

Data Management Planning for qualitative research

QualiFAIR DMP Work Package

Agata Bochynska, Matthew Good, Torgeir Christiansen, Anne Bergsaker

Digital Scholarship Center, University of Oslo Library

01.11.2023,

Questions? Contact QualiFAIR coordinator at agata.bochynska@ub.uio.no

The Data Management Plan (DMP)

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



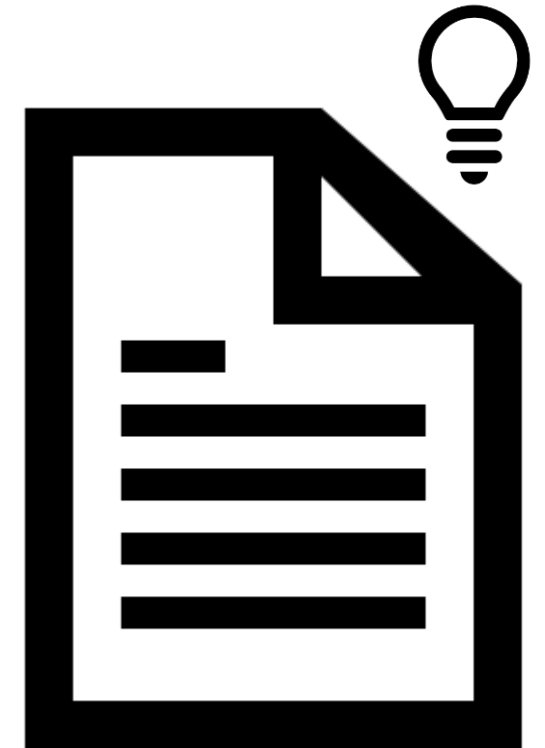
The Data Management Plan (DMP)

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
- [DMPOne](#) 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

Let's see some examples!

[Trinity College Botanic Garden long-term monitoring program](#) (DCC template)

[Realist Evaluation of Adapted Sex Offender Treatment Programs for Men with Intellectual Disability](#)

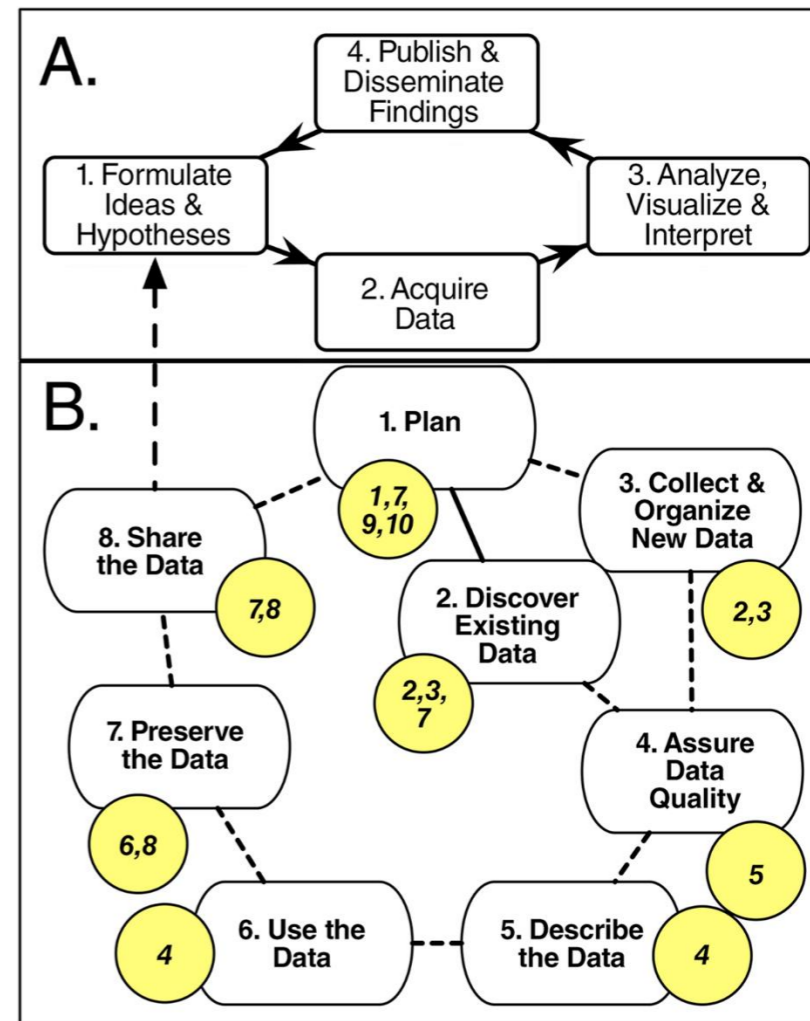
[Bridges in social networks: Harnessing dual identity to improve interethnic relations](#) (ERC Consolidator grant)

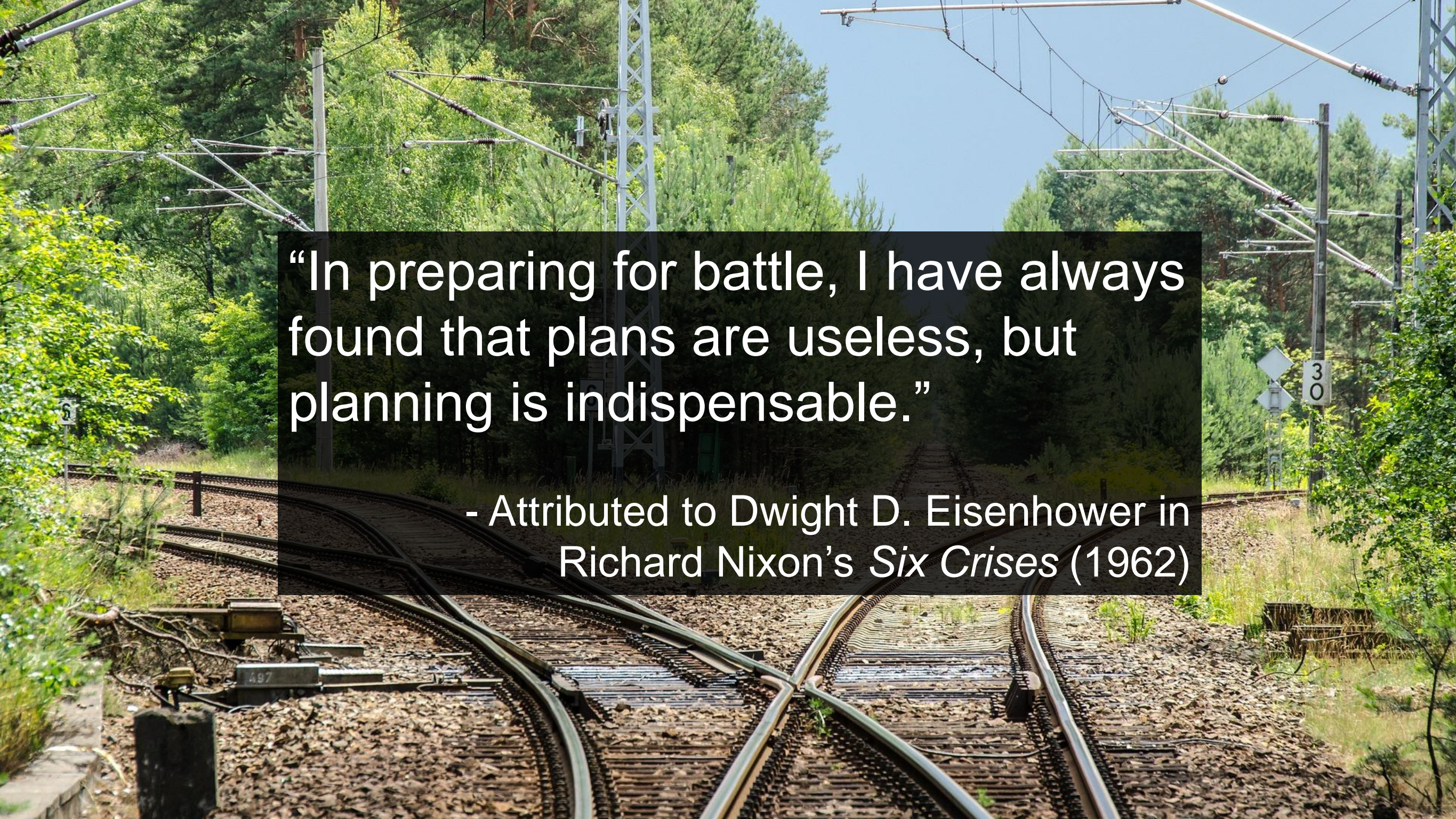
Find more examples in Public DMPs service from DMPonline:

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

10 Simple Rules to Writing a Data Management Plan

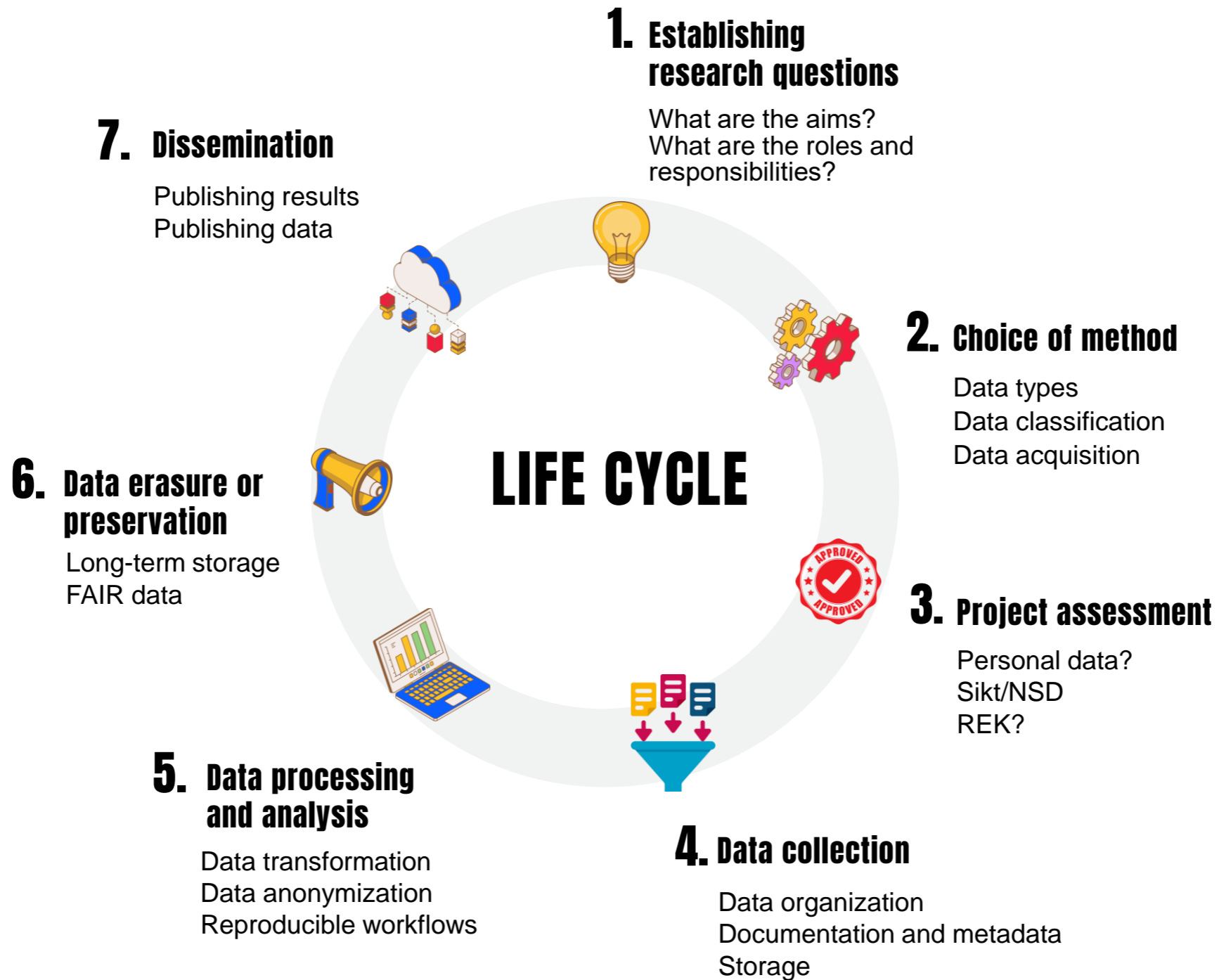
1. Determine research sponsor **requirements**
2. Identify **data** to be collected
3. Define how data will be **organized**
4. Explain how data will be **documented**
5. Describe how **quality** will be maintained
6. Develop a strong **storage** and **preservation** strategy
7. Define project's data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





“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)



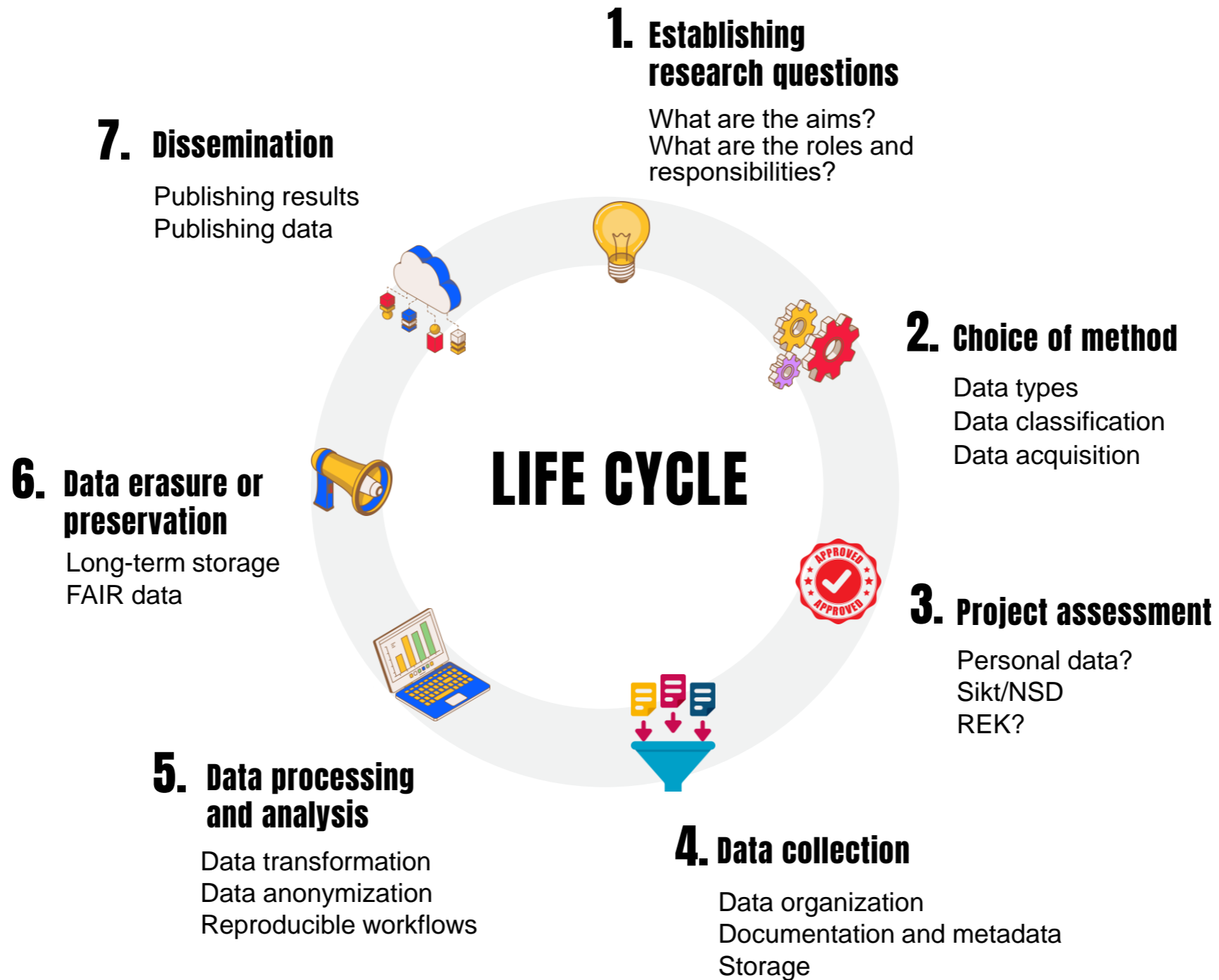
Assigning roles and responsibilities

- 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 make sure the **data management plan (DMP)** is being followed?
- Who will be responsible for **updating the DMP** and how often?



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 **qualitative** research data?

*Qualitative data refers to information that is **not gathered in numerical form**, but rather describes qualities or characteristics. Data are expressed in natural language and often in **textual or visual form**. Collection methods for instance include photography, audio recordings, video, interviews, diary accounts or other unstructured observations.*

Ref: CESSDA, 2022. <https://zenodo.org/record/7777519> .

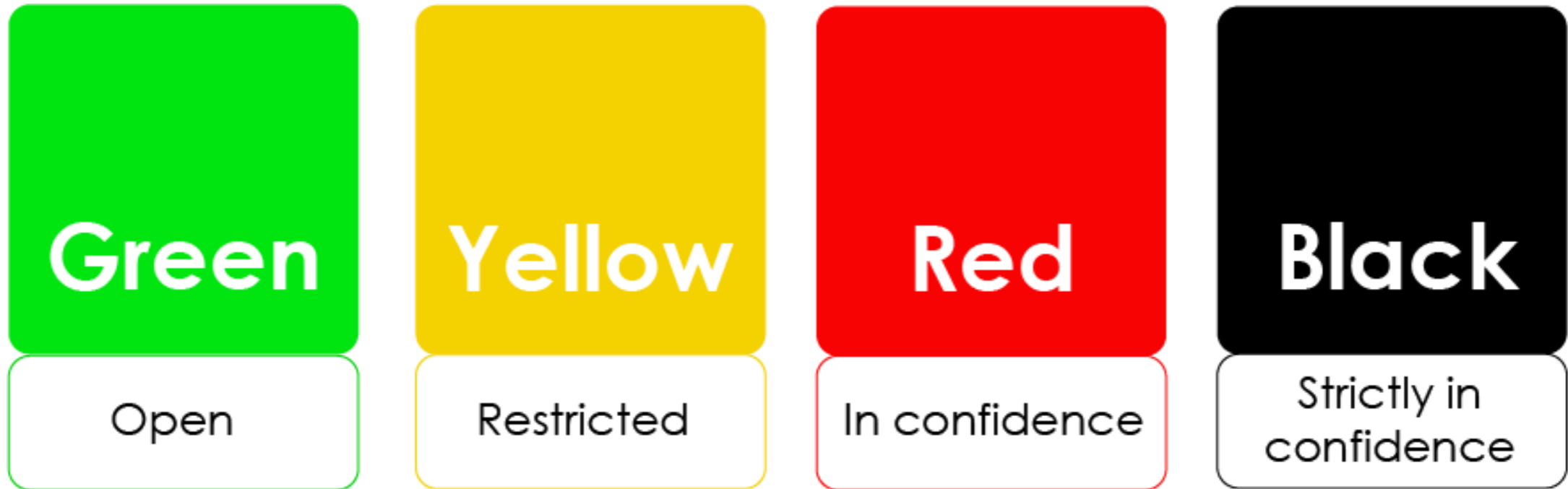
Research data types

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)
- Observational, Experimental, Computational, Records
- Text, Audio, Video, Image
- Personal data
 - General (non-sensitive) personal data, e.g. name, address, voice
 - Sensitive personal data, e.g. health information, ethnicity, religion

Data classification at UiO

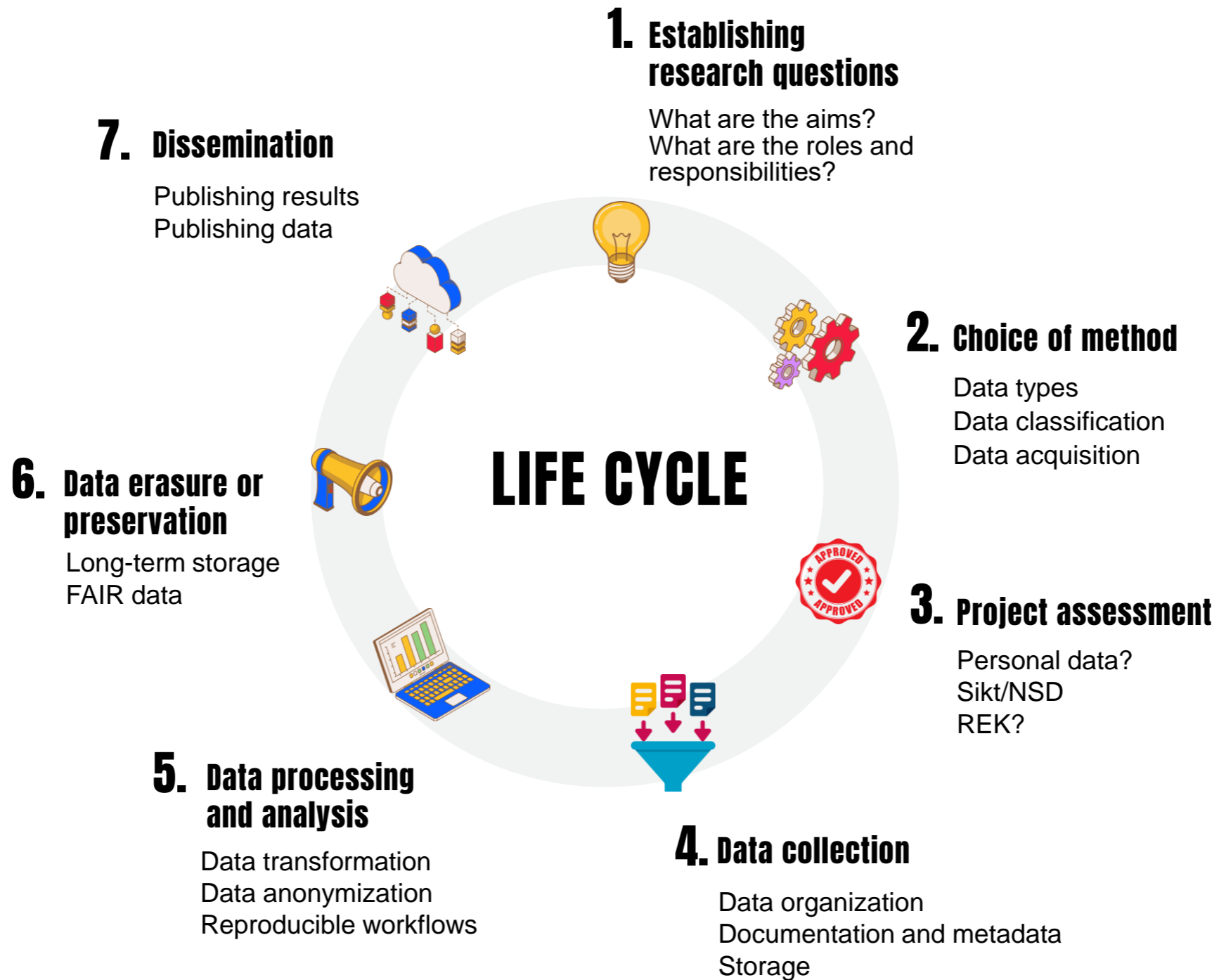


<https://www.uio.no/english/services/it/security/isis/data-classes.html>

2

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

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

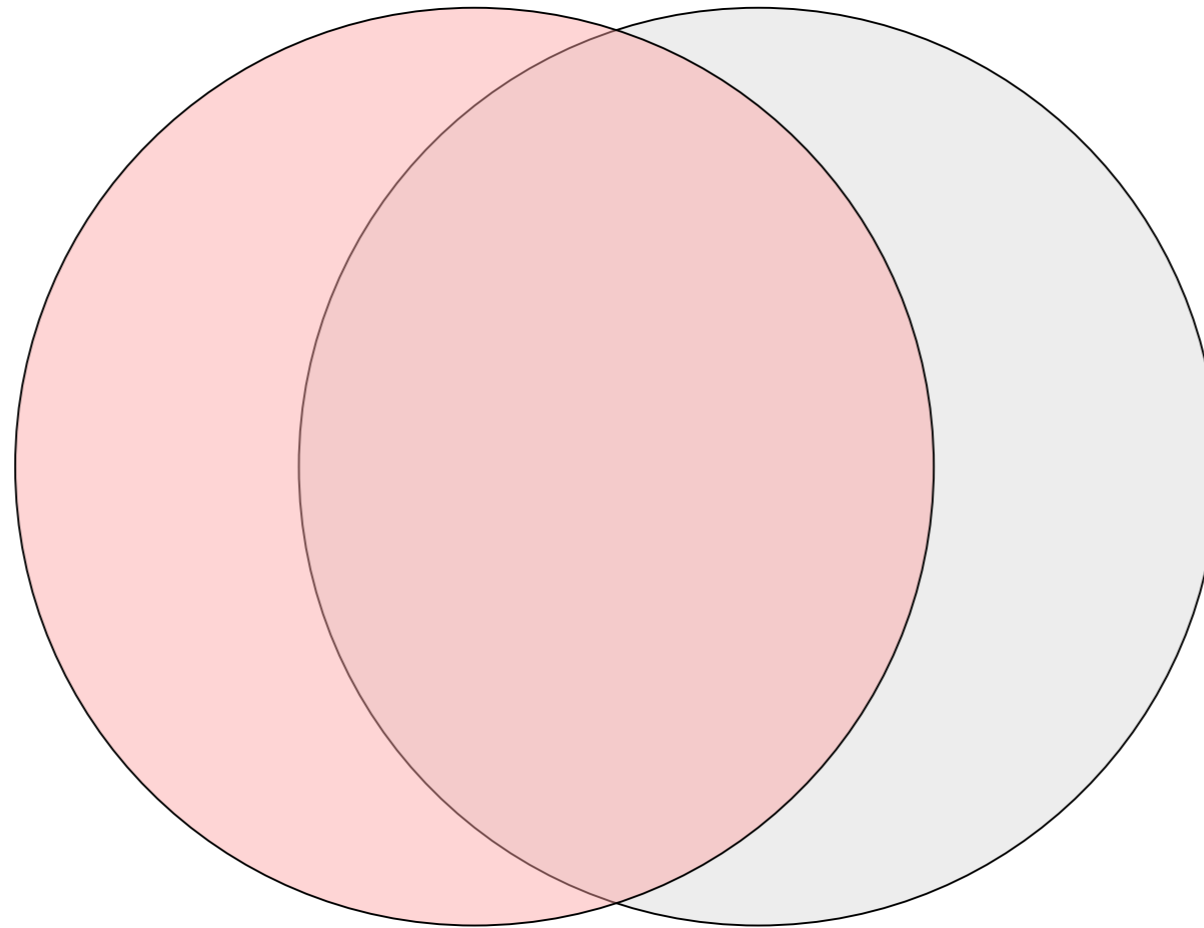


Law

Privacy
(GDPR)

Copyright

Export control



Ethics

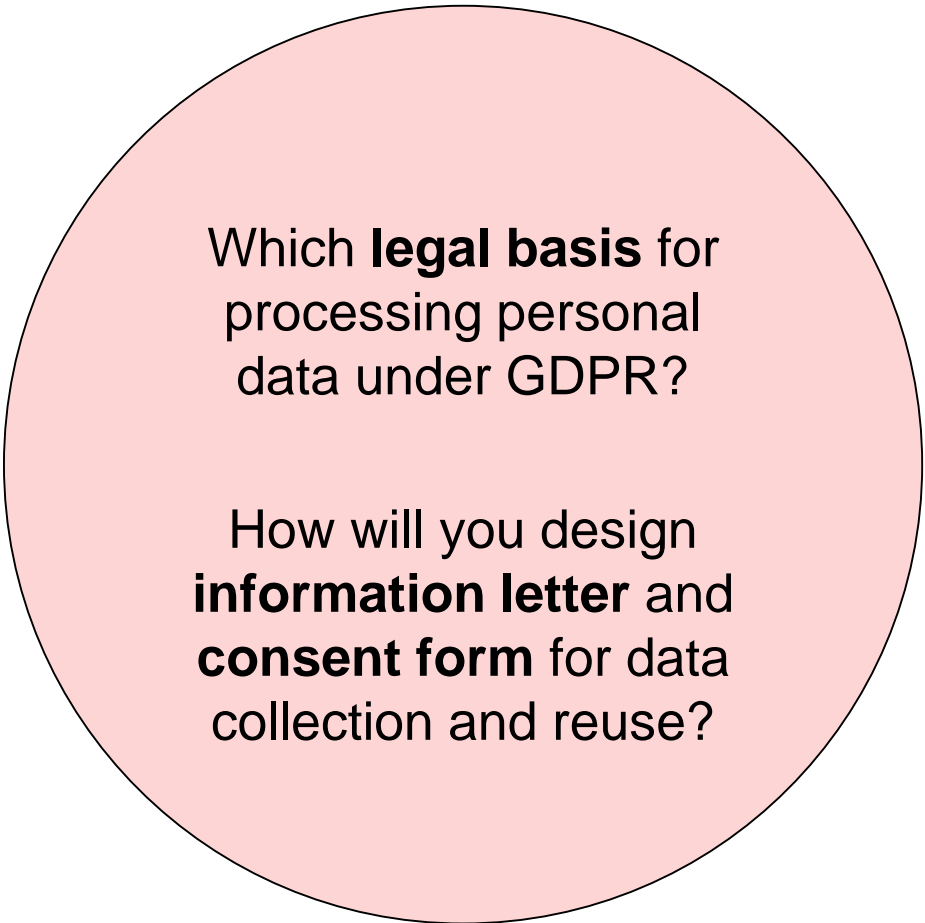
Guidelines
and norms
in your field

Law

Privacy
(GDPR)

Copyright

Export control



Which **legal basis** for processing personal data under GDPR?

How will you design **information letter** and **consent form** for data collection and reuse?

Template for information letter

The privacy regulations contain clear requirements for what the information

Because there are different types of legal bases for processing, we have prepared templates that you can use as a starting point for some of the most common bases in research.

Template for information letter - gathering consent (word-format)

[DOCX, 33kB]



Template for information letter - public interest (word-format)

[DOCX, 25kB]



Department of Teacher Education and School Research

← About the department ← Organization ← Teaching Learning Videolab (TLVlab)

Consent form archive

Consent form archive

Welcome to the consent form archive, a collection of consent forms from different disciplines at the University of Oslo. In the consent form archive you may search by topics related to your own project to find relevant consent forms that may help you to create your own consent form.

- Consent Form Archive
- Consent form archive upload portal
- Feedback
- Guidance

Law

**Privacy
(GDPR)**

Copyright

Export control

**Is personal data
processing in the
project safe?**

Personal data
processing **assessment**
by **Sikt**: Norwegian
Agency for Shared
Services in Education
and Research

Data protection officer

Norwegian

What is a data protection officer?

The data protection officer's job is to contribute to UiO safeguarding the privacy interests of the employees, students, guest researchers, guests and respondents or informants in research projects.

The data protection officer is the point of contact for individuals who have questions about the UiO processing of personal information and about how they can fulfill their rights under the privacy policy.



→ [Read more about the data protection officer](#)

Useful links

- [Privacy at UiO](#)
- [The Speak Up system](#)
- [The Ombud for Students](#)
- [Your rights](#)
- [Privacy declaration for employees at UiO](#)
- [Privacy declaration for applicants, students, alumni and PhD candidates at UiO](#)

Contact

The data protection officer at UiO is [Roger Markgraf-Bye](#). The data protection officer can be reached via e-mail: personvernombud@uio.no

<https://www.uio.no/english/for-employees/support/privacy-dataprotection/data-protection-officer/>

• [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**



Which personal data will be processed?

What are personal data?

What is processing?

General categories of personal data

- Name (also with signature/written consent) ?
- National ID number or other personal identification number ?
- Date of birth
- Address or telephone number
- Email address, IP address or other online identifier ?
- Photographs or video recordings of people ?
- Sound recordings of people ?
- GPS data or other geolocation data (electronic communications) ?
- Background data that can identify a person ?
- Other data that can identify a person ?

Special categories of personal data

- Racial or ethnic origin ?
- Political opinions ?
- Religious beliefs ?
- Philosophical beliefs ?
- Trade Union Membership ?
- Health data ?
- Genetic data ?
- Biometric data ?
- Sex life or sexual orientation ?
- Criminal convictions and offences ?

Law

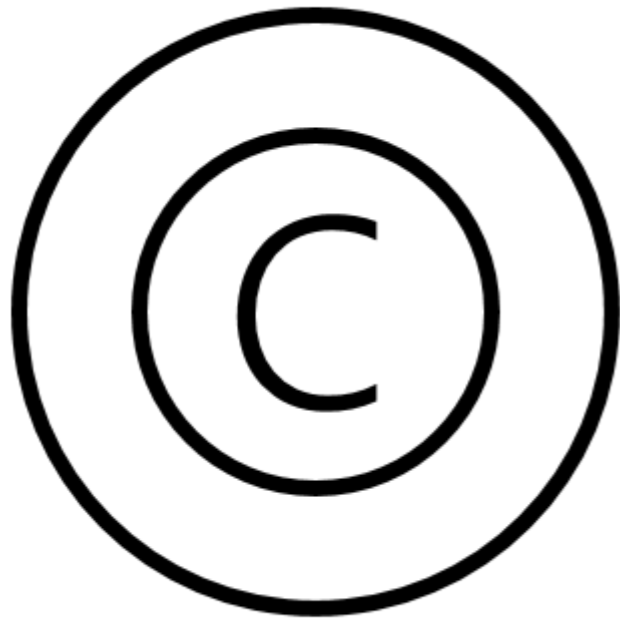
Privacy
(GDPR)

Copyright

Export control



Are there any
copyright
restrictions?



Can limit use and reuse of data

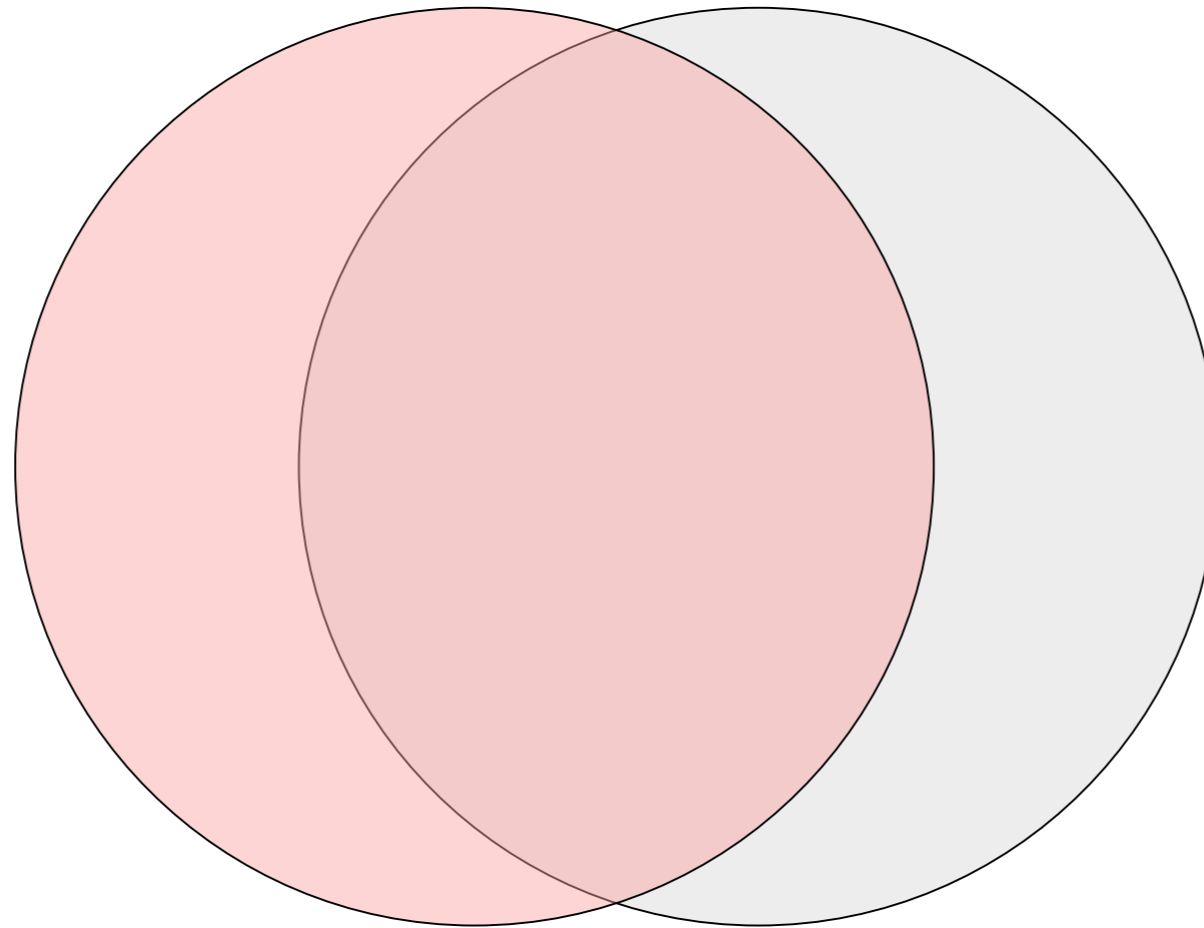
Clarify restrictions early

Law

Privacy
(GDPR)


Copyright

Export control



Ethics

Guidelines
and norms
in your field



Are any ethical
approvals needed?

Medical or health-
related research?

Any particular ethical
considerations?

Ethics
Guidelines
and norms
in your field

RETNINGSLINJER

Guidelines

Medical and health research

Science and technology

Social Sciences and Humanities

Human remains

General guidelines

Guidelines for Research Ethics in the Social Sciences and the Humanities

Published: 5/26/2022

Given by The National Committee for Research Ethics in the Social Sciences and the Humanities (NESH) in 2021 (5th edition). English translation published 2022.

Content

Preface	
Introduction	▼
A) The Research Community	▼
B) Research Participants	▼
C) Groups and Institutions	▼
D) Commisioners, Funders, and Collaborators	▼
E) Dissemination of Research	▼
Attachment	▼
Notes	

A Guide to Internet Research Ethics

Published: 8/6/2019

Issued by the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH) in 2003. Second edition published in Norwegian in 2018 and in English May 2019.

Content

1	Foreword	
2	Research Ethics	
3	Delimitations	
4	Ethical Considerations	▼
5	References	
6	Notes	

Foreword

This guide has been prepared by the National Committee for Research Ethics in the Social Sciences and the Humanities (NESH). The work was initiated by the previous committee (2014–2017) and completed by the current committee (2018–2021, listed below).[1]



<https://www.forskningsetikk.no/en/about-us/our-committees-and-commission/nesh/guidelines-nesh/a-guide-to-internet-research-ethics/>

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](#).

The Science Ombud at UiO



The Science Ombud is an independent and impartial authority for all academic staff at the University of Oslo, for issues and disputes concerning good scientific practice, research integrity and research ethics.

How can we help you?

- Provide you with insight into research ethics, norms, and regulations
- Provide guidance and confidential advice on issues of scientific integrity
- Help resolve cases early and at the lowest possible level

Contact

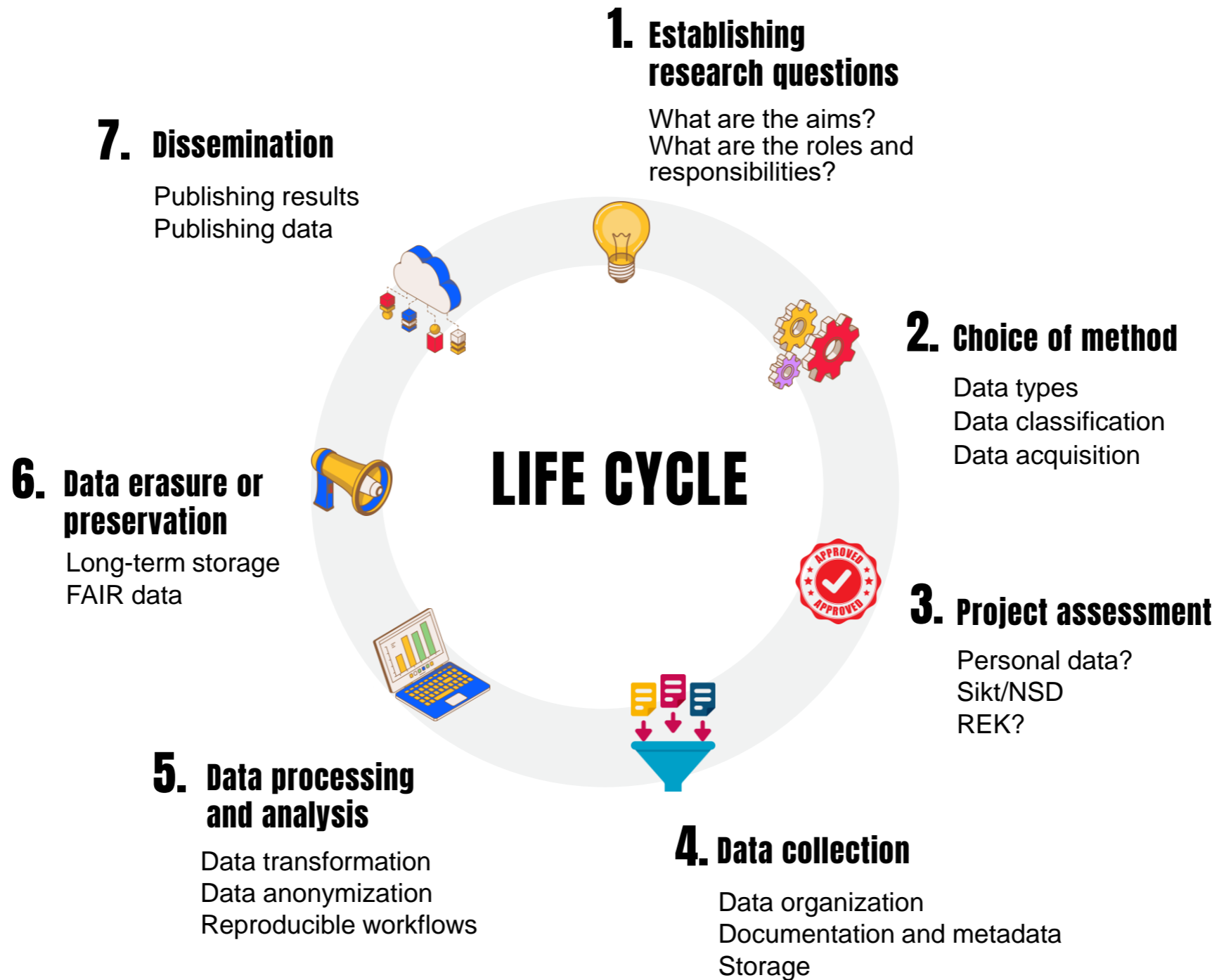
- E-mail: vitenskapsombudet@uio.no
- Tel: 22857034

The Office of the Science Ombud is located in the Lucy Smith building – 6th floor

3

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

For example, is there a need for project assessment (e.g. by Sikt or REK)? Are there privacy or copyright issues? Any issues relating to gathering consent? Any other ethical or legal issues?



Structuring the folder system

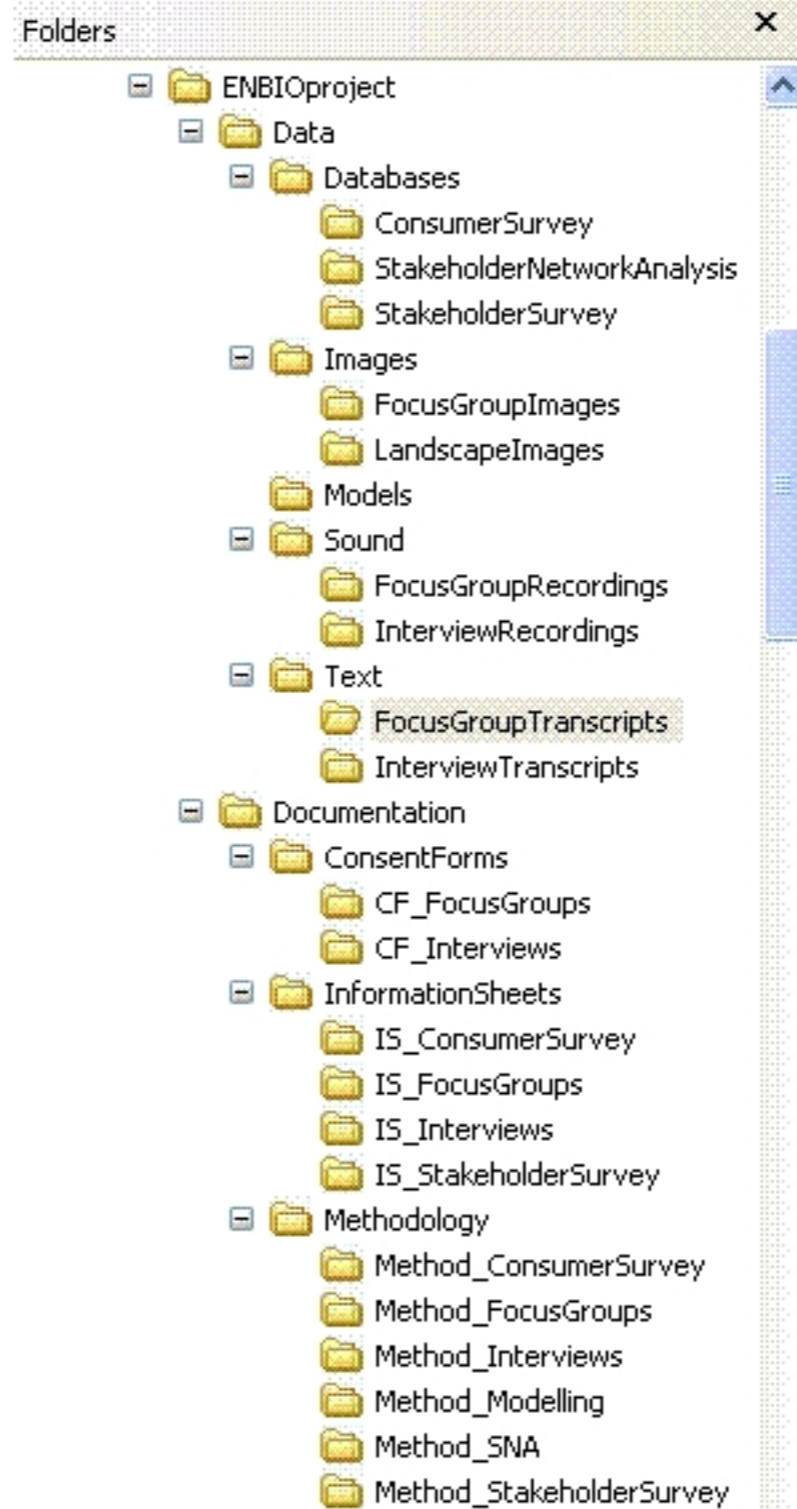
1. Make a folder **hierarchy**
2. Give folders **descriptive** and **informative** names
3. Avoid folders that become too broad or general, create more **subfolders** instead
4. Keep **active** and **finished** parts of your project separate
5. Set aside some time to clean up regularly



Name	Date modified	Type	Size
P-10-01	26.10.2020 16:29	File folder	
SEM 2012	26.10.2020 16:25	File folder	
P-10-03-01.tif	11.06.2012 11:17	TIF File	1 880 KB
P-10-03-02b.tif	11.06.2012 11:25	TIF File	1 880 KB
P-10-03-02c.tif	11.06.2012 11:26	TIF File	1 880 KB
P-10-03-05.tif	11.06.2012 11:38	TIF File	1 880 KB
P-10-03-79b.tif	11.06.2012 12:59	TIF File	1 880 KB
P-10-03-92b.tif	11.06.2012 13:26	TIF File	1 880 KB
P-10-03-109.tif	11.06.2012 13:49	TIF File	1 880 KB
P-10-03-119b.tif	11.06.2012 14:02	TIF File	1 880 KB
P-10-03-193b.tif	11.06.2012 15:45	TIF File	1 880 KB
P-10-03-222b.tif	11.06.2012 16:29	TIF File	1 880 KB
P-10-03-222c.tif	11.06.2012 16:30	TIF File	1 880 KB
P-10-03-226b.tif	12.06.2012 09:33	TIF File	1 880 KB
P-10-03-226c.tif	12.06.2012 09:33	TIF File	1 880 KB
Gasser_2014.pdf	27.01.2015 13:11	Adobe Acrobat D...	5 127 KB
Gernigon_and_Bronner_2012.pdf	30.01.2015 15:07	Adobe Acrobat D...	3 445 KB
Gernigon_et_al_2014.pdf	26.01.2015 16:02	Adobe Acrobat D...	54 447 KB
Glorstad_Clark-2010.pdf	15.09.2014 15:06	Adobe Acrobat D...	10 269 KB
Gronlie et al. 1980 Seismic inversion of Bj...	09.10.2012 13:28	Adobe Acrobat D...	643 KB
Gudlaugsson_et_al_1998.pdf	17.10.2014 12:50	Adobe Acrobat D...	3 411 KB
Harland_and_Gayer_1972.pdf	17.10.2014 11:47	Adobe Acrobat D...	1 474 KB
Høy_and_Lundchien_2011_NBarenstSea_...	17.09.2014 14:10	Adobe Acrobat D...	13 600 KB
Isaksen 1996 Organic geocem Bjornoya.p...	09.10.2012 13:24	Adobe Acrobat D...	1 439 KB
Klausen et al 2015 Triassic Snadd in Baren...	30.03.2015 13:15	Adobe Acrobat D...	14 911 KB
Klausen_et_al_2014_Triassic_Snadd_Fm.pdf	08.09.2014 13:20	Adobe Acrobat D...	5 069 KB
Klitzke_et_al_Barents_Sea_Region_2014-pr...	21.01.2015 15:22	Adobe Acrobat D...	2 951 KB
Gabrielsen et al. 1990_Structural_E_BS_NP...	11.03.2013 15:45	Adobe Acrobat D...	9 099 KB
Gac_et_al_2013_ultra_deep_EBB.pdf	20.01.2015 12:15	Adobe Acrobat D...	1 008 KB
Abstract.doc	15.09.2014 16:04	Microsoft Word 9...	234 KB
Copy of NuAge_090214-120115-run1.xls	22.01.2015 13:06	Microsoft Excel 97...	1 390 KB
CPP_Svalbard_Bjornoya_Copy of CUMUL...	17.12.2014 15:35	Microsoft Excel 97...	2 429 KB
2013 Wintermeeting poster-1_Edina_Final...	07.01.2013 15:12	Adobe Acrobat D...	130 KB
ICPMS data	26.10.2020 16:31	File folder	
Franz Josef presentation.pptx	09.01.2013 12:06	Microsoft PowerP...	11 334 KB

Example of a project
without the folder structure

How it could look:



<https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/organising/>

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

Let's see some examples!

[README file for: The moderating effect of gender equality on education](#)

[README file for: Cartobike transcripts focus group interviews](#)

Document everything data has been through

- Documentation of the collection and coding process
- Field journals
- Interview protocols, surveys and questionnaires
- Codebooks, data dictionaries
- Methodology reports
- Geolocation, orientation (e.g. when collecting a sample)
- Instrument and software settings and calibration



Based on:

https://mantra.edina.ac.uk/documentation_metadata_citation/

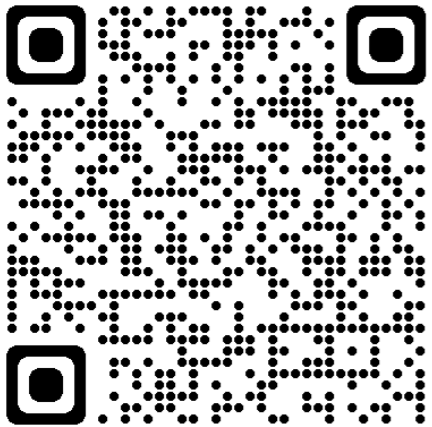
5

For each dataset, write what types of **documentation** would be needed for others to understand what's in it

Where can you store the different types of data?

UiO storage guide:

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



Data storage guide

This guide tells you where you can *store* and *process* information. Click on the headings below to read more about about the different kinds of storage.

Please also see the [classification guide](#) for information about the different information categories.

NB! The numbers in the table refers to notes. The notes corresponding to the numbers are in the bottom of this article.

Storing on a Mac, PC or hard drive

	Open (green)	Restricted (yellow)	In confidence (red)	Strictly in confidence (black)
Privately owned laptop (BYOD)	Yes	<u>11</u>	No	No
Privately owned home computer	Yes	<u>11</u>	No	No

Where can you store different types of data?

Green data

for example, completely anonymous data and/or data without any copyright restrictions

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

for example, transcriptions of audio or video recordings - pseudonymized or with small amount of non-sensitive personal information, interview notes or field notes with small amounts of non-sensitive personal information, copyrighted material

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

for example, recordings, transcriptions or notes with large amounts of personal information or any amount of sensitive personal information

*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?

Red data

- Fully encrypted disc, memory stick or external hard drive
- EduCloud Research platform
- 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

«Snapshot» backup for last 7 days

Additionally, consider:

Size?

Transfer?

Access for external partners?

Hidden information in file metadata?

(e.g. location information in picture files)

UiO's Data Sharing Guide

	Open (green)	Restricted (yellow)	Confidential (red)	Strictly confidential (black)
Intern e-post	Yes	Yes	See below	No
Ekstern e-post	Yes	See below	No	No
"Del"-knapp i Office 365	Yes	Yes	No	No
Teams, OneDrive, SharePoint, Mattermost	Yes	Yes	No	No
FileSender	Yes	Yes	No	No
FileSender + 7-Zip-kryptering	Yes	Yes	Yes	No
Zoom	Yes	Yes	Yes	No
EduCloud Share	Yes	Yes	Yes	No
UiO SafeShare	Yes	Yes	Yes	No
TSD	Yes	Yes	Yes	Yes

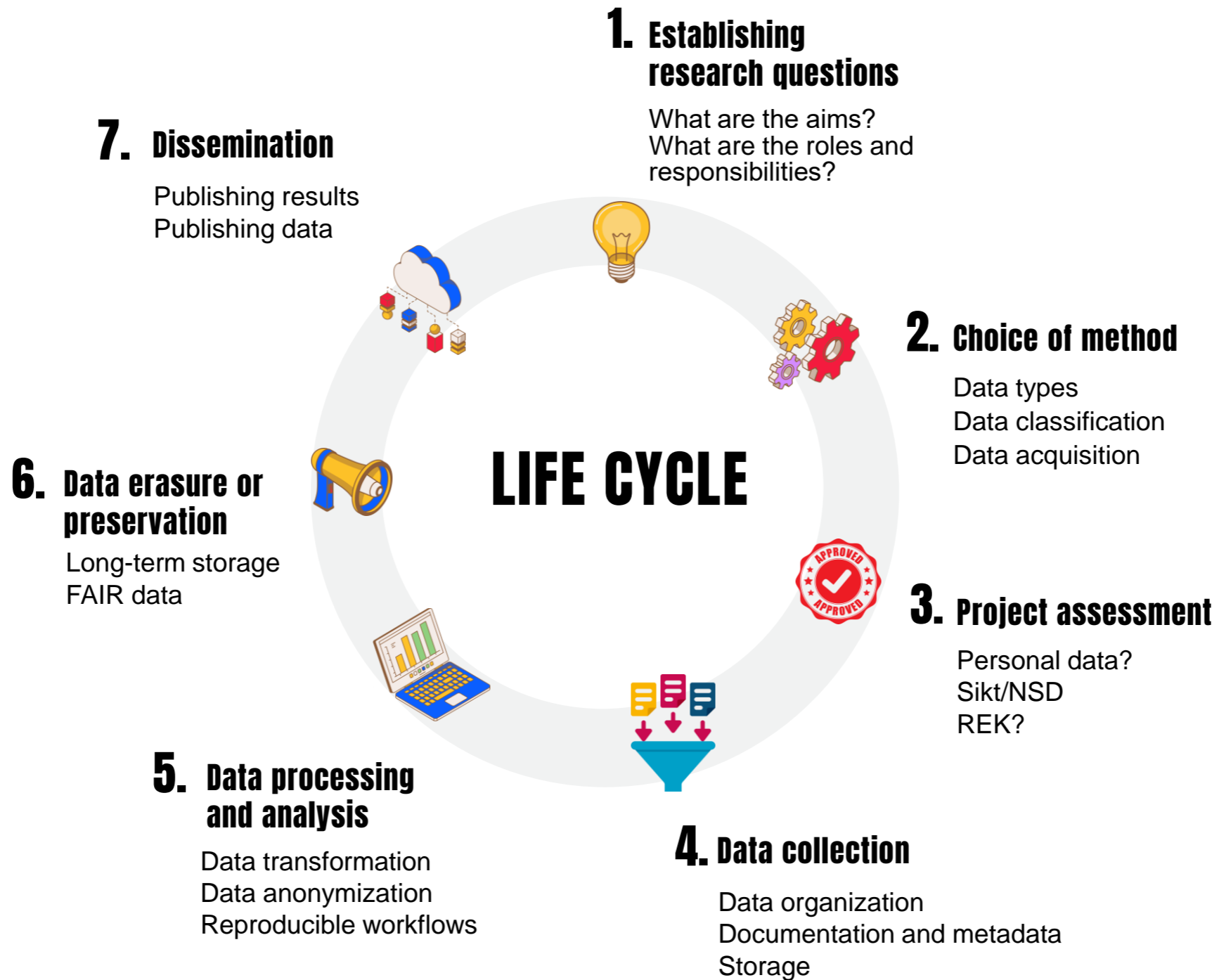
Some things to remember



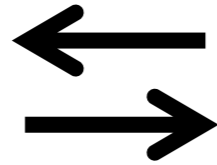
- Use **safe** storage solutions, where data/files are backed up
- Always keep a **copy** of your raw data material safely hidden away (a folder that you never edit)
- Only saving files on your own computer is not safe
- A memory stick or an external hard drive is not safe
- How much **extra work** would it be for you to recreate your data or your work?

6

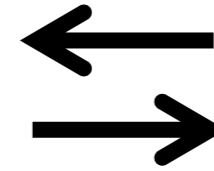
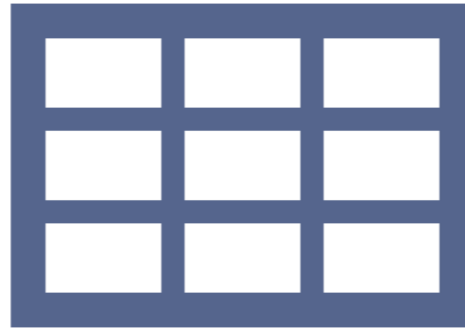
For each dataset, write down where you will **store** it during the project



Raw data



Transcribed, coded or derived data



Analyzed data



Data anonymizing and pseudonymizing

anonymization in the context of
research data with human participants,
i.e. **personal data**

Data anonymizing and pseudonymizing

Anonymization: to irreversibly remove identifying information from the data so that a person cannot be identified based on the data

Pseudonymization: removal or replacement of identifiers with pseudonyms or codes, which are kept separately and protected (note: this is often still perceived as personal data)

Anonymisation and pseudonymisation techniques

1. Replacing personal names with aliases

John Smith -- Respondent1

2. Categorizing proper nouns

Oslo – [city]

3. Changing or removing sensitive information

Dementia – [illness] or [deleted]

4. Categorizing background information

Workplace: UiO – Workplace: university

5. Changing values of identifiers (where possible)

6. Removing metadata from files

Direct and indirect identifiers

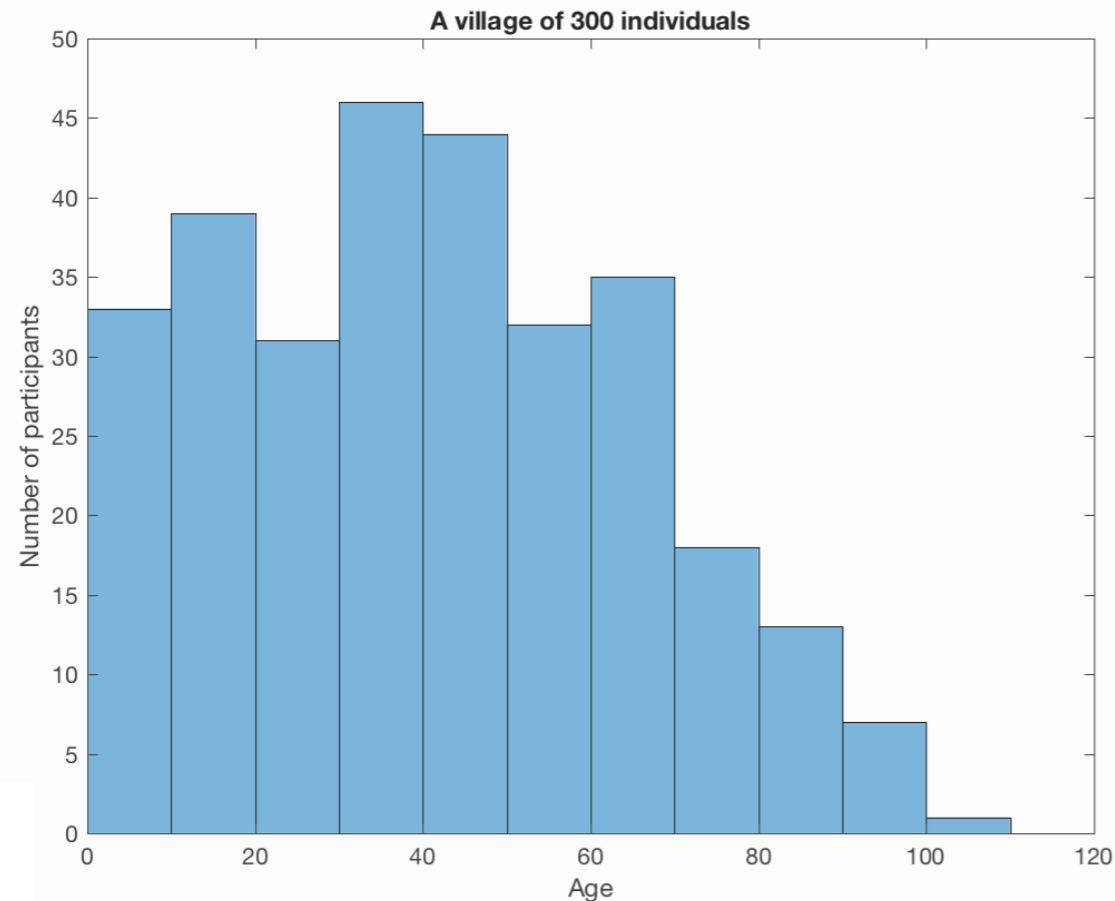
A direct identifier is information that is sufficient on its own to identify an individual.

e.g. a person's name, personal identification number, a facial image, fingerprints

An indirect identifier is information that might fairly easy identify an individual or lead to identification when linked with other available information.

e.g. location data, online identifiers, unusual job title, a position held only by one person at a time, age, education, ethnic background

Indirect identification: a small village of 300 people



People counted together based on their “decade” age

We can single out one very old individual which most likely everybody in the village knows

Data anonymizing and pseudonymizing

Anonymization: to irreversibly remove identifying information from the data so that a person cannot be identified based on the data

Pseudonymization: removal or replacement of identifiers with pseudonyms or codes, which are kept separately and protected (note: this is often still perceived as personal data)

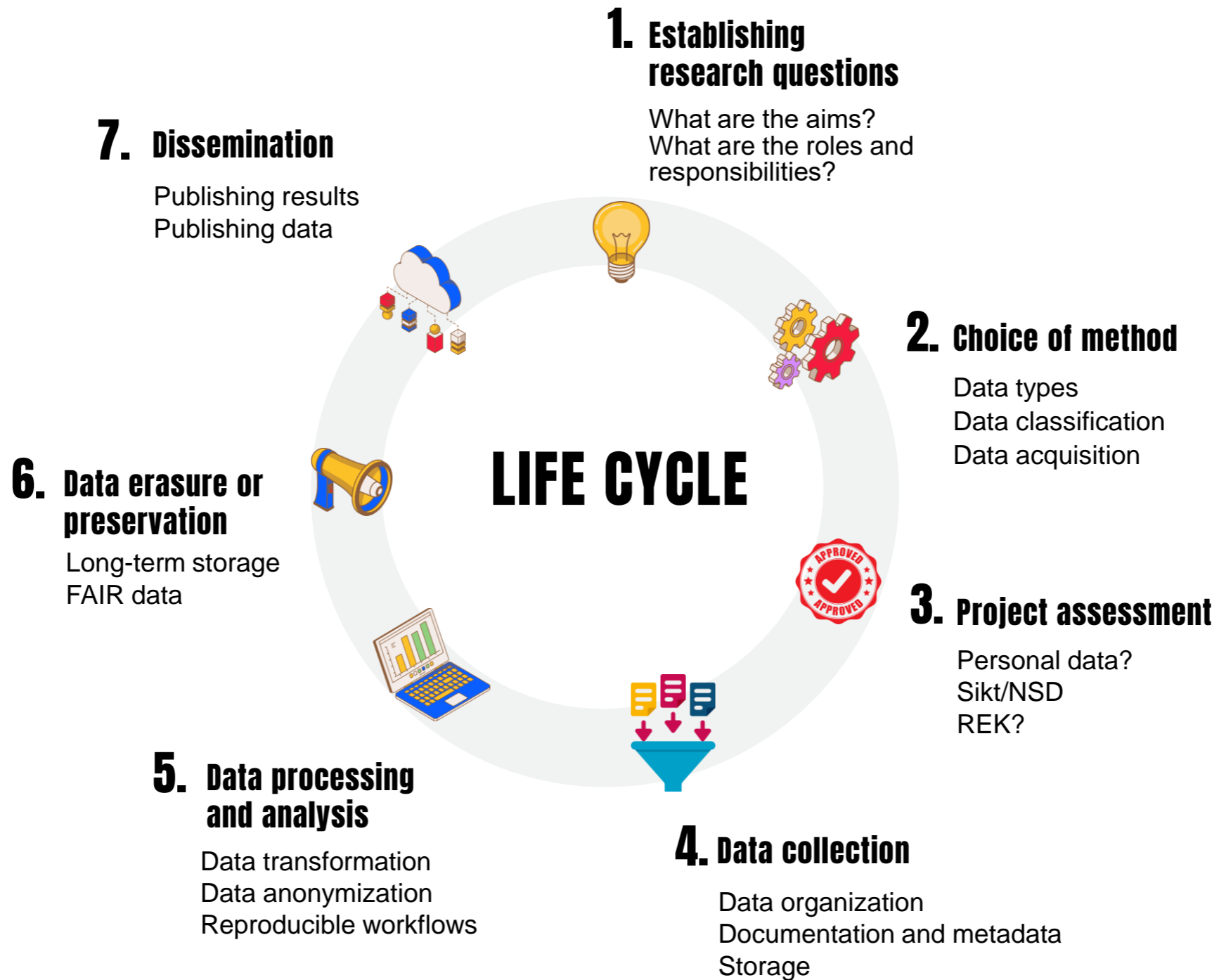
7

Are you going to anonymize or pseudonymize data in your project?

Briefly describe your **anonymization plans** (if any) and/or any planned data transformations.

8

How you will ensure and check for **data quality**?



Can we preserve and
reuse qualitative data?

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

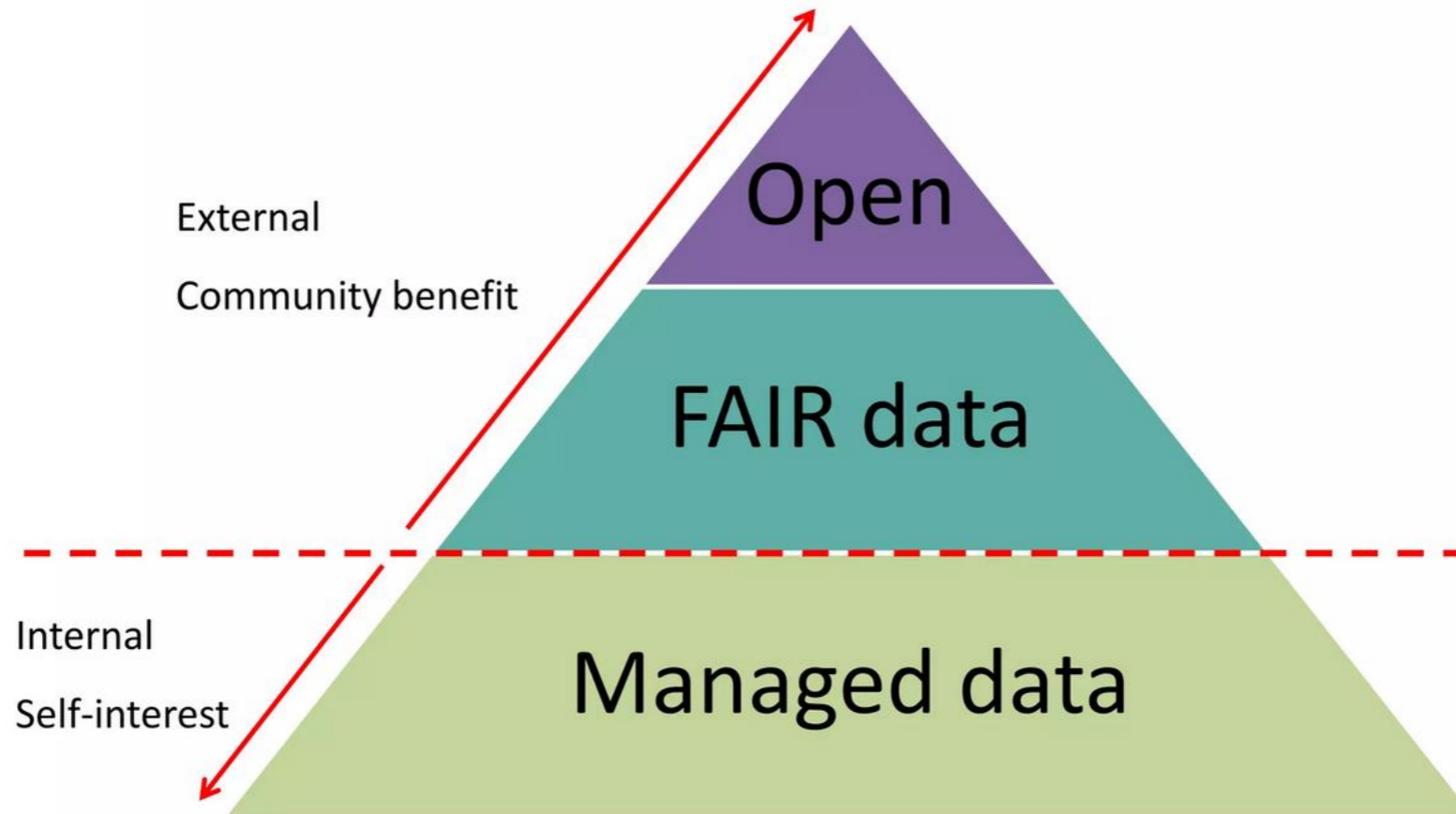
made **freely available**

(but the actual distribution cost should be covered)



I was born in [REDACTED]. My parents were both born and raised in [REDACTED]. My father, [REDACTED], was [REDACTED] and my mother, [REDACTED], was [REDACTED]. They both lived in [REDACTED], on either side of [REDACTED], and there was no chance that they would meet each other.

As **open** as possible,
as **closed** as necessary





Beyond Replication: Secondary Qualitative Data Analysis in Political Science

Florian G Kern ¹ and Katariina Mustasilta ²

Abstract

Shared qualitative data – such as interview or focus group transcripts – can be used for secondary qualitative data analysis (SQDA). Yet, much archived qualitative data remains unused after primary analysis. Applications and guidance on how to employ SQDA are rare. We use an example application of SQDA studying informal institutions and resilience in Sub-Saharan Africa to show: First, SQDA depends on how primary researchers share ‘raw’ qualitative data and additional documentation to understand primary context. Second, deductive and inductive uses of SQDA require varying engagement with primary data. Third, current practices of participant consent often do not consider potential SQDA. Fourth, SQDA is not less time-consuming than primary data research but offers different benefits, such as expanding the comparative sample of cases or avoiding research fatigue of studied communities. Going forward, SQDA requires greater consensus on the instruments (e.g. transcripts and participant consent forms) used by researchers and further applications of hypothesis-testing and hypothesis-generating designs.

Open science and sharing personal data widely – legally impossible for Europeans?

Giske Ursin^{a,b,c} and Heidi Beate Bentzen^{a,d} 

^aCancer Registry of Norway, Oslo, Norway; ^bInstitute of Basic Medical Sciences, University of Oslo, Oslo, Norway; ^cDepartment of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA; ^dNorwegian Research Center for Computers and Law, Faculty of Law, University of Oslo, Oslo, Norway

ARTICLE HISTORY Received 15 October 2021; accepted 15 October 2021

A requirement for having a research paper published in many medical journals is that the authors include a data sharing statement. Although the requirement from the International Committee of Medical Journal Editors is not very strict, simply requiring a statement [1], interpretation varies. Some journals essentially require that data must be *readily available* for other researchers for the paper to be accepted.

While most of us eagerly welcome open science and reuse of data to ensure reproducible science, the EU General Data Protection Regulation (GDPR) provides strong protection of privacy and rather restricts and counteracts open sharing of personal data [2]. Some editors will accept that data are not readily sharable with others than peer reviewers for legal reasons. However, editors of non-European journals will often object to a GDPR-compatible data sharing statement and, consequently and often at the last minute, reject the research paper.

Why is this an issue? How difficult is it for European researchers to share data with researchers in other parts of the world?

supplementary measures in place to protect the data. The European research institution will in collaboration with the data importer need to conduct a thorough assessment of the importer's country's laws to ensure that an EU level of data protection is obtained. Such assessments require sound knowledge of the EU Charter of Fundamental Rights, the GDPR, the Court of Justice of the European Union *Schrems II* judgment, and subsequent guidance from the European Data Protection Board, which comprises all Data Protection Authorities in the European Economic Area (EEA) [4]. Finally, the data exporter must be willing to take the risk that the national Data Protection Authority agrees that all requirements have indeed been met, as fines can be high if the institution makes a mistake. Whenever such transfer is possible to achieve, you are lucky! The only cost is that the legal and administrative work on your end has quadrupled compared to the pre-GDPR era.

The federal challenge

Can you preserve and make
data reusable in your project?

Option 1:

Open archiving if full anonymization possible and/or no copyright

UiO's institutional data archive



General-purpose data repositories

The Zenodo logo consists of the word "zenodo" in a white, lowercase, sans-serif font, centered within a solid blue rectangular background.

zenodo



Open Science Framework



Discipline-specific archives/repositories

Search

Browse ▾

Suggest

Resources ▾

Contact



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Search...

 Search

Option 2:

Access-controlled archive for
research data that need protection

Template for information letter

The privacy regulations contain clear requirements for what the information

Because there are different types of legal bases for processing, we have prepared templates that you can use as a starting point for some of the most common bases in research.

Template for information letter - gathering consent (word-format)

[DOCX, 33kB]



Template for information letter - public interest (word-format)

[DOCX, 25kB]



• [Home](#) • [Research data](#) • [Deposit your research data](#)

Deposit your research data

Sikt accepts and archives research data on people and society and makes the data accessible for reuse. We will assist you with documenting your research data and choosing the right access conditions.

We make sure that the data are safeguarded and accessible for at least 10 years, with a long-term preservation commitment of at least 50 years for high-value and highly curated data.

How to deposit and share your research data

- Sikt's archive is responsible at a national level for archiving and sharing data from the **social sciences, humanities**, and parts of **medical and health research**.
- We can archive all types of digital research data, including data that requires special handling or permissions, such as personal data.
- Data should be shared as openly as possible, and as closed as necessary. That means data can't have more restrictive usage and access conditions than necessary.

<https://sikt.no/en/archiving-research-data>

The Qualitative Data Repository

Search data...

Discover Data

Deposit Data

Just Getting Started?



Option 3:

Open archiving of metadata and
internal long-term storage of data

Published March 24, 2022 | Version 1

Other

🔒 Open

Quality in Nordic Teaching (NCoE) Common QUINT Dataset 2022

Kirsti Klette¹

Show affiliations

Contact person: Maria Dikova¹

Data manager: Bjørn Gulheim¹

Hosting institutions: Kirsti Klette¹; Torgeir Christiansen¹

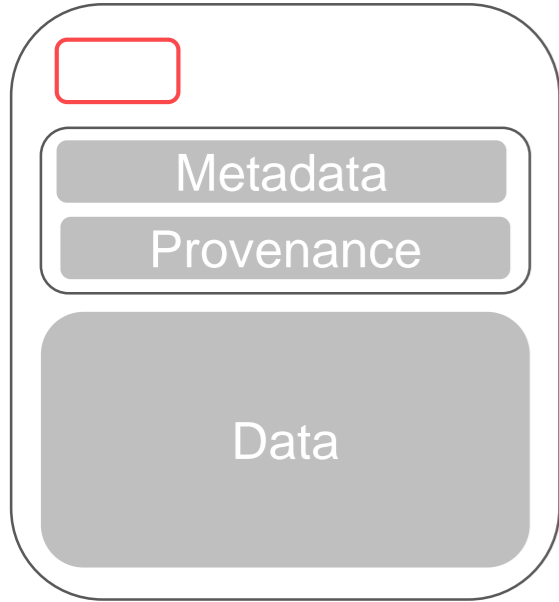
Researcher: Michael Tengberg²

Show affiliations

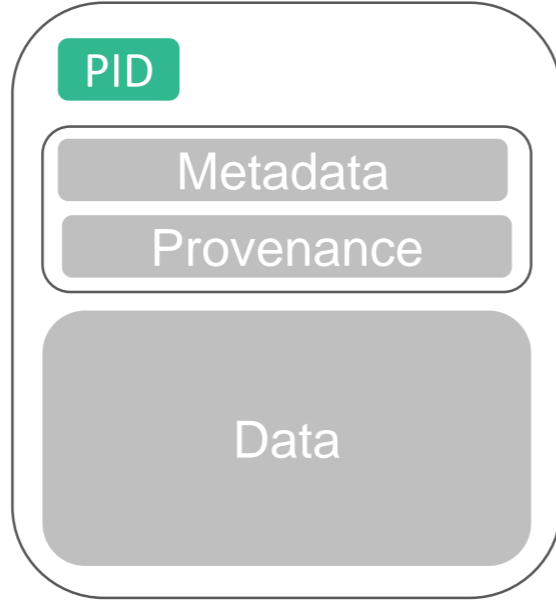
This file describes the Quality in Nordic Teaching (NCoE) Common QUINT Dataset 2022 and related prerequisites and procedures for gaining access to this dataset.

<https://zenodo.org/records/6381818>

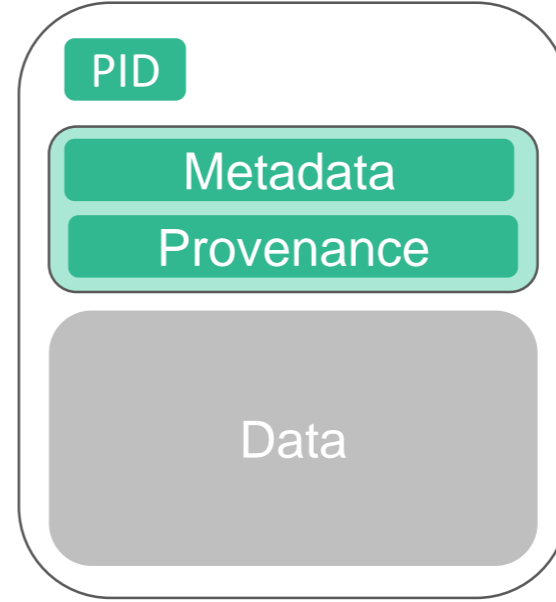
Totally UNFAIR



Findable
Usable for humans



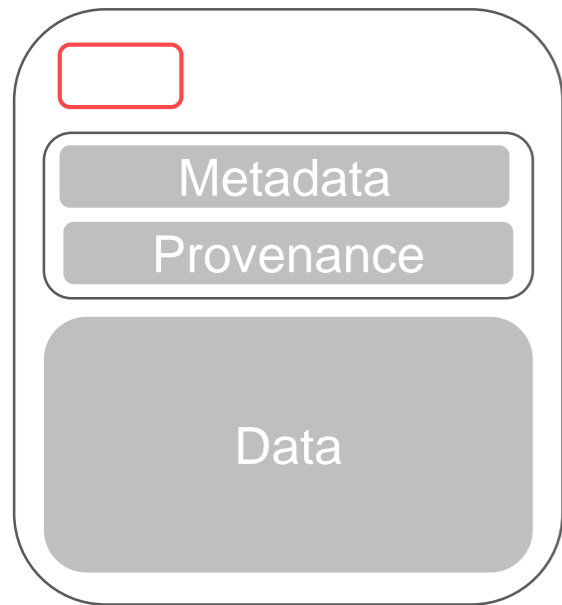
FAIR metadata



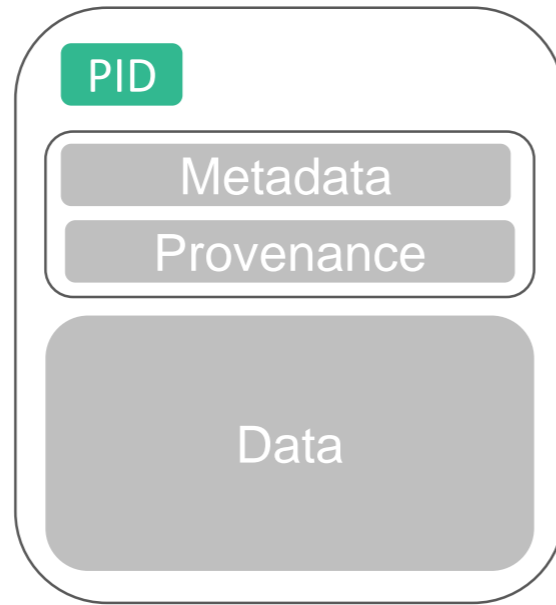
Levels of FAIR

Levels of FAIR

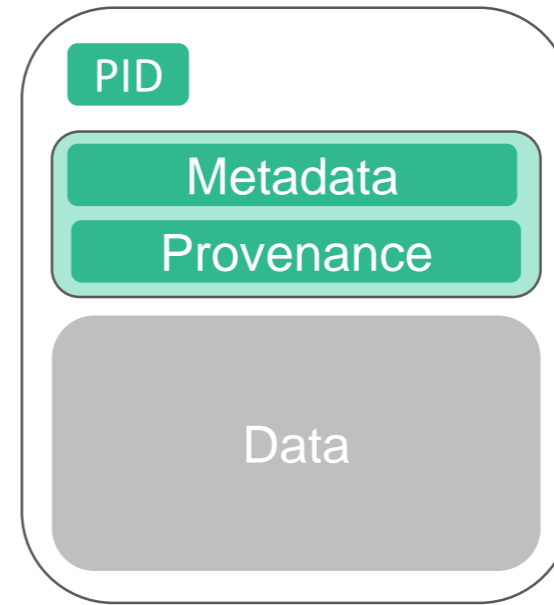
Totally UNFAIR



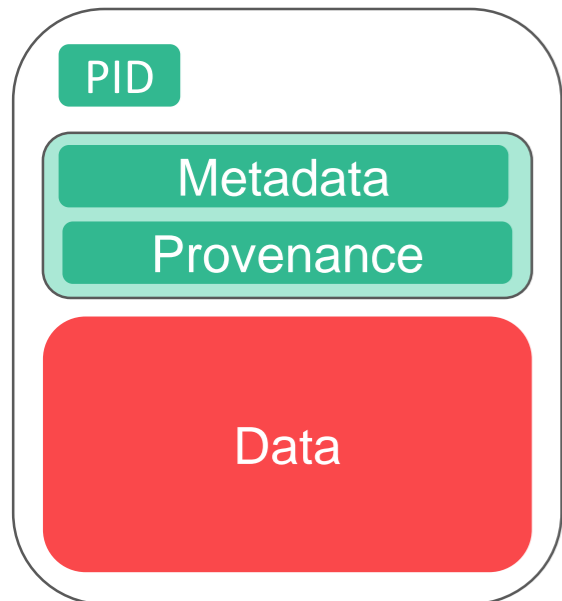
Findable
Usable for humans



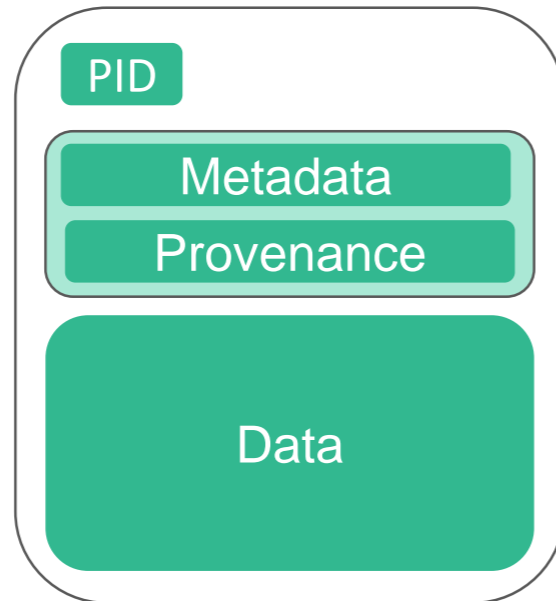
FAIR metadata



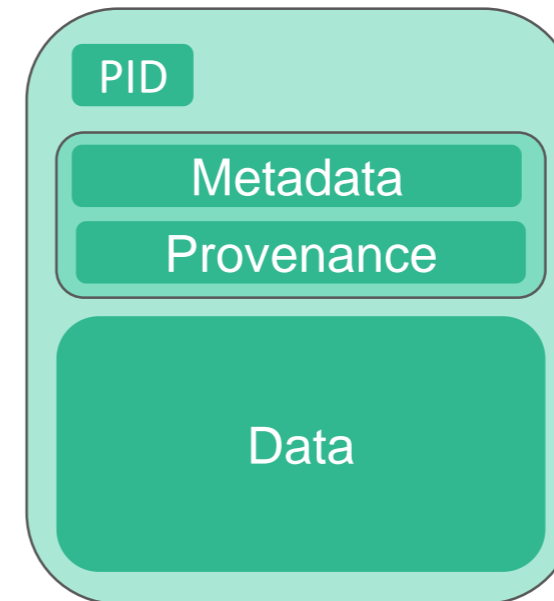
FAIR data
Restricted access



FAIR data
Open access

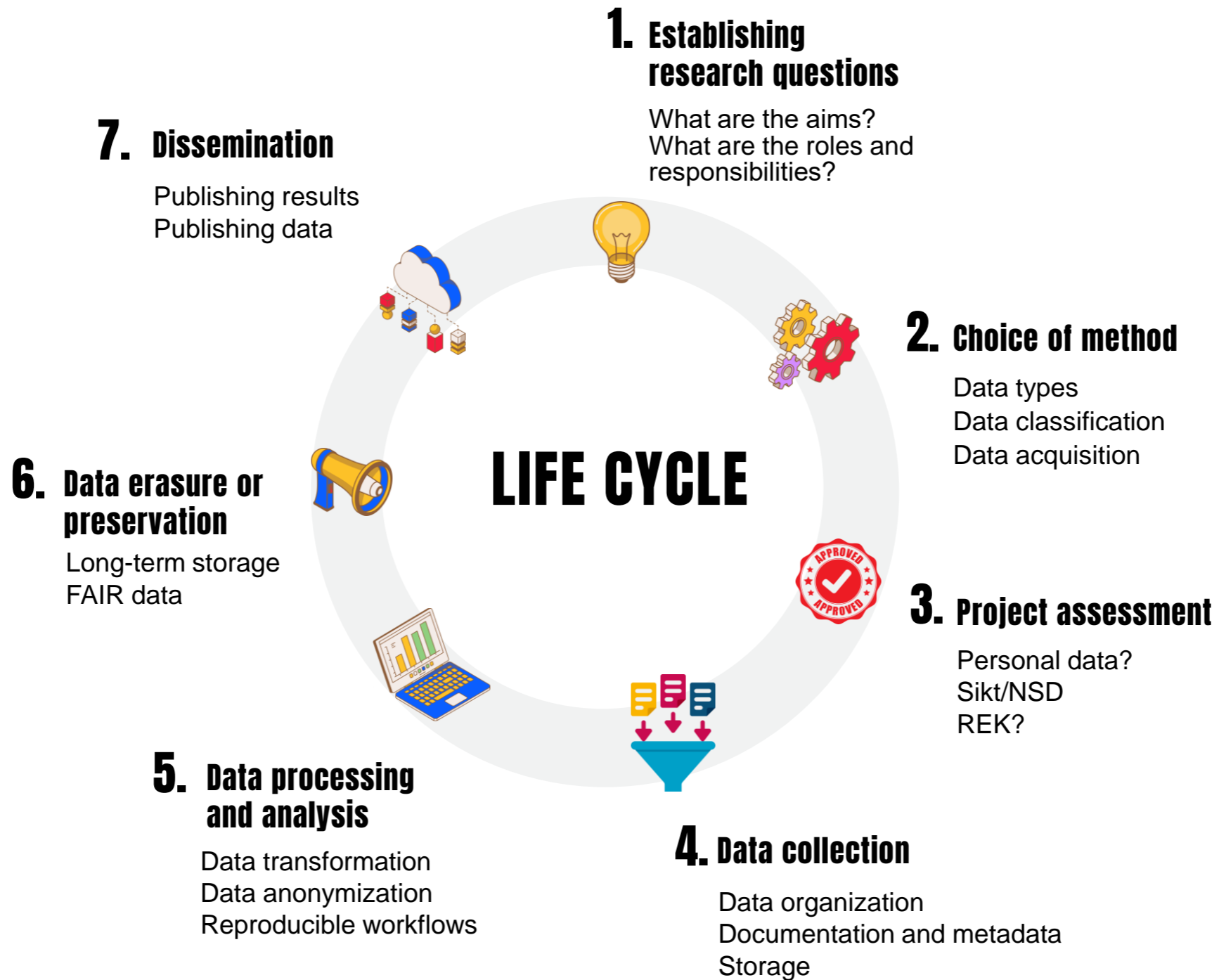


FAIR data
Open access and functionally linked

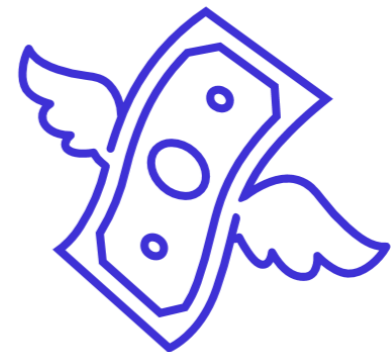


9

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



- Are there costs associated with purchase of **software** or hiring a **data manager**?
- Will you need to purchase new **hardware** or equipment?
- **Storage and maintenance** 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.)



10

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

Thank you!

QualiFAIR DMP Work Package

Agata Bochynska, Matthew Good, Torgeir Christiansen, Anne Bergsaker

Digital Scholarship Center, University of Oslo Library

01.11.2023,

Questions? Contact QualiFAIR coordinator at agata.bochynska@ub.uio.no