HAMK HÄMEEN AMMATTIKORKEAKOULU
HÄME UNIVERSITY OF APPLIED SCIENCES

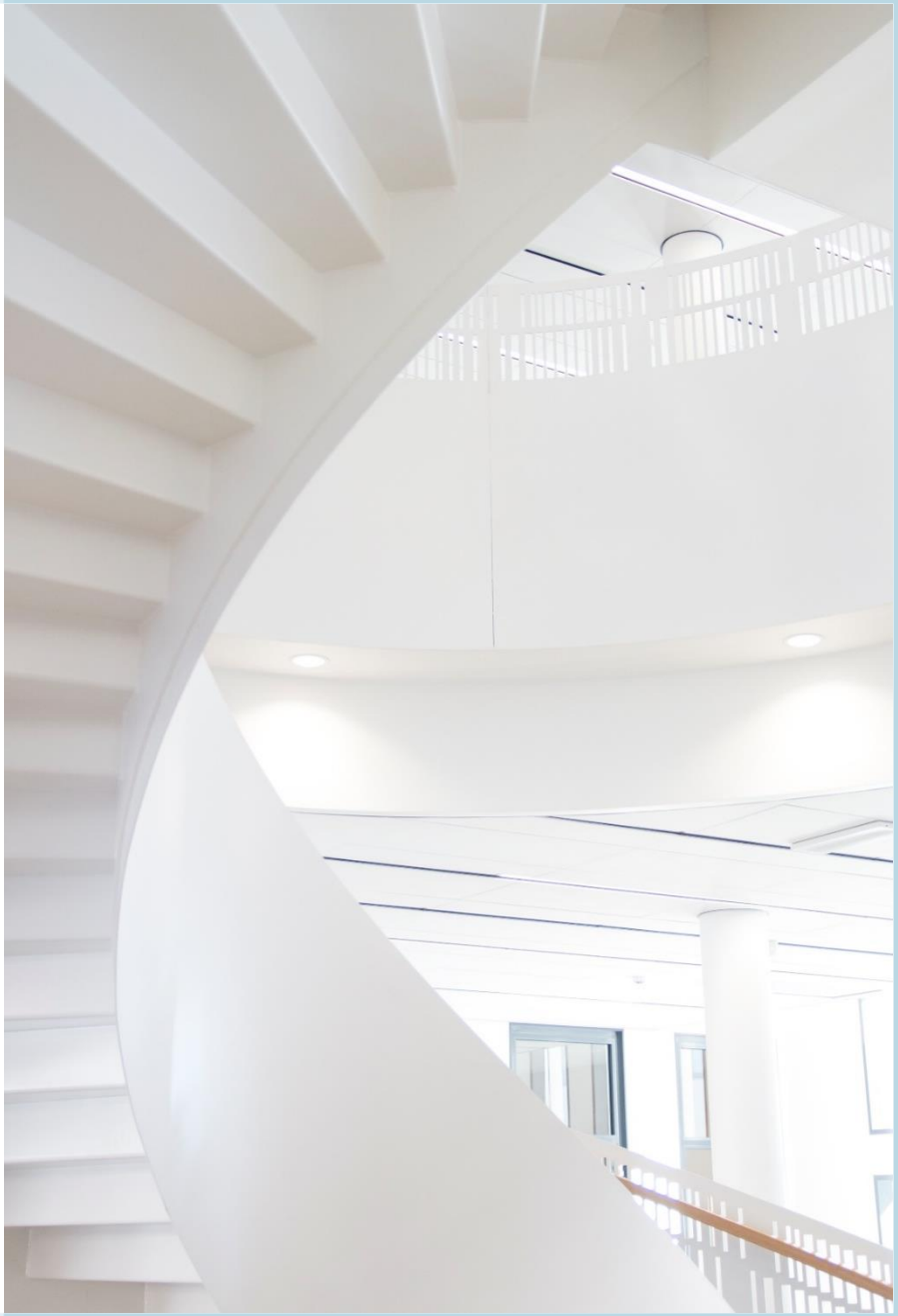
RUN-EU PLUS
Online workshop on FAIR data
2. June 2022

Data management planning: what, why, how?

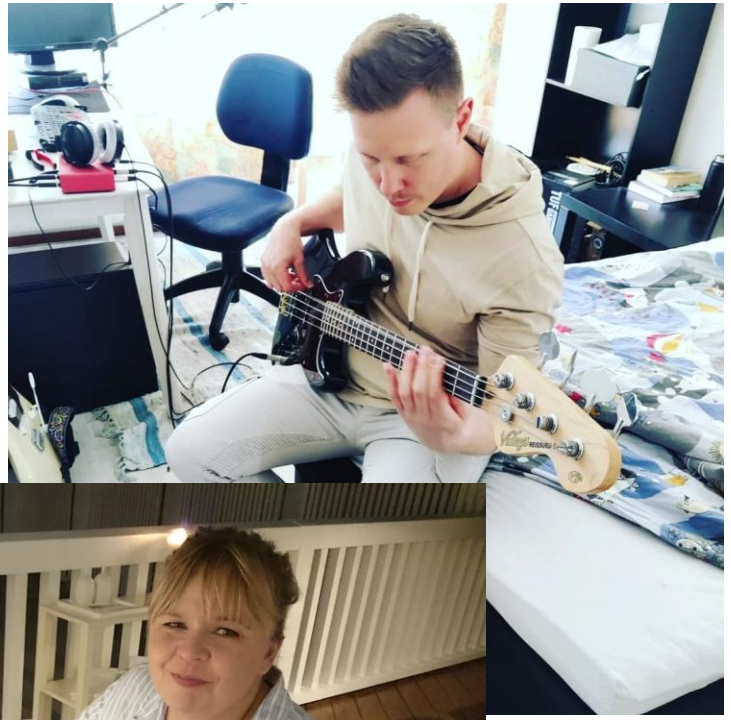
Anna Mikkonen, Nina Hynnä,

Toni Pulliainen

Häme University of Applied Sciences



Presentation of our "HAMK team"



What to expect from this session?

1. The very basics of the data management and data management planning. A brief introduction to the topic.

3. A demo of Finland's DMP tool DMPTuuli.

4. A short break (10 minutes)

5. Hands on! Let's practice together to write a data management plan.

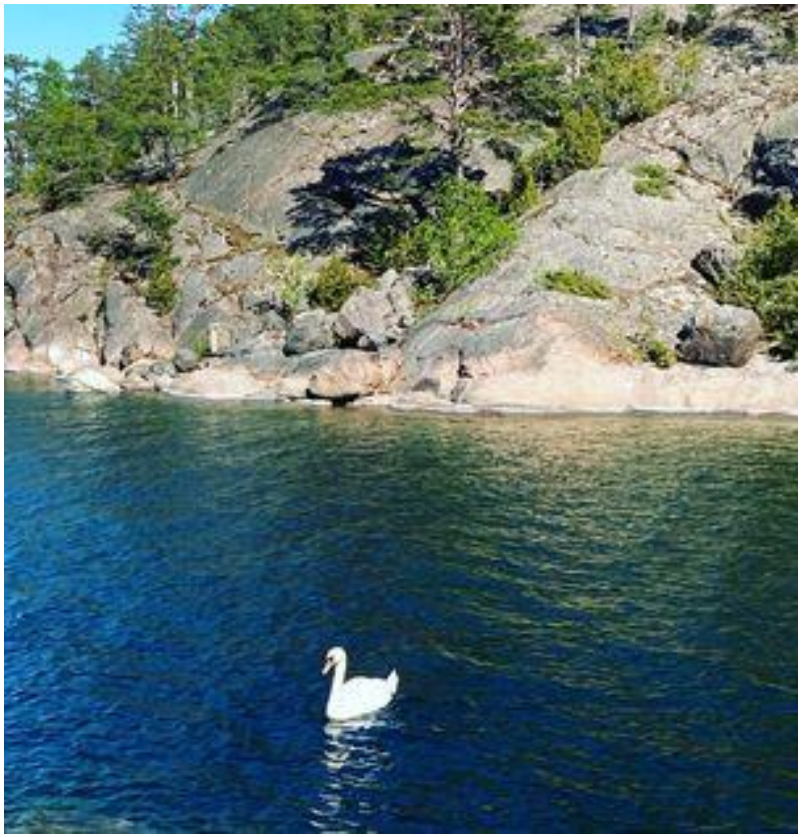
A group exercise in small groups.

Data management at HAMK

- HAMK has a history of actively promoting data management planning (and open science in general) since 2017.
- The level of organizations open science is evaluated in every two years by Finland's Ministry of Culture and Education.
  - Open publishing, open data, open research culture, open education.
- In Finland, open science is strongly nationally coordinated.
  - Declaration for open science and research.
  - Policies of open science and research.
  - Recommendations of open science and research in Finland.
  - https://avointiede.fi/en

Data management at HAMK

- Data management planning is required for each research or development project that has received a positive funding decision.
- In HAMK's theses, data management planning is required.
- As a rule, research data related to published research results is shared and published.



Why is research data managed?

- The processing and storage of data collected for a project is called **research data management.**
- Research data management entails that
  - data and related metadata are created, preserved and organised in a manner which ensures that data remain accessible and reliable, and
  - data protection and security are maintained over the whole data life cycle.
- Research funders (some) require data management.
- Some international science publishers have policies where access to data is a mandatory condition of publication.
- Benefits of reuse and open access.
  - Archiving research data constitutes academic credit for researchers.
  - Citations to archived data also constitute significant potential credit.

- Source: Data management guidelines. Finnish social science data archive. https://www.fsd.tuni.fi/en/services/data-management-guidelines/

Data management planning 1/2

- In the early stages of a project, *a data management plan* must be written.
  - A data management plan
    - describes how research data are collected or created, how data are used and stored during research and how made accessible for others after the research has been completed.
    - ensures that the collection, processing, storage and destruction of data has been planned well and implemented in a sensible manner.
  - If circumstances change, the plan needs to be updated.
- If a concise data management plan has been used for funding application, you will need to expand and specify the plan once the research has started.

Data management planning 2/2

- The data management plan should include the consideration and descriptions of the following:
- **The data**
- **The data storage**
- **Confidentiality and data security**
- **Rights**
- **File formats and programs**
- **Documentation on data processing and content**
- **Life cycle**

Rights and responsibilities

- When several researchers participate in a research project, the responsibilities and rights of the researchers should be agreed on.
- From an archiving point of view, when it comes to research material, it is extremely significant who has the right to decide on the handover of research data for reuse and to decide the terms which apply to such reuse.
- The following list can be used as a checklist.
    - What data will be handed over for reuse?
    - When can the data (or parts thereof) be handed over for reuse? (E.g. a pre-determined schedule, or upon publication)
    - For which purposes is the data going to be handed over (solely for research, or also for educational purposes or studying)?
    - Who has the right to make an archiving agreement concerning the research data?
    - Will any terms and conditions be set for data reuse?
    - If permission is required to use the data, who will decide on granting the permission?

- Source: Data management guidelines. Finnish social science data archive. https://www.fsd.tuni.fi/en/services/data-management-guidelines/copyright-and-agreements/

How to make RDI data accessible for others after the project?

- The data collected and/or created during the project:
  a) storaged and shared/published for reuse (open data)
  b) storaged and shared/published for reuse partly ("as open as possible, as closed as necessary")
  c) storaged and shared for reuse against permission
  d) storaged, not opened (long-term preservation)
  e) deleted.

  According to the data management plan.

Data storage

- Choosing the storage option for your data depends on the collected data, data protection and research integrity issues.
  - Sensitive data should not be kept in cloud services (etc. Teams, Google Drive, One Drive)
- Do you need to share your data with researchers from other institutions? Do you actively analyze the data together with researchers from other institutions?
  - For example, national data storage service for research data IDA in Finland.
- Where can you open your data?
  - National open access repositories
  - International open access repositories (e.g. Zenodo, EOSC)
  - Organization's own data archive
    - For example, HAMK's own repository HAMK Data.
    - •Local secured network drive for storing the data collected and/or created during the projects.

Data description and metadata

- Carefully describing and documenting the content, data collection procedures and variables of research data is essential to ensure the usability of data.
- Without this descriptive information, that is, metadata, research data are simply a meaningless collection of files, values and characters.
- When creating metadata, it is important to focus on describing the dataset itself instead of the results, conclusions and publications based on the data.

- Source: Data management guidelines. Finnish social science data archive. https://www.fsd.tuni.fi/en/services/data-management-guidelines/data-description-and-metadata/

Just to remind you..
Research integrity issues

- Research integrity issues are a vital part of data management planning.
  - If you want to open your data, you must inform your participants before you collect any data about the opening of the data.
  - Ethical review.
    - At the planning stage of the research, researchers should find out whether the research requires ethical review.
    - University websites often have instructions on the matter (you can search with phrases such as 'ethics review board', 'ethics review committee', 'human participants').
  - Informed consent.
- National ethical guidelines.
- International codes and guidelines for research ethics:
  - European Code of Conduct for Research Integrity. (All European Academies ALLEA 2017)
  - Recommendations for the Investigation of Research Misconduct (European Research Integrity Offices ENRIO 2019)

- Describe in the DMP how you handle the research integrity issues.

Just to remind you..
Data protection in research

What do you need to do **before** collecting personal data?

Purpose and minimization of the personal data

- Plan careful. For example: do you have plan to use / open the collected research data after this research?
- Collect only personal date which is relevant to the research questions.

Inform the participants about collecting personal data.

Describe in the DMP how you handle the data protection issues.

A demo of Finland's DMP tool DMPTuuli.

- Let’s have a look at Finland’s DMP tool.

DMP Tuuli

Group exercise 1/2

- The case: adults' experiences of coping with long term Covid symptoms.
- You are a researcher in a university A conducting research on adults' experiences of (mentally) coping with long term covid symptoms. This research project is conducted in a co-operation with your fellow researcher from a university B. You and your researcher colleague have decided to collect the data with two different data collection methods. For the first, you will design and conduct a survey exploring the experiences of coping with long term covid symptoms. In the end of the survey, the participants can leave their personal details if they want to attend to an interview, exploring the same theme more precisely. The interviews will be recorded and transcribed. During the research project, you, and your researcher colleague both analyze the data. Thus, both of you must have access to the collected data. At the end of the research project, the data will be published.

- Padlet https://padlet.com/annamikkonen/9faxnipta3f6c95z

Group exercise 2/2

Together we write a short data management plan for the project in small groups.

1. Introduce yourself to the other group members.
2. Pick a secretary for your group to write the answers to Padlet https://padlet.com/annamikkonen/fw9t8tyv1jpbjgu9
3. Answer the following questions.

- What kind of data will be collected?
- What kind of research integrity issues will be faced?
- Will personal information be collected?
- If yes, what do you have to consider before collecting the data?
- Briefly describe the data management practices for this research project.
  - Where will the data be stored during the project?
  - How will you share the data?
  - Who will have access to the data?
- What do you have to consider to be able to publish the data in the end of the project?

Thank you!

- Any questions, comments?