

# Data Management Plans (DMPs)

*Introduction to Data Management Practices course for NGI SNP&SEQ*

Section 1: Data Management Plans (DMPs)

7/9 10:15-11:00

**Mattias Strömberg** and **Erik Hedman**

NBIS DM Team

[data-management@scilifelab.se](mailto:data-management@scilifelab.se)



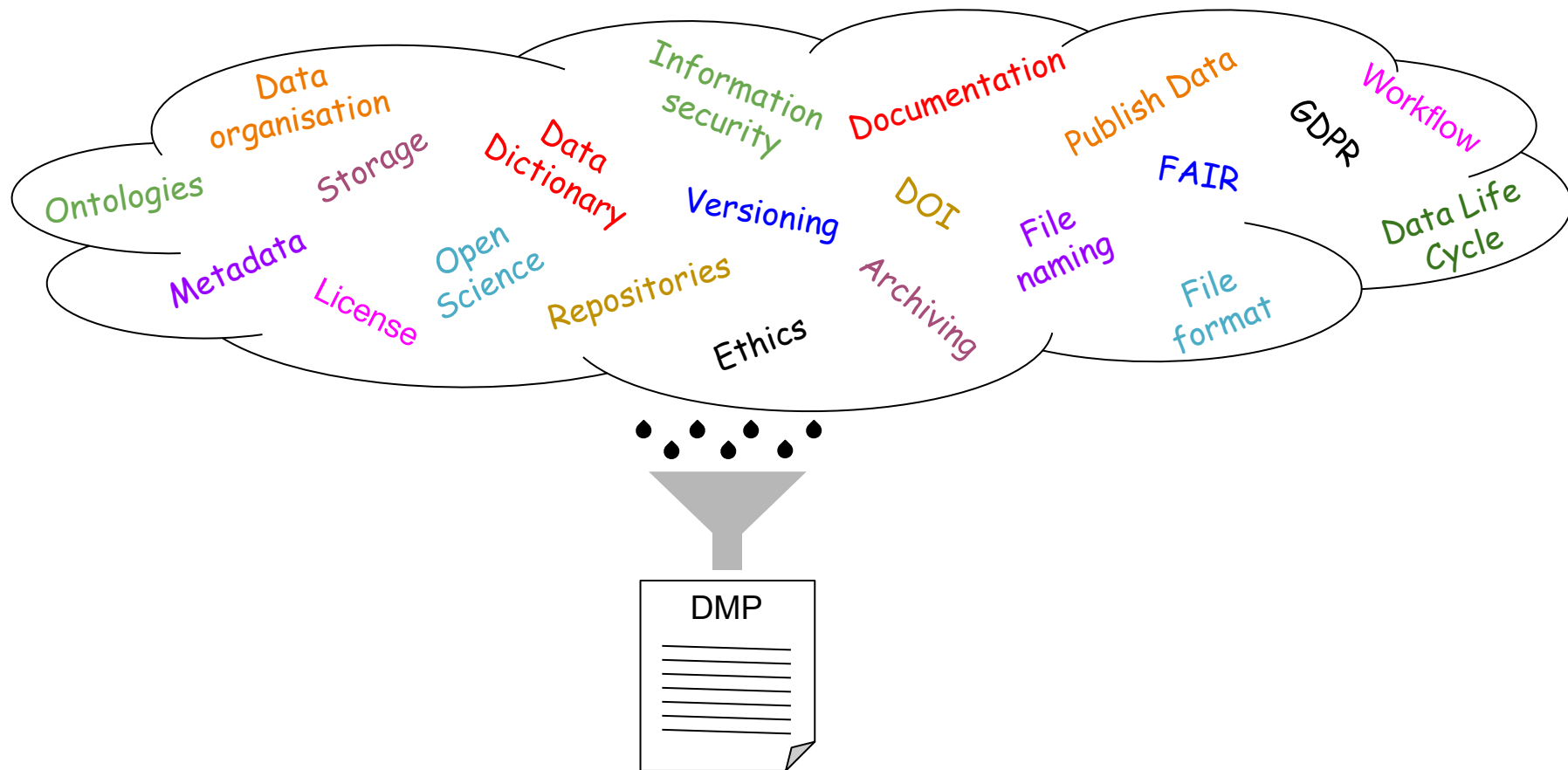
---

## First half (15 min): presentation

- What is a DMP?
- Why and when write a DMP?
- The main parts of a DMP.
- How to write a DMP?
- The role of NGI SNP&SEQ?
- Key points.
- Advertisement of an upcoming DMP event in November.

## Second half (30 min): discussions in smaller groups

- How can we all **facilitate the data journey** for the researchers and **minimize the administrative workload** for NGI SNP&SEQ?



- A document **addressing requirements and practices** for managing the project's data, code and documentation throughout the data life cycle.
  
- Outlines the **data management strategies** in a project, how the data is:
  - collected
  - documented
  - organized
  - preserved



- Identify **gaps** in current data management strategies (reduce time spent later on).
- Facilitate collaboration by setting project wide **standards** (data organisation, naming conventions, metadata standards, vocabularies etc.).
- **Cost** estimation.
- Clarify **responsibilities**.
- Obtain **well-managed** research data (which enables reproducibility and reusability).
- First step towards **FAIRness**.
- **Requirement** from funders and other stakeholders for reasons of **transparency, openness** and **return on investment**.

---

A DMP is a **living document** that will develop throughout the project:

- **Project planning:** Outline the strategies to be able to estimate the resources needed, so this can be included in the proposal for funding.
- **Project start:** Complete with details e.g. about documentation, data quality measures, file and folder strategies, etc.
- **Project end:** Update with e.g. links to published data and details about archiving (what data and where).

## 1. Description of data

What types of data will be created and/or collected, in terms of data format and amount/volume of data?

Reuse and  
production of  
research data

## 2. Documentation

How will the material be documented and described, with associated metadata relating to structure, standards and format for descriptions of the content, collection method, etc.?

Documentation  
procedures and  
standards

## 3. Storage and backup

How is data security, storage and backup of data and metadata safeguarded during the research process?

Storage,  
security and  
backup



## 4. Legal and ethical aspects

How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?

Legal & ethical  
and assurance  
procedures

## 5. Accessibility and long-term storage

- How, when and where will research data or information about data (i.e. metadata) be made accessible?
- In what way is long-term storage safeguarded, and by whom?

Future reuse  
and long-term  
access aspects

## 6. Responsibility and resources

- Who are the responsible persons for data management?
- What resources (costs, labour input or other) will be required for data management?

Responsibilities  
and resources  
required

## DMP templates:

- Provided by funding agencies, e.g. [Swedish Research Council](#) and [Science Europe](#).
- High-level questions, with no guidance on how to answer.
- Use MS Word?



## DMP tools:

- [DMPOnline](#) - Common for universities; Good guidance but not Life Science specific; Free text answer type.
- [Data Stewardship Wizard](#) - Provided by [SciLifeLab](#); Life Science specific guidance; Many questions with answer options.



Example: <https://dsw.scilifelab.se/projects/9e96f8a0-9422-449a-bf72-81dea7a9cd9d>

---

The **journey of sequencing data** in the Data life cycle involve **three key players: sequencing facilities, researchers and data repositories.**

- How can NGI SNP&SEQ in collaboration with NBIS DM-team support researchers in order to **facilitate the sequencing data journey**?
- Which **general documentation** does NGI SNP&SEQ provide on data and metadata formats for sequencing projects?
- For **sensitive human data deliveries**, is there any discussion with the researcher about legal or security DM guidelines?

- A data management plan (DMP) is a document that **describes the data produced** in the course of a research project.
- A DMP allows for **well-managed data**, and funding agencies often requires a DMP for transparency and return on investment.
- A DMP is a **living document**, the first version is written during project planning, and is then updated as the project proceeds.
- There are **standard templates** available e.g. at funder agencies, and **tools to assist** when writing.

## When

Online 22 Nov 10:00-12:00 (more information will be distributed later on)

## Target audience

Researchers, doctoral students, personnel handling research data etc.

## Aim

Highlight the concept and use of DMPs and how funders consider the importance of DMPs.

## Agenda (preliminary)

10:00-10:45 Introduction

- DMPs in relation to research plans, why write DMPs?
- Funders perspectives on DMPs and their requirements (VR, Riksbankens Jubileumsfond, and others).
  - a. *What is your view on DMPs and their importance?*
  - b. *How do you ensure that researchers have DMPs?*

10:45-11:00 Future perspectives on the use of DMPs.

11:05- 12:00 Demos of 2 tools for creating DMPs: DMP Online and DSW.

---

How can we **facilitate the data journey** for the researchers and **minimize the administrative workload** for NGI SNP&SEQ?

**Questions to answer are available in collaborative notes**