UiO : Universitetet i Oslo

GDPR and safe data management

Anne S. Bergsaker and Annika Rockenberger
IT for research, USIT and the University Library

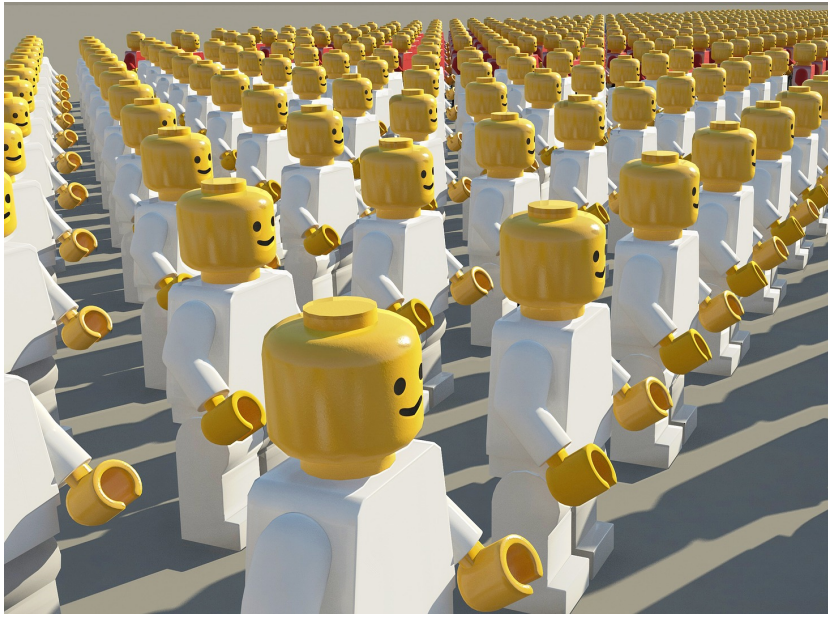
Photo by Simen Kjelling/UiO

GDPR – General Data Protection Regulation

- New (implemented 2018), more precise, common rules for data protection and privacy, set by EU
- ENORMOUS FINES if you do not comply

Privacy and personal data

- Personal data is any information relating to an indentified or identifiable natural person
- Differ between regular personal data and special categories of personal data (sensitive information)



What is considered personal data?

- Name, address, age, phone number, email, social security number (personnummer)
- Contents of exam papers, bachelor or master theses and the student's grades
- Contents in case files, employee or student files
- Contents of email communication between students, or between students and lecturer/supervisor
- Video and sound recordings where someone is recognizable
- Pictures of staff or students published on the university website
- Activity logs in computer systems where logs can be connected to specific users, e.g. when someone is logged into a service

What are considered special categories?

- Information relating to health
- Genetic and biometric data, used to identify a specific person
- Ethnicity
- Political, philosophical or religious views
- Sexual orientation or sexual relations
- Workers' union memebership

Consent

A consent is valid if and only if it is:

- Voluntary
- Specific
- Informed
- Unambiguous
- Actively given
- Documented
- Retractable

If you wish to use your data for a new purpose, then a new consent must be given by all participants.

You have to ask for consent if you wish to not have to delete the data at the end of your project.

What is considered as handling of personal data?

- Collecting
- Registering
- Storing
- Combining
- Using
- Transferring
- Publishing
- Deleting

De-identified and anonymous data

De-identified data means that identifying information has been replaced by a code with a corresponding key. De-identified data is still considered to be personal data. If you delete the key, de-identified data can in some cases become anonymous data.

Anonymous data means that no one can recognize the people in the data set. Anonymous data is not personal data. Personal data can in some cases be anonymized by deleting identifying information.

De-identified and anonymous data

De-identified data means that identifying information has been replaced by a code with a corresponding key. De-identified data is still considered to be personal data. If you delete the key, de-identified data can in some cases become anonymous data.

Anonymous data means that no one can recognize the people in the data set. Anonymous data is not personal data. Personal data can in some cases be anonymized by deleting identifying information.


Can your data be de-identified?
Can your data be anonymized?

Re-identification – when de-identification is not permanent

«I know where you were last summer»: Single people re-identified based on their use of public city bikes

Kilde: https://vartree.blogspot.com/2014/04/i-know-where-you-were-last-summer.html

What is your responsibility now, before you have collected your data?

- Ensure that the privacy of all participants will be taken care of safely and securely
- Be aware of and keep track of all personal data that will be collected in your study
- Creat consent forms and provide participants with information about the study and their data protection
- Report the study to NSD, at least 30 days before data collection starts
- Not start collecting data before you have been given permission from NSD
- Conduct a risk and vulnerability analysis of any IT-solutions you use to handle personal data, that is not a service offered by UiO

Reporting the study to NSD

https://nsd.no/personvernombud/en/notify/meldeskjema?eng

What is your responsibility during the project?

- Answer any questions from respondents regarding privacy
- Ensure that data is deleted or properly anonymized if consent is withdrawn
- Ensure that no personal data is used in any other way than what the consent specified
- Ask for a new consent if the data is used for a different purpose
- Verify that terms in agreements with any external information providers, such as registry owners or partners at other institutions are followed
- Report any irregularities that may occur during the processing of personal data

What is your responsibility after the project is finished?

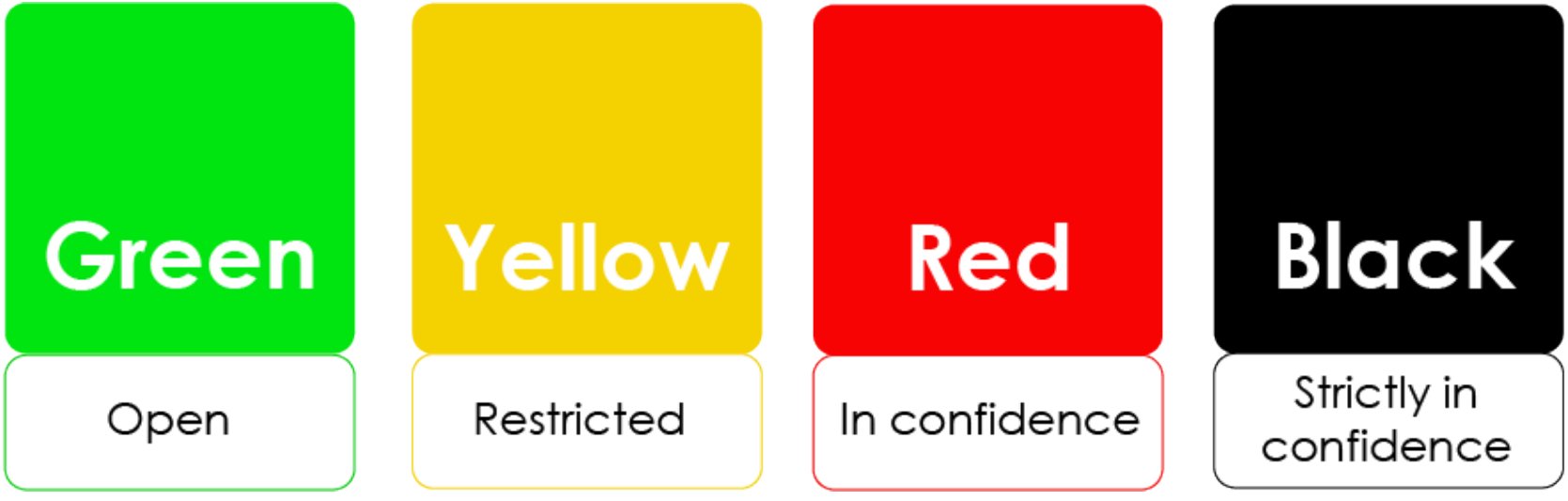
- Determine which personal information about respondents or informants shall be deleted and which information will be kept
- Ensure that all personal information about respondents or informants which will not be retained after the end of the project is properly deleted
- Ensure that personal data that will be retained after the end of the project is properly anonymized
- Ensure that personal information that will be retained after the end of the project is properly stored
- Ensure that the master's theses containing confidential information are properly categorized in DUO

Help and extra material can be found on the UiO website

https://www.uio.no/english/studies/examinations/privacy-protection/

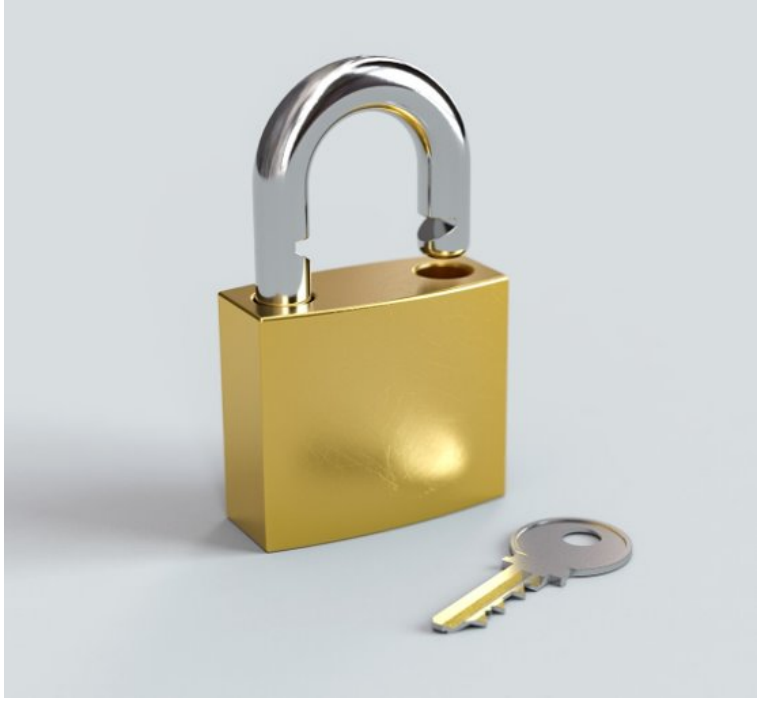
DATA CLASSIFICATION

Classifying your data



What is green data

- This class is used if 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
- Completely anonymous data
- Data without any financial or commercial value



What is yellow data?

- This class is used if it could cause a certain damage 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**

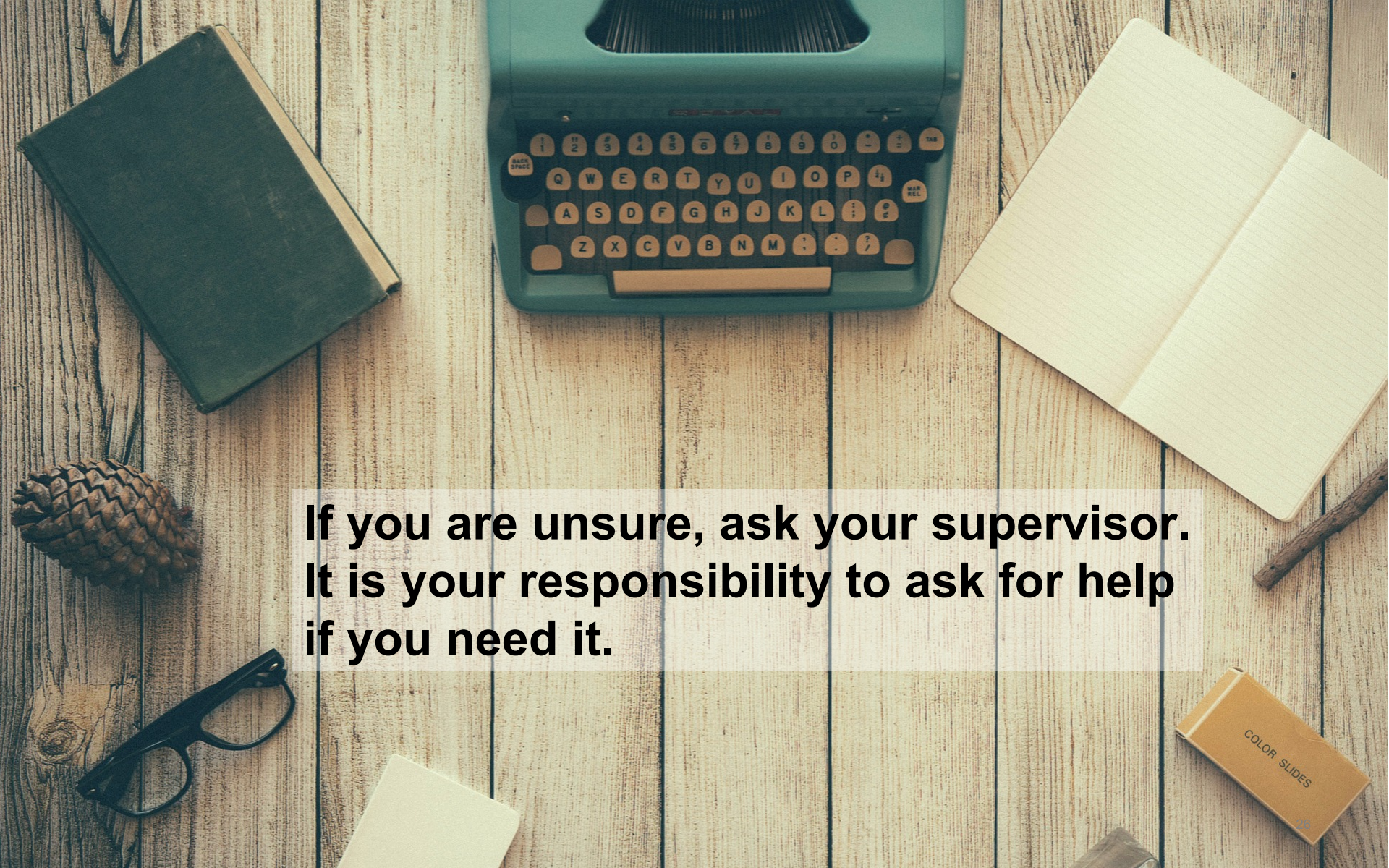
What is red data?

- This class is used if it will 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
- Information relating to e.g. safety systems in buildings or IT systems
- Health related data

What is black data?

- This class is used if it could cause significant 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


What type of data will you have?
COLOR SLIDES


If you are unsure, ask your supervisor.
It is your responsibility to ask for help
if you need it.
COLOR SLIDES

Your data vs your thesis

- Not all the data you collect should necessarily be included in your thesis
- You can include data that participants have consented to publishing, or that may be properly anonymized
- You can have personal information in the thesis, but then you HAVE TO specify this when you submit in DUO, so that the thesis does not become openly available.

Where can you store the different types of data?

- Green data: anywhere you want, but ideally **not only** on your private laptop (very vulnerable)
- Yellow data:
  - encrypted hard drive
  - UiO cloud solution (UiO OneDrive, UiO Google Drive)
  - UiO home directory (M:)
  - **UiO storage hotel**(«lagringshotell»)
  - Not your private laptop or phone, unless they are encrypted and safely used

UiO commercial cloud solutions

- Microsoft OneDrive + Office365

https://www.uio.no/english/services/it/store-collaborate/o365/

- Google Drive + Docs, Sheets, etc

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

G Suite for Education

Where can you store the different types of data?

- Red data:
  - For data that is red, but not highly sensitive:
    - UiO Storage hotel
    - UiO maintained encrypted computer without any automatic syncing with home directories or cloud solutions
  - Red, highly sensitive data:
    - Services for sensitive data - TSD

- Black data: **TSD**

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

Storage hotel

- Your supervisor will help you get access
- Remote disc where you can safely store and access your data
- Safer than the cloud solutions, so you can store up to red data

Guidelines on accessing storage hotel:
https://www.mn.uio.no/geo/english/services/it/help/storage/1_data-storage-at-the-department/lh-howto.html

Storage solutions and 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)
- Only saving files on your own computer is not safe!
- A memory stick or an external harddrive is not safe!
- How much extra work would it be for you to recreate your data or your work?

Save all the files!

- «For every result, keep track of how it was produced» (Sandve et al, 2013)
- Save processed data at important stages in the project
- Save processed data and analysis/methods for every table and graph

Make a system

- File system
  - Hierarchy
  - Divide by data type, collection data, data subject, etc.
  - Informative folder names
  - Avoid overly broad folders
- File names
  - Short and sweet
  - Informative
  - Dates backwards (YYMMDD)
  - Avoid special characters or spaces
  - Use _ or – or capital letter to divide words

Useful links

Data management at ISS:
https://www.sv.uio.no/studier/master/iss/Datahandtering/data-management-at-iss.html

Your GDPR related responsibilities during your thesis work:
https://www.uio.no/english/studies/examinations/privacy-protection/

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

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