



UiO : University of Oslo Library

# Module II: Data management planning

Matthew Good, Kirsten Borse Haraldsen, Aino Rinhaug, Agata Bochynska  
Digital Scholarship Center



29.april 2021

## Agenda

- What is a data management plan (DMP)?
- Why use one?
- Ten steps to writing a data management plan
- Examples



## Rules for today

- Interrupt me if I speak too fast!
- Feel free to write your questions in the chat, we will answer them at the end
- We will be using [menti.com](https://www.menti.com) – Go to [menti.com](https://www.menti.com) and use code 2809 1686

## **Menti – Go to [menti.com](https://menti.com) and use code 2809 1686**

- Which faculty or department do you work in?

# What is a Data Management Plan (DMP)?



- Describes how you plan to manage research data
- A tool to keep an overview of the project's data
- “Living” document that should be updated as needed
- An agreement between participants
- A form of documentation for your project

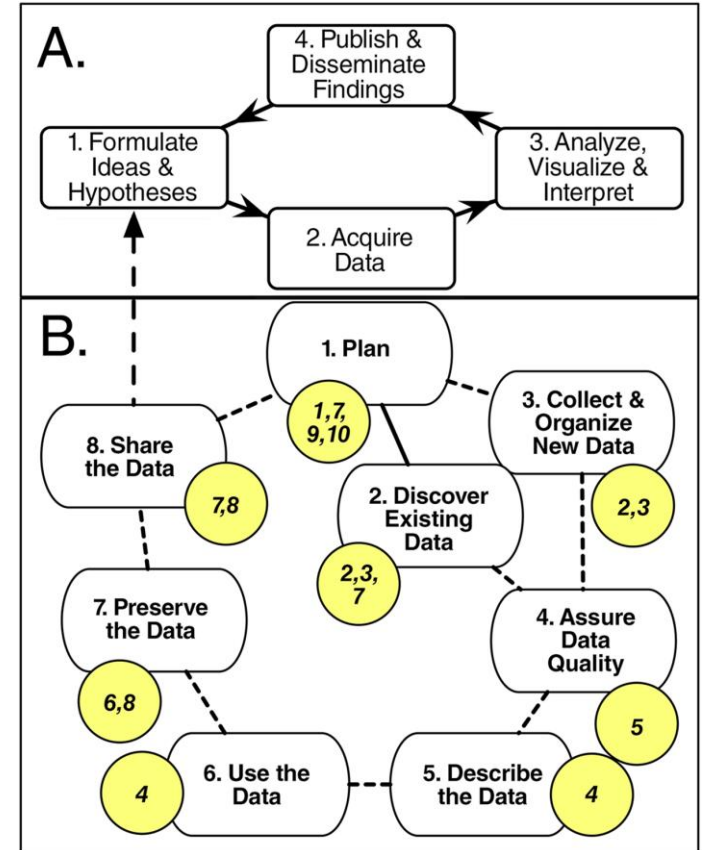
# Why use a Data Management Plan (DMP)?

- Comply with sponsor requirements
- Save time during and after project
- Easier to share and publish data
- Make your data reusable and reproducible
- Ensures security and quality of data
- Helps identify resource requirements, time dependencies or problems
- Sharing data can give you credit in research assessment schemes




# 10 Simple Rules to Writing a Data Management Plan


1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget





# 1. Sponsor Requirements


 **DMP template or content requirements?**


 Storage requirements?


 Metadata and documentation requirements?

 Publishing data openly?

 Archiving requirements?

 Which repositories are approved?

 Licensing requirements?

 Other legal requirements?





## A truncated list of DMP templates

If you:	Consider using:	Link
Have human and society data	NSDs DMP tool	<a href="https://www.nsd.no/lag-en-datahandteringsplan">https://www.nsd.no/lag-en-datahandteringsplan</a>
Are working in an EU project	Digital Curation Center's DMPOnline	<a href="https://dmponline.dcc.ac.uk/">https://dmponline.dcc.ac.uk/</a>
Have life science data	ELIXIR's template	<a href="https://elixir.no/Services/data-management/data-stewardship-wizard">https://elixir.no/Services/data-management/data-stewardship-wizard</a>
Use Sigma2's resources	Sigma2's easyDMP	<a href="https://www.sigma2.no/data-planning">https://www.sigma2.no/data-planning</a>
Want a simpler or text-based template	UiO's DMP template	<a href="https://www.uio.no/english/for-employees/support/research/research-data-management/data-management-plan/index.html">https://www.uio.no/english/for-employees/support/research/research-data-management/data-management-plan/index.html</a>

## UiO Requirements and Policy

The University of Oslo aims to manage research data according to the **FAIR and CARE principles**.

Data should be:

- *made openly available for further usage*
- *made available at an early stage*
- *provided with a data management plan*
- *have metadata and be documented*
- *be securely archived*
- *be provided with licenses for access, reuse and redistribution*
- *made freely available, but the actual distribution cost should be covered*



# FAIR Principles

Make your data:

- **F**indable
- **A**ccessible
- **I**nteroperable
- **R**eusable

## Findable

- Descriptive metadata
- Persistent Identifiers

## Accessible

- Determining what to share
- Participant consent and risk management
- Access status

## Interoperable

- XML standards
- Data Documentation Initiative
- CDISC

## Reusable

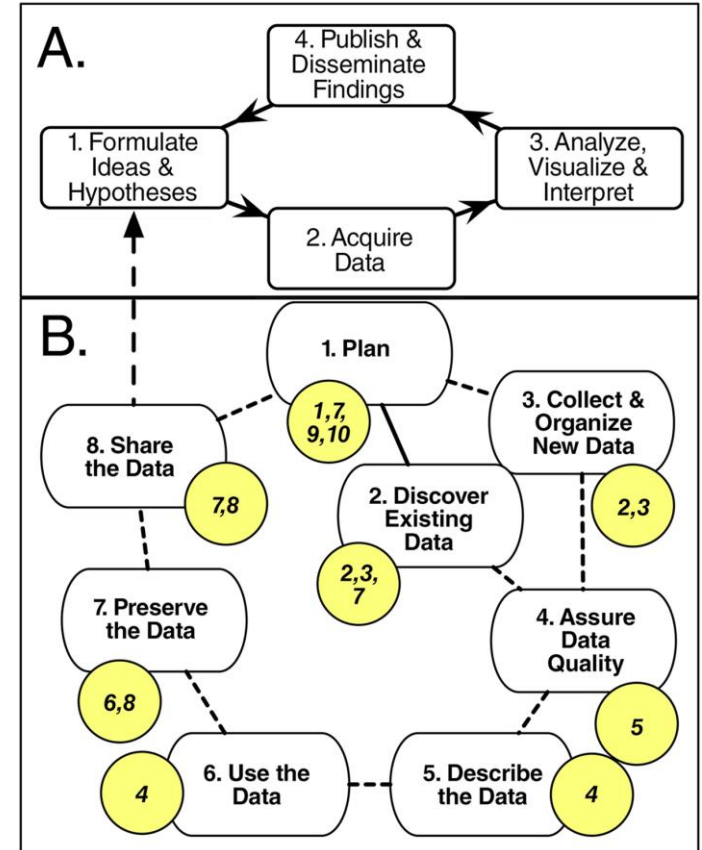
- Rights and licence models
- Permitted and non-permitted use

<http://datafairport.org/>



# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
- 2. Identify data to be collected**
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget



## 2. Identify data to be collected



### Origin of Data

Existing data or new data?

If data is reused: What is the scope, volume, format?

How are different data sources integrated?



### Types of Data

What type(s) of data will be collected?

What is the scope, quantity and format of the material?

What is the total amount of data collected (MB/GB/TB/PT)?



### Data Collection

How, where and when will data be collected?

What resources are required?

Who is responsible for data collection?

## **Menti – Go to [menti.com](https://menti.com) and use code 2809 1686**

- What kinds of data do you work with?

## Example

**Type:** accelerometer output

**File format:** .csv

**Sensor:** AX3 Axitivity

**File size:** ~ 15 x 1 MB

**Placement:** tape to the base  
of the skull

**Classification:** green

**Purpose:** measure head  
movement of audience and  
musicians

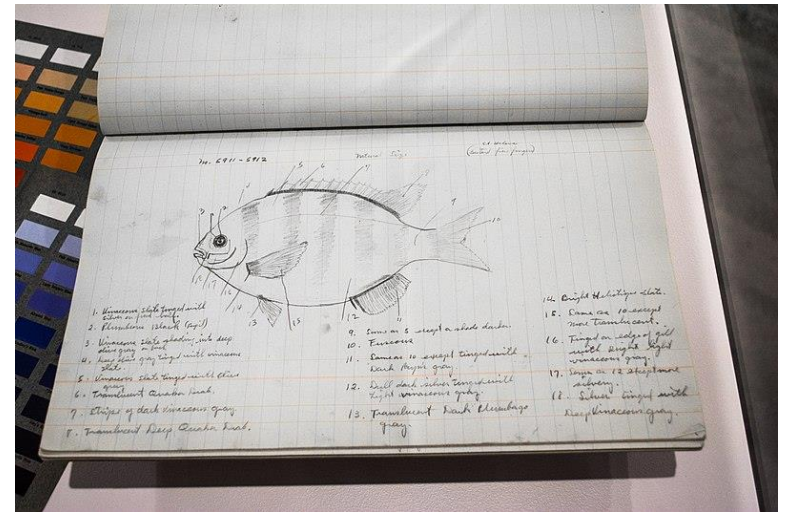
**Setting:** musical  
performance at a music  
festival on (date)

# Administrative and Physical Data

Important to think about how non-digital or administrative data will be organized

Examples can include:

data management plan, applications forms, field notebooks, lab notebooks, physical artifacts, protocols, etc.

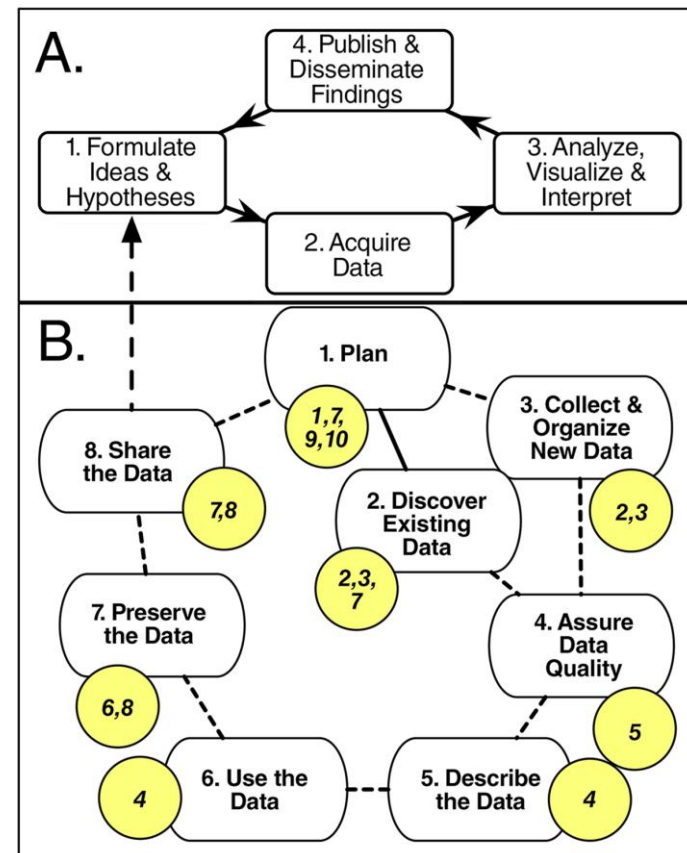


by Tim Evanson, CC-BY-SA



# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
- 3. Define how data will be organized**
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget

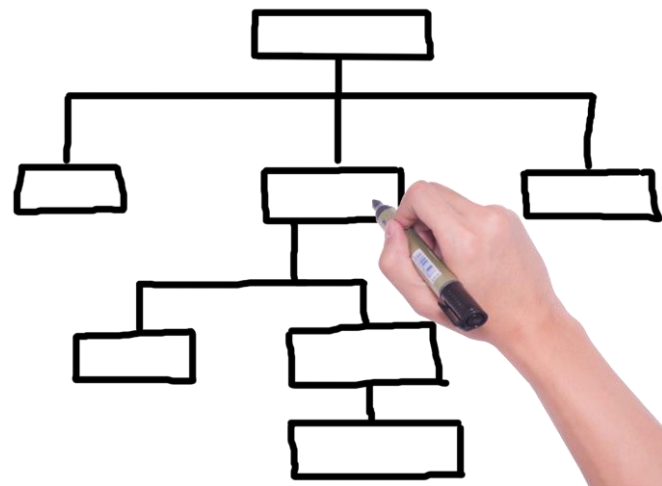


### 3. How will data be organized?

- Make a structure and stick to it!
- Come to an agreement within your team
- Follow-up and revise as necessary

#### Questions to answer

- What file formats will you be using?
- How will you structure the folder system and/or database?
- How will you name files?
- How will you take care of older versions and articles?



## Organizing your files



### What to do:

- Adapt a **file naming convention (FNC)**
- Include in your directory a **readme.txt file**, that explains your file naming format and other codes and abbreviations in use.



### What *not* to do:

Research\_data\_2020  
Project\_data 2019  
Notes\_2018%  
Observations\_br\_16\_1\_\*

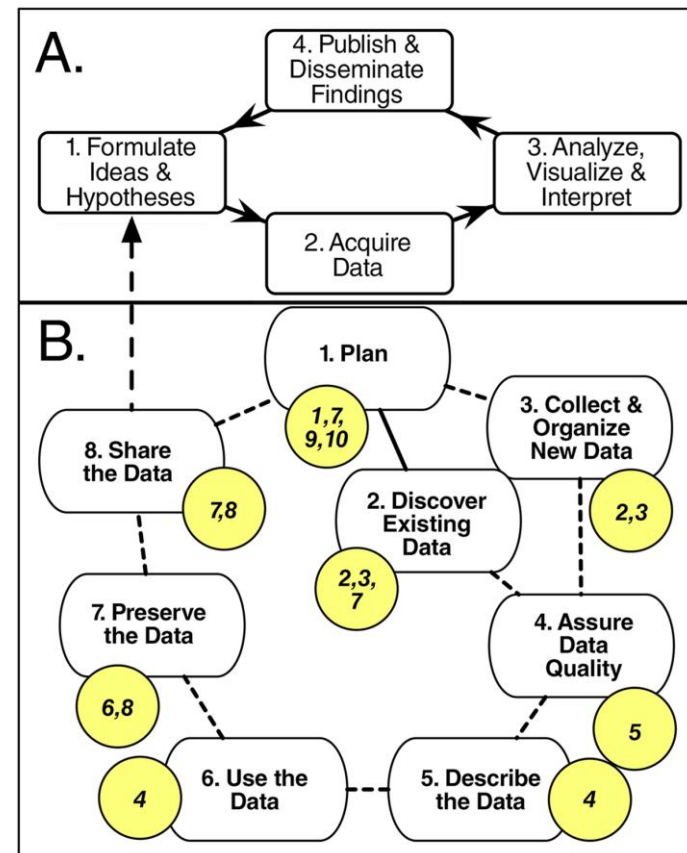
## File naming convention (FNC)

- Is consistent and descriptive in naming and organizing files:
  - *Project or experiment name or acronym*
  - *Location/spatial coordinates*
  - *Researcher name/initials*
  - *Date or date range of experiment*
  - *Type of data*
  - *Conditions*
  - *Version number of file*
  - *Three-letter file extension for application-specific files*

[Source: Best practices for file naming | Stanford Libraries](#)

# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
- 4. Explain how data will be documented**
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget



## 4. How will data be documented?

### Documentation

- What information is needed for comprehensibility, reusability, and reproducibility?
- What kind of documentation?
  - Instrument settings and calibration
  - Protocols and scripts
  - Field journals
  - Lab notebooks
  - Readmes

### Metadata

- What kind of metadata will be provided with the data?
- How will metadata for each object be created?
- (If applicable) What metadata standard(s) will you use?

## Example README file:

```
*AUTHOR_DATASET_ReadmeTemplate - Notisblokk
Fil Rediger Format Vis Hjelp
This DATSETNAMEREADME.TXT file was generated on YYYY-MM-DD by NAME
<help text is included in angle brackets, and can be deleted before saving>

GENERAL INFORMATION

1. Title of Dataset:

2. Author Information
   A. Principal Investigator Contact Information
      Name:
      Institution:
      Address:
      Email:

3. Date of data collection (single date, range, approximate date) <suggested format YYYY-MM-DD>:

4. Geographic location of data collection <latitude, longitude, or city/region, State, Country, as appropriate>:

5. Information about funding sources that supported the collection of the data:
```

Cornell University template & guide

<https://data.research.cornell.edu/content/readme>

## **Dataography: Metadata – “data about data”**

*Basic (descriptive):*

- Who created the data*
- What the data file contains*
- When the data were generated*
- Where the data were generated*
- Why the data were generated*
- How the data were generated*

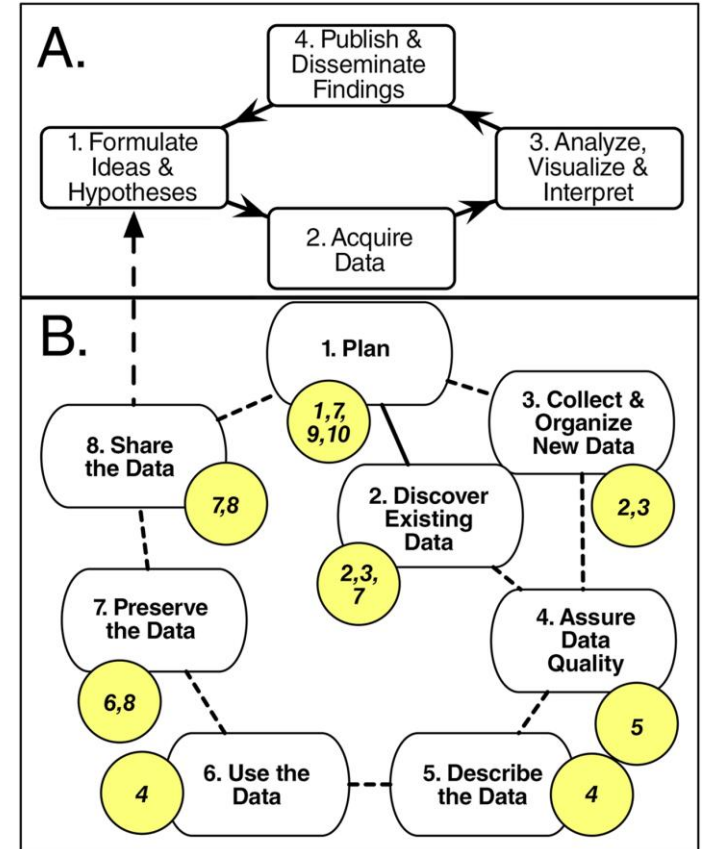
*Advanced (structured):*

- Repositories may require you to format your metadata using a **metadata standard, ontologies and/or controlled vocabulary.***
- Metadata structures = **«schema»***
- The metadata are often reported in a **machine-readable language (XML)***

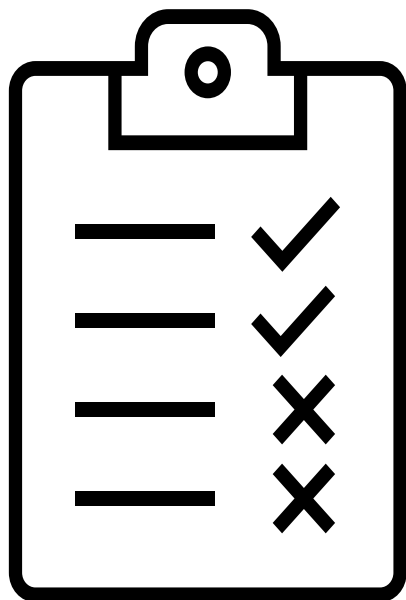


# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
- 5. Describe how quality will be maintained**
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget



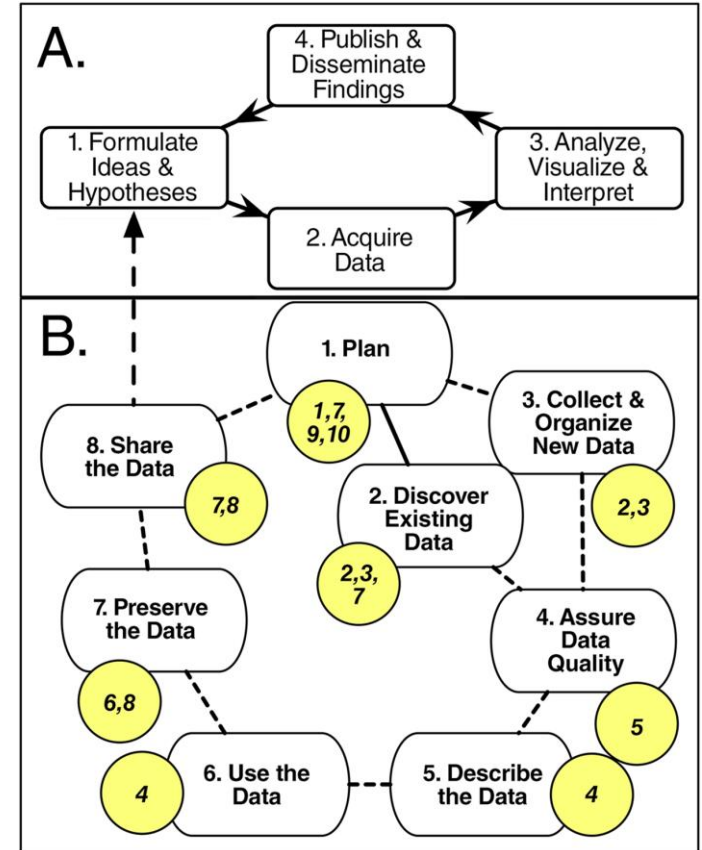
## 5. How will data quality be assured?



- Dependent on the type of research project and sponsor requirements
- Examples include:
  - training activities,
  - instrument calibration and verification tests
  - double-blind data entry
  - statistical and visualization approaches to error detection
  - multiple researchers that can review transcripts

# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
- 6. Develop a strong storage and preservation strategy**
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget



## 6. How will data be stored and preserved?

- Data classification will guide your choice of storage
- What storage solution is most appropriate?
- How will need for access be managed and assessed?
- Backups?
- Version control?
- Short-term and long-term costs?

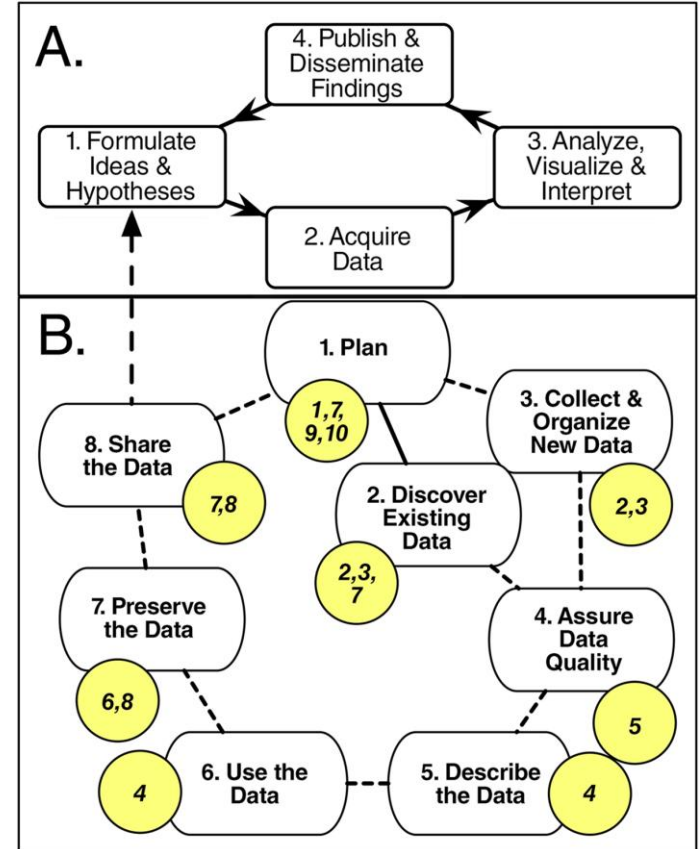


# Where can you store the different types of data?

- **Green data:**
  - anywhere you want, but ideally **not only** on your private laptop (very vulnerable!)
- **Red data:**
  - Fully encrypted disc, memory stick or external hard drive
  - Services for sensitive data – TSD
  - UiO Storage Hotel\*
  - UiO Vortex\*
- **Black data:**
  - Services for sensitive data – TSD\*
- **Yellow data:**
  - encrypted hard drive
  - UiO cloud solutions (UiO OneDrive, UiO G-Suite, ...)
  - UiO home directory (M: )
  - UiO storage hotel («lagringshotell»)
  - UiO e-mail; UiO Teams; UiO Nettskjema; UiO Canvas...
  - **Not** your private laptop or phone, unless they are encrypted and safely used

# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
- 7. Define project's data policies**
8. Describe how data will be disseminated
9. Assign roles and responsibilities
10. Prepare a realistic budget



## 7. What are the project's data policies?

How will you comply with legal requirements such as privacy, confidentiality and intellectual property rights?

- e.g. music performances and artistic works are copyrighted

How will you ensure ethical principles for research are followed throughout the project?

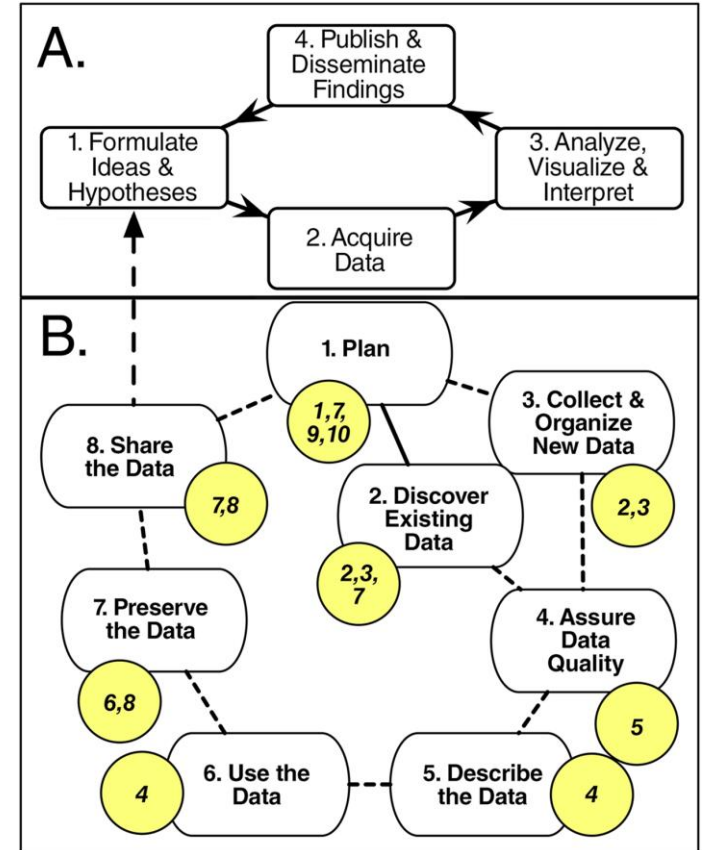
- e.g. CARE principles or other research norms (i.e. REK)

What kind of licenses and restrictions do you intend to use for sharing your data (if any)?

- e.g. creative commons (CC) licenses.

# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. **Describe how data will be disseminated**
9. Assign roles and responsibilities
10. Prepare a realistic budget



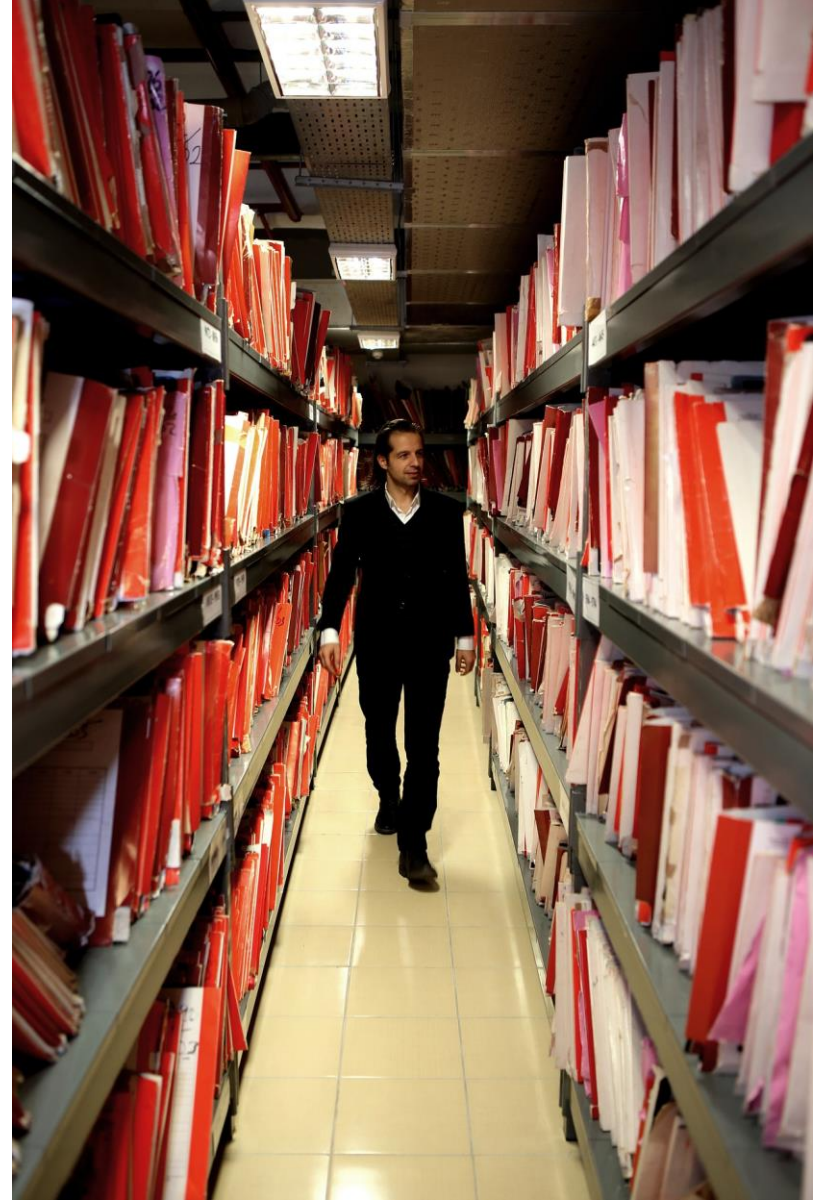


## 8. How will the data be disseminated?

- Personal website or email?
- Published in an open/closed repository/archive?
- Submitted as a part of a journal article?
  - Check the journal’s “data policy”
- Published in a data paper (e.g. GeoScience Data Journal)?
  
- What are the requirements for your chosen option?
  - Persistent identifiers (e.g. DOI), metadata, documentation, non-proprietary file formats, FAIR

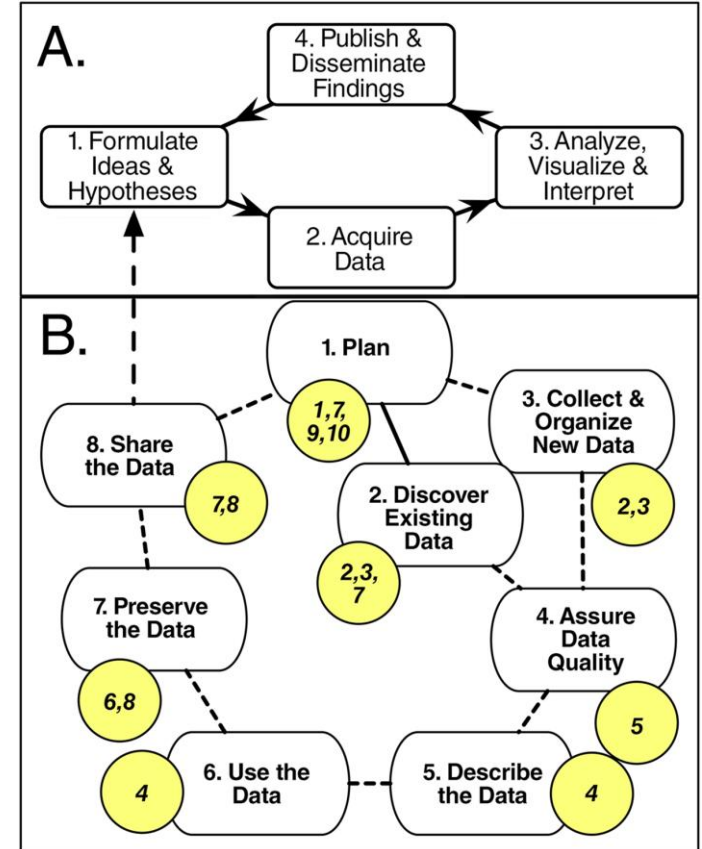
## Examples of archives:

- ✓ [Zenodo](#) EU's archive
- ✓ [NSD](#) national archive
- ✓ [The Norwegian e-infrastructure for research & Education \(NIRD\)](#) national archive
- ✓ [DataverseNO](#) national archive
- ✓ [re3data.org](#) a global list of archives



# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
- 9. Assign roles and responsibilities**
10. Prepare a realistic budget



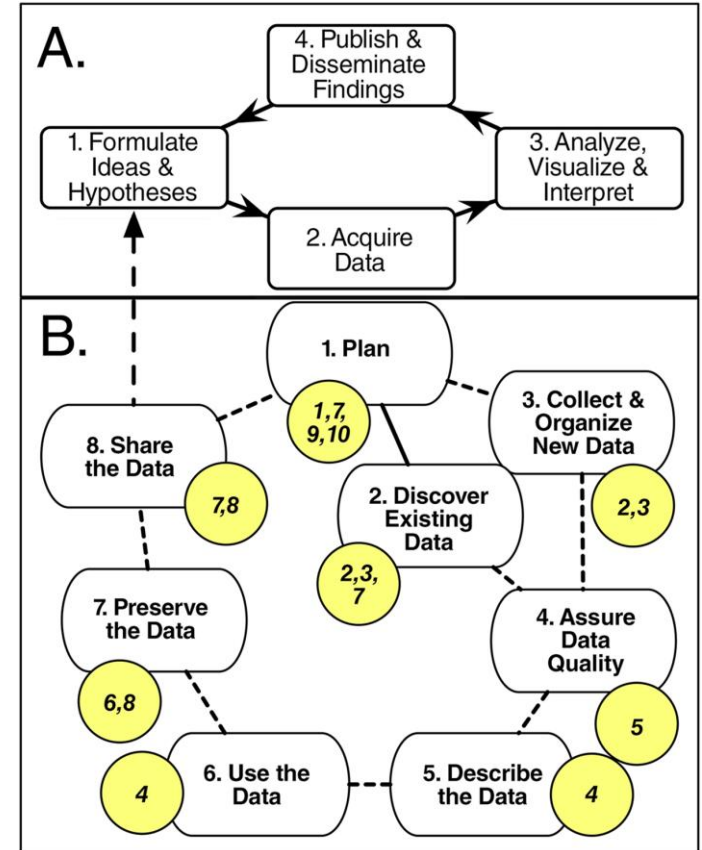
## 9. Assigning roles and responsibilities

- What are the roles of each participant in the project (not just in data management)?
- Who else outside of the project will be contributing?
- Who is responsible for data management during the project?
- Who is responsible for data management and archiving after the project is finished?
- Who will make sure the DMP is being followed?
- Who will be responsible for updating the DMP and how often?



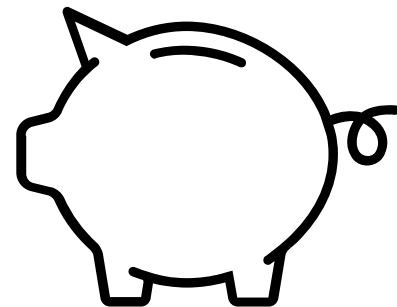
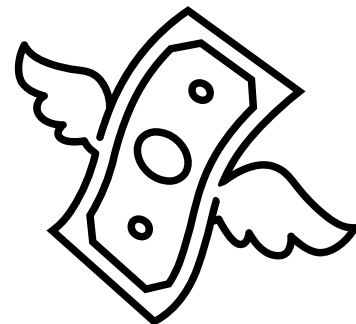
# 10 Simple Rules to Writing a Data Management Plan

1. Determine research sponsor requirements
2. Identify data to be collected
3. Define how data will be organized
4. Explain how data will be documented
5. Describe how quality will be maintained
6. Develop a strong storage and preservation strategy
7. Define project's data policies
8. Describe how data will be disseminated
9. Assign roles and responsibilities
- 10. Prepare a realistic budget**




## 10. Prepare a realistic budget

- Are there costs associated with purchase of software or hiring a data manager?
- Will you need to purchase new hardware or equipment?
- Storage costs: How much will it cost during the project (and after!)
- Is there a cost for archiving?
- What other resources are required? (e.g. HPC, licenses, ELN, etc.)







“In preparing for battle, I have always found that plans are useless, but planning is indispensable.”

- Attributed to Dwight D. Eisenhower in Richard Nixon's *Six Crises* (1962)



## **Menti – Go to [menti.com](https://menti.com) and use code 2809 1686**

- What kinds of problems have you had working with your data?
- How could a DMP help you in your project?



## Examples of DMPs

Synthetic Chemistry / EPSRC from University of Glasgow  
[https://www.gla.ac.uk/media/Media\\_418166\\_smxx.pdf](https://www.gla.ac.uk/media/Media_418166_smxx.pdf)

PhD Ecology / Lincoln University, New Zealand  
<https://riojournal.com/article/8664/>

[Public DMPs \(dcc.ac.uk\)](https://www.dcc.ac.uk/)

[Example DMPs and guidance | DCC](#)

**Thank you!**

**Questions?**

**Email us [research-data@uio.no](mailto:research-data@uio.no)**

**Book an appointment with us:**

**[Universitetsbiblioteket i Oslo |](#)  
[Scheduling and Booking Website](#)  
[\(simplybook.it\)](#)**



## Upcoming sessions...

- Data Organization, Metadata and Documentation  
10/05 9:00-10:30
- Data Classification and Storage Selection  
12/05 9:00-10:30
- Ethics, Privacy and Data Protection  
19/05 9:00-10:30
- Sharing and Archiving Research Data  
20/05 9:00-10:30
- Data Discovery  
21/05 9:00-10:30

Sign up here: <https://www.ub.uio.no/english/courses-events/courses/other/research-data/>

# Shameless advertising! Skills development @ UiO



Carpentry @ UiO



CODE REFINERY

**Carpentries**

[uio.no/carpentry](http://uio.no/carpentry)

**CodeRefinery**

[coderefinery.org](http://coderefinery.org)

**UiOs Digital Scholarship Center**

<https://www.ub.uio.no/english/writing-publishing/dsc/>

## Resources

UiO resources for data management – including the policy:

<https://www.uio.no/english/for-employees/support/research/research-data-management/>

FAIR principles: <https://www.go-fair.org/fair-principles/>

Further readings:

About DMP from Nature: <https://www.nature.com/articles/d41586-018-03071-1>

Michener, W.K. (2015) Ten Simple Rules for Creating a Good Data Management Plan. *PLoS Computational Biology* 11(10): e1004525.  
<https://doi.org/10.1371/journal.pcbi.1004525>