DMP WORKSHOP

Research Data Management Planning Workshops Data Support, Spring 2024

1

WORKSHOP PROGRAM

Intro
1. Data
3. Documentation and Metadata
6. Resources and Responsibilities
+ Writing and Discussion

2. Ethical and legal compliance
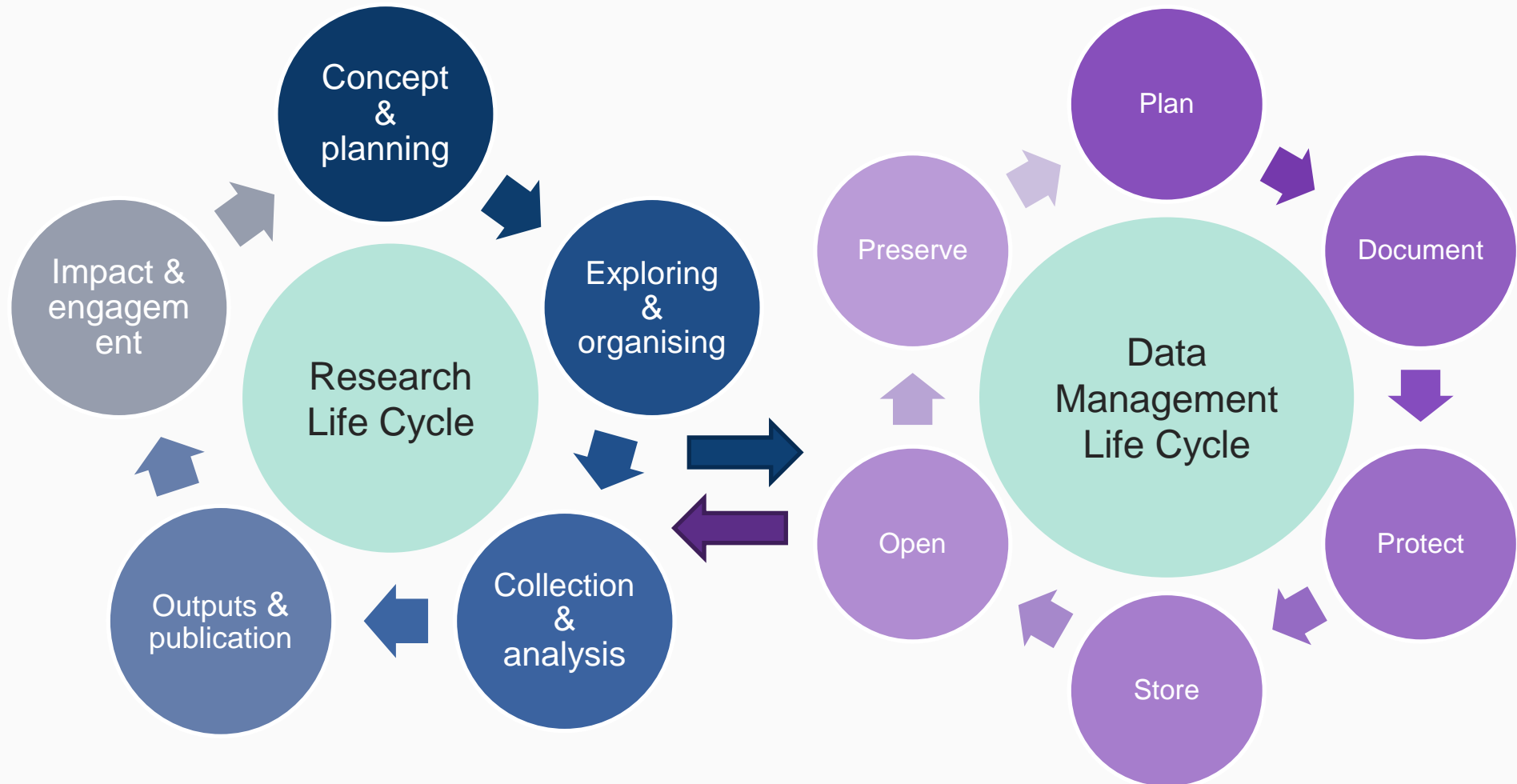+ Writing and Discussion

4. Storage during the research project
+ Writing and Discussion

5. Opening, publishing and archiving
+ Writing and Discussion

HELPFUL TOOLS

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY

RDM RESEARCH GUIDE

- Research Data Management Guide https://www.helsinki.fi/en/research/services-researchers/data-support-services/research-data-management

- Might be handy to keep the page open in another window while outlining a DMP with Tuuli

FINNISH SOCIAL SCIENCE DATA ARCHIVE

- FSD's Data Management Guidelines

https://www.fsd.tuni.fi/en/services/data-management-guidelines/

- Handy to keep the page open in another window while outlining a DMP with DMPTuuli

CONTENTS OF DATA MANAGEMENT GUIDELINES

Why Are Research Data Managed and Reused?
- Part of good scientific practice
- Research funders require data management
- Benefits of reuse and open access
- Research data life cycle

Data Management Planning
- Data management plan
- The data
- Rights
- Confidentiality and data security

Copyright and Agreements
- Copyright
- Agreements between researchers
- Agreement concerning the transfer of rights
- Authorship in scientific publications and research data

Informing about Personal Data Processing
- Clarity requirements and form of the information
- Timing for providing the information
- Content of the information
- Exceptions to the obligation to provide information

File Formats and Software
- Converting files
- Digitisation of research data
- Data matrices and statistical packages
- Textual data

Processing Quantitative Data Files
- Recording a data matrix
- Variable names and labels
- Values and value labels
- Missing data

Finnish DMP evaluation guidance:
10.5281/zenodo.4729831

Use www.dmptuuli.fi

DMP AS A DOCUMENT

KUVA: JOEL GRANDELL
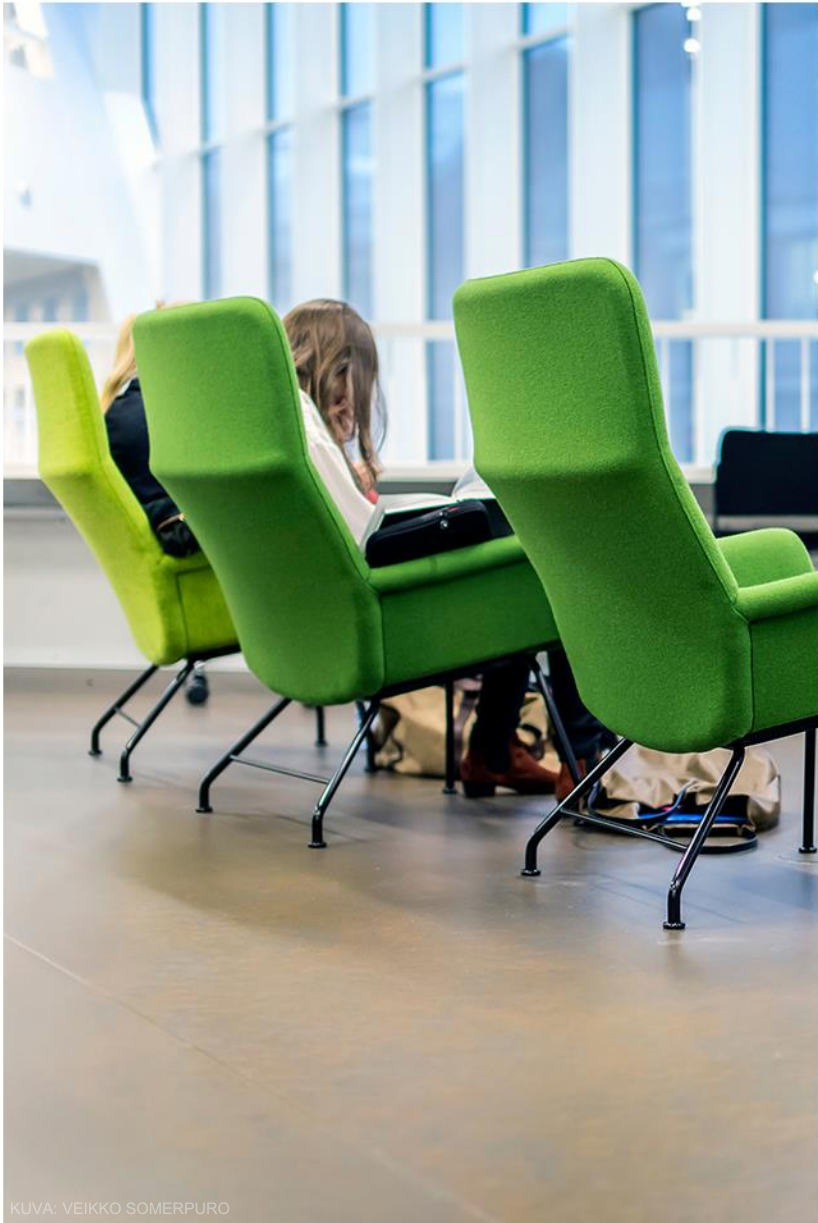
DATA MANAGEMENT PLAN (DMP)

- **Read all the questions first!**

- Length 1-3 pages (abstract and potential table excluded)

- Even if the research project does not produce any data, the DMP must be returned with a clear statement of the situation

- If you have had funding from Research Council of Finland, DMP has to follow their [guidelines](#); answer questions which are applicable

  - If some questions are not applicable, justify why!

KUVA: VEIKKO SOMERPURO

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY

THE DO'S AND DON'T'S OF A DMP

… be vague or general

… ramble

… re-write your research plan

… write a "good DMP" for the sake of bureaucracy

… write the DMP for Data Support or RCoF

… write one thing and do another

… be precise and unambiguous

… stay on point

… focus on the data and its lifecycle

… think through and write out your data management process

… write the DMP for the sake of your current and future self

… practice what you preach

DESCRIBE YOUR DATA

1. GENERAL DESCRIPTION OF THE DATA

1.1 What kinds of data is your research based on? What data will be collected, produced or reused? What file formats will the data be in? Additionally, give a rough estimate of the size of the data produced/collected.

Where to start? List your data types in a table or as bullet point list! Data sheet model link

Data type	Source	File format	Sensitivity (controller)	Size
Questionnaire	collected	.csv, .txt	Yes (HUS)	1 Gb
Analysis of the questionnaire	produced	.csv (.xlsx)	No	100 Mb
DNA samples	reused from Biobank	samples	Yes (Biobank)	N=1000

- Explain if any special or uncommon software are needed to view or use the data.
- Clearly describe how data volume or its accumulation has been calculated.

Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831

1. GENERAL DESCRIPTION OF THE DATA

1.2 How will the consistency and quality of data be controlled?

- **Clearly recognize possible error sources** during the data lifecycle, to ensure the quality of data.

- **Describe appropriate practices**
  - Data capture
  - Validation/monitoring
  - Versioning
  - Logs, etc

Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831
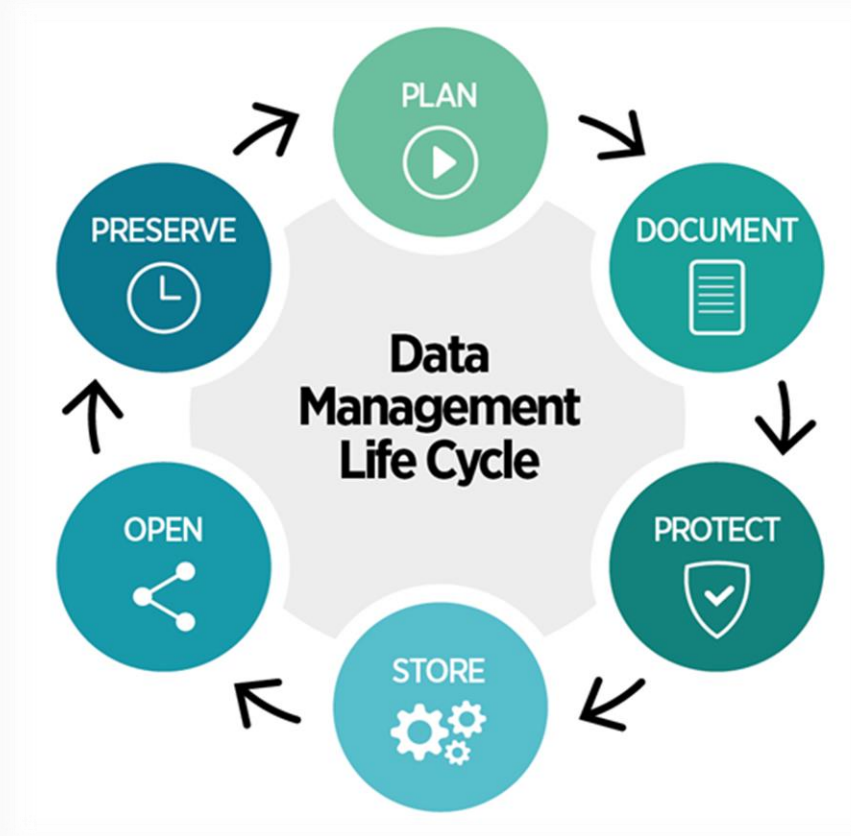
DOCUMENTATION AND METADATA

3. DOCUMENTATION AND METADATA

3.1 How will you document your data in order to make the data findable, accessible, interoperable and re-usable for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

- o Documentation enables verification and data re-use

- o List the **metadata standards** used for each data type. Refer to documentation requirements of data repositories/archives planned to use. Or search standards from https://www.dcc.ac.uk/guidance/standards/metadata

Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831

Research Data Management Planning Workshops Data Support, Spring 2024

3. DOCUMENTATION AND METADATA

3.1 How will you document your data in order to make the data findable, accessible, interoperable and re-usable for you and others? What kind of metadata standards, README files or other documentation will you use to help others to understand and use your data?

○ Describe how the documentation protocol is agreed, **if no standard is available for a data type**.

○ Outline **who is/are responsible for the documentation** during the data lifecycle (collection, analysis, storing, publishing, etc.)

Teksti

RESOURCES AND RESPONSIBILITIES
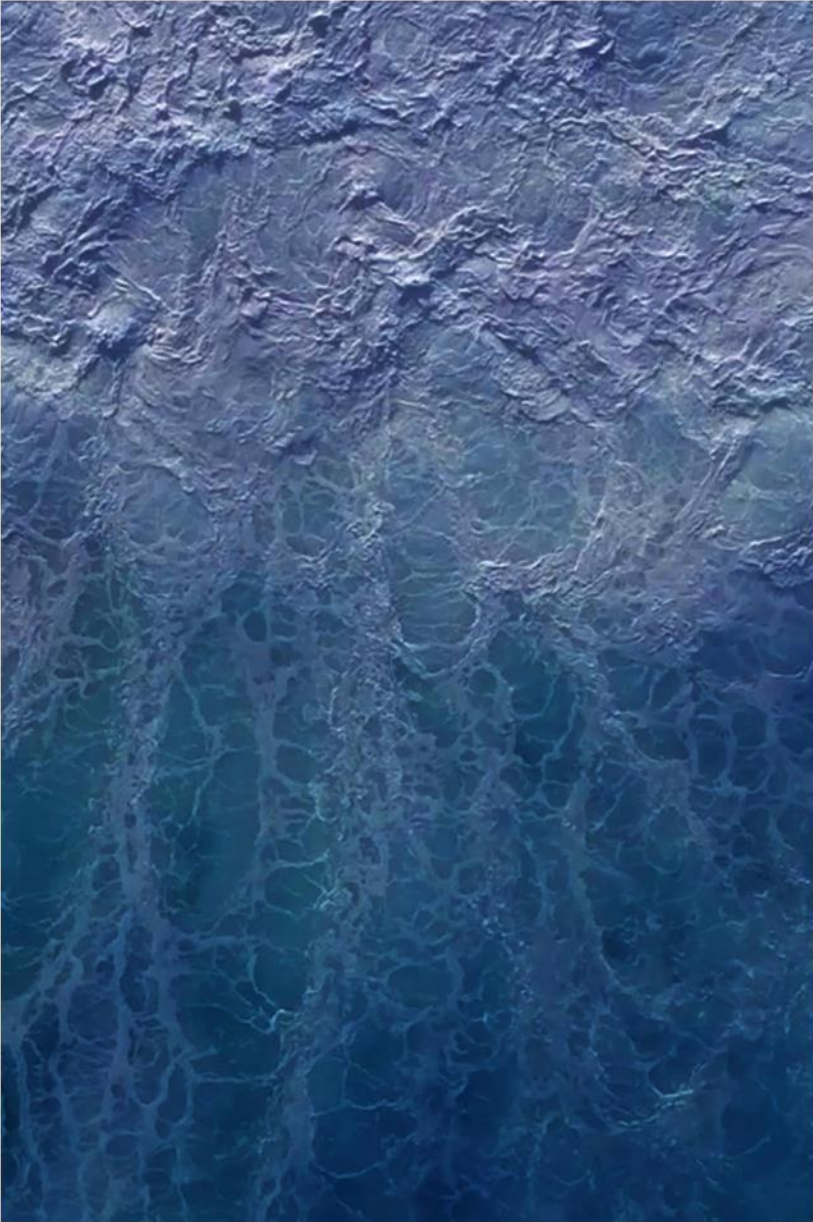
KUVA: JOEL GRANDELL

6. RESOURCES AND RESPONSIBILITIES

6.1 Who (for example role and institution) will be responsible for data management?

- **Outline roles and responsibilities** described in the DMP and name the individuals where possible: e.g. data management / stewardship, data capture, metadata production, data quality, storage and backup, data archiving, and data sharing
- **State who is responsible** for the data resulting from the project after the project has ended.
- **State the procedure for transferring** these responsibilities (in case the person is expected to leave the project).
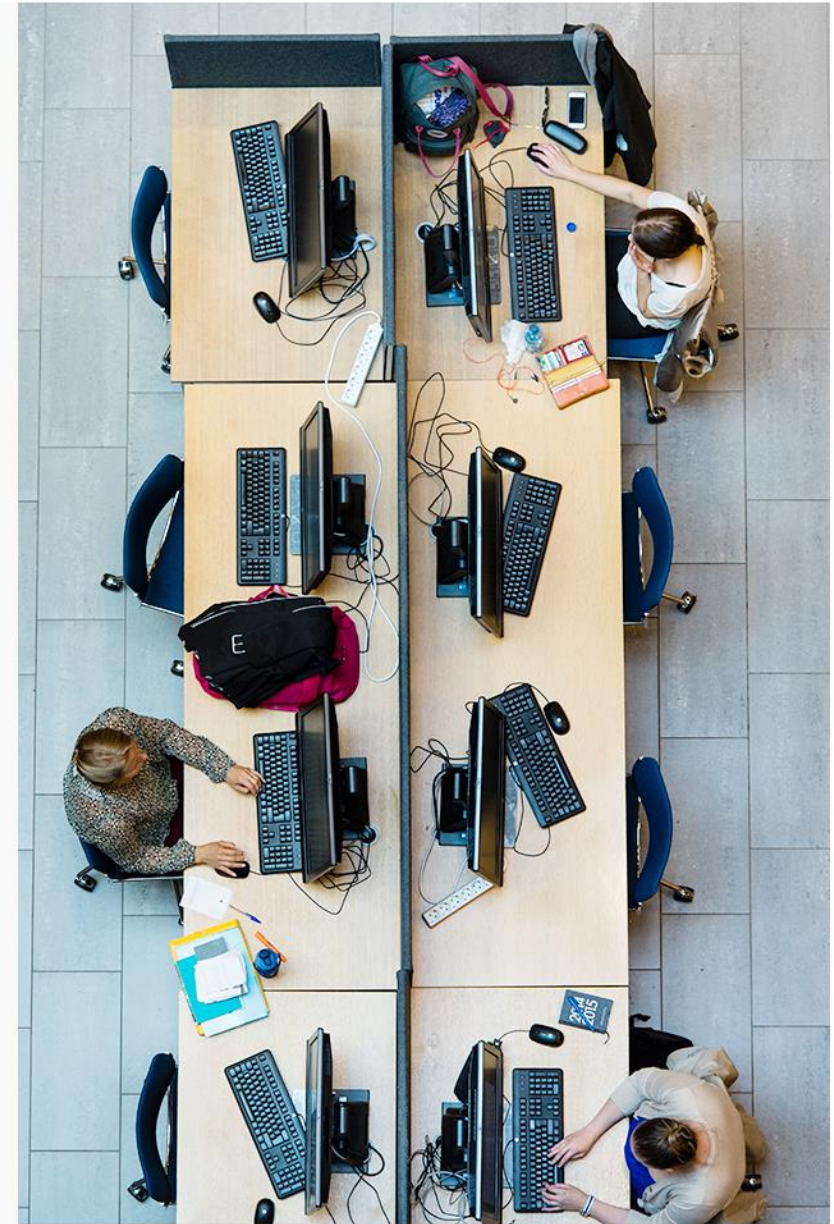
6. RESOURCES AND RESPONSIBILITIES

6.1 Who (for example role and institution) will be responsible for data management

- **Explain how data management responsibilities are co-ordinated** in collaborative projects.
- **Indicate who is responsible for implementing the DMP** and updating it during the project

6. RESOURCES AND RESPONSIBILITIES

6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?

- **List the required resources and facilities** for data management

  - (e.g. storing environment, computational facilities, hardware, staff time for preparing data for sharing, deposit, and repository charges) and refers to the specified financial costs in the budget, according to funder requirements.
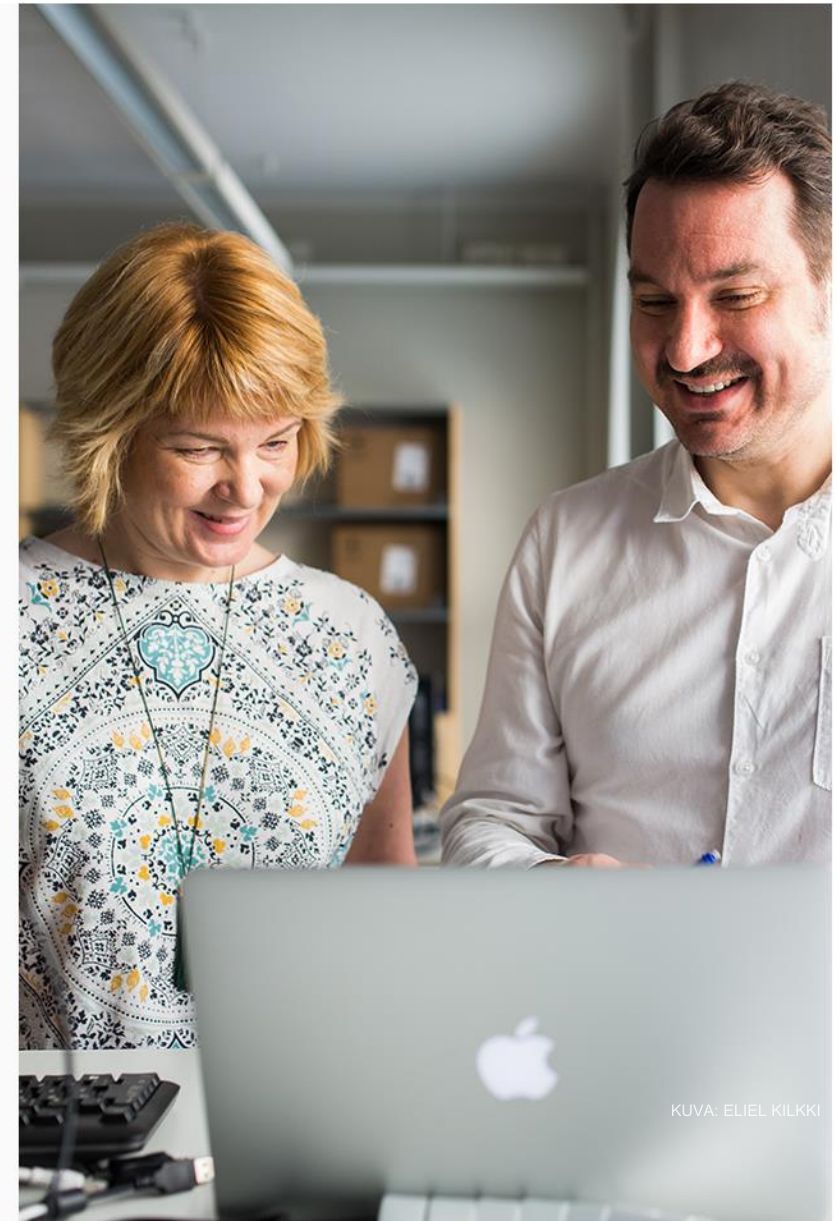
KUVA: ELIEL KILKKI

6. RESOURCES AND RESPONSIBILITIES

- *6.2 What resources will be required for your data management procedures to ensure that the data can be opened and preserved according to FAIR principles (Findable, Accessible, Interoperable, Re-usable)?*

- Provide **estimates of time and money needed to prepare the data** for sharing, publishing, preservation (data curation).
  - Open Aire [RDM costing tool](#)
- Describe **investments to expertise**, like how lawyer, data steward, transcription service, IT expert's consultancy is purchased, or are these experts hired to the project.
- Outline **what kind of resources is needed on training** data management skills.

Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831

ETHICAL AND LEGAL COMPLIANCE

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

2.2 How will you manage the rights of the data you use, produce and share?

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY

2. ETHICAL AND LEGAL COMPLIANCE

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

You understand **what is personal data** and **recognize when you are processing personal data**

Roles
- Who is/are the **data controller(s)**, or is there a joint controllership?
  - UH is usually the data controller when you have employment relationship with UH
- Who is/are the **processors**? Or is the processing of personal data outsourced to third parties?

Identify
- the legal basis (usually "research carried out in public interests")
- rights to process the data
  - secondary use of the data?
  - need for the Data Protection Impact Assessment (DPIA)? Ethical review?
- secure storing place (related to the point 4 Storage and backup)

2. ETHICAL AND LEGAL COMPLIANCE

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing.)

- **Research Ethics Committee in the Humanities and Social and Behavioural Sciences and the Research Ethics Committee of the Faculty of Medicine**

  - In medical research: a statutory statement from the ethics committee of the hospital district in which the person responsible for the research is working or in whose area the research will primarily be conducted, or from the National Committee on Medical Research Ethics.

- **Research Ethics Committee on Animal Research**

- **Research Ethics Committee in the Natural, Biological and Environmental Sciences and Engineering**

- The ethical review process must be completed before the collection of research data can begin.

- Although scientific publishers often require the ethical review, it is not always necessary → A description of the Finnish research system is sufficient, it can be obtained from the Secretary of the Ethics Committee.

2. ETHICAL AND LEGAL COMPLIANCE

2.1 What legal issues are related to your data management? (For example, GDPR and other legislation affecting data processing)

- Describe
  - the ways to protect the data and planned safeguards (pseudonymization, anonymization etc.)
  - how to take into account the data subject rights (especially informing)
- Mention how data protection will be taken into account when opening and sharing of the data
  - Will data be archived for future use? Any ideas for a location or service?

2. ETHICAL AND LEGAL COMPLIANCE

2.2 How will you manage the rights of the data you use, produce and share?

Data rights and agreements

Explain

- Who is/are the rights holder(s) of the data?

- The licenses (CC BY, MIT for software etc.) or other terms and conditions, e.g., for re-use of data

- If applicable, explain how intellectual property rights (IPR = copyright, patents, etc.) will be managed and what agreements are required, e.g., for transfer of rights.

2. ETHICAL AND LEGAL COMPLIANCE

2.2 How will you manage the rights of the data you use, produce and share?

For projects with multiple partners and data owners, explain their roles and responsibilities regarding data.

- with non-UH partners you need agreements at least about research collaboration;

- If handling personal data happens outside UH, then data processing agreements (DPA),

- If handling data or material happens outside UH, agreement about data/material transfers (DTA, MTA)

If personal data is transferred or disclosed outside EU/EEC area check

- is there adequacy decision with the country

- if not, then use standard clauses to agreements

2. ETHICAL AND LEGAL COMPLIANCE

- If your project doesn't comply with the funder's data sharing/opening policy, or
- If questions 2.1 and 2.2 do not apply to your project, **please explain why.**

- Sometimes it is not possible to share and open data for legal or ethical reasons. You could still open your metadata ☺

- Data providers might ask Privacy Notice though project / data doesn't include personal data
  - they want to be aware how data is handled, processed or archived (possible sharing and re-using data)
  - → DMP could be better document to show this

TRANSFER OF RIGHTS

As part of the agreement with the Research Council of Finland (formerly the Academy of Finland) you will confirm that the transfer of rights to the university has been made.

- Since 1.1.2022 new employees have signed this as part of their employment contract

- If you want to sign the transfer of rights only for this project, the template is available on Flamma

- If you want to sign the transfer of rights for all your UH projects, contact the HR of your faculty

More information about the process on Flamma:

- **Research agreement and legal matters:** https://flamma.helsinki.fi/en/group/tutkimuksen-tuki/tutkimussopimus-ja-lakiasiat

- **Instructions on concluding an agreement:** https://flamma.helsinki.fi/en/group/tutkimuksen-tuki/sopimukset-sopimusohje-ja-neuvottelut

KUVA: MIKA HUISMAN

HELSINGIN YLIOPISTON KIRJASTO
HELSINGFORS UNIVERSITETS BIBLIOTEK
HELSINKI UNIVERSITY LIBRARY

DATA PROTECTION GUIDE FOR RESEARCHERS ON FLAMMA

⟨ SUPPORT FOR RESEARCH AND INNOVATIONS

RESEARCH AGREEMENT AND LEGAL MATTERS

Instructions on concluding an agreement

Data protection guide for researchers

Material and data specific guidance

% Share   ☆ Add to favourites

Data protection guide for researchers

Legal counsels at Research Services assist UH employees in data protection issues related to research. Please explore the material and links below for answers to general questions about data protection before contacting the legal counsels at researchlawyers@helsinki.fi.

On issues or questions about the management of your research data, you can get assistance from:

DATA SUPPORT or **HELPDESK**

Data protection guidance for researchers

The purpose of this site is to provide the University's researchers with information on how the requirements of the EU's General Data Protection Regulation and other general data protection legislation must be taken into account when processing personal data in research. Researchers must also take into account any field-specific requirements (e.g., legislation on medical research) as well as principles and guidelines concerning research ethics.

If you process personal data in research, you must do the following things before you begin your research work:

- A data management plan (in which you reflect on the implementation of general principles and on potential risks to research subjects)
- If necessary, an impact assessment concerning data protection
- A data protection statement (submit it to tietosuoja@helsinki.fi), an information sheet for research subjects and, if applicable, consent forms
- Contracts with possible collaboration partners
- Acquire possible researchpermit
- Please also consider whether an ethical review is required. An ethical review of research

Data Protection Principles

Data Protection Principles of the Univeristy of Helsinki

Forms

Data Protection Impact Assessment

Does my research need a data protection impact assessment?

Data Protection Impact Assessment

Privacy Notices (i.e. Data Protection Statement)

Privacy Notice for Research (fin)

Privacy Notice for Research (eng)

Privacy Notice for Research (sve)

Disclosing data outside the EU/EEA?

Transfer assessment and planning safeguards

More Information

Office of the Data Protection Ombudsman

STORAGE OPTIONS

CHECKLIST FOR YOUR DATA

- How much storage space is needed (is the "scale" GB, TB or PB)?

- What type of data is stored and is it sensitive?

- How long the data needs to be stored?

- What software is needed to process the data?

- Who needs access to the data and is sharing it outside UH necessary?

- What is the available budget?

- Consider the whole lifecycle of the data, including long term plans

STANDARD STORAGE OPTIONS

Home and group folders

- For low and medium risk data

- Backup and version control

- Access control

- In UH premises

Umpio

- For high risk data

Office 365 (Onedrive, Teams)

- For collaboration

Redcap

- Survey tool (also for sensitive data)

Data storage and sharing table [link](#)

UH Servers & Computing services
- Virtual and physical servers
- UH high performance computing

CSC services
- IDA
- Servers & HPC
- SD Connect & Desktop (sensitive data)

OPENING, PUBLISHING, ARCHIVING

5. OPENING, PUBLISHING AND ARCHIVING THE DATA AFTER THE RESEARCH PROJECT

*5.1 **What** part of the data can be made openly available or published? **Where** and **when** will the data, or its metadata, be made available? (1/2)*

o Describe **how the data, metadata, or software will be opened after the project**

o Define the **name of the platform (**repository, data catalogue, data journal).

o Explain **why a chosen solution** is ideal for you

o Collection of repositories: https://fairsharing.org/databases/ and https://www.re3data.org/

5. OPENING, PUBLISHING AND ARCHIVING THE DATA AFTER THE RESEARCH PROJECT

*5.1 **What** part of the data can be made openly available or published? **Where** and **when** will the data, or its metadata, be made available? (1/2)*

o Ensure that data will get **a persistent identifier** (DOI, URN, Handle...)

o Explain, if applicable, why data opening is not possible or it is limited and who can access the data under which conditions (for example, only members of certain communities or via a sharing agreement).

o Note that any sensitive data cannot be opened!

o It is recommended to open the findability metadata if the data itself cannot be opened

5. OPENING, PUBLISHING AND ARCHIVING THE DATA AFTER THE RESEARCH PROJECT

*5.1 **What** part of the data can be made openly available or published? **Where** and **when** will the data, or its metadata, be made available? (2/2)*

o Explain, where possible, **what actions will be taken to overcome or to minimise data sharing restrictions.**

o Indicate **which specific tools or software potential users may need** to access, interpret, and re-use the data (for example specific scripts, codes, or algorithms developed during the project, version of the software).

Research Data Management Planning Workshops Data Support, Spring 2024

5. OPENING, PUBLISHING AND ARCHIVING THE DATA AFTER THE RESEARCH PROJECT

*5.2 **Where** will data with long-term value be archived, and for **how long**?*

o Categorize datasets that need different **length of preservation**:

A) data to be destroyed after the project. Describes how the data will be disposed after preservation period.

B) data to be archived for a verification period, e.g. 5-15 years.

C) data to be archived for potential re-use, e.g. for 25 years;

D) data to be preserved and curated for tens or hundreds of years
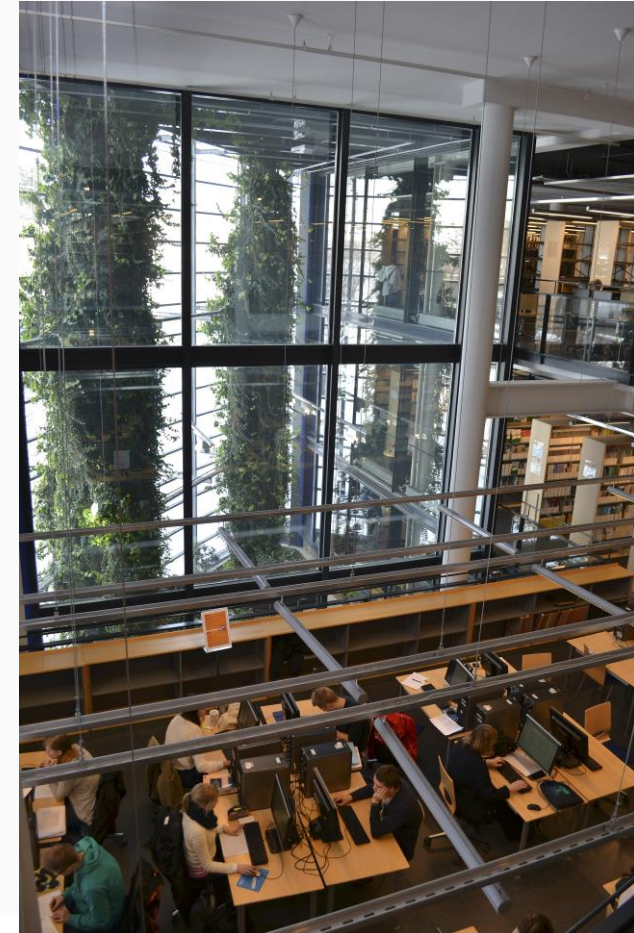
Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831

5. OPENING, PUBLISHING AND ARCHIVING THE DATA AFTER THE RESEARCH PROJECT

*5.2 **Where** will data with long-term value be archived, and for **how long**?*

o Describe how **management, preservation and admission to the datasets will be secured** when needed because of verification, agreements or other reasons.

o Provide the **name of the archive or trustworthy repository** – or the way to curate and preserve data – that will be used to make data available for re-use.

o Acknowledge the impact of legal, ownership, agreements, funders', institutions', and publishers' demands on data preservation.

Finnish DMP evaluation guidance:
https://doi.org/10.5281/zenodo.4729831

Zoom meetings available after the training

o Discuss about your data management plan with a data management expert

o 30-60 min time slots

o Pick the time for the meeting from the table:
https://bit.ly/DMP_meetings_2024
Password: 4202_dmp

o Send the draft of your DMP and project abstract to datasupport@helsinki.fi one office day before (tell in the message you have booked a meeting)

TURN TO UH DATA SUPPORT!

In this research project I will collect data from...

IMPROVE YOUR

DMP

BY SENDING IT TO DATASUPPORT@HELSINKI.FI

Tuuli