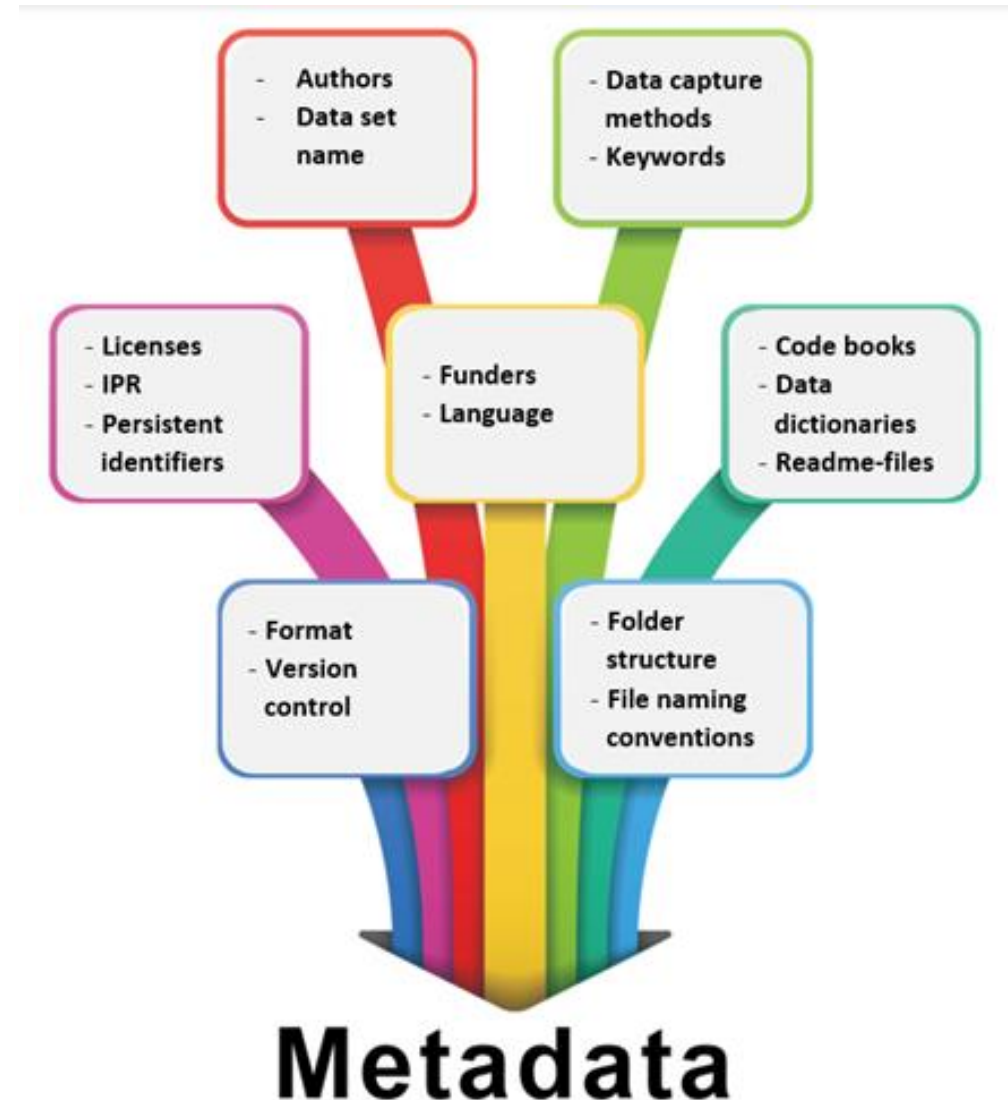


What to describe?

The project-level documentation explains:

- For what purposes was the data created?
 - Background information of your project
- What does the dataset contain?
 - Interviews, questionnaires, pictures...
- How was data collected?
- Who collected the data and when?
- How was the data processed?
- What possible manipulations were done to the data?
 - Is the data anonymised?
 - Are there new versions?
- What were the quality assurance procedures?
- How can the data be accessed?

Source: <https://www.essda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/Documentation-and-metadata>



Kuva 2. Tutkimusaineiston metadata

Public description - metadata

- Public description includes information
 - about the data collection
 - about the content of the data
 - on the authors of the research data
 - about the terms of reuse and licenses
 - permanent identifier ([DOI](#), URN)
- You can use Fairdata services metadata tool [Qvain](#) to describe the research data.
 - Metadata stored through Qvain is published in the research data search service [Etsin](#).
- If you store your data in an archive, the archive defines a metadata format. The archive also provides permanent identifiers.

Public metadata

Example of public metadata
in The Finnish Social Science
Data Archive (FSD) [Aila](#)

Display study description in English ▾ .

FSD3403 Ninth-Graders' Views on Happiness 2019

[Overview](#)

[Detailed description](#)

[Questions](#)

[Publications](#)

[Download data](#)

Authors

- Kulha, Saija (Tampere University. Faculty of Education and Culture)
- Huttunen, Sari (Tampere University. Faculty of Education and Culture)

Keywords

adolescents, family life, friends, happiness, hobbies, lower secondary education, natural environment, parents, satisfaction

Abstract

This dataset consists of 28 texts written by ninth-graders on what affects their happiness. The data were collected as part of research that charts ninth-graders' connection with nature through a questionnaire (the Nature Relatedness Scale) and writings on happiness. The data were originally collected as a paper questionnaire during the school day. The dataset consists of the writings as copied by the researchers and converted into digital format (text files).

The dataset is (B) available for **research, teaching and study**.

[Download the data](#)

Study description in other languages

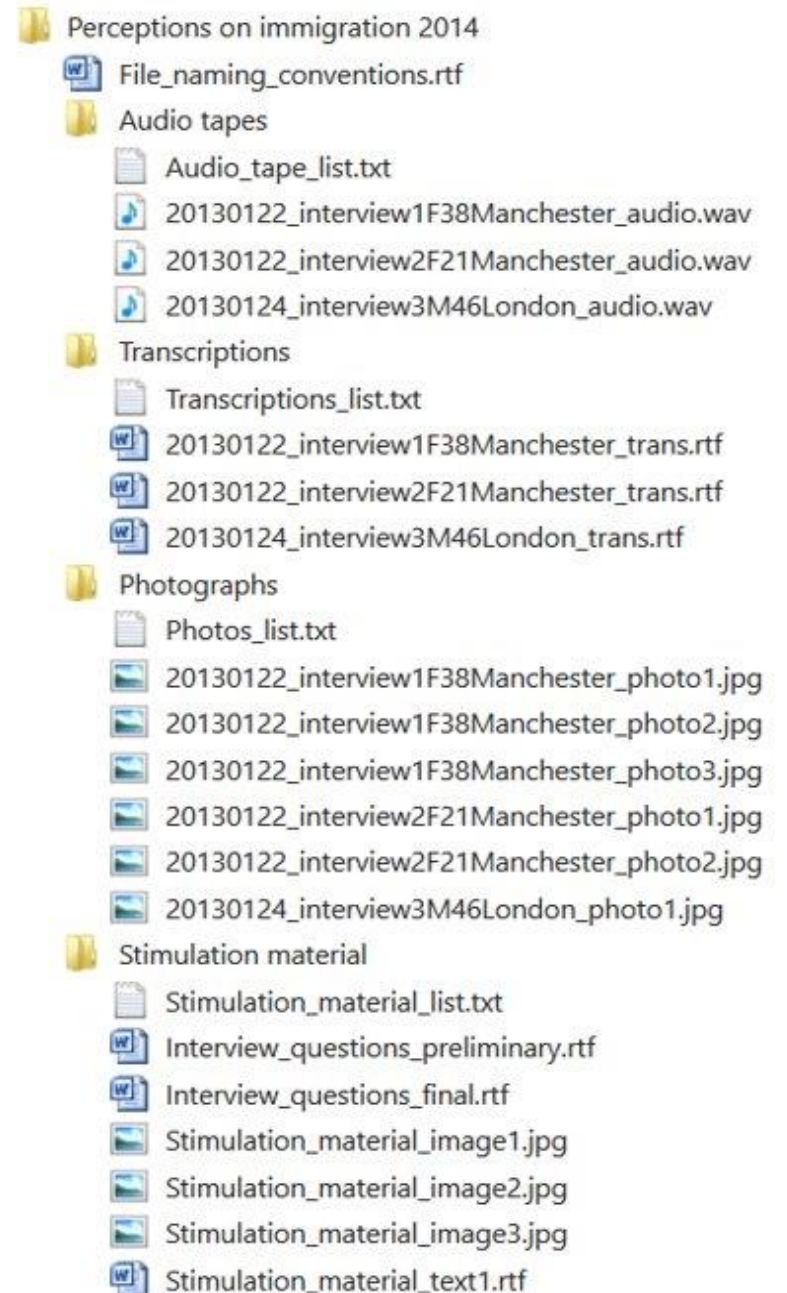
- [in Finnish](#)

Related files

- No other files available

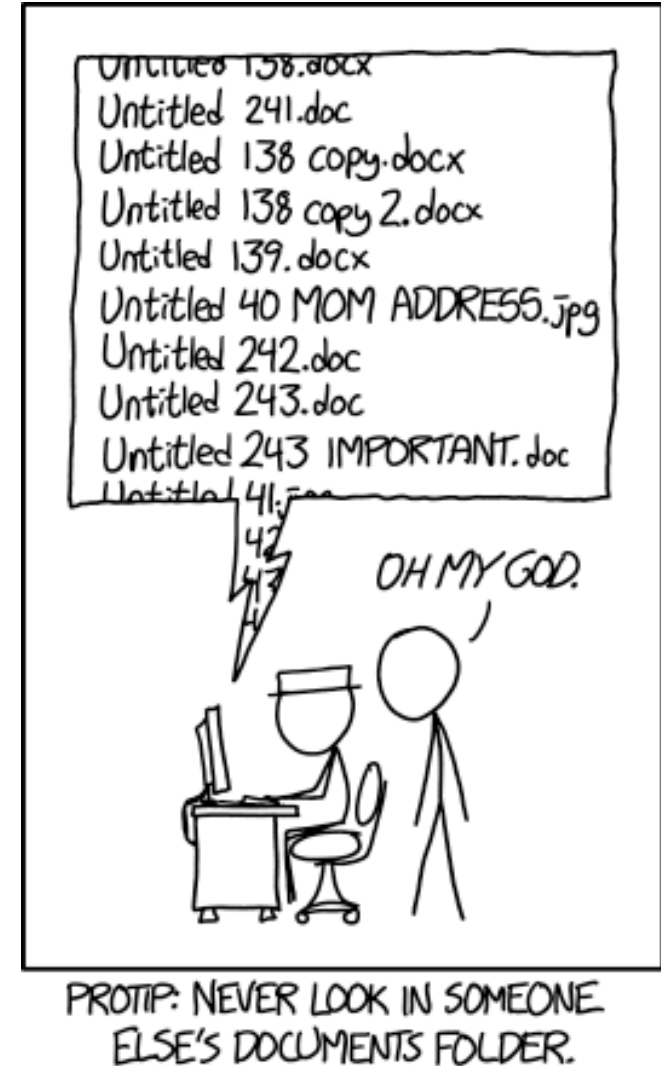
Keep your files organized

- Organized and consistent folder structures
 - Use subfolders
 - Organize by material type
 - The description can also be stored in connection with the research data or in a separate file (text files, [README files](#)).
- Also remember
 - information sheets, research permits, privacy notices, documentation of the processing of personal data, data management plan, funding application, confidentiality agreements and other agreements and documents
 - Notes



Name your files

- Use sensible file names
 - Short
 - Informative
 - Consistent
 - Caution! No personal information in file names
- Make versions visible
 - e.g. v01, v02, v03
- Use standards
 - e.g. Standard for date > 20211013_somedata.docx



<https://xkcd.com/1459/>

Sharing your data



Why to share my data?

- Sharing the data on principle: "*as open as possible, as closed and necessary*".
- The thesis author, and especially the doctoral researcher, also has the opportunity to share the data fully, partially or only metadata.
 - You can use [Qvain](#), a metadata tool for Fairdata services, to publicly describe the research data.
- Benefits of opening the data:
 - More visibility and references to your research.
 - Credit to the researcher. Also a merit for the student for working life.
 - Verifiability of research results and replicability of research.
 - Overlapping work is reduced, and it's also a matter of research participants' time. Should students reuse existing data?
 - You meet the requirements of our higher education community, several funders and scientific publishers.

Data sharing in practice

- Inform the participants of the research and request permission to open the data, i.e. an information sheet, a privacy notice and a consent to participate to study are required.
- Make sure there are no ethical or legal restrictions to data sharing.
- Describe your data adequately.
- Use file formats that can be opened without a special software.
- Anonymize your data. Check FSD's [guidance on anonymisation](#).
- Contact data archive, such as FSD, in advance.
 - FSD also takes data collected for the master's thesis.
 - Examples of data archives: [FSD](#), [Language Bank](#), [Zenodo](#)

Links and guides

- [Research Data Management -guide](#) (Tampere University Library)
 - [Data Management Guidelines](#) (Finnish Social Science Data Archive)

 - [Data protection path of research](#) (Tampere University)
 - [Data protection passport](#) (TUNI intra)

 - [Quick guide to information security](#) (TUNI intra)
 - [Responsible conduct of research](#) (Tampere University)

 - [Researcher's check list for publishing research data](#) (Responsible Research)
- Tampere Higher Education Community's policies
 - [Open Science and Research Policy](#)
 - [Data Protection Policy](#) (TAU intra)
 - [Information Security Policy](#)

Help available



Research Data Services comprehends:

- Library
- IT-services
- Research services
- Record management
- Legal services
- Data protection office
- Finnish Social Science Data Archive (FSD)
 - We organize research data management trainings.
 - We provide instructions and resources about data management.
 - We comment data management plans.

Contact us: researchdata@tuni.fi