SCIENCE AND EDUCATION FOR SUSTAINABLE LIFE
Data management as part of open science, FAIR data and good research practice – an introduction

We will use Mentimeter.

- Answering the Menti-questions is voluntary.
- We will only show and save the aggregated results, disclosing no names.
- Mentimeter AB will store your network address for two weeks. As a visitor, you never have to give your e-mail address to Mentimeter.

Webinar house rules

- Write all questions in the chat and we will get to them!
- This webinar is recorded and the presentation will be available from the DMS web page. The Q&A-section will not be included in the recording.
- You may turn your camera on if you wish, but please keep your microphone turned off.

Welcome to DMS webinars (autumn 2022)

- **28th of September (Wednesday 10-11.30)**
  Data management as part of open science, FAIR data and good research practice – an introduction

- **26th of October (Wednesday 10-11.30)**
  The principle of public access to official documents – can anyone ask for my research material?

- **16th of November (Wednesday 10-11.30)**
  Research data and Swedish law, GDPR and immaterial rights

- **6th of December (Tuesday 10-11.30)**
  Kom igång med din datahanteringsplan

- **7th of December (Wednesday 10-11.30)**
  Get started with your data management plan
Data management as part of open science, FAIR data and good research practice – an introduction

Hanna Östholm, Kerstin Belin and Simon Hallstan (DMS)
28th of September 2022
Contents
1. Presentation
2. Defining research data and data management
3. Why should I improve my data management?
4. Data management as part of good research practice
5. Will improved data management benefit my research?
   Break
6. Open science and research data management
7. What is FAIR data?
8. Data publishing
9. SLU Data Management Support
10. Questions
Defining research data and data management
What is research data?

Research data refers to information collected to be examined and is considered as a basis for reasoning, discussion, or calculation (H2020 definition)

Research data is data collected, observed, generated, created or obtained as part of a research or environmental monitoring and assessment activity, and on which an argument, theory, hypothesis, conclusion or any another research output is based.
Types of research data
Research data management (RDM)

The effective and responsible handling of research data

Illustration CC BY DMS 2020 with icons CC BY Prosymbols (Flaticon.com).
Mentimeter: Time to confess!

Go to www.menti.com and use the code 7661 5345

Instructions

Go to www.menti.com
Enter the code
7661 5345

Or use QR code
Plan data management

- Address data management in proposals/grant applications/consider your funder’s data management and data publishing requirements
- Consider legal and ethical aspects that may affect data management (ethical review, consents, GDPR etc.).
- In collaboration research activities, make agreements regarding responsibilities for data (storage, long term preservation, publishing etc.)
- Use a data management plan (DMP)
- Contact the data management support function

6th of December: Kom igång med din datahanteringsplan
7th of December: Get started with your data management plan
Collect, organise and store data

- Use SLU storage that is covered by backup and security routines
- Choose open (non proprietary), widely used and machine readable file formats
- Apply a logical and consistent system for organising and naming files
- Format, label and describe data, using common standards and terminologies/controlled vocabularies
- Maintain documentation on a study level (context, methods etc.), file level and data element level (variables, allowable values, abbreviations etc.)
Process and analyse data

- Make a working-copy of collected data for processing and analysis, keeping original raw data intact
- Document data actions and steps in data analysis (cleaning, validation, aggregation, coding, etc.)
- Save and backup iterations of data and apply a clear and consistent versioning system
- Document and preserve code and scripts
- If specialised software is used for data processing or analysis, export data to an open format
Archive and preserve data

- Save data in open and widely used file formats
- Keep data files in good order and be prepared that access to data can be requested during data collection and before the project is finalised
- Do not disperse, remove or dispose of research data without a prior legal decision to do so
- Secure the future of data by archiving them at SLU!
- Archive data with associated documentation to enable interpretation and reuse in the future

26th of October: The principle of public access to official documents – can anyone ask for my research material?
Share and publish data

- At an early stage, consider what can and will be shared, how and under what conditions

- Publish data in a data repository under the principle “as open as possible, as closed as necessary”, and receive a persistent identifier (e.g. DOI)

- Include documentation that describes data files and data variables

- Also publish questionnaires, protocols, code etc. that may facilitate data interpretation, reproducibility and reuse

- Add a data availability statement in your article and cross link between article and dataset

16th of November: Research data and Swedish law, GDPR and immaterial rights
Discover, reuse and cite data

• Before deciding to collect new data, investigate if there are existing data that could be used to answer your research question
• Assess reliability and quality
• Check terms of use and obtain permissions if required (legal/ethical)
• Account for the origin/location of data and how they were applied in your study
• Cite properly
Why should I improve my data management?
Something you already do – a skill to improve

Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse by the community after the data publication process.

Good research data management helps you achieve goals…
# Increase your research efficiency
# Keep your research safe
# Improve your research integrity
# Enhance visibility of research output
# Enable collaboration
# Comply with legislation and funder policies
# Demonstrate responsible practice
# Work towards open science and FAIR data
Data management as part of good research practice
Research practice

Prepare
- Question
- Method
- Hypothesis

Propose
- Project plan and budget
- Publishing plan
- Application

Investigate
- Collect data
- Read up

Analyse
- Process
- Rethink question
- Conclude

Present outcome
- Conferences
- Scientific journals
- Data, code etc.

Report results
- Archive
- Funder

Illustrations: CC BY DMS 2020 with icons CC BY ProSymbols (Flaticon.com); CC 0 (DMS 2022).
Research practice

Discover, reuse and cite data

Plan data management

Collect, organise and store data

Process and analyse data

Share and publish data

Archive and preserve data

Prepare
- Question
- Method
- Hypothesis

Propose
- Project plan and budget
- Publishing plan
- Application

Investigate
- Collect data
- Read up

Analyse
- Process
- Rethink question
- Conclude

Present outcome
- Conferences
- Scientific journals
- Data, code etc.

Report results
- Archive
- Funder

Illustrations: CC BY DMS 2020 with icons CC BY Prosymbols (Flaticon.com); CC BY DMS 2022.
Good research ethics

... tell the truth about your research;
... strive to conduct research without harming people, animals or the environment;
... do not make unauthorised use of the research results of others
... be fair in your judgement of others’ research;

... review and report the basic premises of your studies;
... openly account for your methods and results, as well as for your commercial interests and other associations;
... keep your research organised, e.g. through documentation and filing

Ref. Good research practice (vr.se) p12.
Illustrations CC BY Peggy_Marco (Pixabay) with text from VR.se
Professional ethics

State employees
- Democracy
- Legality
- Objectivity (impartiality)
- Freedom of opinion
- Respect for everybody’s equality, freedom and dignity
- Efficiency and service

SLU additions
- Scientific approach
- Creativity
- Openness
- Responsibility
SLU Data management policy

… for digital material, both raw data and aggregated, that is produced, processed and used by research and environmental analysis at SLU

… to support and guide SLU employees in the effective and compliant handling of data and in the transition to open science

… and covering….

• Basic principles for data management at SLU
• Legal aspects on research and environmental analysis data
• Practical aspects: Data management plans, storage, information security and archiving
• Roles and responsibilities for employees, heads of departments and SLU
Legal obligations

• Research data are considered official documents and must be archived at the university

• Research data can be requested by anyone, according to the principle of public access to official documents

• If you process personal data – sensitive or not – you have to inform the participants

• Personal data and sensitive data must be processed with caution, but they also must be archived

Poorly managed research data are often difficult to archive properly

If we cannot hand out official documents that were requested, SLU will be liable

Incorrect information to participants on how personal data will be handled can render data unusable

If personal/sensitive data get in the wrong hands, SLU will be liable
Ethical requirements

• You need to know what you will, can and cannot do with the data you manage

• You need to know what has been done with the data during the research project

Could you show what has been done with the data, in case your results are questioned?

May I share the data with other researchers?

What data storage is OK to use?

Is it allowed to publish data?

What could you promise when you inform the participants in a study?
Will improved data management benefit my research?
Would improved data management benefit your research?

Go to www.menti.com and use the code 7661 5345

Instructions

Go to www.menti.com
Enter the code 7661 5345

Or use QR code
Open science and research data management
What is open science?

“Making science more accessible, inclusive and equitable for the benefit of all”

UNESCO (United Nations Educational, Scientific and Cultural Organization)
Open science – global core values and principles

VALUES
- Quality and integrity
- Collective benefit
- Equity and fairness
- Diversity and inclusiveness

PRINCIPLES
- Transparency, scrutiny, critique and reproducibility
- Equality of opportunities
- Responsibility, respect and accountability
- Collaboration, participation and inclusion
- Flexibility
- Sustainability

“The COVID-19 pandemic has also demonstrated the critical importance of fulfilling the human right to benefit from the advances of scientific progress as particularly highlighted in Article 27 of the Universal Declaration of Human Rights.”

Shamila Nair-Bedouellé, Assistant Director-General for Natural Sciences, UNESCO

Open Science: fulfilling the human right to share in scientific advancement and its benefits | UNESCO

Sustainable Development Goals | United Nations

UNESCO Recommendation on Open Science
CC BY-SA 4.0 UNESCO
Open science and research data
– snapshots from national and local level

The Swedish government has commissioned the National Library of Sweden (Kungliga biblioteket) to develop national guidelines for open science.

About our work on open access to research data
The Swedish Research Council coordinates and promotes Sweden's work on introducing open access to research data. The goal is to complete the transition no later than 2026.

SLU’s Strategy 2021–2025
Digitalisation will also contribute to open science, where information and the communication of research results and open data are made available, refined and adapted for stakeholders and society at large.

Screenshots from:
The National Library has been tasked with developing national guidelines for open science
About our work on open access to research data
SLU’s strategy 2021–2025
SLU data management policy (not yet published)
SLU policy for scientific publishing
Open science in practice
- making the outputs and the process of research more open

"Open science is just science done right."
(Jon Tennant | Tidningen Curie)
Open data and FAIR data

“Open data and content can be **freely used, modified and shared by anyone for any purpose.**”
(http://opendefinition.org)

Principle for publically funded research:
“as open as possible, as closed as necessary”
(Council of the European Union)

“Data can, and should be FAIR even when access is restricted.”
(Horizon Europe Programme Guide)
What is FAIR data?
FAIR principles

Findable Accessible Interoperable Reusable

Complete FAIR principles

**To be Findable:**

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (defined by R1 below)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

**To be Accessible:**

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

**To be Interoperable:**

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data

**To be Reusable:**

R1. (meta)data are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards

From Wilkinson, M.D. et al (2016) CC BY
FAIR principles in short

**Findable**
Make your data findable by ensuring it:
- Has a persistent identifier
- Has rich metadata
- Is searchable and discoverable online

**Accessible**
Make your data accessible by ensuring it:
- Is retrievable online using standardised protocols
- Has restrictions in place if necessary

**Interoperable**
Make your data interoperable by using:
- Common formats and standards
- Controlled vocabularies

**Reusable**
Make your data reusable by ensuring it:
- Is well-documented
- Has clear licence and provenance information

Screenshots from How to make your data FAIR (openaire.eu) (CC BY)
Suggested steps towards FAIR data

• Plan data management with sharing and reuse in mind.
• Format and describe data using common standards and vocabularies.
• Save data in open, widely used, machine-readable file formats.
• Think about what explanatory information others will need to be able to reuse data and provide this as documentation.
• Publish data including documentation in a trusted data repository, with clear terms of use. If data cannot be open, publish data with restricted access (for request from the SLU archive).
Data publishing
Benefits of data publishing

- Reduced duplication of effort, accelerated science and innovation
- Democracy and transparency
- Research integrity
- Visibility and recognition
- New collaborations
Requirements for open and FAIR data - examples

• Governing bodies
  – Sweden: transition to open access to research data to be completed by 2026
  – Council of the European Union: ”as open as possible, as closed as necessary”

• Funders
  – Data publishing, FAIR data, open science

• Publishers
  – Publishing data that underpin research articles
  – Data availability statements

• Institutions
  – SLU data management policy (open science, FAIR data, data publishing etc.)
Publishing in data repositories

• Data repositories
  – General purpose
  – Domain specific
  – Institutional

• Open or less open
  – Full dataset for download
  – Description of dataset, access on request

• With or without published article

• Publish AND archive!

-re3data - registry of data repositories

Examples of data repositories

SND
Swedish National Data Service

DRYAD
GenBank

Harvard Dataverse

GBIF
Global Biodiversity Information Facility

Zenodo
Swedish National Data Service (SND)

SND is a national infrastructure for making research data available, run by a consortium of Swedish universities including SLU.

- Free of charge
- Secure storage
- Compliance with most funders’ and journals’ data sharing requirements
- Supports many of the FAIR principles
- Provides datasets with a unique identifier and persistent link (DOI)
- Allows a rich and standardised description of datasets including a selection of controlled vocabularies
- Datasets are discoverable through the SND research data catalogue as well as Google, Web of Science etc.
- Data submissions from SLU are reviewed and curated by Data Management Support (DMS) before being published. DMS also assist in the data archiving process.
Swedish National Data Service (SND)
- example

Published paper

**Biological Conservation**
Volume 267, November 2021, 109963

Flower strips enhance abundance of bumble bee queens and males in landscapes with few honey bee hives


* Swedish University of Agricultural Sciences, Department of Ecology, SE-750 07 Uppsala, Sweden
* Swedish Rural Economy and Agricultural Society, SE-291 09 Kristianstad, Sweden
* Queensland Department of Agriculture and Fisheries, Toowoomba, Queensland, Australia

Data accessibility statement

Data and R script are available online at the Swedish National Data Service at
https://doi.org/10.5878/s2bx-3j34. Persistent link to dataset

Published paper: https://doi.org/10.1016/j.biocon.2021.109363

Published dataset

Published dataset: https://doi.org/10.5878/s2bx-3j34

Citation


Citation with persistent identifier/link

Published paper: https://doi.org/10.1016/j.biocon.2021.109363

Published dataset: https://doi.org/10.5878/s2bx-3j34

Citation with persistent identifier/link
If you need support...
Data Management Support: Team competence

• A division within the SLU university library

• Background in research, libraries, archives, education, system development and public institutions

• Domain expertise on environmental data

Website www.slu.se/dms
Team inbox dms@slu.se
www.slu.se/dms
Support and services from the SLU Library

• Scientific literature (journals, books, databases etc.)
• Publication strategies and visibility
• Open access publishing and CC-licensing
• Publication of theses, reports etc.
• SLUpub – SLU publication database
• Language and academic writing support
• Information retrieval, structured literature searching, systematic reviews, reference management
• Research analysis and visualization of publication and citation patterns

https://www.slu.se/library/
Thank you for Zooming in!

Sign up for more at Workshops and webinars | Externwebben (slu.se)

Welcome to DMS webinars (autumn 2022)

28th of September (Wednesday 10-11.30): Data management as part of open science, FAIR data and good research practice – an introduction

26th of October (Wednesday 10-11.30): The principle of public access to official documents – can anyone ask for my research material?

16th of November (Wednesday 10-11.30): Research data and Swedish law, GDPR and immaterial rights

6th of December (Tuesday 10-11.30): Kom igång med din datahanteringsplan

7th of December (Wednesday 10-11.30): Get started with your data management plan