



Introduction to Research Data Management

Linde Tuybens & Robin Björklund

RDM Team – Research Affairs Office

ADS DOCTORAL DAYS

19 & 20 OCTOBER 2026

Research Data Management (RDM) Team

- Part of the **Research Affairs Office** at **RIVA – Research Innovation & Valorisation Antwerp**
- Offer advice on all things related to research data management, including data management plans, storage, Nagoya, and Open Science (Open Data, FAIR data)
- Contact us via: rdm-support@uantwerpen.be



- **Siham Benramdane** (Life Sciences & Medicine)
- **Robin Björklund** (Natural Sciences & Engineering)
- **Linde Tuybens** (Social Sciences & Humanities)
- **Dunya Nasser** (RDM inbox)

Research Data

*“Research data are all **digital or physical data** - regardless of the manner in which these data are collected or stored - used or analyzed to **support research findings**, validate research results or underlie a scientific reasoning, discussion or calculation in the study.*

*Research data cover the entire spectrum of **raw data to processed and analyzed data** included or discussed in a publication. These data can be **generated data, derived or composite data**, as well as **self-generated data and data provided by third parties.**“*

RDM Policy of University Antwerp

YOUR RESEARCH DATA



**"No data" to manage?
Think again!**

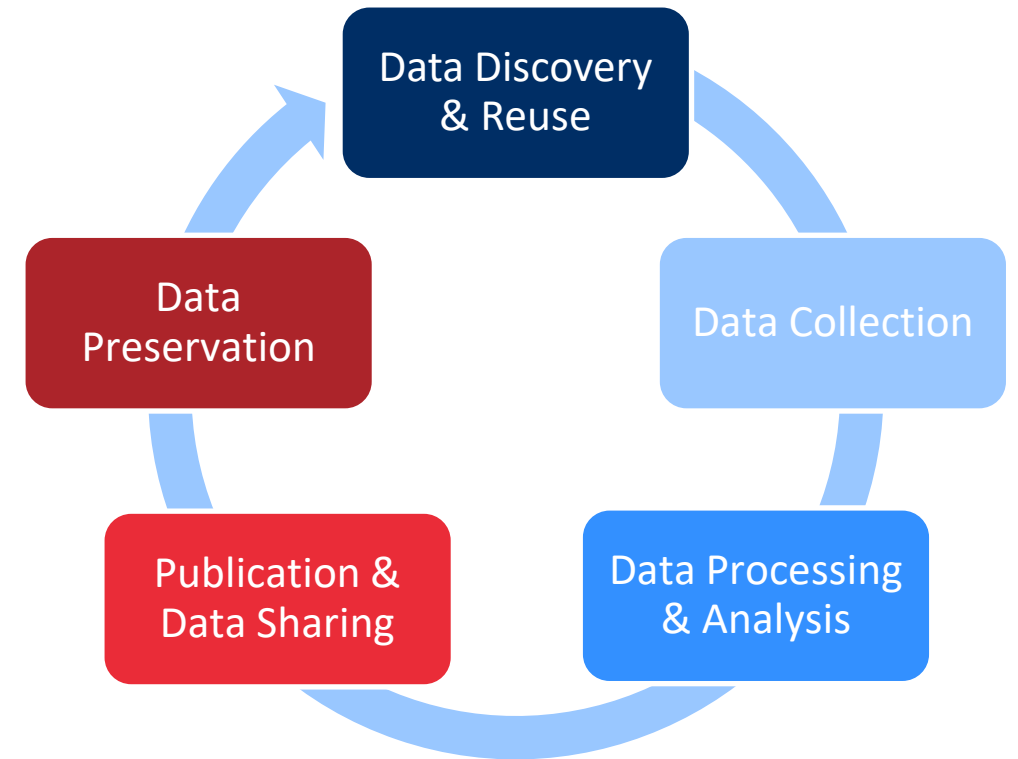
[Cheat sheet for humanities
& theoretical disciplines](#)

Research Data Management (RDM)

A set of actions and practices to take care of research data during the course of your research and beyond

- An integral part of doing research
- Includes all steps of the Research Data Lifecycle
- All steps are bound by conditions and regulations
 - Legal, ethical, financial & technical

RDM = Risk management



Research Data Management – Why?

- **Avoid losing data**
- **Work more efficiently, save time and resources**
 - Find the data you need easily
 - Understand the content of your data in the future
- **Protect sensitive information**
- **Ensure data quality**
- **Share your data**
 - Ensure reproducibility
 - Facilitate collaboration
 - Give your data and research more impact
 - Support Open Science, improve your CV

Research data



Secure



Sustainable



Findable, Understandable,
Reusable

**Your data is valuable,
It is also research output!**

Key concepts of RDM

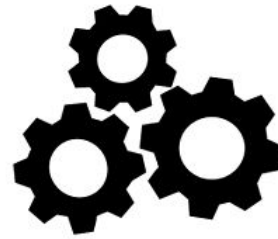
F
indable



A
ccessible

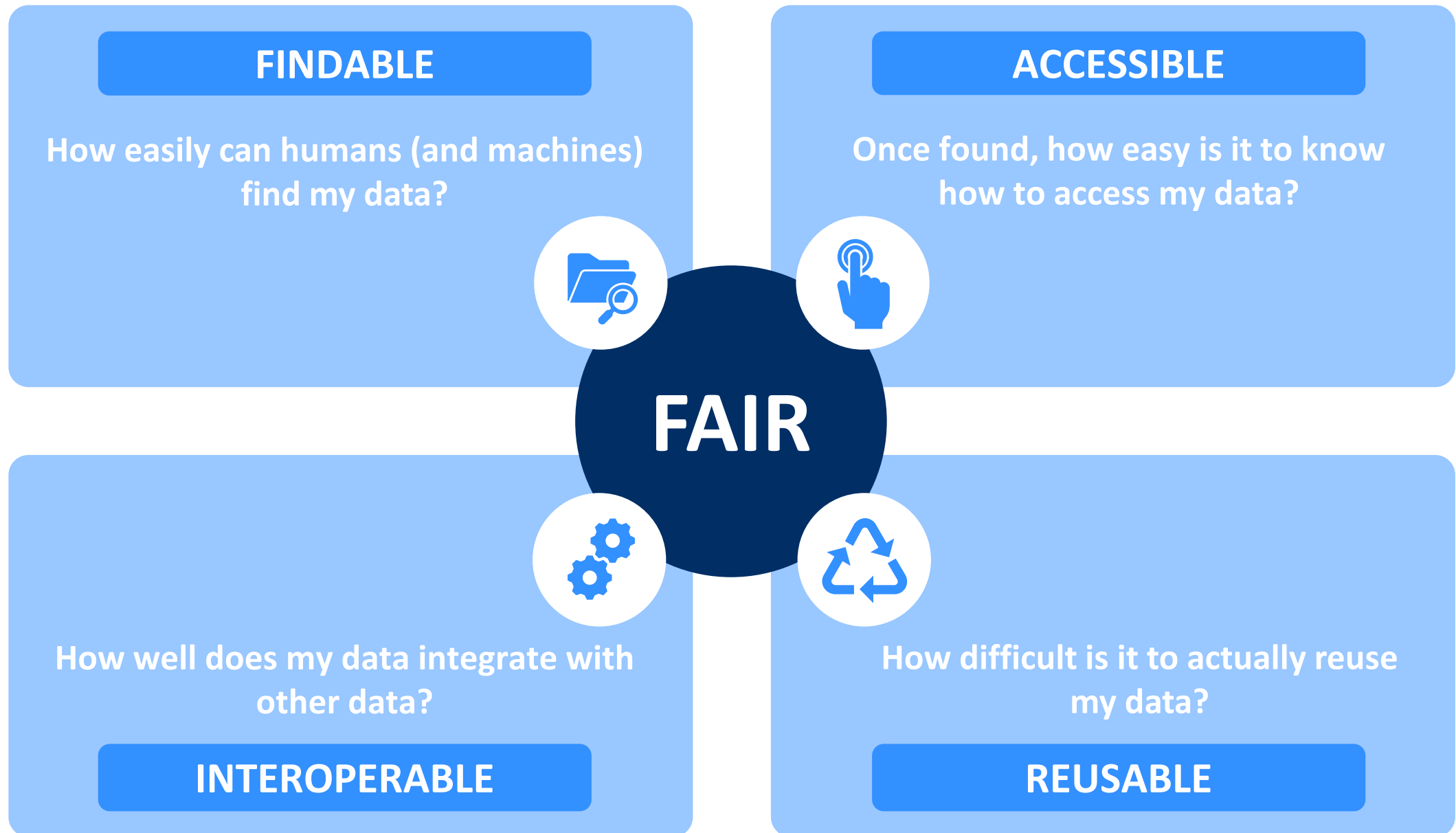


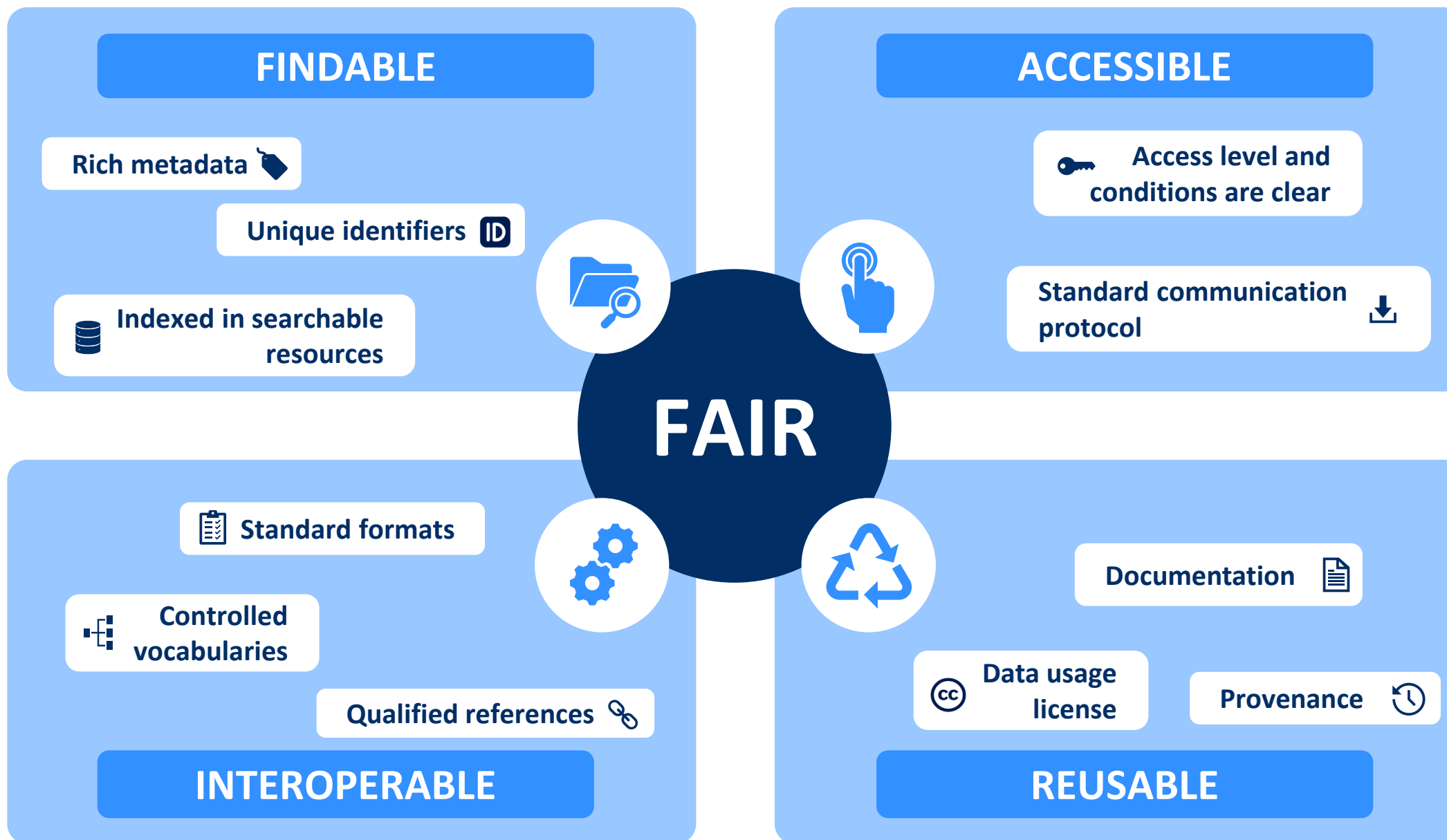
I
nteroperable



R
eusable

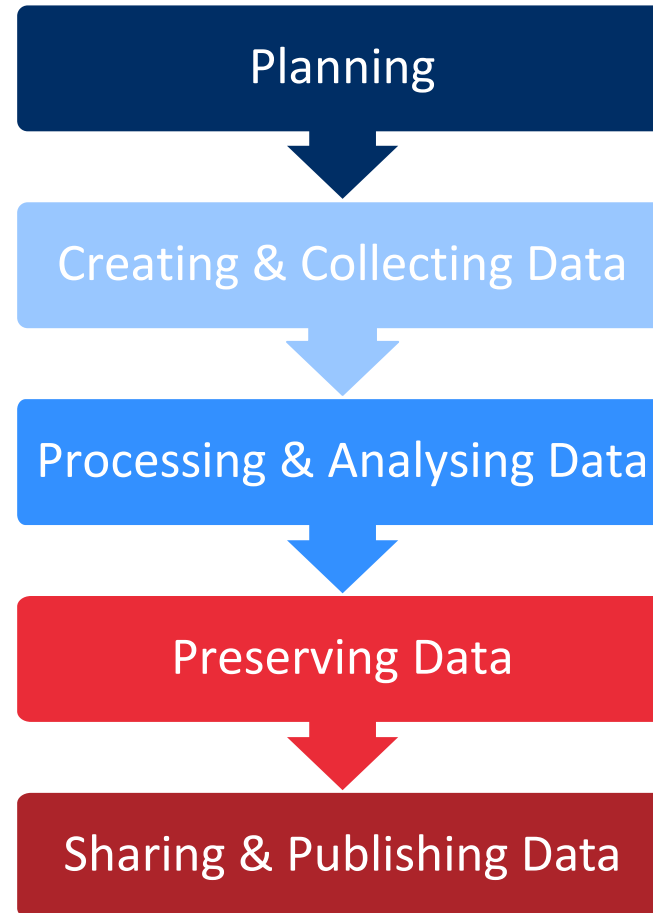






RESEARCH DATA MANAGEMENT

RDM STAGE



RDM PRACTICES

1. PREPARING A DMP

2. ORGANISING & DOCUMENTING DATA

3. STORING DATA & DATA SECURITY

4. PRESERVING & SHARING DATA



Data Management Plan

Formal document describing how research data will be managed during the research lifecycle



DMP Components

1. **Project Information**
2. **Data summary**
 - Overview of data
 - Reuse of existing data
 - Ethical issues, Personal data, Third Party agreements, Other legal issues (IP)
3. **Documentation and Metadata**
4. **Data storage and backup during project**
5. **Data preservation after the project**
6. **Data sharing and reuse**
7. **Responsibilities**

DATA SUMMARY



DOCUMENTATION
& METADATA



DATA STORAGE &
BACKUP



DATA
PRESERVATION



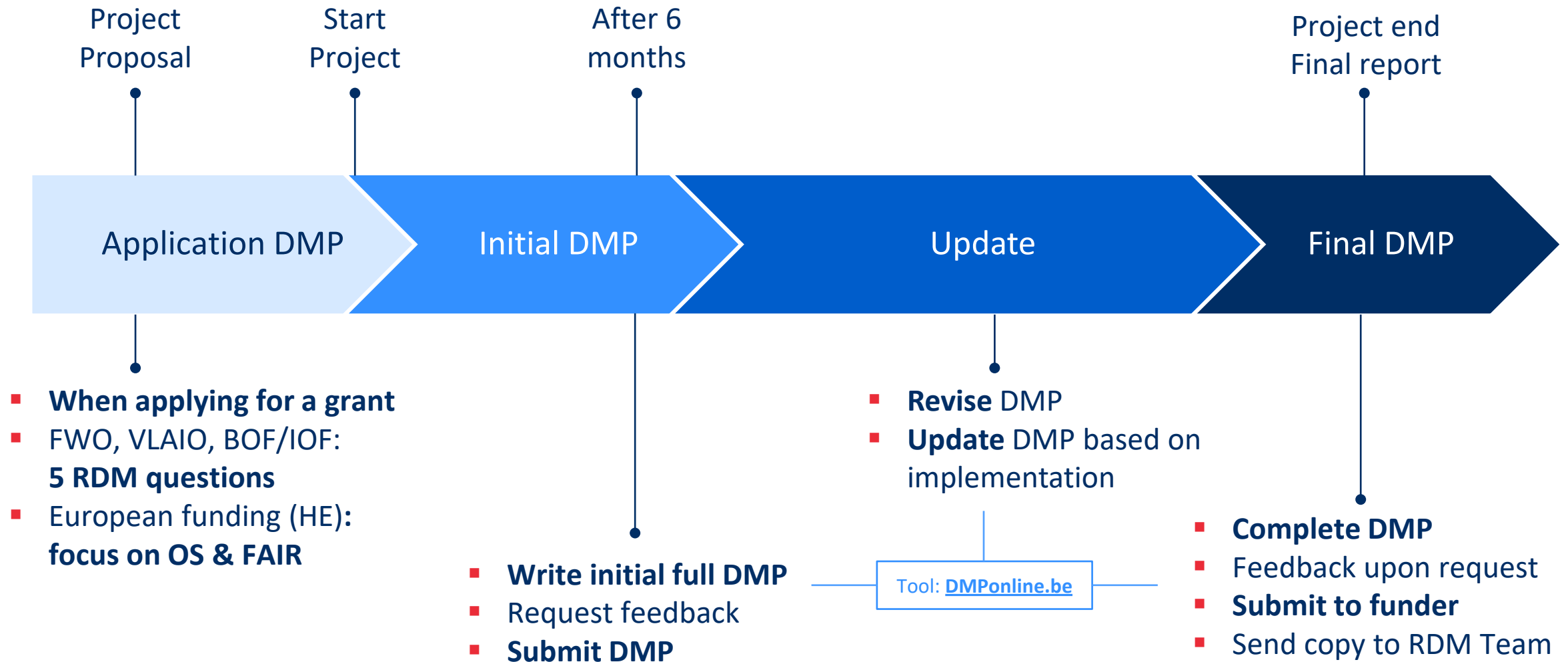
DATA SHARING
& REUSE



RESPONSIBILITIES
& RESOURCES



DMP Flow



Research data: other considerations

- **Handling personal data – GDPR**

- Privacy Register
- DPIA

- **Ethical issues**

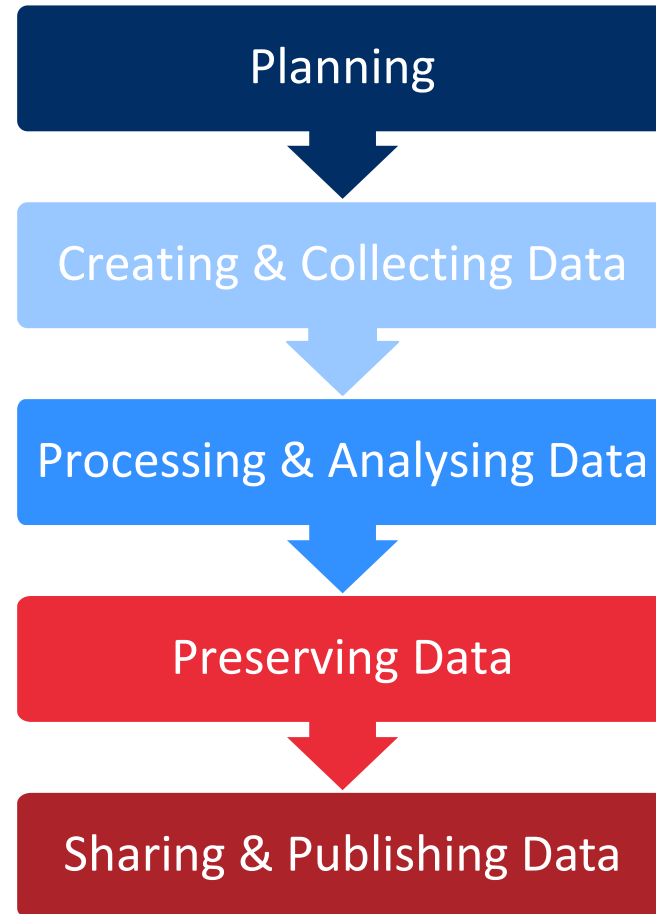
- Human participants (e.g. interviews/surveys) → EA SHW
- Patient data → CME
- Animal experiments → ECD
- Potential for dual use or misuse → MiHRS
- You may need to seek ethical approval from an Ethical Committee

- **Valorisation potential** → contact Valorisation Office (RIVA)

- **3d party agreements** → e.g. company or hospital that transfers data to you (with MTA/DTA)



RDM STAGE



RDM PRACTICES

1. PREPARING A DMP

2. ORGANISING & DOCUMENTING DATA

3. STORING DATA & DATA SECURITY

4. PRESERVING & SHARING DATA

File Organization

- **Maintain a consistent folder structure**
 - Convention within team/research group/field?
 - Avoid duplicate files
 - No generic folders (“my stuff”)
 - Keep a “raw data” folder for each data type or experiment
- **Use clear / consistent file names**
 - Convention within team/research group/field?
 - Include anything relevant to your project
 - Keep it short (<32 characters)
 - No special characters or spaces
- **Version control: manual or preferable automatic**
 - e.g. for coding:

software.py
software_v2.py
software_v10_final_for_real_v2.py



OneDrive

Versioning software:

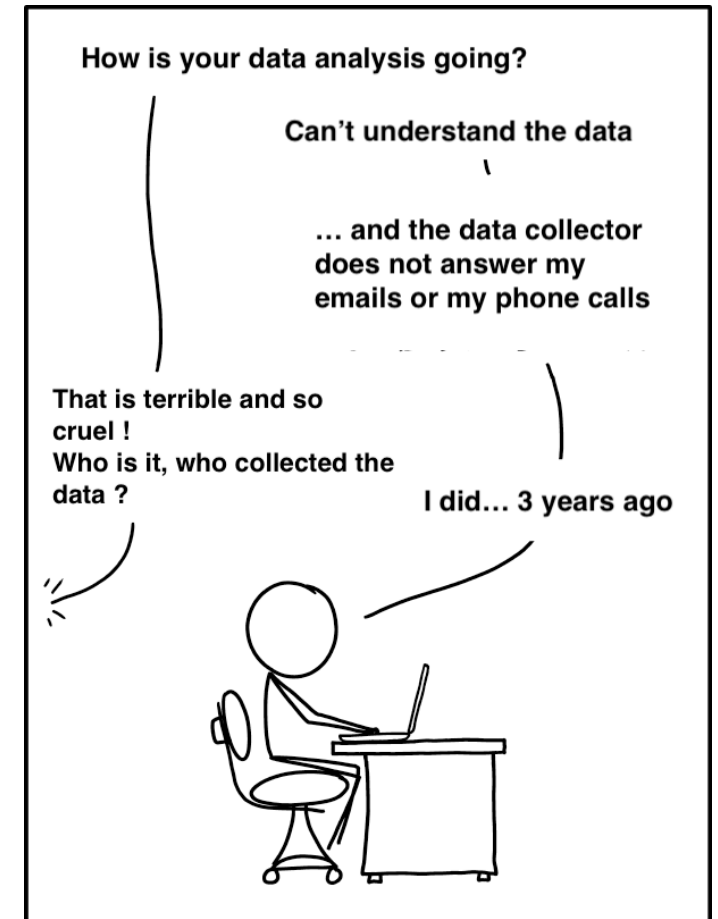


git



Documentation: What to document?

- Contextual information (project, funder,...)
- Relations to other research output
- Organisation and structure of the dataset, relation between files
- Data collection info (units of measurement, response rate, ...)
- Data methodology (analytical procedures, SOPs, protocols, ...)
- Codebook and vocabularies, e.g. explaining variables
- Quality assurance procedures
- Known problems
- Software used to generate data
- Machine configurations
- Any notes/annotations you make
- Licensing and usage rights
- ...

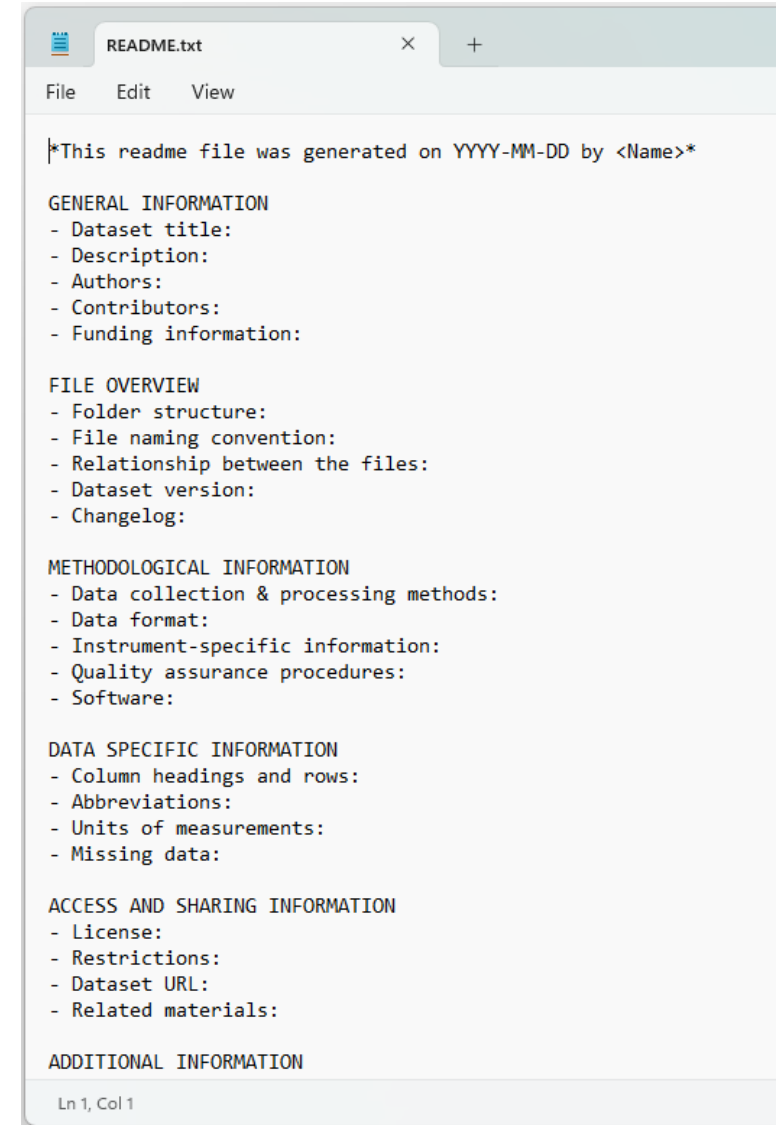


**Your first collaborators
are your future selves,
be nice to them !**

Documentation: How to document?

- **README.txt file**
 - Naming convention
 - Good entry point to a dataset or any folder structure
 - Briefly explains contents of each file
- **Data dictionary/codebook**
- **In-file documentation**
 - Commented code
 - 1st tab of Excel
- **Lab journal/notes**
 - Consistent Excel file
 - Electronic Lab Notebook

→ Keep these documents alongside your datasets



The screenshot shows a text editor window titled 'README.txt'. The content is a template for a README file, organized into sections with bullet points. The sections are: GENERAL INFORMATION, FILE OVERVIEW, METHODOLOGICAL INFORMATION, DATA SPECIFIC INFORMATION, ACCESS AND SHARING INFORMATION, and ADDITIONAL INFORMATION. The first line of the template is a placeholder for the date and author: '*This readme file was generated on YYYY-MM-DD by <Name>*'.

```
*This readme file was generated on YYYY-MM-DD by <Name>*
```

GENERAL INFORMATION

- Dataset title:
- Description:
- Authors:
- Contributors:
- Funding information:

FILE OVERVIEW

- Folder structure:
- File naming convention:
- Relationship between the files:
- Dataset version:
- Changelog:

METHODOLOGICAL INFORMATION

- Data collection & processing methods:
- Data format:
- Instrument-specific information:
- Quality assurance procedures:
- Software:

DATA SPECIFIC INFORMATION

- Column headings and rows:
- Abbreviations:
- Units of measurements:
- Missing data:

ACCESS AND SHARING INFORMATION

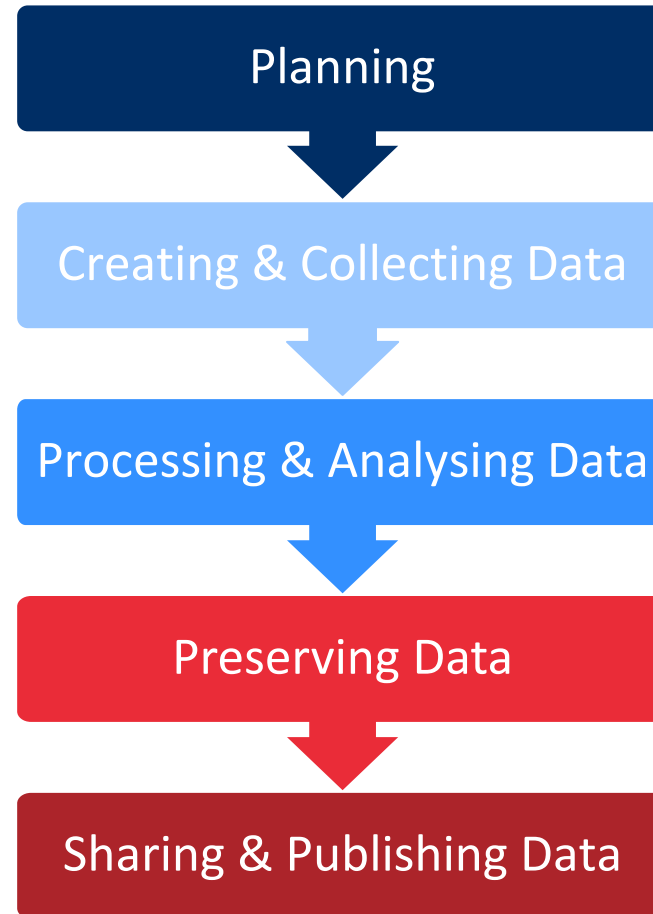
- License:
- Restrictions:
- Dataset URL:
- Related materials:

ADDITIONAL INFORMATION

Ln 1, Col 1

RDM STAGE

RDM PRACTICES



1. PREPARING A DMP

2. ORGANISING & DOCUMENTING DATA

3. STORING DATA & DATA SECURITY

4. PRESERVING & SHARING DATA

Data Storage

- **Ensure safe storage of your data**
 - **Good practices for storage**
 - Use solutions maintained by UA ICT
 - Otherwise: more than one storage location for critical data
 - Take special care in storing confidential data
 - Think about who is allowed read/write access
 - Perform regular back-ups
 - **Storage solutions by UA ICT**
 - Cloud solution (Microsoft 365)
 - **SharePoint Online-site (for sharing & collaboration)**
 - **Teams-site (for communication & collaboration)**
 - OneDrive for Business (personal cloud storage)
 - Central storage facilities with automated backup
 - Shared network drive (N drive)
 - Personal network drive (H drive)
- Additional fileserver storage: €120 per TB/year



New: UAntwerpen Connect - Flex



- PERSONAL LAPTOP
- EXTERNAL CLOUD SERVICE
- WETRANSFER



[PINTRA | ICT GUIDE STORAGE](#)

Security of Data Files

- **Access control**

- Encryption (BitLocker, VeraCrypt, FileVault)
- Password protection
- Access restriction by assigning user rights
'no access', 'read only', 'read and write' or 'administrator'

- **Procedural arrangements**

- Non-disclosure agreements

- **Secure transfer of files**

- Use [Belnet FileSender](#) for large files (up to 5 TB)

- **Authorisation and authentication**

- Very selective access rights to sensitive information
- Multifactor authentication

- **Destroying data in a consistent and reliable manner when (no longer) needed**

**PROTECT YOUR DATA
WITH ORGANISATIONAL
MEASURES**

e.g., access control,
agreements with external
parties, data protection
policies, researcher training

**PROTECT YOUR DATA
WITH TECHNICAL
MEASURES**

e.g., anonymise,
pseudonymise,
encrypt your data,
use safe storage

Research Tools & Software

Choose them conciously!

TOOLS YOU USE

-  Data collection
-  Storage & sharing
-  Analysis software
-  AI tools
-  Code & repositories

- ✓ Follow community standards
- ✓ Consider open-source alternatives

BEFORE USING A TOOL, ASK:

-  Where is my data stored?
-  Who has access?
-  Is it compliant with policy & GDPR?
-  Is data used to train models?
-  Can I export, preserve & share?
-  Can others reproduce my work?

KNOW YOUR TOOL



- Convenient ≠ compliant
- Evaluate use per individual study
- Remain skeptical
- Run tool locally & control where data is stored and kept

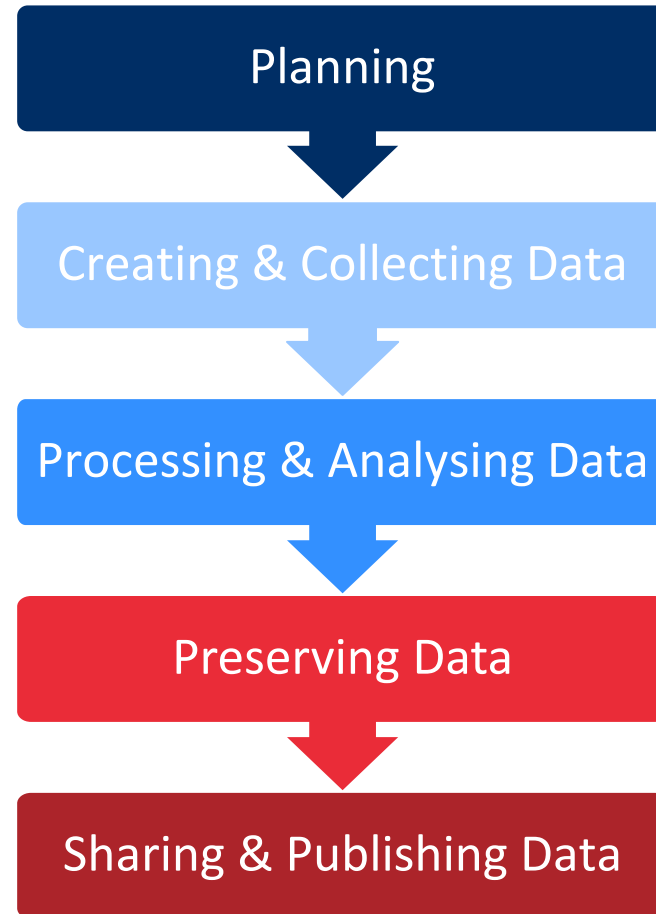
AI TOOLS



- ✓ Check institutional guidance
- ✓ Avoid uploading sensitive data
- ✓ Verify outputs and references
- ✓ You remain responsible

 [PINTRA | INFOCENTER AI](#)

RDM STAGE



RDM PRACTICES

1. PREPARING A DMP

2. ORGANISING & DOCUMENTING DATA

3. STORING DATA & DATA SECURITY

4. PRESERVING & SHARING DATA

Data Preservation

The main reasons for data preservation are

- Ensuring that your research can be verified and reproduced
- Maintaining data for future reuse (e.g. further research/teaching)

WHAT? - SELECT DATA TO KEEP

- Data related to publications
- Data that cannot be recreated or at great cost/effort
- Data legally obliged to keep
- Use open and standard file formats

HOW LONG?

- Check funder policy
- At UAntwerp, the default preservation period is **10 years**
- Potential deviations
 - Retaining clinical data longer due to legal requirements (20 or 25 years)
 - GDPR: keep no longer than necessary → Destroy files to protect privacy

WHERE?

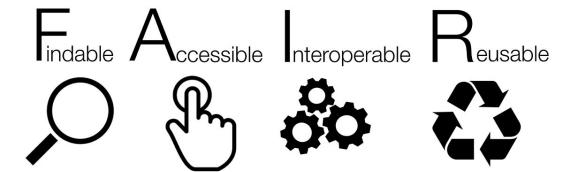
- Trusted external repository
- Central fileservers
 - SharePoint/Teams only for active research data
- Coming soon for sensitive data:



Data Sharing: Repositories

Advantages

- No handling of data reuse queries and managing data access.
- Guarantees in terms of sustainable access to data.
- Make your data discoverable and citable.
- Make your data FAIR



Wide variety of data repositories

- Most have the option to publish your dataset using a persistent identifier
- Some provide the service of long-term preservation
- Some host data from various disciplines and others are domain- or discipline specific

ACCESSIBLE

≠

OPEN

April 30, 2020 (v1)

Dataset

🔓 Open

High-resolution future climate data for species distribution models in Europe

De Troch, Rozemien ; Termonia, Piet ; Van Schaeybroeck, Bert 

Description This dataset contains a set of 13 climatological variables (Variable, VariableName) at a spatial resolution of 1x1km for Europe (nx = 13147, ny = 6071) for historical (ClimatePeriod) and future climate conditions. These variables are a subset of the so-called bioclimatic variables that are often part of global gridded datasets (e.g...

Uploaded on April 30, 2020

👁 683

📄 623

March 13, 2023 (v1)

Dataset

🕒 Embargoed

Flemish Political Ambition Survey

Bram Wauters ; Robin Devroe ; Audrey Vandeleene 

Dataset based on a survey about political ambition among a random sample of the youth population (aged 18-35) in Flanders (Belgium), N = 1,000

Uploaded on March 13, 2023

👁 32

📄 0

November 12, 2018 (v3)

Dataset

🔒 Restricted

Data for development, evaluation and application of NicheNet

Browaeys, Robin ; Saelens, Wouter ; Saeys, Yvan 

Here you can find the datasets used for development, validation and application of NicheNet, a method that links ligands to target genes to model intercellular communication. The software to use NicheNet (i.e. R package nichenetr) is available at: <https://github.com/browaeysrobin/nichenetr>
Short description of individual files: * data for d...

Uploaded on November 12, 2018

2 more versions exist for this record

👁 425

📄 6

“As open as possible, as closed as necessary”

- There might be **access controls** or **restrictions**
- Available **to whom** and **under what conditions?**
- FAIR data $\neq$ OPEN data
- Reasons not to share data openly:



Personal data



Confidential data



Sensitive data



Third-party data



Copyright



Valorisation

Preferred method for depositing data (FAIR-aligned)

- Publish in a domain specific or a generic **repository or database**
- Make sure the dataset has its own **permanent identifier** (doi, handle, ...)
- Use repositories that link with **ORCID**
- Create at least a metadata record with study level metadata and add domain specific metadata if applicable
- Use standard formats and vocabularies
- Include **documentation** (README.txt)
- Include a **license** (CC-0 or CC-BY, Open Source)
- Include a **data availability statement** in your publications
- Register the dataset in **IRUA**

Wrap up



University of Antwerp RDM Policy

- **Write a Data Management Plan**
 - Mandatory for all (partially) publicly funded research, recommended for others
 - Proper data management as shared responsibility of all parties involved
- **“As open as possible, as closed as necessary”**
- **FAIR principles**
- Preserve all relevant research data for **min. 10 years**
- **Register** published research datasets in **IRUA**
- Maximum use of **PIDs** (ORCID, ROR,...)





How to get started?

- **Know your data!**
- Discuss RDM practices with your promotor and **plan ahead!**
 - e.g. tools, storage, security measures
- **Start documenting** all your research steps
- Set up your **ORCID** if you do not have one (<https://orcid.org/>)
 - Add to it to your account on our institutional repository ([instructions](#))
- Make use of our **DMP tool** (<https://dmponline.be>)
- Reach out if you have RDM questions: rdm-support@uantwerpen.be
 - Check out our [support pages on RDM](#)
 - Check out our [ADS course on DMP](#)

Useful links for UAntwerpen guidance

- [Research Data Management | University of Antwerp](#)
- [RDM | Pintra](#)
- [ORCID | Library | University of Antwerp](#)
- [IRUA UAntwerpen Institutional Repository](#)
- [Personal data & Privacy | DP² Pintra](#)
- [Storage | ICT Pintra](#)
- [Infocenter AI](#)
- [Infocenter Cybersecurity](#)



**Unless stated otherwise, content in these slides is distributed under a
Creative Commons Attribution 4.0 International license**