Plan your research data management

Elisa Pierfederici, Edina Pózer, Ivana Malovic

Digital Scholarship Center, University of Oslo Library

16.04.2024

Questions? Contact us at research-data@uio.no

Agenda

- Research data types
- Research data management
- Data Management Plans

Break

- Data classification and storage
- Data organisation

Break

- Documentation and metadata
- Summary and final questions
- Wrap up and further resources

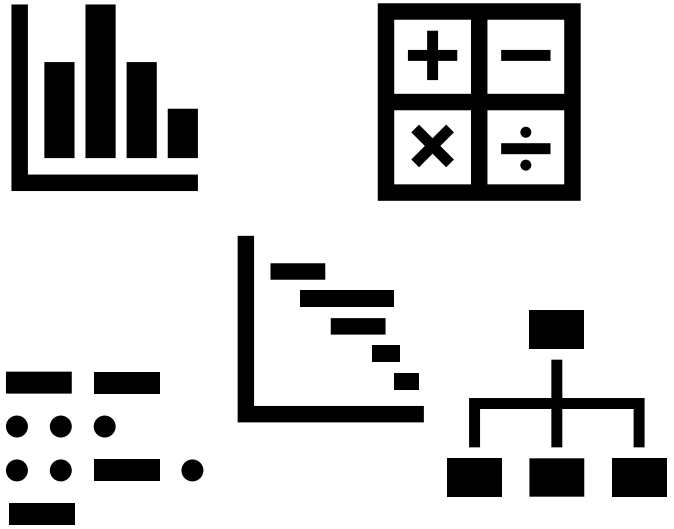
What are research data?

What data are you working with?

Mentimeter

71

Qualitative
(descriptive)



Quantitative
(numerical)

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

What is personal data?

Any piece of information about an identified or identifiable physical person.

In other words, any data about living people from which they can be identified.

Direct and indirect identifiers

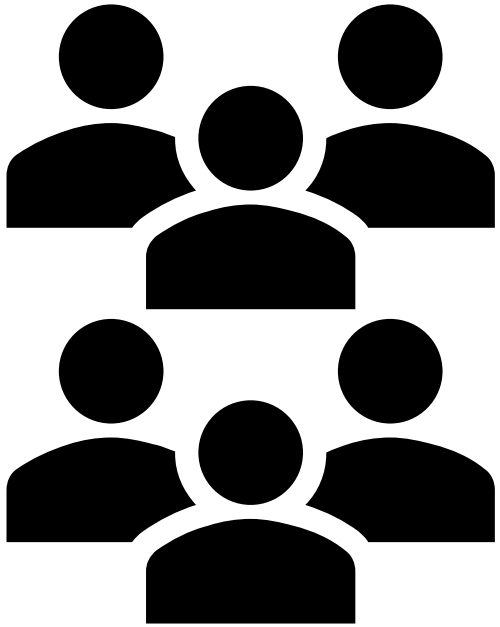
A direct identifier is information that is sufficient on its own to identify an individual.

e.g. *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

Anonymous data or not?



What constitutes as **general personal data**?

- Name, address, age, phone number, e-mail and Norwegian national identity number
- Video- and audio recordings where individuals are recognizable
- Identifying images of individuals
- Contents of email communication
- Content of case documents, investigations or considerations
- Content of exam papers or grades
- IP addresses (that can identify individuals) and logging of activity in computer systems (where the logs can be associated with individuals)

Special categories of personal data (sensitive data)

- health information and health related conditions
- genetic or biometric information which can be used to identify a physical person
- ethnic or racial origin
- political, philosophical or religious perceptions and beliefs
- sexual orientation or sexual relationships
- trade-union membership

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

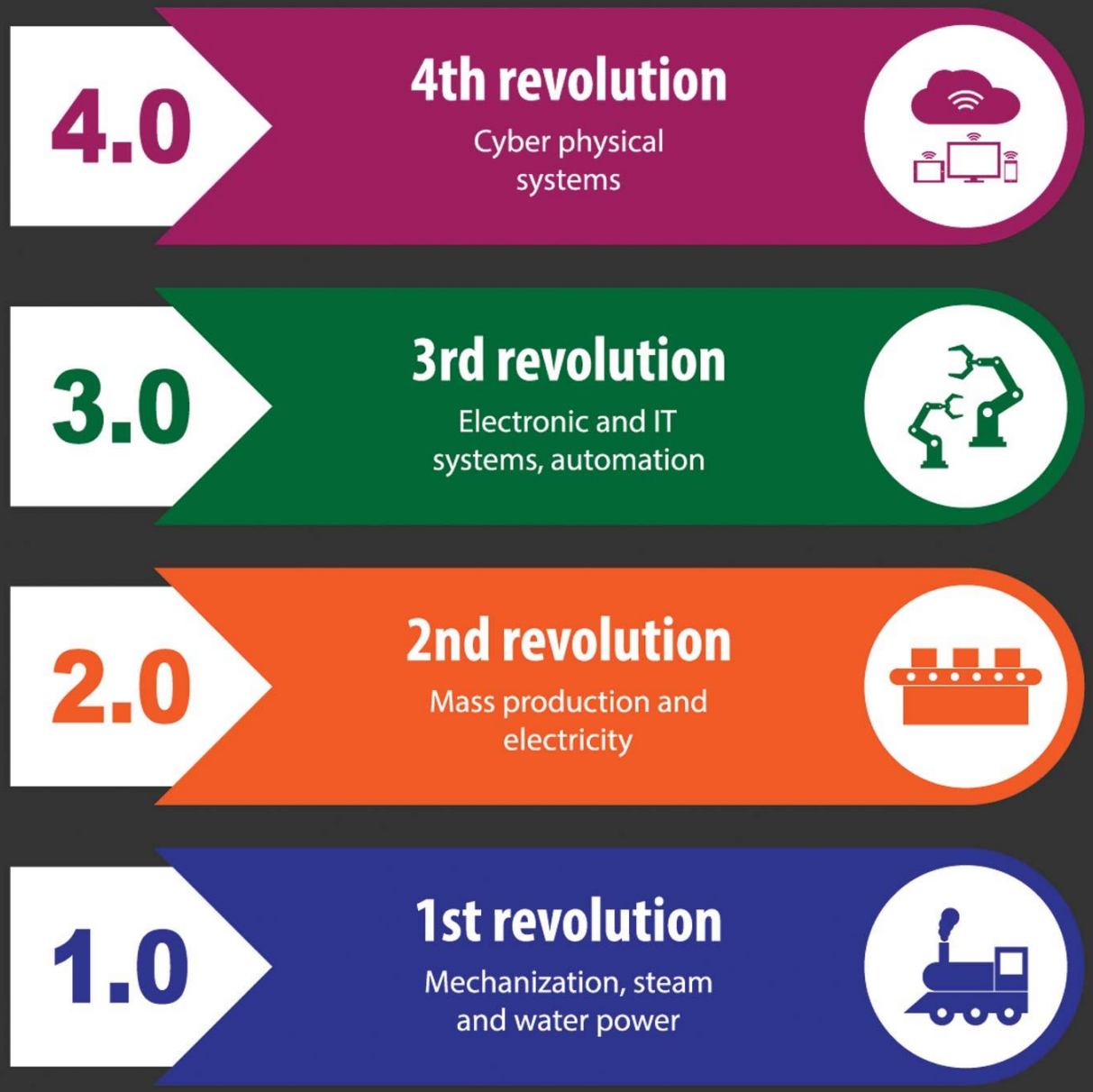
Exercise 1: Data types

Think – Pair – Share

Think about what types of data you are or will be collecting in your research projects.

Share and discuss with your neighbor!

3 min



https://www.britannica.com/topic/The-Fourth-Industrial-Revolution-2119734

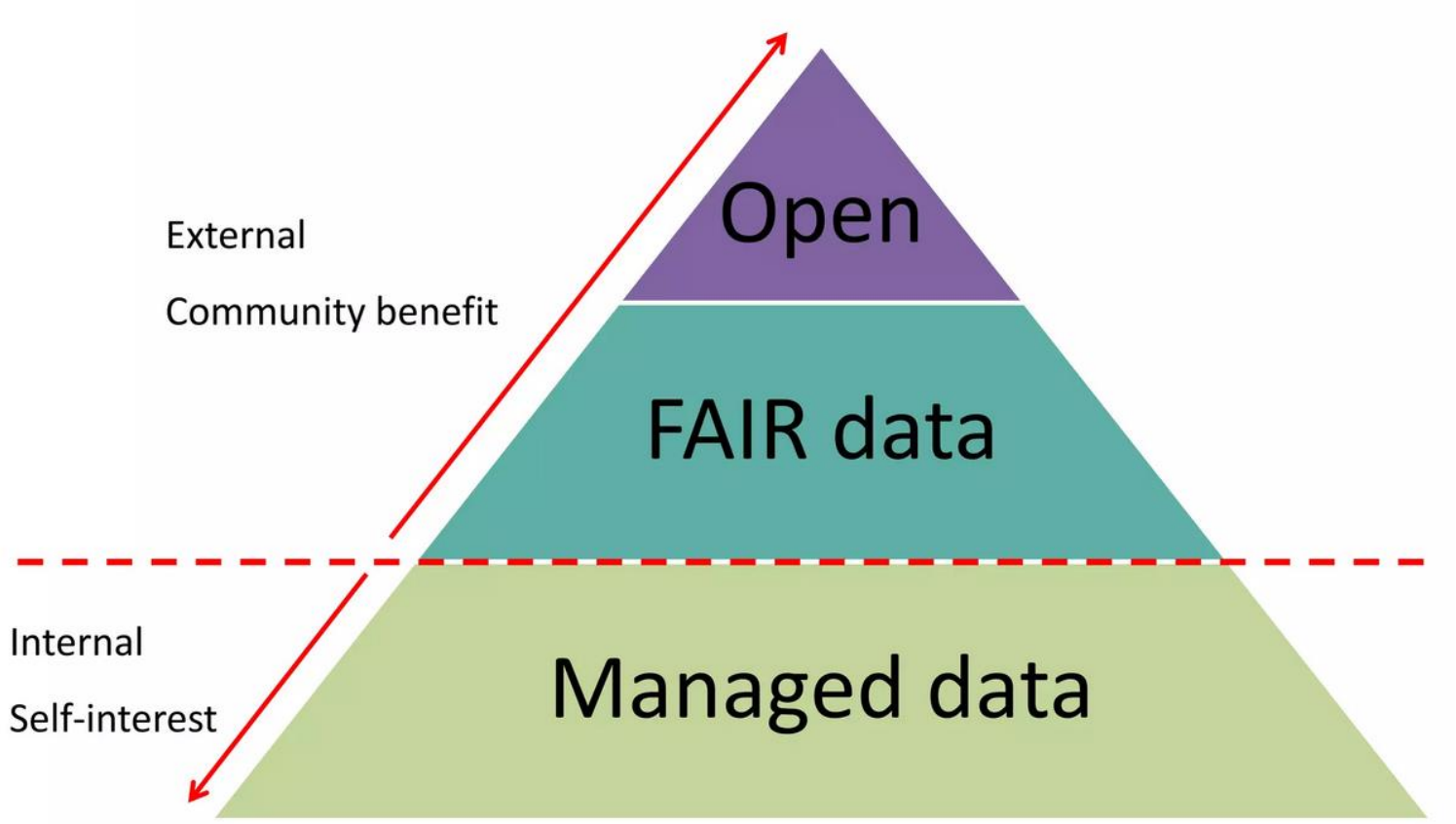
“**Open Science** has the potential of making the scientific process more **transparent**, **inclusive** and **democratic**. It is (…) a true game changer in bridging the science, technology and innovation gaps and fulfilling the **human right to science**.”

UNESCO Recommendation on Open Science

https://youtu.be/I3Wkvx_ZaFo
https://www.unesco.org/en/natural-sciences/open-science

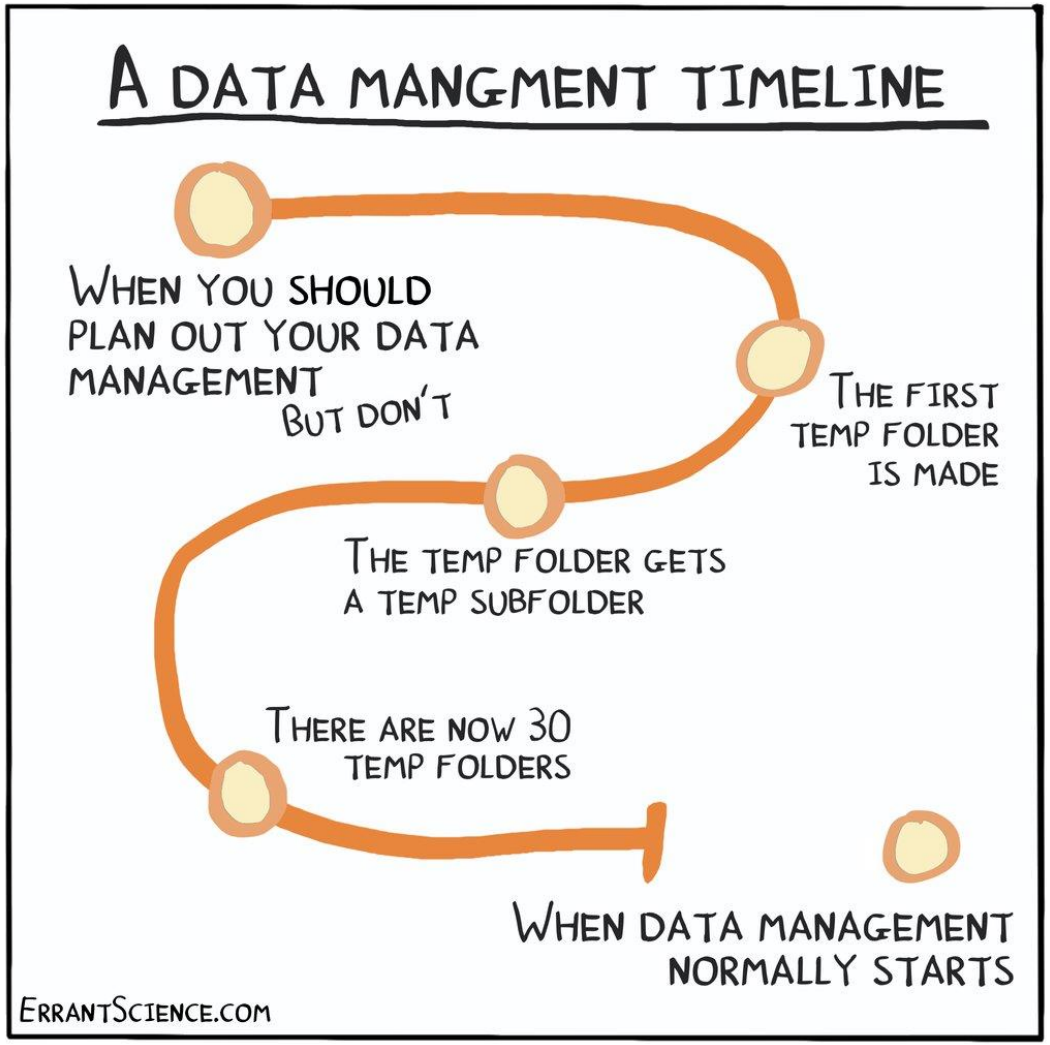
Source: *10.5281/zenodo.3332807*

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



Source: https://www.slideshare.net/sjDCC/open-fair-data-and-rdm

Research Data



Protection
Organization
Documentation (and metadata)
Classification
Short-term storage
Licensing
Sharing
Long-term preservation
Dissemination (publishing)
Reusing

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

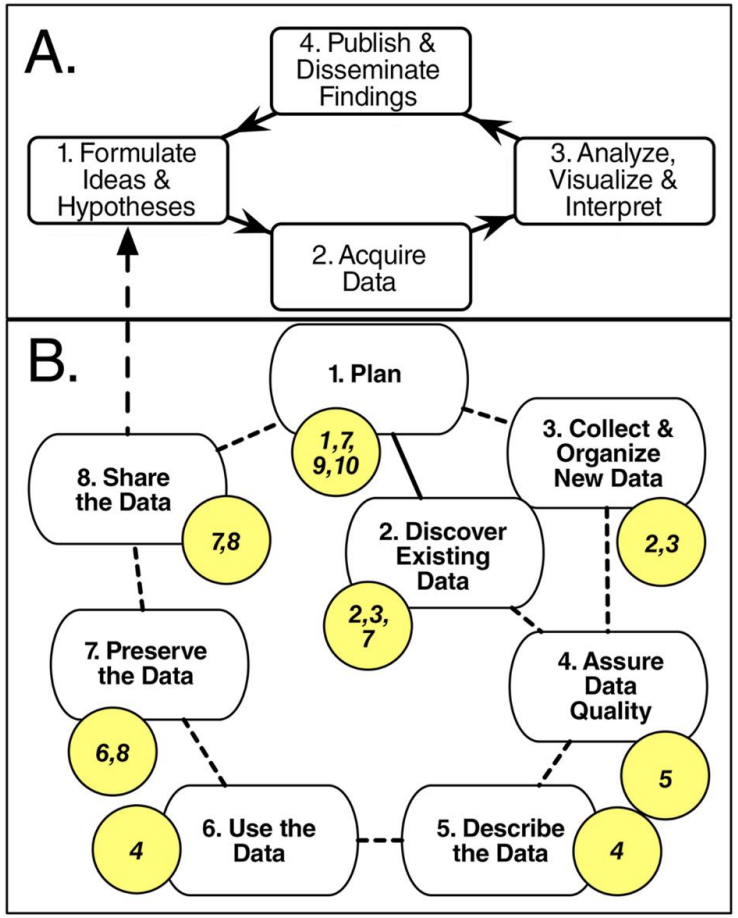
Interactive tools and web forms that assist with writing a DMP

- Sikt (formerly NSD) - single and easy to use template but with limited functionalities
- DMPonline and DMPTool - popular generic tools supporting multiple templates; can be used with DMPs that will be made openly available in the future
- Data Stewardship Wizard (ELIXIR) - life science data but not only

https://www.uio.no/english/for-employees/support/research/research-data-management/write-a-dmp.html

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**



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

Let’s see an example!

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

Find more examples in Public DMPs service from DMPonline:

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

Exercise 2: DMP

Write and discuss

Write down keywords for most important elements that you would like to include in your Data Management Plan.

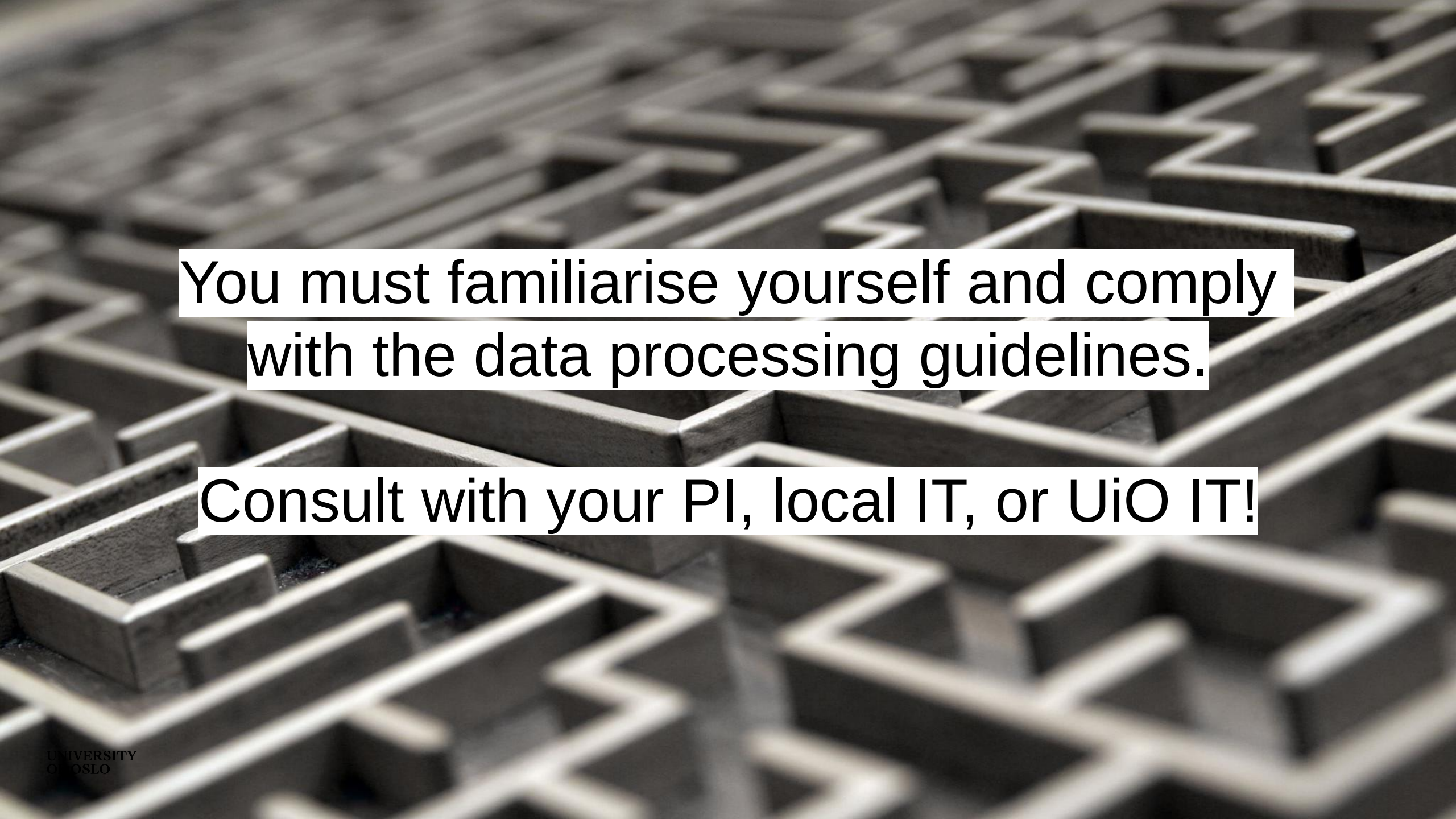
4 min

Data classification and storage selection

All UiO information and data shall have **an unambiguous and identifiable responsible person**

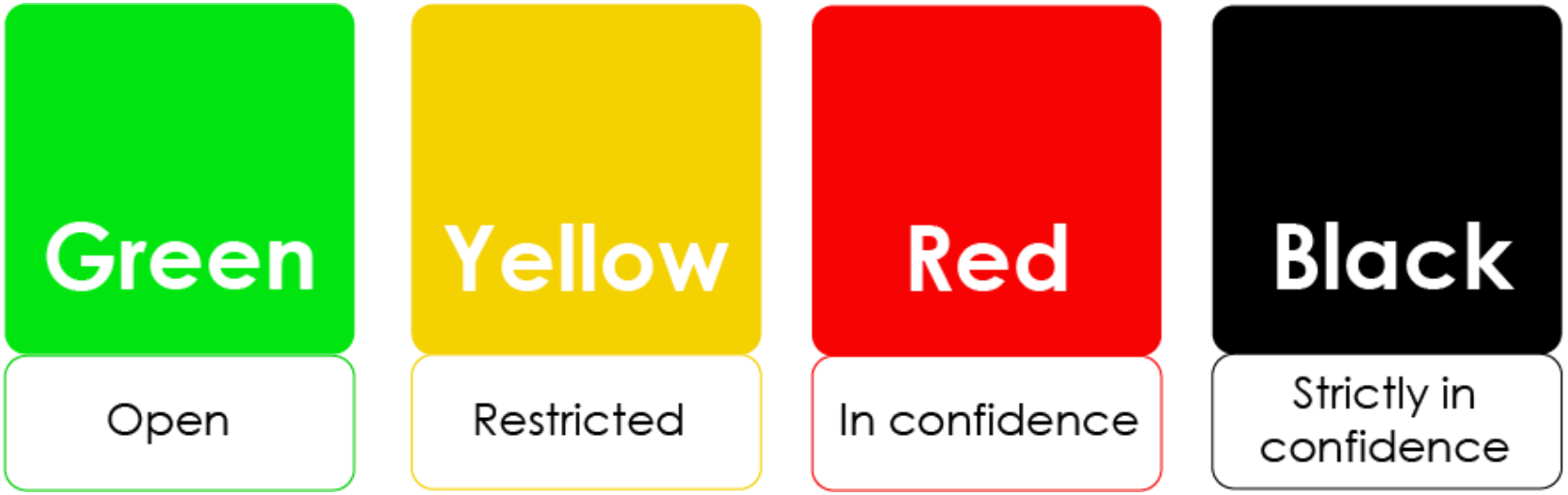
Anyone should **easily** be able to **find out who is responsible** for the data

This person is **also responsible for the assessment** to which **category** data belong


You must familiarise yourself and comply with the data processing guidelines.
Consult with your PI, local IT, or UiO IT!

Classifying your data

Data classification at UiO is based on how harmful for the university, its partners, public interests, or individuals it would be if the data become exposed to third party



https://www.uio.no/english/services/it/security/lsis/data-classes.html

Where can you store the different categories of data?

Data storage guide:

https://www.uio.no/english/services/it/security/lsis/storage-guide.html

Green data

Open or freely available to the general public

- This class is to be used if the university or its partners are **not subject to any harm** if the information is exposed to third parties.

Green data

Open or freely available to the general public

- Examples are:
  - Webpage presenting a department or a class, published openly on the internet
  - Course material which is openly published, but marked with a certain license and/or copyright
  - Research data that does not need any protection (the researcher is responsible for this assessment)
  - Teaching material that does not need any protection (the teacher is responsible for this assessment)

Green data

Open or freely available to the general public

- Recommended storage solutions:
  - **Anywhere you want**, but ideally **not only** on your private laptop (very vulnerable!)

Yellow data

Restricted information not open for everyone

- This class is to be used if the university or its partners may be subject to **limited harm** if the information is exposed to third parties.
- The information needs a certain protection, and may be accessible to people both within and outside the university, provided that the access is limited and controlled per user.

Yellow data

Restricted information not open for everyone

- Examples may be:
  - Certain work documents
  - Information which is to be kept from the public
  - Many types of personal data
  - Grades
  - Work by students
  - Examination answers
  - Unpublished research data and corresponding works

Yellow data

Restricted information not open for everyone

- Recommended storage solutions:
  - **Not** your private laptop, unless you comply with the guidelines for use of private computer
  - Encrypted memory stick or external hard drive
  - Encrypted laptop maintained by UiO
  - UiO cloud solutions
  - UiO Storage Hotel («lagringshotell»)
  - UiO personal researcher storage
  - UIO shared storage for your unit
  - UiO e-mail; UiO Teams; UiO Zoom; UiO Nettskjema; UiO Canvas, UiO Vortex, …

Red data

In confidence information which the university is obliged to protect by law, agreements and other regulations

- This class is used if it **could cause harm** to the university, its partners, public interests, or individuals if the information is exposed to third parties.

Red data

In confidence information which the university is obliged to protect by law, agreements and other regulations

- Examples may be:
  - certain types of sensitive personal data
  - personnel files
  - certain information about for example protection and safety of buildings and IT systems
  - information about a person's health

Red data

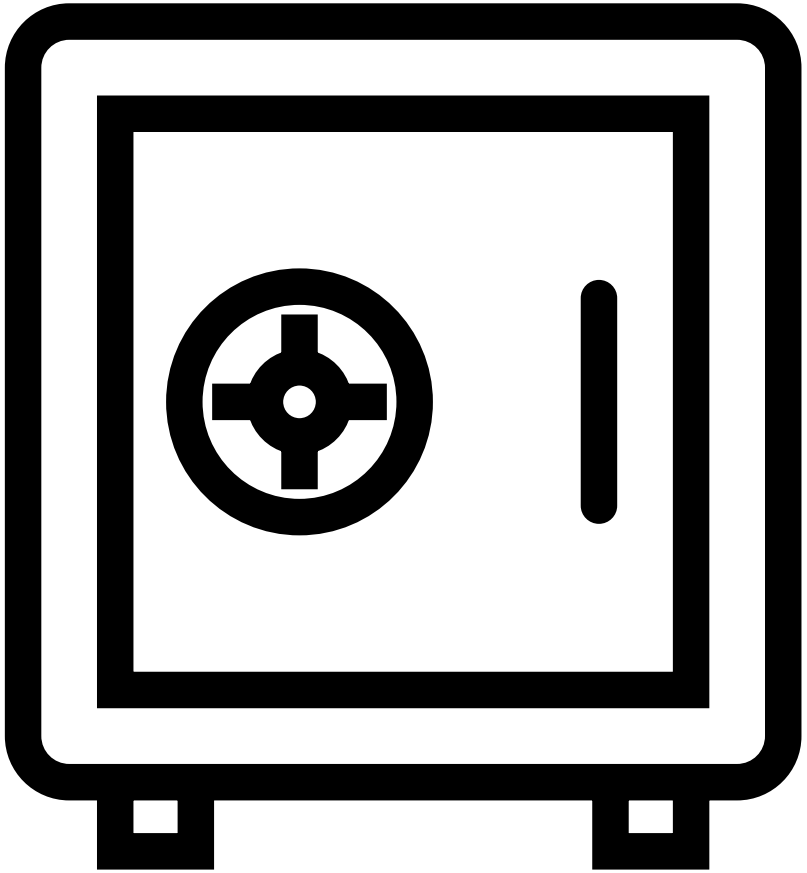
In confidence information which the university is obliged to protect by law, agreements and other regulations

- Recommended storage solutions:
  - Laptop with a fully encrypted disk, an encrypted memory stick or an encrypted external hard drive.
  - Services for sensitive data – TSD
  - *Educloud Research*
  - In Vortex, on your shared research network drive or in the UiO Storage Hotel, if you make an individual risk assessment.

Black data

Strictly in confidence category encompasses the same type of information as «In confidence (red)», but where special circumstances makes it necessary to protect the information even more.

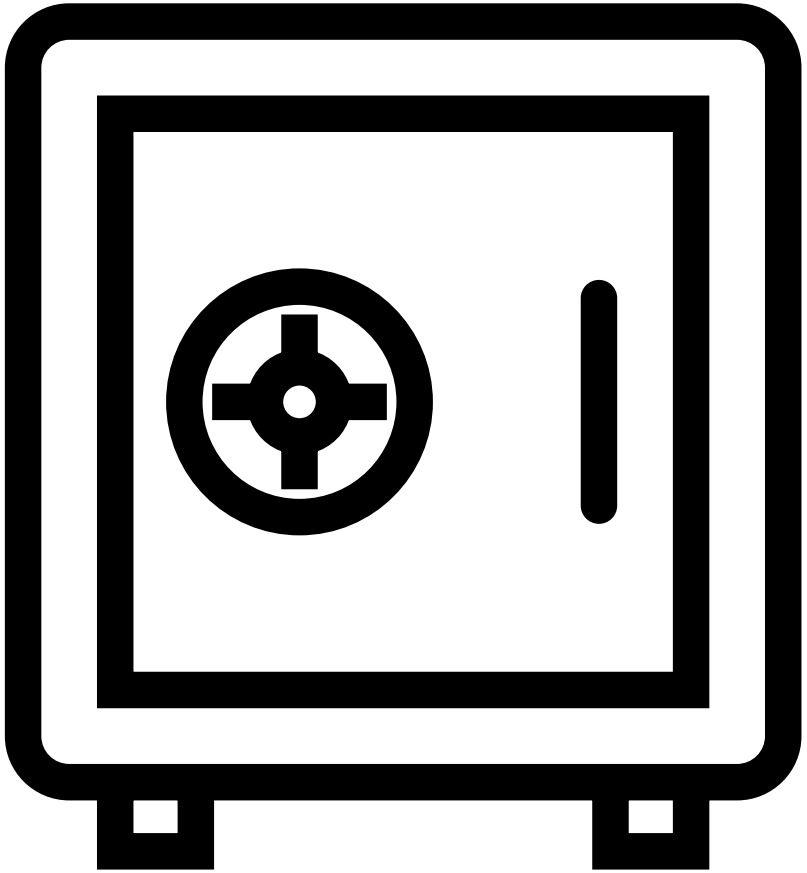
- This class is used if it **could cause considerable harm** to the university, its partners, public interest, or individuals, if the information is exposed to third parties.



Black data

Strictly in confidence category encompasses the same type of information as «In confidence (red)», but where special circumstances makes it necessary to protect the information even more.

- Examples may be:
  - Large amounts of sensitive personal data
  - Large amounts of data about people's health
  - Research data and datasets of huge economic value

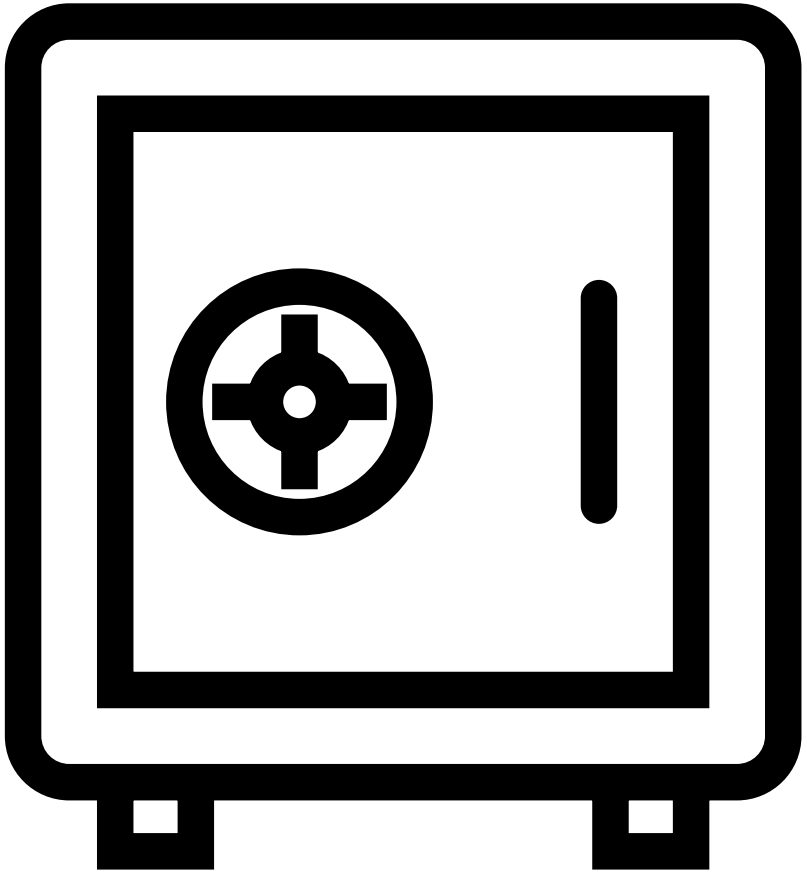


Black data

Strictly in confidence category encompasses the same type of information as «In confidence (red)», but where special circumstances makes it necessary to protect the information even more.

- Placement of data and information in this category should be done in cooperation with the lawyers at USIT and the IT security manager.

- Recommended storage solutions:
  - UiO TSD Services for sensitive data



Classifying your data

Always put the information in a sufficiently safe class.

If you are not sure whether to your information is red or yellow, choose red.

Exercise 3: Data classes

Menti

What are correct data classes in these cases?

8 min

Menti.com → code: 67 37 28 1

Overview of all cloud services evaluated or under consideration at UiO

- https://www.uio.no/tjenester/it/ny-tjeneste/vurderte_skytjenester.html

(in Norwegian)

- Services with the UiO agreement: https://www.uio.no/english/services/it/all/

Backup

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

https://www.uio.no/english/services/it/store-collaborate/backup/

Data organisation

Designing a data file structure

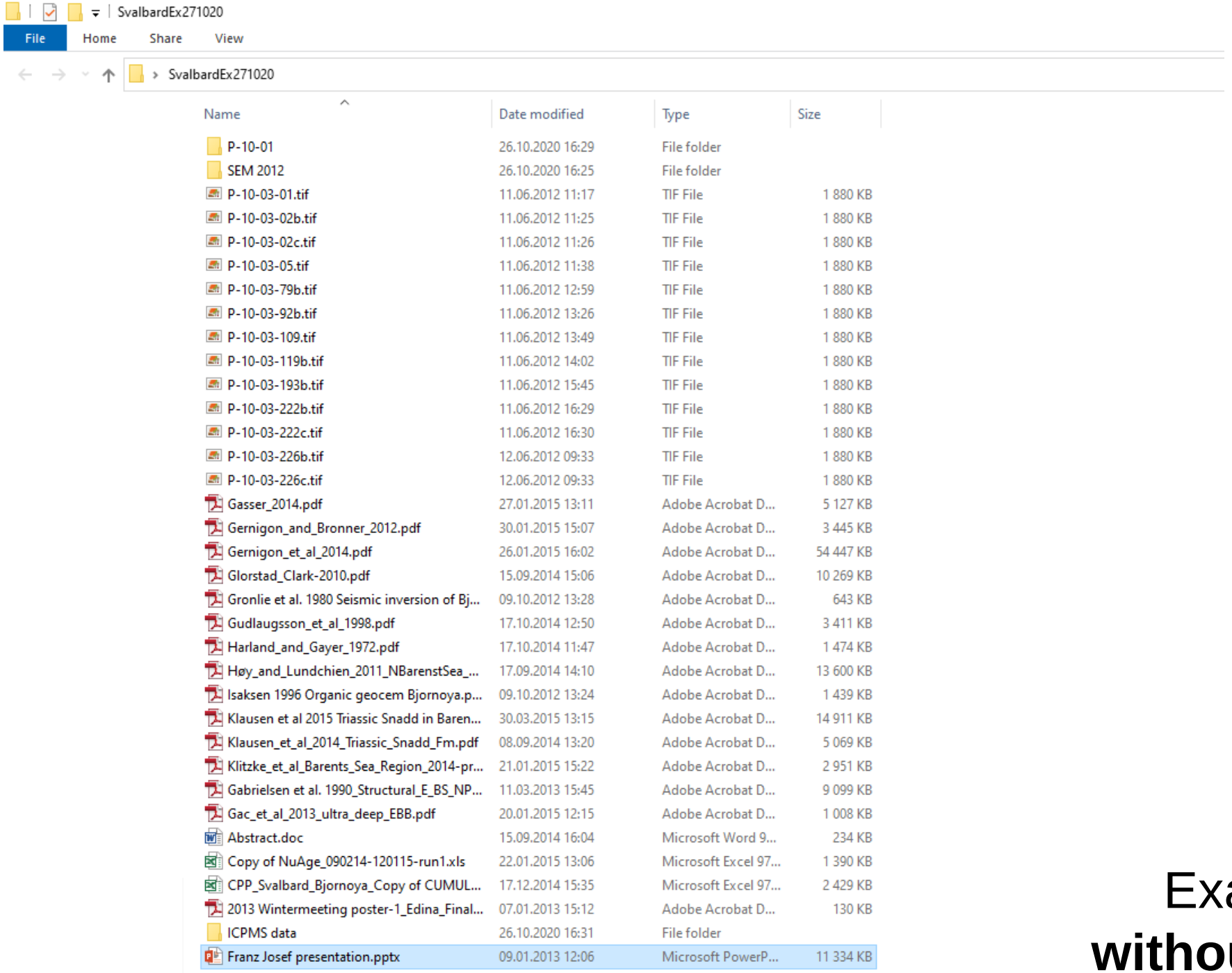
Make a folder hierarchy

Give folders descriptive and informative names

Avoid folders that become too broad or general, create more subfolders instead

Keep active and finished parts of your project separate

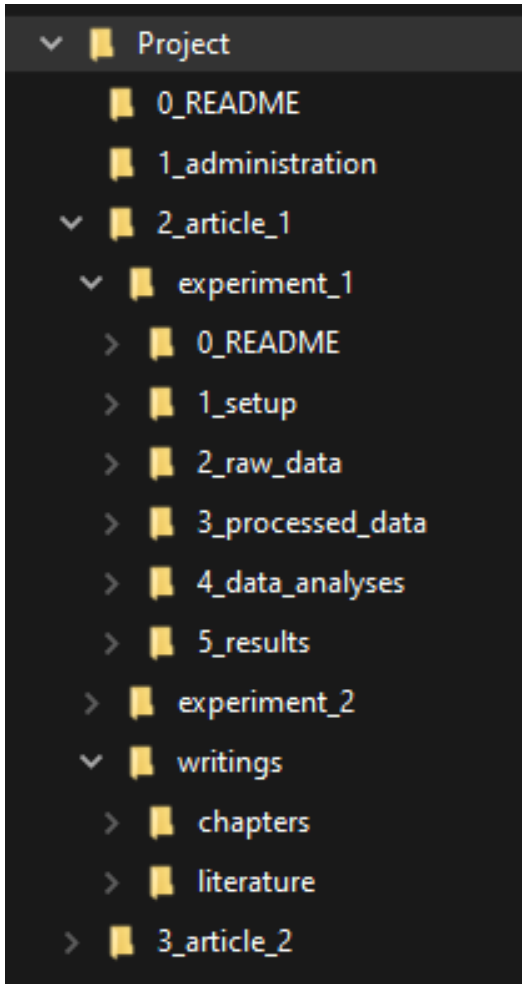
Set aside some time to tidy regularly



Example of a project **without** folder structure

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
    └── ...
```



Research project with a proper data file structure. Image taken from CodeRefinery, Lesson on Reproducible Research. Shared under CC-BY 4.0
https://coderefinery.github.io/reproducible-research/02-organizing-projects/

Exercise 4: Folder structure

Think – Share – Compare

Think about the ideal folder structure for your research project. Then discuss in pairs or as a smaller group.

4 min

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

Do not use \ / ~ ! @ # $ % ^ & * ( ) ` ; : < > ? . , [ ] { } ' " |

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

Do not use \ / ~ ! @ # $ % ^ & * ( ) ` ; : < > ? . , [ ] { } ' " |

20220503_DOEProject_DesignDocument_Smith_v2-01.docx

Versioned data

Versioning refers to saving **new copies** of your files when you make changes so that later you can go back and **retrieve** specific **versions** of your files

DataFileName_1_0 = original document

DataFileName_1_1 = original document with minor revisions

DataFileName_2_0 = document with substantial revisions

Version control

Perfect for **collaboration**

Tracking and managing changes to a file or set of files over time

You will be able to **recall any version** at any time

It can be used for documents, software development, large websites etc.

Version control table

Manual versioning of changes in your files

Version	Date	Changes
v01	2023-02-21	KM created the presentation
v02	2023-04-05	JS added the section on informed consent
v03	2023-05-17	KM deleted the section on file formats

Marcoux et al 2023 https://zenodo.org/records/10210111

Change log:

- Added
- Changed
- Removed

Version control system

Git: Free and open-source version control system

https://git-scm.com/

https://github.com/

https://youtu.be/gY2JwRfin1M

Version control system

Git: Free and open-source version control system

GitHub: is an internet hosting service for software development and version control using Git

https://git-scm.com/

https://github.com/

https://youtu.be/gY2JwRfin1M

Standardised (open) file formats

- Recognized as having best guarantees in terms of usability, accessibility and sustainability in the future
- Examples:
  - Text: .txt; .pdf (PDF/A) **not** .docx
  - Spreadsheet: .tsv, .csv **not** .xlsx
  - Audio: .wav, .aiff, .flac .mp3 **not** .m4a, .ape, .ogg, .wma
  - Image: .tiff, .png, .jpg **not** .psd, .gif, .bmp
  - Video: .mp4 **not** .avi, .mov
  - Statistical analysis: .R, .Rdata, .dat, .sps, .do **not** .por, .sav, .dta
- Trusted data archives often have guidelines and recommendations for this:
  - DataverseNO: https://site.uit.no/dataverseno/deposit/prepare/#preferred-file-formats
  - Data Archiving and Networked Services (DANS): https://dans.knaw.nl/en/file-formats/

Data documentation

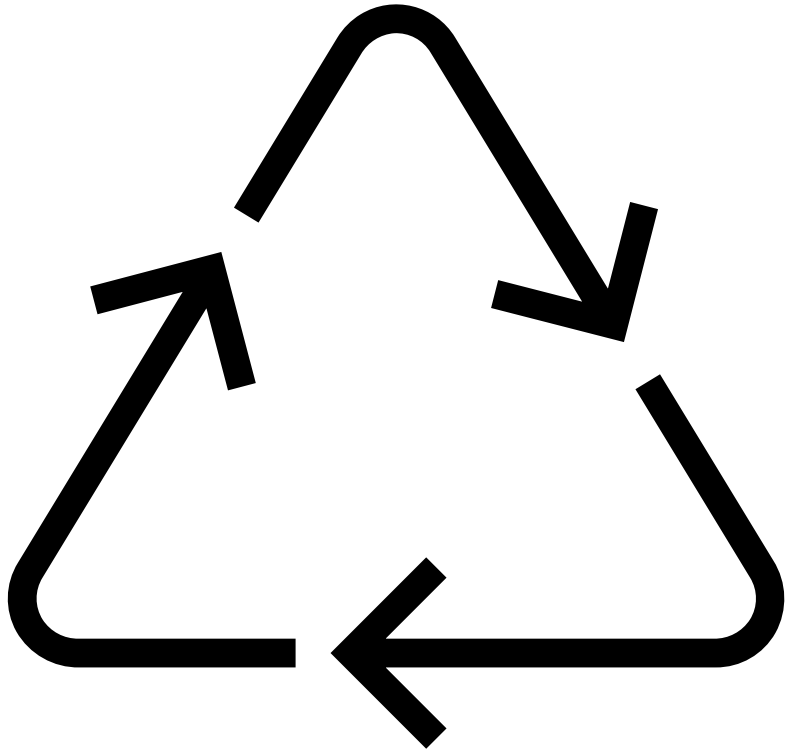
Documentation

- Systematically documented research data is key to making the data publishable, discoverable, citable, and reusable (and FAIR)
- Clear and detailed documentation **improve the overall data quality**

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/Documentation-and-metadata

Why documentation?

- Helps others (and «future you») to understand your project and reuse your data:
  - all you need to know about your data
  - all your collaborators need to know about your data
  - all you need to know about your collaborators data
  - all anyone else who wants to work with your data needs to know about it
- You must have documentation when you archive your data



Rockenberger 2020, Shut Up And Write Documentation: README.txt https://zenodo.org/record/3778273

What documentation?

- The **project-level documentation** explains the aims of the study, what the research questions/hypotheses are, what methodologies were being used, what instruments and measures were being used, etc.
- **Data-level or object-level documentation** provides information at the level of individual objects such as images or variables in a database/table or transcripts, etc.

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/Documentation-and-metadata

Examples of documentation

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

https://mantra.ed.ac.uk/preparingyourdataforarchiving/

Data-level documentation: **Codebook**

- Explains all variables and their codes in the dataset
- It typically contains:
  - variable names, variable labels, variable codes, variable formats, missing data (in quantitative research)
  - codes, code definitions, examples of what to include with a given code (in qualitative research)
- Can be also called **Data Dictionary**

Data-level documentation: **Codebook**

Example

Sex: biological sex self-reported by the participant (0 – Female, 1 – Male)

Gender: self-reported gender (0 – Female, 1 – Male, 2 – Other)

Age: age in months at the date of test (automatically calculated in the system as the difference between the date of test and the date of birth)

	A	B	C	D	E	F	G
1	subjid	sex	gender	recruitment	condition	group	age
2	10001	1		2	9999	9999	272
3	10002	1	1	1	2	2	187
4	10003	1		1	9999	9999	291
5	10004	1	1	1	1	1	300
6	10005	1	1	1	1	1	288
7	10006	1		6	0	9999	254
8	10007	1	1	4	1	2	233
9	10008	0	2	9999	0	0	266
10	10009	1	1	2	1	1	195
11	10010	1	2	2	1	1	207
12	10011	1	1	1	1	2	234
13	10012	1		2	1	2	231
14	10013	1	1	0	1	2	257
15	10014	0	2	0	0	0	309
16	10015	0	2	1	2	2	254
17	10016	0	2	5	0	9999	161
18	10017	1		6	0	9999	
19	10018	1	1	1	1	2	314
20	10019	1	1	5	0	0	188
21	10020	1		0	1	9999	
22	10021	1	1	4	9999	9999	
23	10022	1		1	1	2	294
24	10023	1	1	0	9999	9999	
25	10024	1	1	1	1	1	291
26	10025	1	1	2	1	1	214

Controlled Vocabularies

- Standardized and organized arrangements of preferred terms and phrases used to index and, subsequently, retrieve data and other content
- Provide a consistent way to describe data and their variables
- Common types include term lists, authority files, and thesauri

https://guides.lib.utexas.edu/metadata-basics/controlled-vocabs

Controlled Vocabularies

	A	B	C	D	E	F	G
1	participant number	age	lifestyle	religion	sex	salary group	occupation group
2	1	45	1	0	1	A	B
3	2	54	2	1	1	A	D
4	3	36	1	0	2	B	A
5	4	47	1	NA	2	C	C
6	5	30	2	0	0	B	A
7	6	32	3	0	1	D	B
8	7	52	3	0	1	A	NA
9	8	34	1	1	2	A	A
10							

	A	B	C	D	E	F	G
1	ID	yob	habits	beliefs	gender	income	profession
2	101	1971	2	1	1	A	1
3	102	1983	4	0	2	B	2
4	103	1977	3	0	1	C	4
5	104	1974	1	2		A	2
6	105	1987	1	1	2	A	1
7	106	1994	3	0	1		1
8	107	1990	2	2	2	C	2
9	108	1988	1	3	1	B	3
10	109	1982	4	1	0	C	4
11							

How to create documentation?

- **More information** is better than less
- **Structured information** is better than unstructured information
- **Re-use templates** for good documentation
- It's become a convention to **create multiple README-files**, both for project-level documentation and for data-level documentation

Rockenberger 2020, Shut Up And Write Documentation: README.txt https://zenodo.org/record/3778273

How to create documentation?

- README.txt-files:
  - **Simple text files** that that provide information about any elements (or any workflows or processes these elements have been used for) in the same directory
  - **The first files to open**
  - **Provide a map** for exploring your files
  - Create **one** README.txt file **per folder**
  - Content should **help navigation** through digital files and folders
  - A **project-level** README.txt should give the **general project information** and a very coarse overview of file contents and locations
  - A **data- (object-) level** README.txt would be **more specific** as to what each file contains

Briney 2014, Wrapping Up A Project, http://dataabinitio.com/?p=344

Documentation vs. metadata?

"While data **documentation** is meant to be read and understood by **humans**, **metadata** (which are sometimes a part of the documentation) are primarily meant to be processed by **machines**."

https://howtofair.dk/how-to-fair/metadata/

Exercise 5: Documentation

Shut Up and Write

Write about the origin of the data for 2 min.
Next 2 min write about data storage with focus on a data class, format, size…

4 min

Protection
Organization
Documentation (and metadata)
Classification
Short-term storage
Licensing
Sharing
Long-term preservation
Dissemination (publishing)
Reusing

Courses in research data management and sharing

Spring 2024

Time and place: Apr. 17, 2024 9:00 AM – 12:00 PM, University of Oslo, Blindern, Georg Sverdrups Hus, DSC

Share, archive, and reuse research data

In this half-day course in research data management module 2, participants will gain insight into the potential in archiving and reusing research data in line with the FAIR (Findable, Accessible Interoperable, Reusable) principles.

Time and place: May 22, 2024 9:00 AM – 12:00 PM, Niels Henrik Abels hus - Rom 209

Planlegging av forskningsdatahåndtering

I dette halvdagskurset i forskningsdatahåndtering (modul 1), vil deltakerne få kjennskap til data organisering, lagring, dokumentasjon og skriving av datahåndteringsplaner.

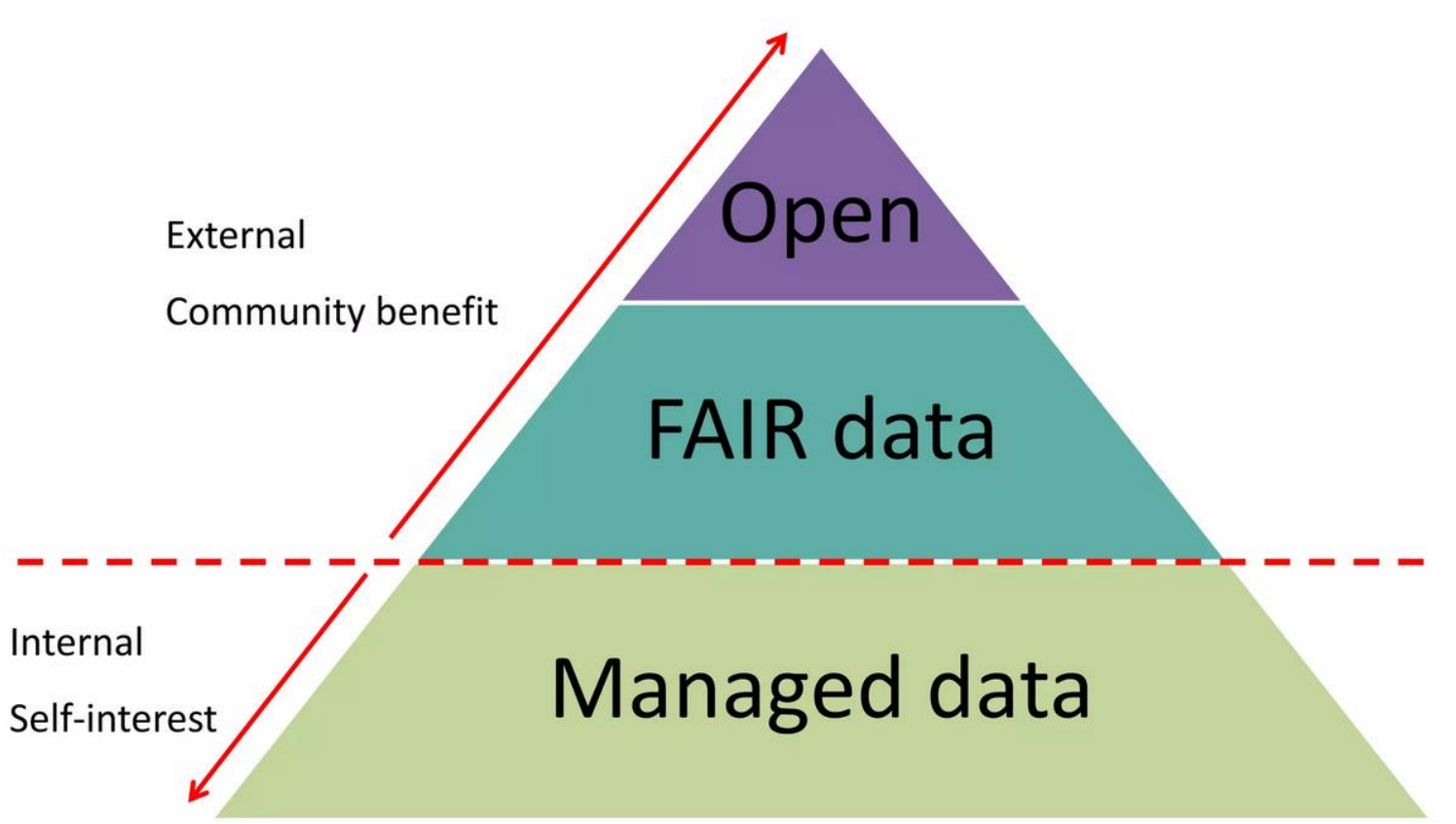
Time and place: May 24, 2024 9:00 AM – 12:00 PM, UiO, Blindern, Georg Sverdrups Hus, Senter for digital forskersøtte (DSC)

Del, arkiver og gjenbruk forskningsdata

I dette halvdagskurset i forskningsdatahåndtering modul 2, vil deltakerne få kjennskap til potensialet ved deling av forskningsdata i tråd med FAIR prinsippene om gjenfinnbare, tilgjengelige, interoperable og gjenbrukbare data.

https://www.ub.uio.no/english/courses-events/courses/other/research-data/

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



Source: https://www.slideshare.net/sjDCC/open-fair-data-and-rdm

Need help with research data?

Send an email to research-data@uio.no

More resources?

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

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

Menu

← Libraries and centres

https://www.ub.uio.no/english/libraries/dsc/

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.

Open and reproducible research

Learn about how to make your research more open and reproducible and get involved in initiatives and communities that are interested in sharing and improving research at UiO.

More and more researchers and students across disciplines are implementing open research practices, preregistering their hypotheses, methods, and analysis plans and sharing research materials, data and analysis scripts. Digital Scholarship Center can help you learn about and implement these practices in your own research as well as advise on the policies and requirements from funders.

Open Science Lunch →

Every last Thursday of the month we meet at noon to discuss topics related to open research.

ReproducibiliTea@UiO →

Join us for a Journal Club where we read and discuss papers on open research and meta-science.

Norwegian Reproducibility Network →

Join a broader community that aims to promote and enable rigorous, robust and transparent research practices in Norway

Courses and workshops →

Click here for the list of upcoming and previous courses and workshops on open and reproducible research at UiO.

https://www.ub.uio.no/english/libraries/dsc/open-repro-research/

Open Science Lunch

Each last Thursday of the month at 12:00 we invite you to join us for a lunch seminar to hear about how to make your research more open. We will discuss research transparency and visibility, open publishing, data sharing, and more!

Time and place: Apr. 25, 2024 12:00 PM – 1:00 PM, Zoom

Researcher Assessment

Join us for this Open Science Lunch to learn about the work on reforming research assessment in Norway and hear about experiences from implementing the CoARA commitments.

https://www.ub.uio.no/english/courses-events/events/open-science-lunch/

UNIVERSITETET
I OSLO

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 →

https://www.ub.uio.no/english/libraries/dsc/open-repro-research/reproducibilitea/

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!

UNIVERSITETET I OSLO | THE CARPENTRIES | DATA CARPENTRY | software carpentry | library carpentry

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.

https://www.ub.uio.no/english/courses-events/courses/other/Carpentry/

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!

UNIVERSITETET I OSLO | THE CARPENTRIES | DATA CARPENTRY | software carpentry | library carpentry

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

https://www.ub.uio.no/english/libraries/dsc/carpentry-uio/index.html

Det senteret for digitalforskerstøttes nyhetsbrev,
en del av Universitetsbiblioteket i Oslo

The Digital Scholarship Centre's Newsletter,
part of the University of Oslo Library

DSC NEWS

Senter for digitalforskerstøtte
Digital Scholarship Centre

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

Give us your feedback!

https://nettskjema.no/a/423496

Thank you!

Elisa Pierfederici, Edina Pózer, Ivana Malovic
Digital Scholarship Center, University of Oslo Library
16.04.2024

Questions? Contact us at research-data@uio.no