

Data Management Plans Workshop

Elisa Pierfederici

Digital Scholarship Center

University of Oslo Library

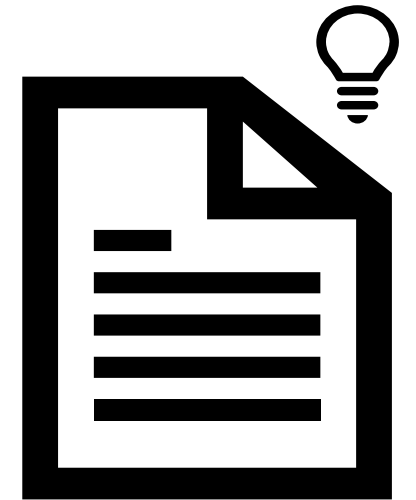
19.04.2023,

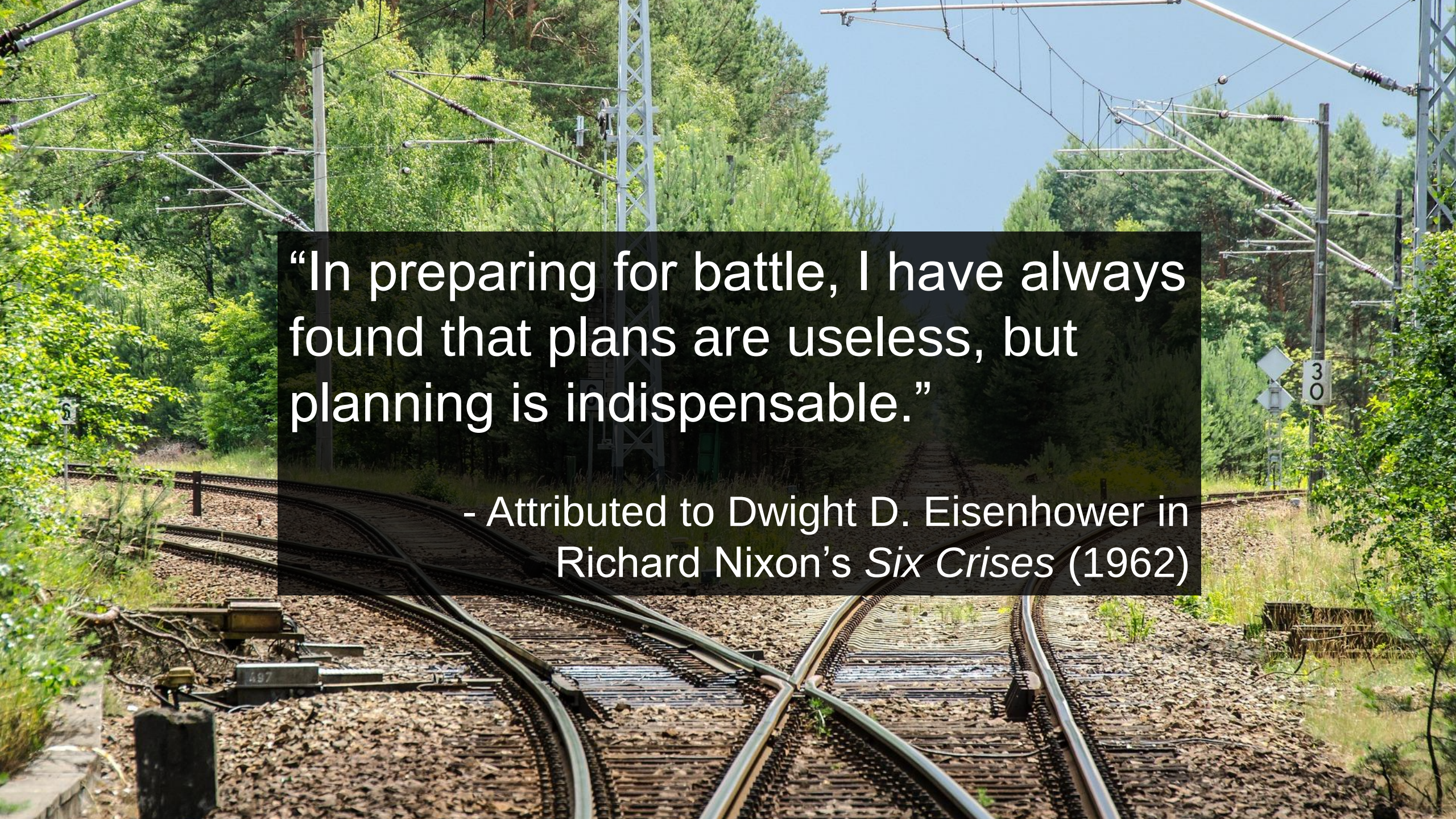
Contact us at research-data@uio.no

The Data Management Plan (DMP)

The Data Management Plan (DMP)

- Is a **living** document that accompanies the research project
- Specifies the **types** of data that will be generated
- Describes how you plan to **manage** your data (organize, document, classify, store)
- Conveys whether and how the data can be **shared**
- **Agreement** between project members
- A **tool** to keep overview over data





“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)

The Data Management Plan (DMP)

Text-based templates for writing DMPs

- UiO's template (in [Norwegian](#) and [English](#)) - simple Word-based template with links to UiO resources
- [Horizon Europe template](#) - recommended for use with Horizon Europe projects

Interactive tools and web forms that assist with writing a DMP

- [Sikt \(formerly NSD\)](#) - single and easy to use template but with limited functionalities
- [DMPonline](#) and [DMPTool](#) - popular generic tools supporting multiple templates; can be used with DMPs that will be made openly available in the future



Horizon Europe template

1. Data Summary

2. FAIR data

- 2.1. Making data findable, including provisions for metadata
- 2.2. Making data accessible
- 2.3. Making data interoperable
- 2.4. Increase data re-use

3. Other research outputs

4. Allocation of resources

5. Data security

6. Ethics

7. Other issues



Horizon Europe

Data Management Plan Template

Version 1.0
05 May 2021

DCC template (DMPonline)

Data Collection

Documentation and Metadata

Ethics and Legal Compliance

Storage and Backup

Selection and Preservation

Data Sharing

Responsibilities and Resources

[Project Details](#)[Contributors](#)[Plan overview](#)[Write Plan](#)[Share](#)[Download](#)

DCC Template

This plan is based on the "DCC Template" template provided by Digital Curation Centre.

The default DCC template

Template version 0, published on 15 June 2020

Instructions

The DCC default template

Data Collection

- What data will you collect or create?
- How will the data be collected or created?

Documentation and Metadata

- What documentation and metadata will accompany the data?

Ethics and Legal Compliance

- How will you manage any ethical issues?
- How will you manage copyright and Intellectual Property Rights (IPR) issues?

Storage and Backup

- How will the data be stored and backed up during the research?
- How will you manage access and security?

Selection and Preservation

- Which data are of long-term value and should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?

Data Sharing

- How will you share the data?
- Are any restrictions on data sharing required?

Responsibilities and Resources

- Who will be responsible for data management?
- What resources will you require to deliver your plan?

Sikt (NSD) template



Project information:

Research responsible institution

Project duration

Purpose

Utility

Funding

Ethical guidelines

Intellectual property rights

Related resources

Description

Dataset information:

Data type

Data about people?

Categories of personal data

Security classification

Data collection period

Collection devices

Data quality

Method

Size

Format

Software

Metadata standard

Naming conventions

Storage

Transfer

Archiving

Degree of openness

License

Archive

Costs

Project information

Research responsibility

Funding

Ethical and legal issues

Related resources

Data packages

XX



Classification

Collection

File management

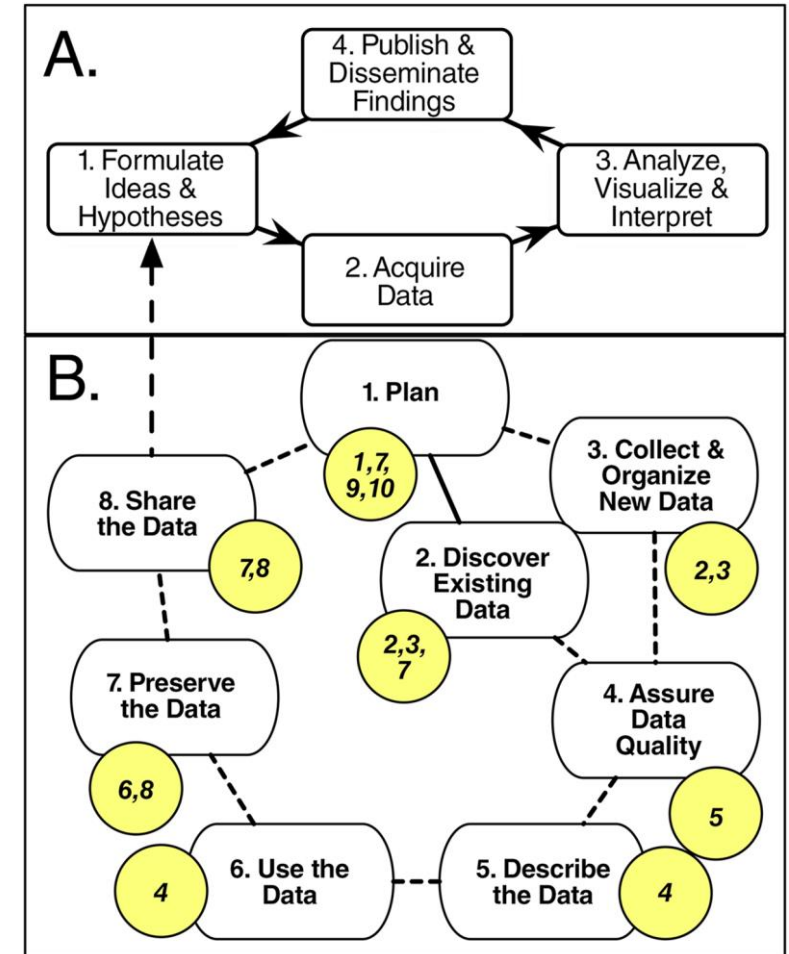
Metadata

Storage and transfer

Archiving

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**



Let's see an example!

[DCC template \(DMPonline\) example](#)

Find more examples in Public DMPs service from DMPonline:

https://dmponline.dcc.ac.uk/public_plans

Project Details

Contributors

Plan overview

Write Plan

Share

Download

DCC Template

This plan is based on the "DCC Template" template provided by Digital Curation Centre.

The default DCC template

Template version 0, published on 15 June 2020

Instructions

The DCC default template

Data Collection

- What data will you collect or create?
- How will the data be collected or created?

Documentation and Metadata

- What documentation and metadata will accompany the data?

Ethics and Legal Compliance

- How will you manage any ethical issues?
- How will you manage copyright and Intellectual Property Rights (IPR) issues?

Storage and Backup

- How will the data be stored and backed up during the research?
- How will you manage access and security?

Selection and Preservation

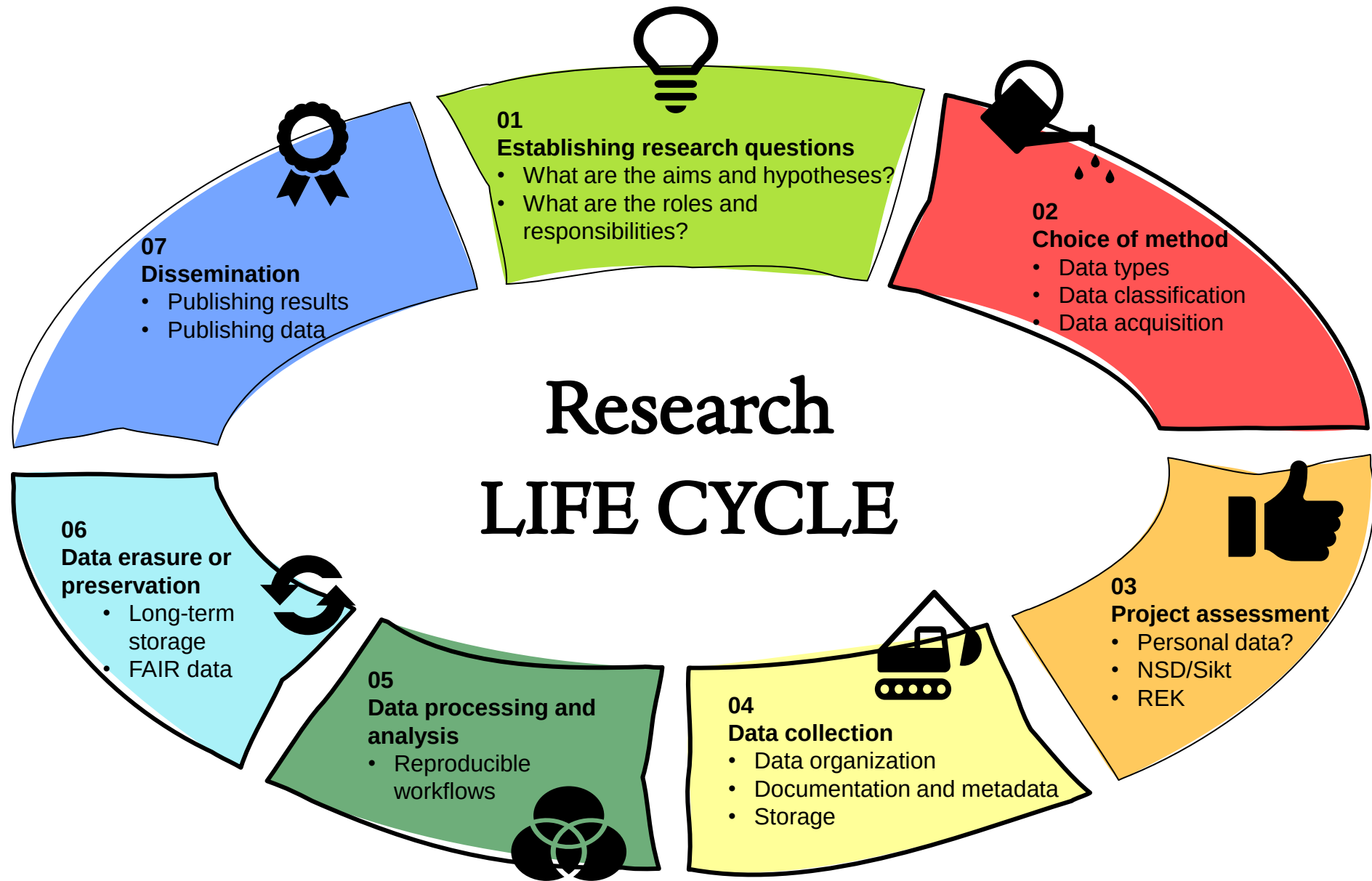
- Which data are of long-term value and should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?

Data Sharing

- How will you share the data?
- Are any restrictions on data sharing required?

Responsibilities and Resources

- Who will be responsible for data management?
- What resources will you require to deliver your plan?



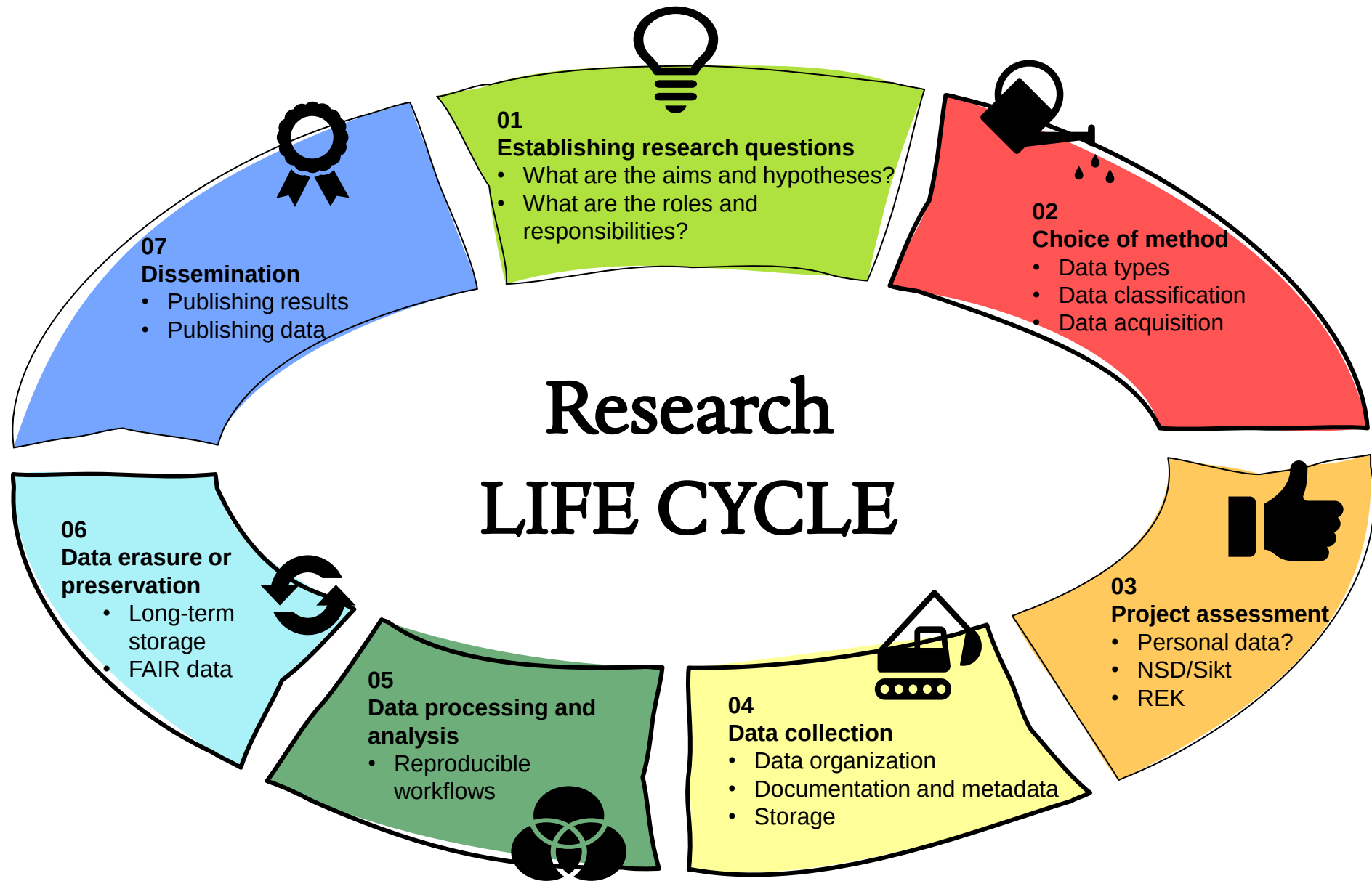
Assigning roles and responsibilities

- What are the roles of each participant in the project?
- 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 be responsible for Data Management Plan?



1

Take a look at the **Responsibilities** section in the handout



What is research data?

Research data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

Ref: Borgman, Christine L. 2015. *Big Data, Little Data, No Data : Scholarship in the Networked World*. Cambridge, MA: MIT Press.

What is research data?

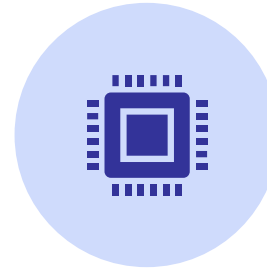
Research data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

Ref: Borgman, Christine L. 2015. *Big Data, Little Data, No Data : Scholarship in the Networked World*. Cambridge, MA: MIT Press.



Observational data

results from recognising, noting, or recording facts or occurrences of phenomena



Computational data

products of executing computer models, simulations, or workflows



Experimental data

results of procedures in controlled conditions to test or establish hypotheses or to discover or test new laws



Records

records of almost any phenomenon or human activity, can be treated as data for research

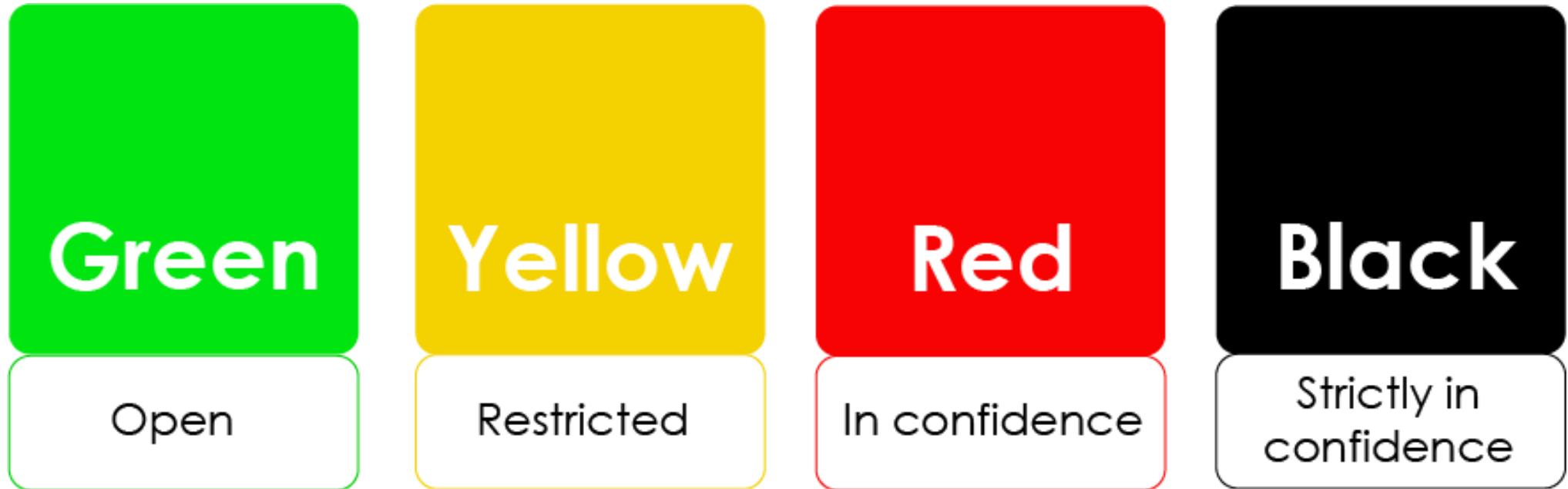
What is research data?

Research data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

Ref: Borgman, Christine L. 2015. *Big Data, Little Data, No Data : Scholarship in the Networked World*. Cambridge, MA: MIT Press.

- Primary (new) or Secondary (existing)
- Qualitative or Quantitative
- Personal data
 - General (non-sensitive) personal data
 - Sensitive personal data

Classify your data

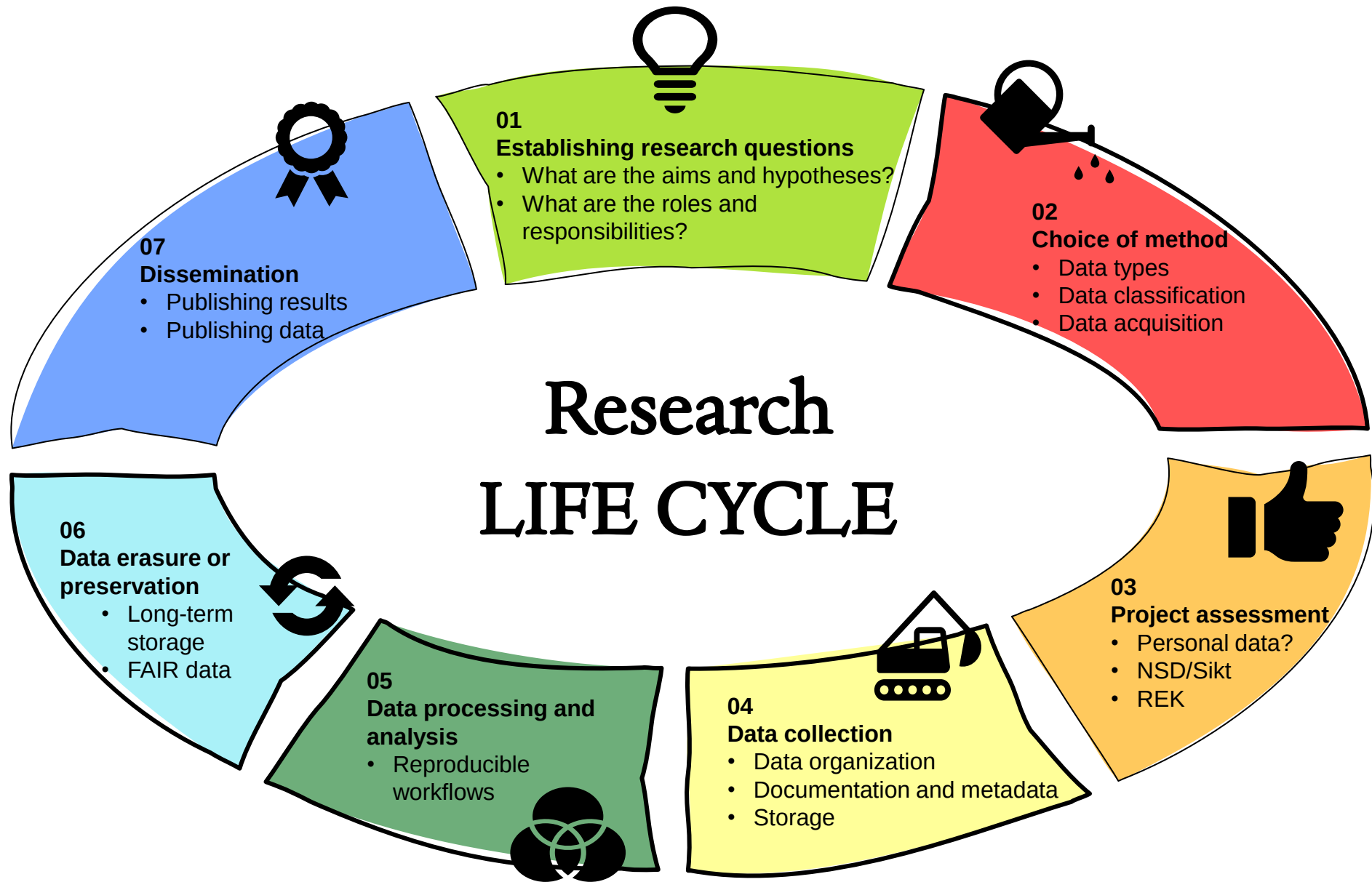


<https://www.uio.no/english/services/it/security/isis/data-classes.html>

2

Write down all **datasets** you will be collecting and working with in your research project

Specify **data types** and **classes** for each dataset



• Forside • Personverntjenester • Personverntjenester for forskning • Meldeskjema for personopplysninger i forskning

Notification Form for personal data

Learn about what personal data is, who should send in a notification form, and what you need to have ready in advance.

**Fill out the
notification form**



**About information
and consent**



**Notify changes to the
project**



REK

Regionale komiteer for medisinsk og helsefaglig forskningsetikk (REK)

Komiteene for klinisk utprøving av legemidler og medisinsk utstyr ([REK KULMU](#))

Aktuelle meldinger

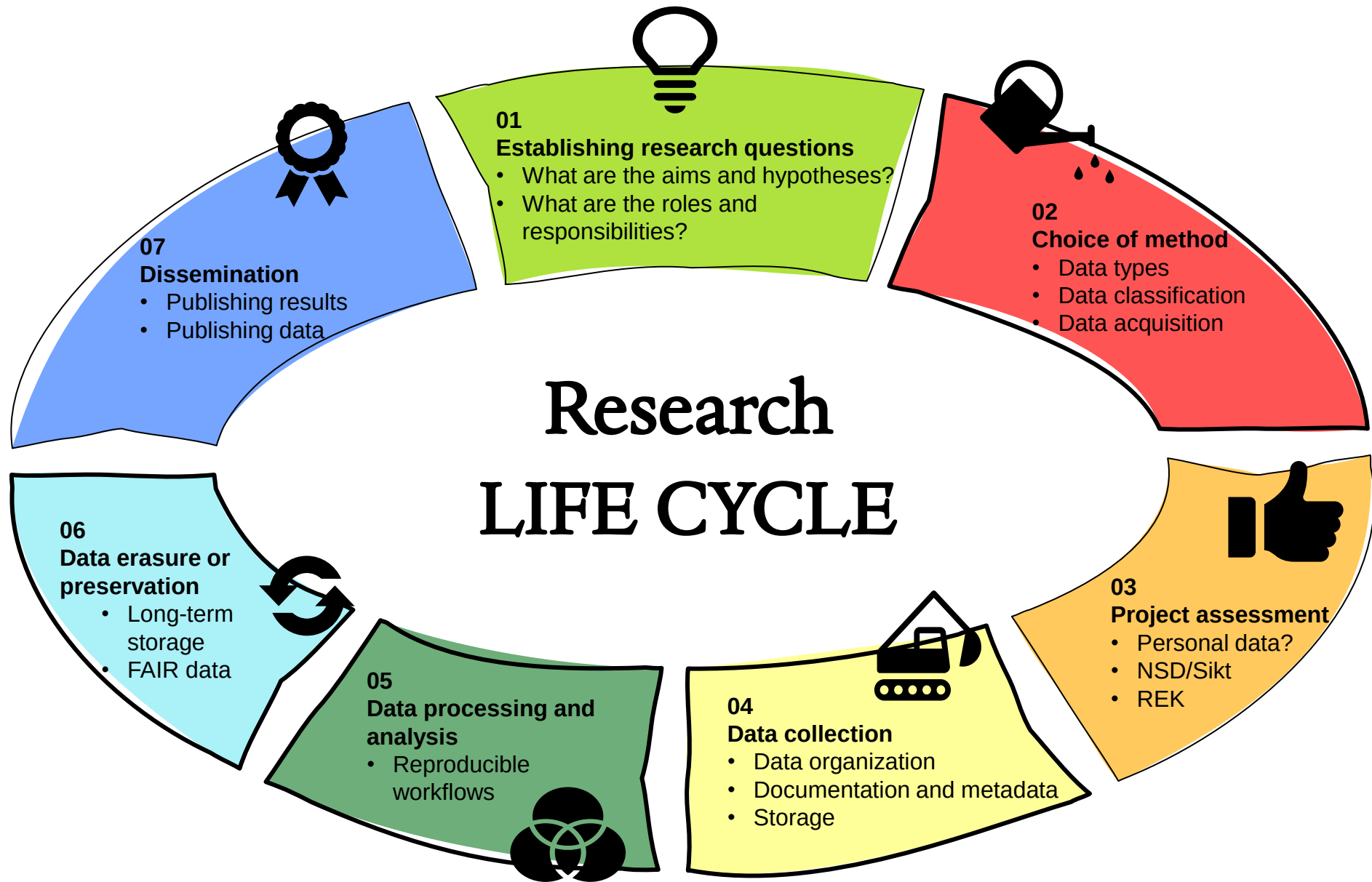
Varsel om endring i søknadsskjema:

REK har gjort en vurdering om å endre antall tegn i prosjektsøknadens punkt 2.1 Prosjektbeskrivelse fra 4000 til 1000 tegn.
Kontroller gjerne at dere er innenfor tekstbegrensningen før frist, da søknaden ikke kan sendes inn med flere enn 1000 tegn.

Se [REK KULMUs nettside](#) for informasjon om klinisk utprøving av legemidler, medisinsk utstyr og in-vitro diagnostisk medisinsk utstyr i henhold til EU-forordningene [536/2014 CTR](#), [2017/745 MDR](#) og [2017/746 IVDR](#).

3

Write a few keywords for the **ethical** and **legal** issues relating to your project



Structuring the folder system

1. Make a folder **hierarchy**
2. Give folders **descriptive** and **informative** names
3. Avoid folders that become too broad or general, create more **subfolders** instead
4. Keep **active** and **finished** parts of your project separate
5. Set aside some time to clean up regularly



How it could look:

```
project_name/
├── README.md           # overview of the project
├── data/               # data files used in the project
│   ├── README.md      # describes where data came from
│   └── sub-folder/     # may contain subdirectories
├── processed_data/     # intermediate files from the analysis
├── manuscript/         # manuscript describing the results
├── results/            # results of the analysis (data, tables, figures)
├── src/                # contains all code in the project
│   ├── LICENSE         # license for your code
│   ├── requirements.txt # software requirements and dependencies
│   └── ...
└── doc/                # documentation for your project
    ├── index.rst
    └── ...
```

File naming conventions (FNC)

- Short names (but long enough that they still make sense)
- The most general information first, then add details to the name
 - Underscore to separate words, DO NOT use space in file names!
- Dates backwards (YYYYMMDD)
- Numbers (e.g. version number) should have the same number of digits, use e.g. 01, not just 1.
- Version number at the end

20220503_DOEProject_DesignDocument_Smith_v2-01.docx

4

Sketch down some initial thoughts to how you would **structure** the folders and **name** the files

Documentation



Project-level documentation

- aim of the study
- research questions/hypotheses
- methodologies/instruments /measures

Data-level or Object-level documentation

- variables in a database/table or transcripts
- images

README.txt-files

GENERAL INFORMATION

SHARING/ACCESS INFORMATION

DATA & FILE OVERVIEW

METHODOLOGICAL INFORMATION

DATA-SPECIFIC INFORMATION



README.txt-files

Document everything your data has been through

- Field journal
- Lab journals and experimental protocols
- Scripts for analysis
- Questionnaires, codebooks, data dictionaries
- Software syntax and output files
- Methodology reports
- Geolocation, orientation (e.g. when collecting a sample)
- Instrument settings and calibration



5

For each dataset, create a list of what **documentation** would be needed to make another researcher understand what's in it

Storing data during the project

UiO storage guide:

<https://www.uio.no/english/services/it/security/isis/storage-guide.html>



Where can you store different types of data?

Green data

- anywhere you want, but ideally **not only** on your private laptop (very vulnerable!)

Where can you store different types of data?

Yellow data

- Encrypted memory stick / external hard drive
- Encrypted laptop maintained by UiO
- UiO cloud solutions ([UiO OneDrive](#), [UiO G-Suite/Google drive](#))
- UiO Storage Hotel («lagringshotell»)
- UiO e-mail; UiO Teams; UiO Zoom; UiO Nettskjema, UiO Canvas, UiO Vortex
- **Not** your private laptop, unless you comply with [the guidelines for use of private computer](#)

Where can you store different types of data?

Red data

- Fully encrypted disc, memory stick or external hard drive
- EduCloud
- Services for sensitive data – TSD
- UiO storage hotel («lagringshotell»)*
- UiO Vortex*
- UiO Nettskjema*
- UiO Zoom (but see [routines!](#))

*see Storage guide for details!

<https://www.uio.no/english/services/it/security/isis/storage-guide.html>

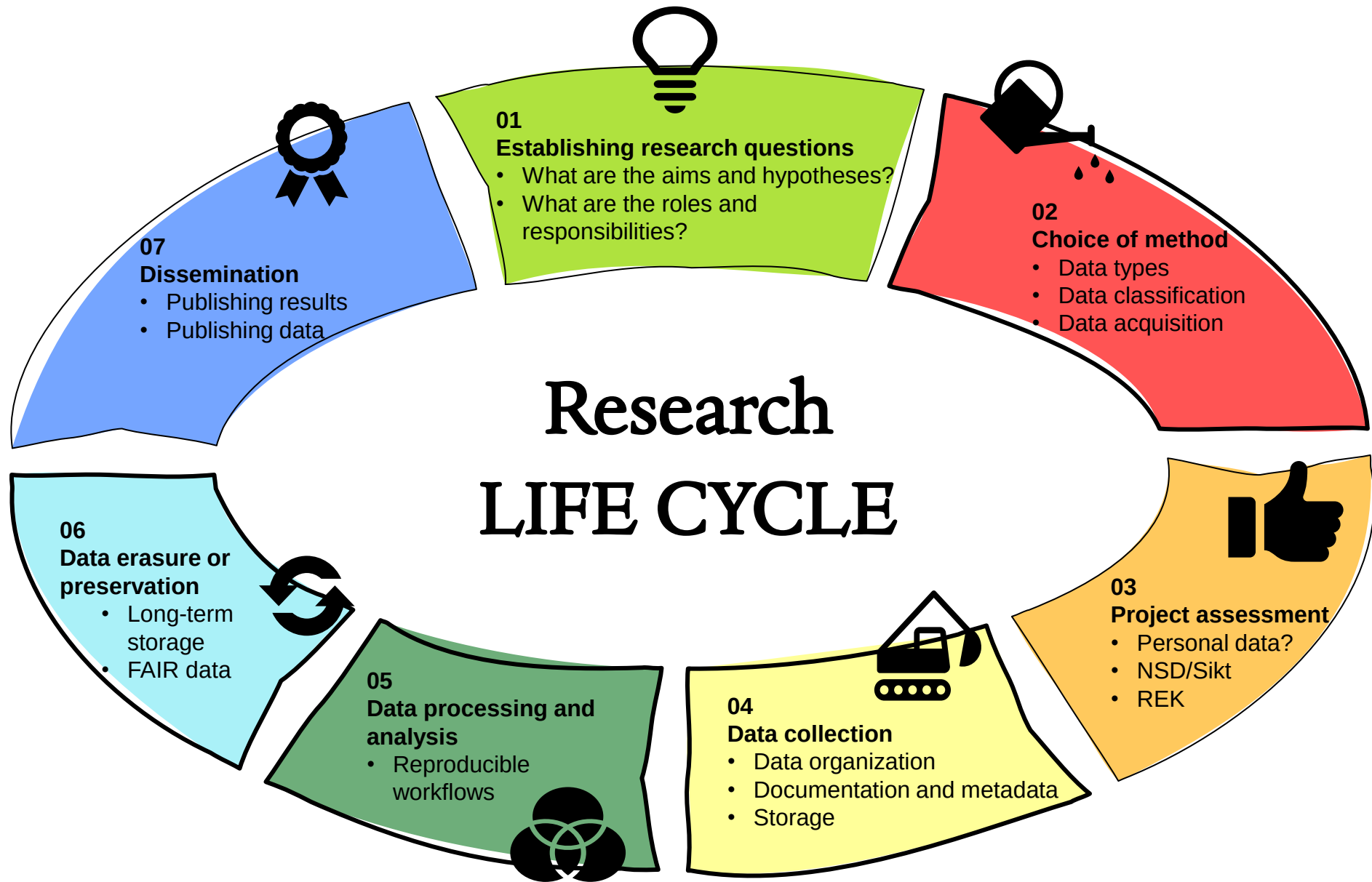
Where can you store different types of data?

Black data

- Services for sensitive data – TSD

6

For each dataset, write down where you will **store** it during the project



F_{indable}



A_{ccessible}



I_{nteroperable}



R_{eusable}



Research Data Policy

Research data at the University of Oslo shall:

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

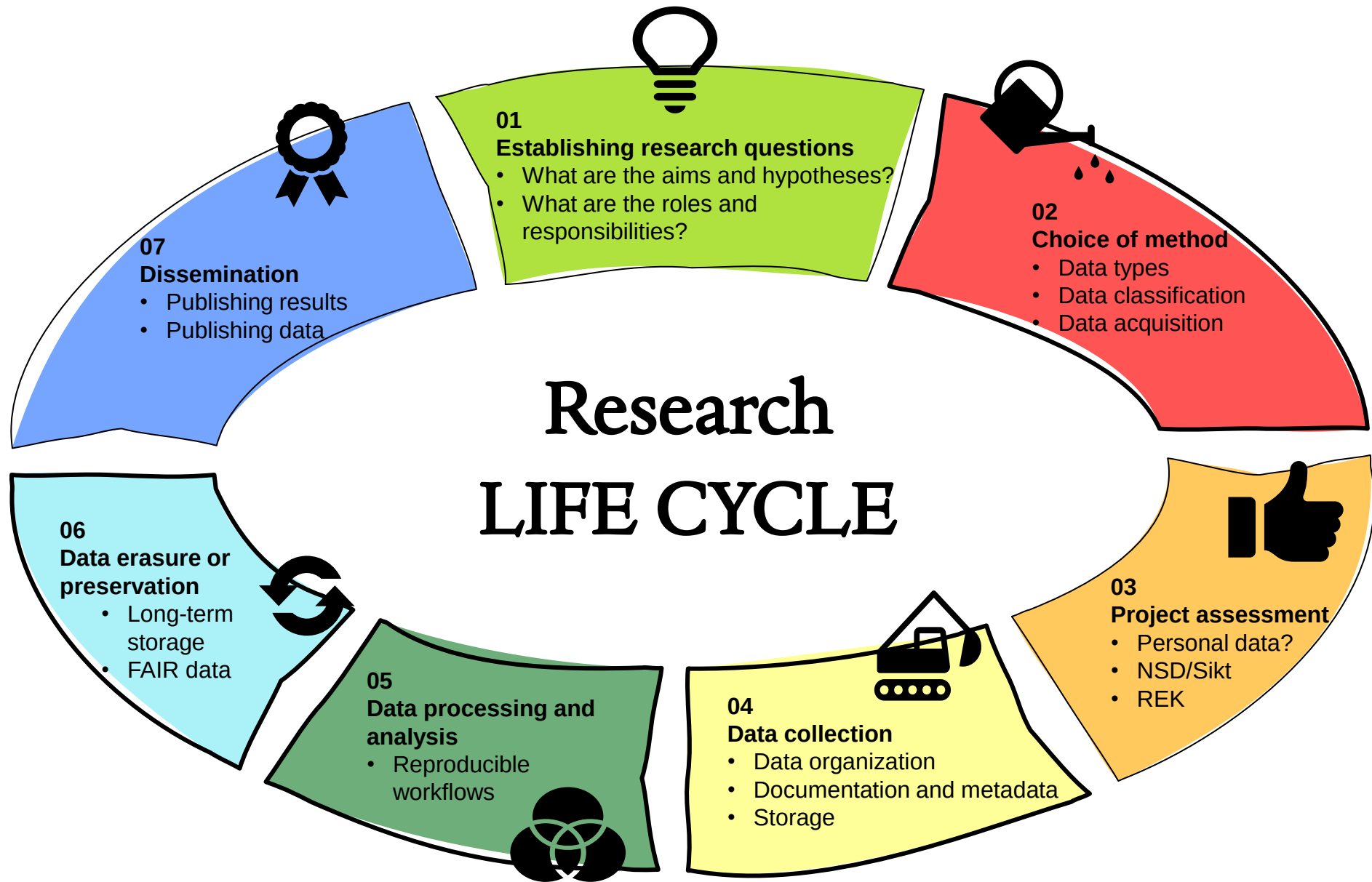


Source: <https://www.uio.no/english/for-employees/support/research/research-data-management/policies-and-guidelines/>

As open as possible,
as closed as necessary

7

For each dataset, write down if it (or its parts) will be **shared** and/or **preserved** long-term



8

Are there any additional **costs** or **resources** associated with the project?

Courses in research data management and sharing

April 12th – 27rd 2023

Data classification and storage selection

Apr. 21, 9:00 AM, Zoom

Copyright and licensing

Apr. 21, 1:00 PM, Zoom

Sharing and archiving research data

Apr. 26, 9:00 AM, Zoom

Finding and reusing research data

Apr. 27, 9:00 AM, Zoom

For employees

☰ All content

◀ [Research support](#)

Research data management

[Norwegian](#)

Welcome to UiO's data management pages maintained by the research data group at [Digital Scholarship Center](#).

Data management plans (DMPs) >

Data classification and storage >

Data organization >

Data documentation and metadata >

Data sharing and publishing >

Finding and reusing data >

[Norwegian
version of this
page](#)

Digital Scholarship Centre

At the Digital Scholarship Centre (DSC) you get guidance on how you can make the best possible use of digital tools and methods in your research and communication activities.

Open Access



Information about open access publishing, publisher agreements, self-archiving, requirements, and guidelines.

Open and reproducible research



Make your research more transparent and reproducible.

Research Data Management



Managing your data both during and after a research project.

Text-mining



Information about digital tools for searching, mining, and analysing textual data.

Systematic search



Information about systematic literature searching, how to get started, and how to get help.

Visualisation



Use of visual methods to explore, communicate and understand data.

Carpentry@UiO



Offers workshops in foundational digital skills such as coding and data management.

Reference management



Styles, tools, and information on reference management.

[Norwegian
version of this
page](#)

Digital Scholarship Centre

At the Digital Scholarship Centre (DSC) you get guidance on how you can make the best possible use of digital tools and methods in your research and communication activities.

Åpen og reproducerbar forskning | Visualisering | Carpentry@UiO | Åpen tilgang | Tekstutvinning

Senter for Digital Forskerstøtte

Digital Scholarship Centre

DSC NEWS

Desember 2022

Digiforskbloggen | Forskningsdatahåndtering | Systematisk litteratursøk | Referansehåndtering

<https://sympa.uio.no/ub.uio.no/subscribe/dsc-news/subscribe>

Thank you!

Elisa Pierfederici

Digital Scholarship Center

University of Oslo Library

13.03.2023,

Contact us at research-data@uio.no

<https://nettskjema.no/a/337151>



Supplementary slides



University of Oslo Library

Digital Scholarship Center and Open Research

Open Science Lunch 2023

Each last Thursday of the month at 12.00 we invite you to join us virtually for an online open lunch to hear about how to make your research more open.

February 23rd

Open your data with DataverseNO

March 30th

Public health reporting: an open source approach

April 27th

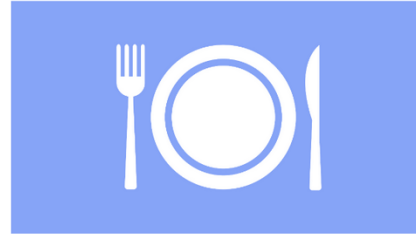
Open knowledge resources: Store norske leksikon and Wikipedia

May 25th

Keep the rights to your work: UiO's Rights Retention Policy



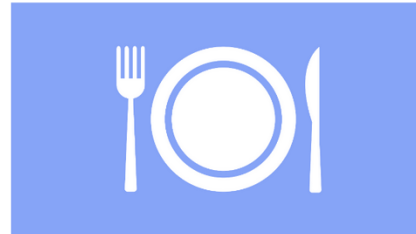
Upcoming



Time and place: Mar. 30, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Public health reporting: an open source approach

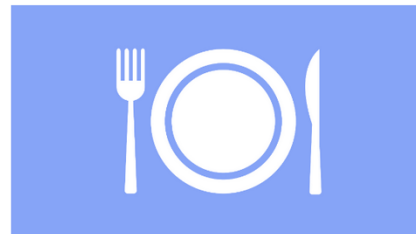
Join us for a discussion on how open source tools can enable automated, reproducible and scalable public health reporting.



Time and place: Apr. 27, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Open knowledge resources: Store norske leksikon and Wikipedia

Join us for a discussion on using, maintaining and contributing to freely available open knowledge resources.



Time and place: May 25, 2023 12:00 PM–1:00 PM, Hybrid: Georg Sverdrups hus and Zoom

Keep the rights to your work: UiO's Rights Retention Policy

Learn about how you can retain the rights to your published work with the new Rights Retention Policy at UiO.



Carpentry@UiO

Carpentry@UiO is a community of people who are passionate about learning, teaching, and sharing best practices and digital skills for making the research process more reproducible and effective. If you want to get involved, or join one of our workshops, check us out!



The Unix Shell

Shell speeds up repetitive and tedious processes. It is also essential skills needed to use high-performance computing (HPC) resources.



Version Control with Git

Git helps you to keep track of what you've done, for a better collaboration and for yourself in future. In the workshop we use GitHub as well.



Programming in Python

Python is now widely used in scientific computing with various powerful packages. Carpentry@UiO runs workshops for participants with no programming experience ("Plotting and Programming in Python" lesson) and for participants at intermediate level ("Programming with Python" lesson, episodes 10-12).



R for Reproducible Scientific Analysis

R is commonly used for statistical analysis, but it is also a powerful programming language. Workshops on R focuses on teaching best practices for scientific computing: breaking down analyses into modular units, task automation, and encapsulation. Workshops on R may use lessons from Data Carpentry instead.



Using Databases and SQL

Databases include powerful tools for search and analysis, and can handle large, complex data sets. The lesson will show how to use a database to explore research data by using SQL.



Carpentry@UiO

Carpentry@UiO is a community of people who are passionate about learning, teaching, and sharing best practices and digital skills for making the research process more reproducible and effective. If you want to get involved, or join one of our workshops, check us out!



Learn, teach, and share digital skills and best practices

Be a part of an interdisciplinary community

Make use of and contribute to community-built teaching materials

ReproducibiliTea

Journal Club

**JOIN IN AND DISCUSS WITH FELLOW
STUDENTS AND RESEARCHERS**

**OPEN RESEARCH, REPRODUCIBILITY
and RESEARCH IMPROVEMENT**



Join us

Everyone is welcome to join us - whether you are an enthusiast of open and reproducible research, a skeptic, or a cautious explorer. Currently, all meetings are hybrid with the possibility of joining on-site at Blindern or via Zoom. Grab a cup of tea (coffee?) and join us!

Subscribe to our mailing list



How to create a documentation?

- DataverseNO guide

<https://site.uit.no/dataverseno/deposit/prepare/>

- DataverseNO template

<https://zenodo.org/record/7454000#.Y8PzLC8w1QI>

- Examples of README files

<https://dataverse.no/dataset.xhtml?persistentId=doi:10.18710/LNCK8O>

<https://dataverse.no/dataset.xhtml?persistentId=doi:10.18710/Q3FZAN>

<https://dataverse.no/dataset.xhtml?persistentId=doi:10.18710/TYPJXU>

Relevant links at UiO web pages

UiO research data policy:

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

Data management guidance:

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

Courses in data management:

<https://www.ub.uio.no/english/libraries/dsc/training/>

UiO data storage guide:

<https://www.uio.no/english/services/it/security/isis/storage-guide.html>

Data archiving:

<https://www.ub.uio.no/english/writing-publishing/data-archiving/index.html>