

Data classification and storage selection

Agata Bochynska and Ivana Malovic
The University of Oslo Library



Agenda

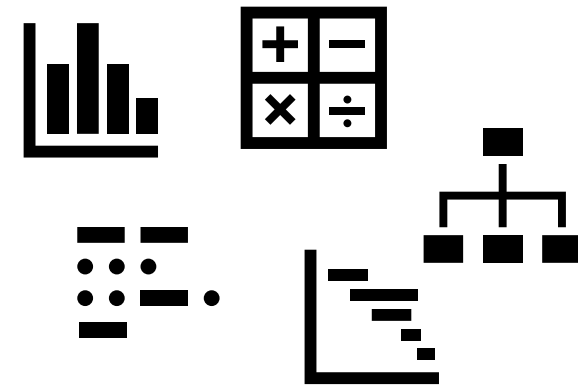
- What types of data do you have?
- Data categories and their colour-coding
- Data storage based on data category
- Q&A session



Types of research data

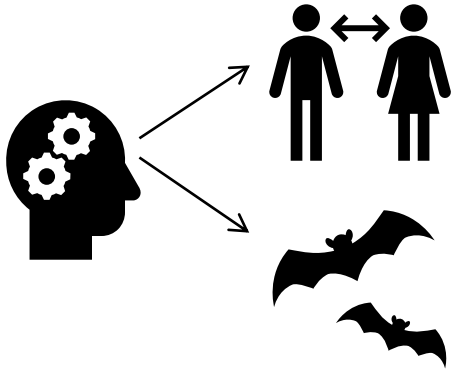


Qualitative (descriptive)

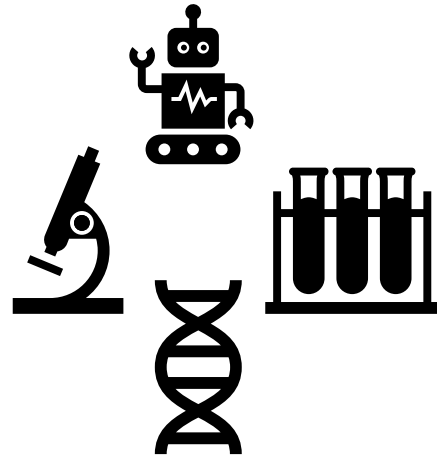


Quantitative (numerical)

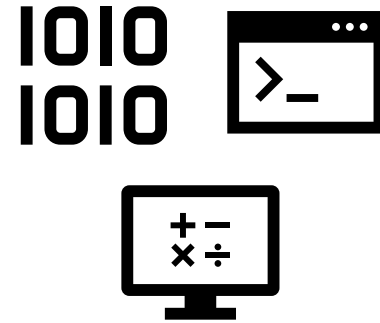
Types of research data



Observational



Experimental



Computational



Records

Types of research data

Personal data or not?

Types of research data

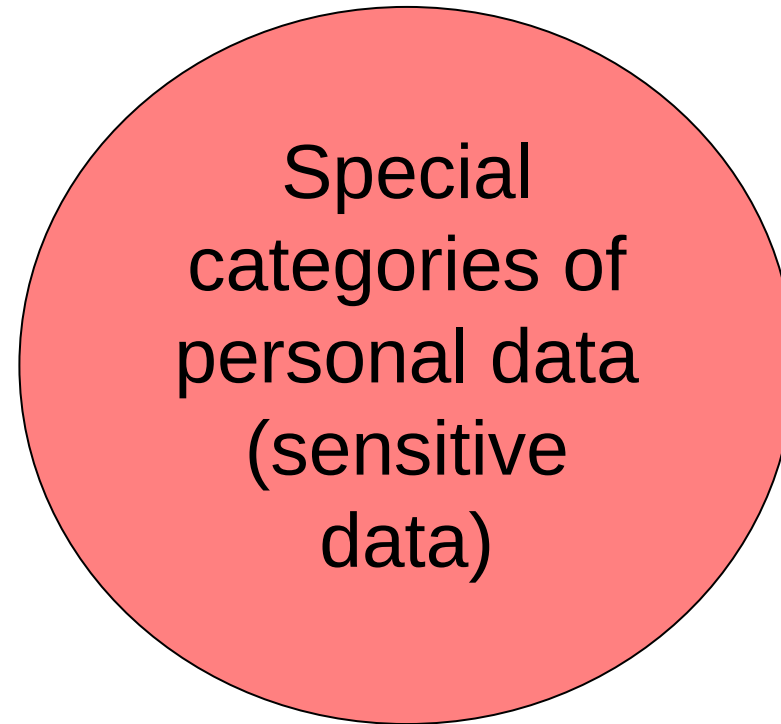
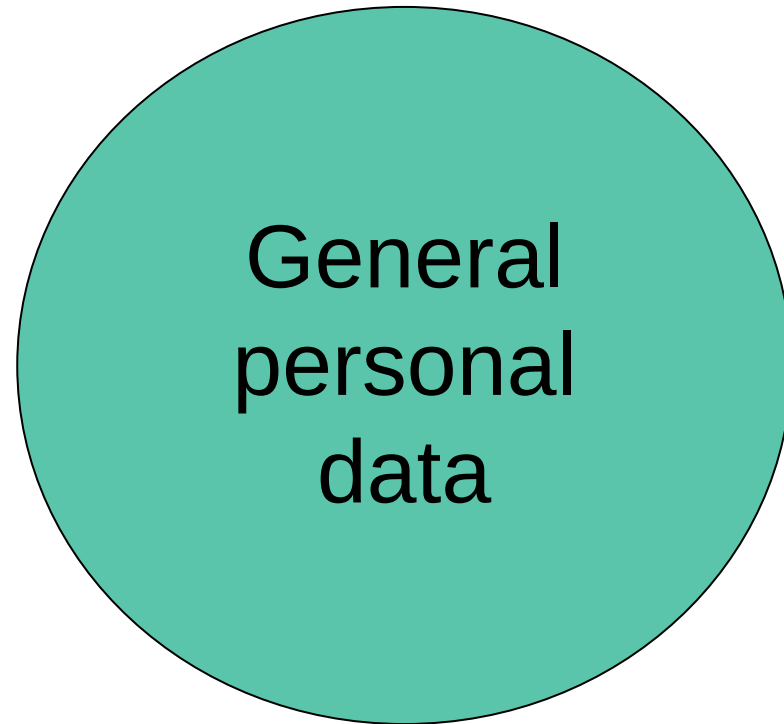
What is personal data?

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.

Types of research data

What is personal data?



Types of research data

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
- Content of case documents, investigations or considerations
- Content of exam papers or grades
- Contents of email communication
- Logging of activity in computer systems where the logs can be associated with individuals

Types of research data

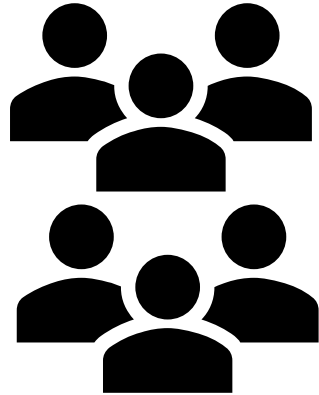
Special categories of personal data (sensitive data)



- 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

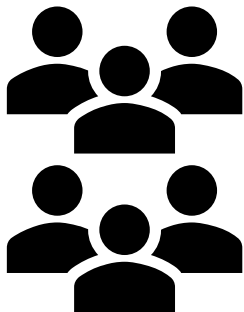
Types of research data

Anonymous data or not?



What are anonymous and de-identified data?

Anonymous data means that NO ONE can identify persons in a dataset. Anonymous data are not seen as personal data. Data can in some cases be anonymised by deleting any information that contributes towards identification.



De-identified data means that information used for identification have been replaced by a code with an associated key. De-identified data are seen as personal data. If you delete the associated key, de-identified data can be anonymous.



On ownership

- All UiO information shall have **an unambiguous and identifiable owner**.
- Anyone should **easily** be able to **find out who is responsible** for keeping the information updated, maintained and correctly labeled.
- **The «information owner» is responsible for the assessment** used to place the information in its given category.
- In those cases where an owner cannot be identified, the University Director is responsible for the information.

On ownership – How to decide which class to use?

- **Always put the information in a sufficiently safe class.**
- If you are not sure whether to your information is red or yellow, choose red.

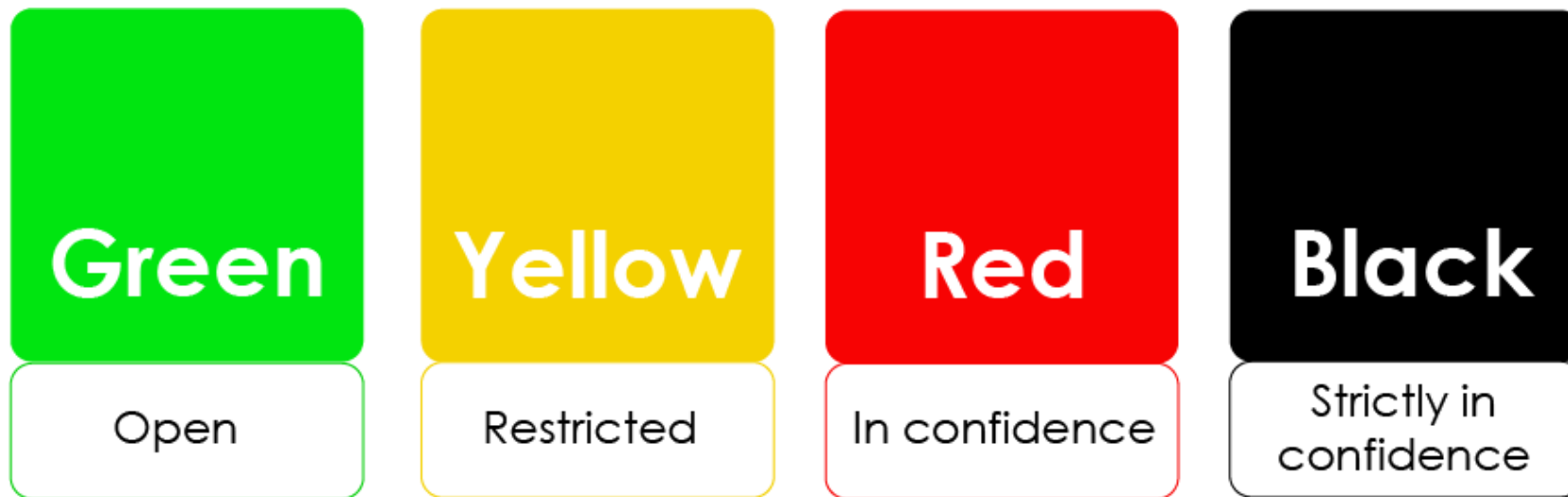




DATA CLASSIFICATION

Classifying your data

- Data classification at UiO is based on how harmful for public interest, the university or individuals it would be if the data becomes exposed



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

What is **green** data

- **Open or freely available** to the general public
- This class is used when it **does not cause any harm** to the institution if the information becomes known to unauthorized persons.
 - Any research that is published openly. Information that is not public, needs to be removed first.
 - Teaching material that is not subject to copyright and does not need any protection
 - Completely anonymous data
 - Statistics



What is **yellow** data?



- **Restricted** - basically information which is not open for everyone
- This class is used when it **could cause limited harm** to the institution if the information becomes known to unauthorized persons.
 - Documents that you don't want everyone else to read
 - De-identified data where the key is locked away and stored safely away from the data
 - Data sets containing minor amounts of non-sensitive personal data
 - E-mail attachments that don't contain information that needs protection
 - Unpublished research material
 - Texts and pictures under copyright
 - Examination answers and grades

What is **red** data?

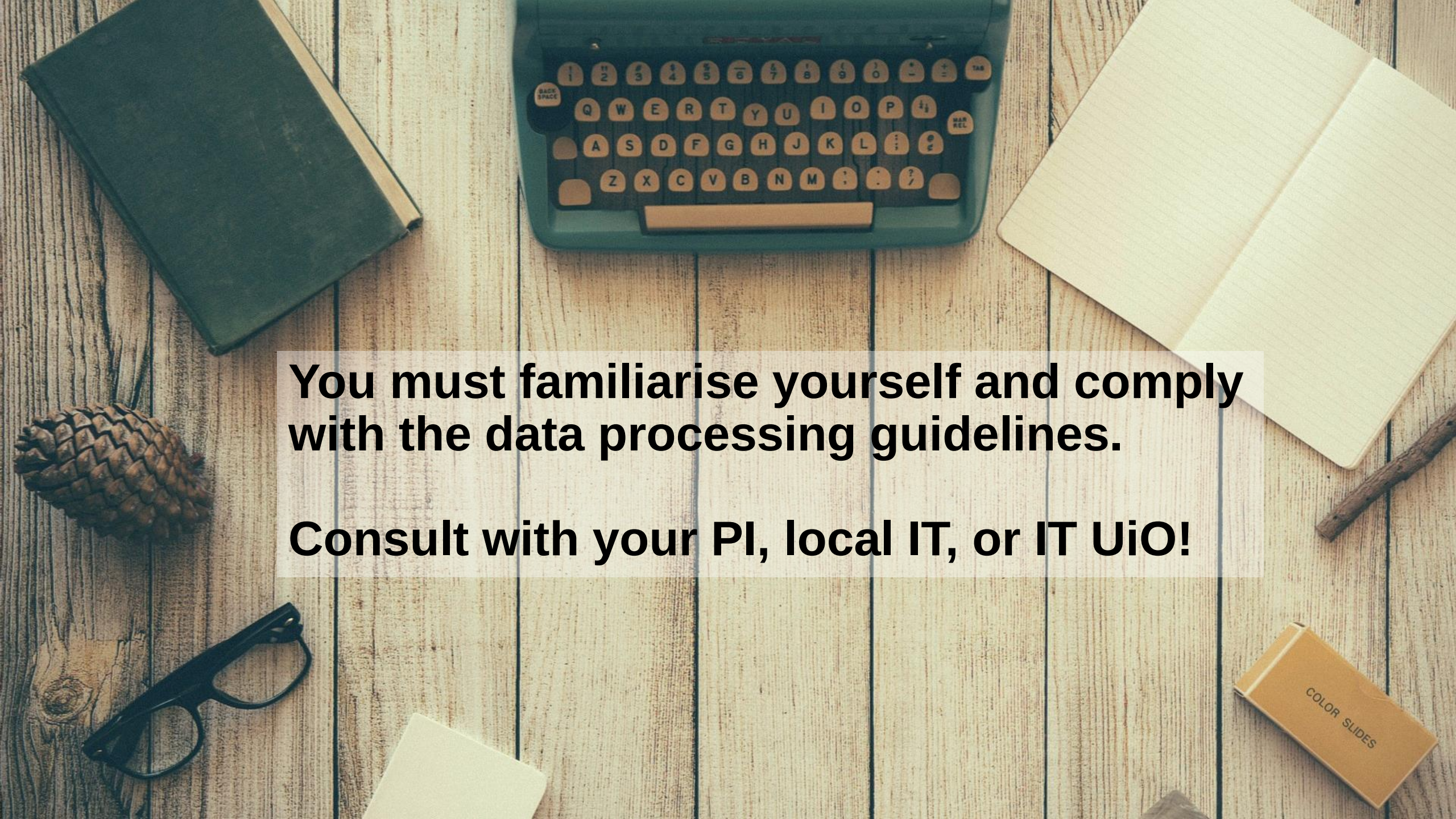
- **In confidence** - the university is obliged to protect by law, agreements and other regulations
- This class is used if it **could cause harm** to public interests, the university or individuals if the information becomes known to unauthorized persons.
 - De-identified data where the key is available and stored with the data
 - Special categories of personal data
 - Personnel files
 - Information relating to e.g. safety systems in buildings or IT systems
 - Health related data



What is **black** data?



- **Strictly in confidence** - special circumstances makes it necessary to protect the information even more
- This class is used if it **could cause considerable harm** to public interests, the university or individuals if the information becomes known to unauthorized persons.
 - Research data that requires extra protection
 - Large quantities of sensitive personal data
 - Large quantities of health related data
 - Research data that is of great financial or commercial value



You must familiarise yourself and comply with the data processing guidelines.

Consult with your PI, local IT, or IT UiO!

A photograph of a large server room with rows of server racks. The racks are filled with server units, and the perspective is looking down a long aisle. The text "STORAGE SELECTION" is overlaid in the center.

STORAGE SELECTION

Where can you store the different types of data?

Data storage guide:

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



Where can you store the different types of data?

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



Where can you store the different types of data?

- **Yellow data:**

- Encrypted memory stick / external hard drive
- Encrypted laptop maintained by UiO
- UiO cloud solutions (UiO One Drive, UiO Google Suite for Education, ...)
- UiO storage hotel («lagringshotell»)
- UiO personal researcher storage
- UiO e-mail; UiO Teams; UiO Zoom; UiO Nettskjema; UiO Canvas, UiO Vortex, UiO Wiki...
- **Not** your private laptop, unless you comply with the [guidelines for use of private computer](#)



UiO commercial cloud solutions

- UiO OneDrive (Microsoft 365)

<https://www.uio.no/english/services/it/store-collaborate/ms365/>



- UiO Google Suite for Education

<https://www.uio.no/english/services/it/store-collaborate/gsuite/>



Where can you store the different types of data?

- **Red data:**

- Fully encrypted disc, memory stick or external harddrive
- Services for sensitive data – TSD
- Educloud Research
- UiO Storage Hotel*
- UiO Personal researcher storage*
- UiO Vortex*



*see Storage guide for details!

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

Where can you store the different types of data?

- **Black data:**
Services for sensitive data – **TSD**
 - «Snapshot» backup for last 7 days



Educloud Research

= «TDS light»

- Project-oriented self-service platform that gives easy and fast access to:
 - A work environment for research projects that have cross institution and cross border users
 - Storage that is available from anywhere, from any device and shared between project users and project platforms
 - A low threshold HPC system that offers batch job submission and interactive nodes, and that has access to the same storage as the Windows and Linux resources available to the project
- For **green**, **yellow** and **red** data
- Free for UiO researchers at least until January 2024
- More information at:

<https://www.uio.no/english/services/it/research/platforms/edu-research/>

<https://www.uio.no/english/services/it/research/platforms/edu-research/more/>

<https://www.uio.no/english/services/it/research/platforms/edu-research/order.html>

<https://www.uio.no/english/services/it/research/platforms/edu-research/help/> (User Manual)



Sign up - Educloud research

Choose one option



Request a new project

Only for project administrators and currently only available for users with a Norwegian electronic ID.

Go to Nettskjema



Apply for access to a project

Currently only available for users with a Norwegian electronic ID.

Log in (ID-porten)

Back

Skjemaet lagrer personinformasjon.

Svarene lagres i Tjenester for sensitive data (TSD). [Les mer](#) ▼

Registrering av nytt prosjekt i Educloud Research

0 %



Obligatoriske felter er merket med stjerne *

Informasjon om deg som prosjektadministrator

Du som fyller ut dette er enten Principal Investigator (PI) i selve prosjektet eller har fått fullmakt fra PI. Oppgavene som prosjektadministrator skal utføre når prosjektet er registrert i EduCloud Research er innmelding og utmelding av brukere i prosjektet, endre rettigheter, utvide ytelser mm. Dersom du fyller ut dette skjemaet med fullmakt fra PI må du bruke din egen BankID



Sign in

Username

Password

[Forgot your password?](#)

One time code

[Need a new one time code?](#)

Sign in

Do you not have an account yet?

Get started with Educloud



 Projects •

Support

 Settings

 Help

Projects

1 project

Name		Status	
ec32		Active	Select project

 Projects

Project

 Users

 Fox users

 Windows users

 Desktop •

Support

 Settings

 Help

Desktop
Connect to your virtual desktop

Available desktops

 Educcloud VMware Connect



VMware Horizon

You can connect to your desktop and applications by using the VMware Horizon Client or through the browser.

The VMware Horizon Client offers better performance and features.



Install VMware
Horizon Client



VMware Horizon
HTML Access

☐ Check here to skip this screen and always use HTML Access.

To see the full list of VMware Horizon Clients, [click here](#).

For help with VMware Horizon, [click here](#).





VMware Horizon

Enter your Sentry user name and One-Time Code

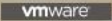
Login

Cancel

[Help with VMware Horizon](#)

[Privacy Policy](#)

[Download the Horizon Client for Windows \(x64\)](#)

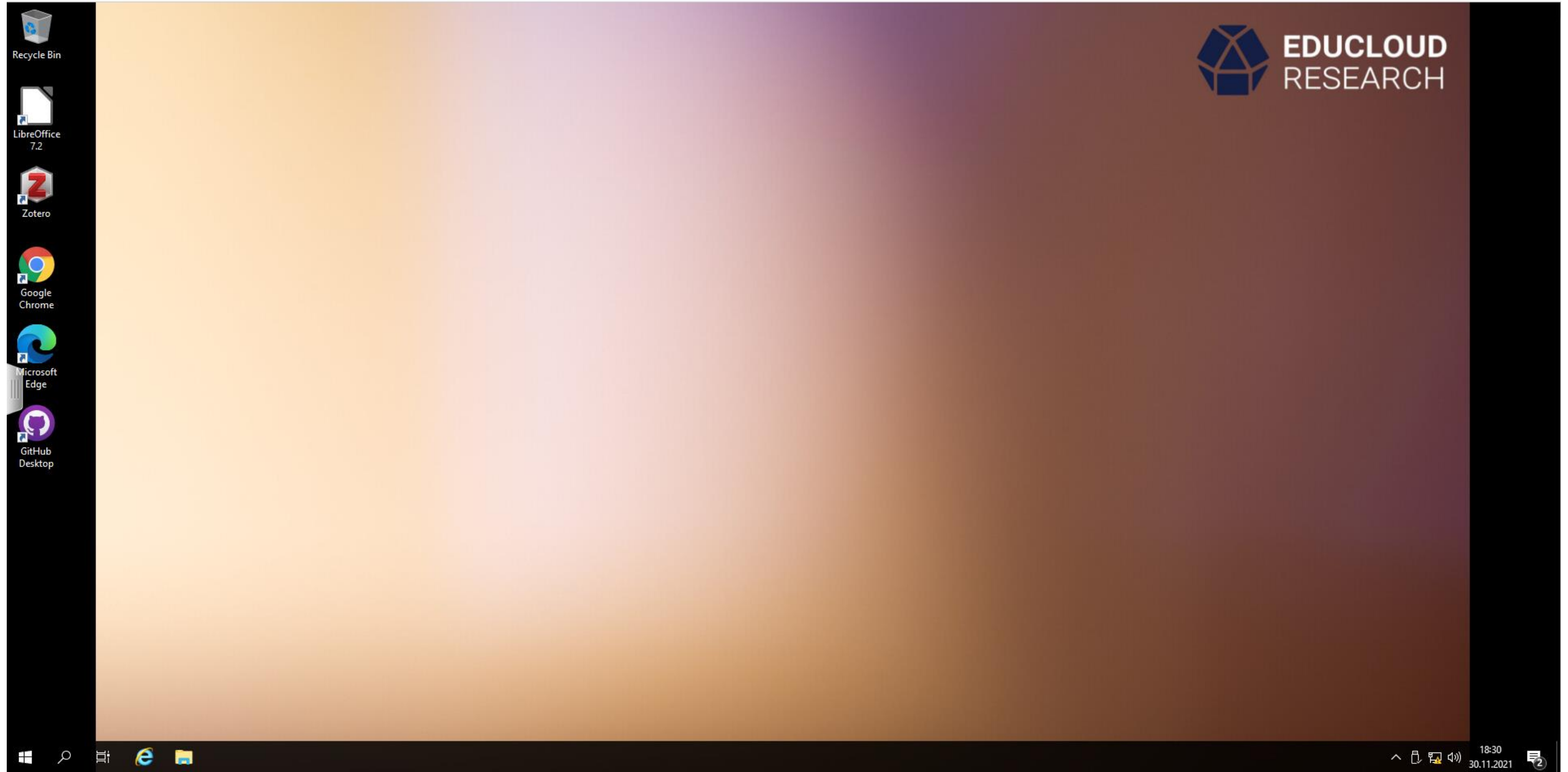




🔍 Search



EC Windows
Desktop





File Explorer

File Home Share View

Quick access

Quick access

- Desktop
- Downloads
- Documents
- Pictures
- This PC
- Network

Frequent folders (4)

- Desktop This PC
- Downloads This PC
- Documents This PC
- Pictures This PC

Recent files (0)

After you've opened some files, we'll show the most recent ones here.

4 items

Where can you find an overview of all cloud services at UiO that are evaluated or under consideration?

https://www.uio.no/tjenester/it/ny-tjeneste/vurderte_skytjenester.html
(in Norwegian)



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?

Courses in research data management and sharing

November 6th – 27th

Upcoming Courses

Copyright and licensing

Nov. 13, 9:00 AM, Zoom

Sharing and archiving research data

Nov. 16, 9:00 AM, Zoom

Finding and reusing research data

Nov. 17, 9:00 AM, Zoom

Upcoming Workshops

Workshop on Data Management Plans

Nov. 20, 12:00 PM, Georg Sverdrups hus: DSC-Oasen

Workshop on Data Documentation

Nov. 23, 9:00 AM, Georg Sverdrups hus: DSC-Oasen

Workshop on Archiving in DataverseNO

Nov. 27, 12:00 PM, Georg Sverdrups hus: DSC-Oasen

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.

Digiforsk blog →

A blog about work and resources at UiO where digital research methods are involved. Posts in both Norwegian and English.

Research data management

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](#) >

Glossary

An overview of [often used terms](#) in research data management.

Laws and ethics

[Information on privacy and data protection at UiO](#)

[Intellectual property rights at UiO](#)

[National guidelines from Research Ethics Committees](#)

Courses and support

[Currently offered courses](#)

Need advice? Feel free to [book an individual appointment](#) or contact us at **research-data@uio.no**

Are you a data manager? Join our [data managers network!](#)



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.

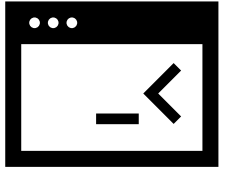


Using Databases and SQL

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.

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

Courses from the IT-department



Events and Courses

Intermediate SPSS

Nov. 20, 9:15 AM, Ole-Johan Dahls hus

Intermediate SPSS

Nov. 21, Ole-Johan Dahls hus

Intermediate SPSS

Nov. 22, Ole-Johan Dahls hus



Materials developed as a part of the *Skills development for research data* project:
<https://www.ub.uio.no/english/about/projects/rdm-skills/>

Thank you!

Contact us at:

research-data@uio.no