



Introduction to Research Data Management

Linde Tuybens

ADS Doctoral Day

11.03.2025

Research Data Management (RDM) Team

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



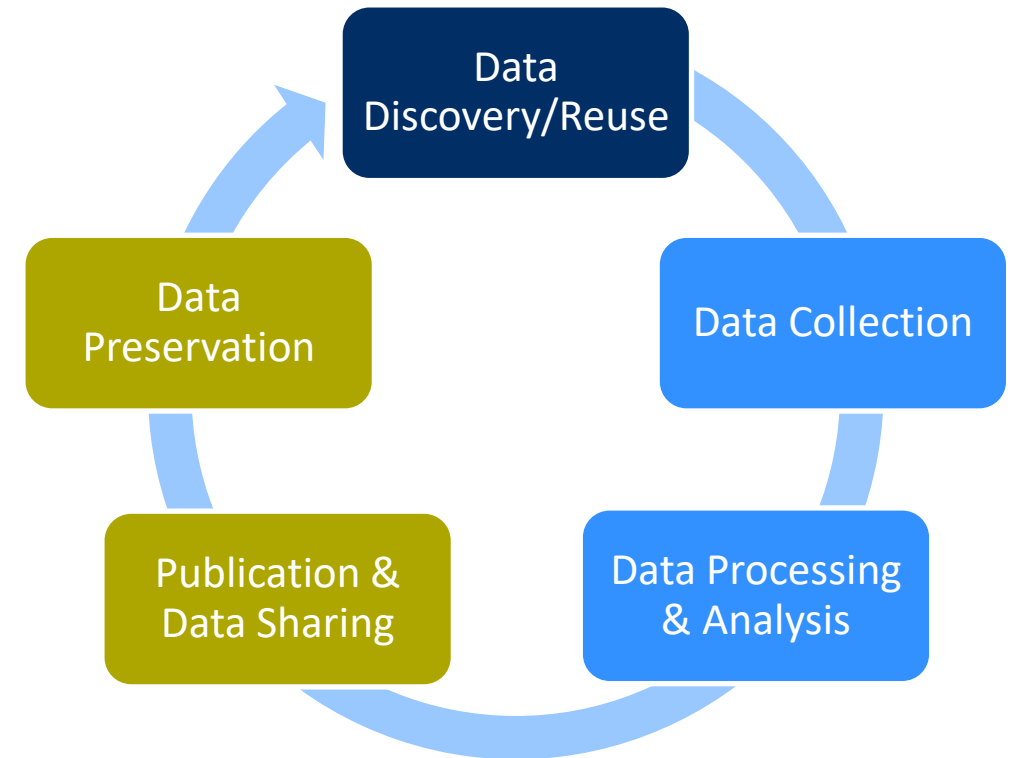
- **Siham Benramdane** (Life Sciences & Medicine)
- **Robin Björklund** (Natural Sciences & Engineering)
- **Linde Tuybens** (Social Sciences & Humanities)
- **Dunya Nasser / Niels Vervoort**

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 a research output!**

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.**“*

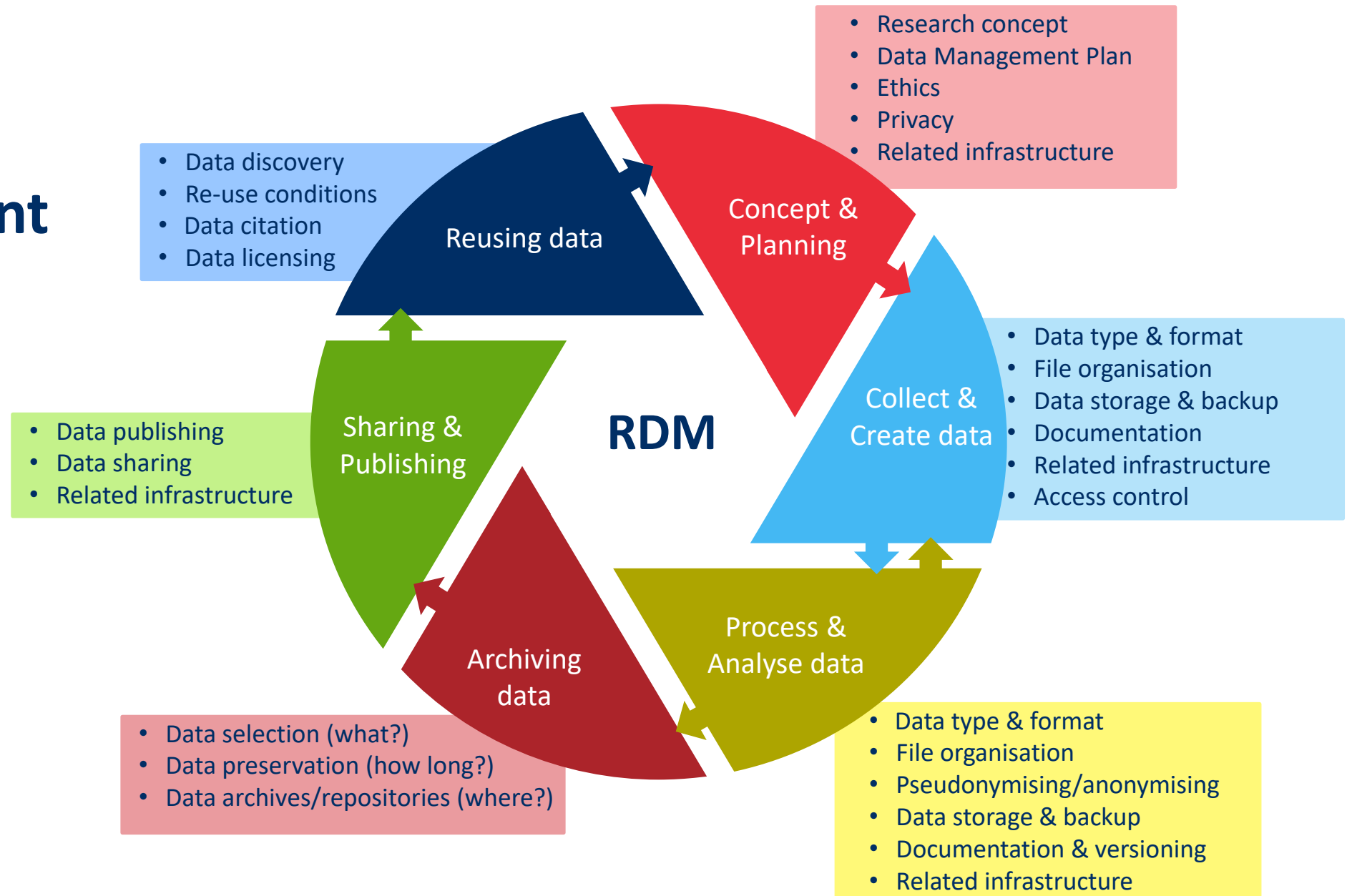
RDM Policy of University Antwerp



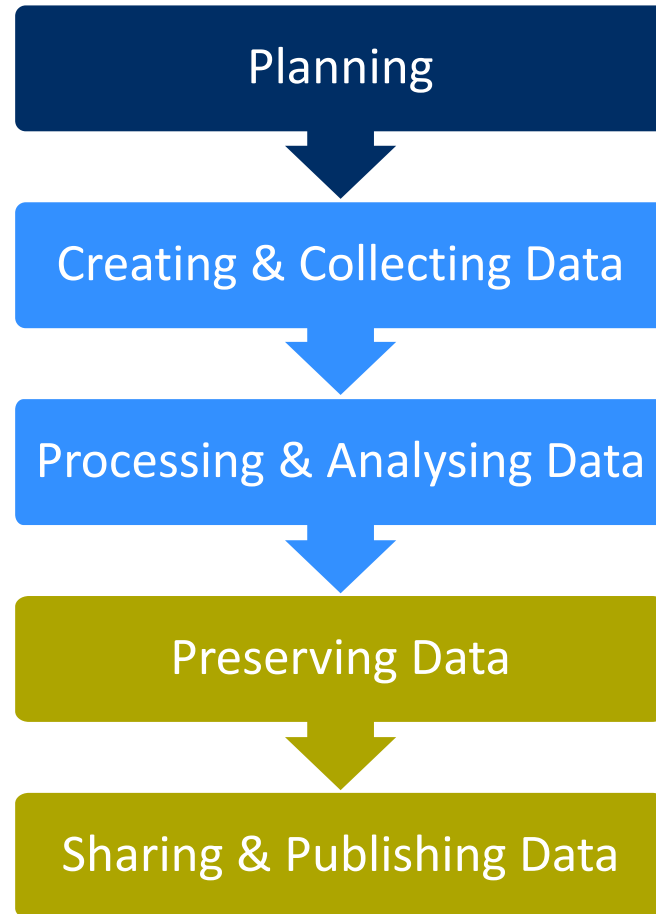
**Your
research
data**

Key concepts of RDM

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 processed during the research lifecycle

29/1 3200

GEWICHTSTABEL

Datum	Leeftijd	Gewicht
25-1-79	4 wk	3220.49
29-1-79	15-45	3400
15-1-79	6 wk	3560
2-2-79		3880
23-2-79	14 wk	4940
		5240
		6060

Data Management Plan Content

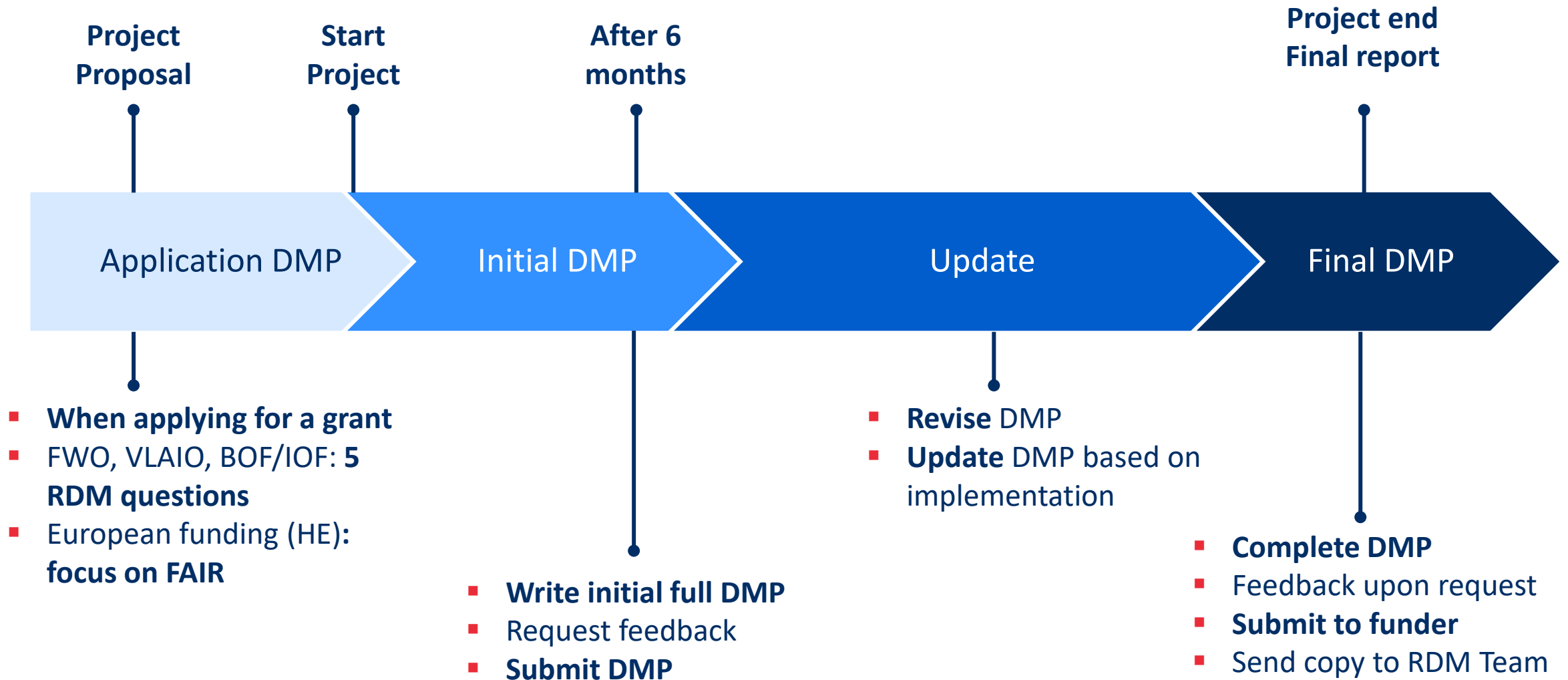
- **Contextual information**
- **Data description**
- **Ethical and legal issues**
- **Documentation and Metadata**
- **Data storage and backup during project**
- **Data preservation after the project**
- **Data sharing and reuse**
- **Responsibilities**

Purpose of a Data Management Plan?

The DMP is a text document that:

- Describes which research data you will use and/or generate during your research and how you **plan** to manage them.
- Helps you **track** your research data management practices throughout your project.
- Acts as **checklist** to ensure you have considered all important questions about data.
- Helps you plan how your research data will be **preserved and shared** once the project is finished.
- Provides an **opportunity to reflect** on your research data, possible obstacles in sharing said data, and ways to overcome them.



















DMP Flow



Funder Templates

Templates are provided by a funder.

Templates for data management plans are based on the specific requirements listed in funder policy documents. DMPonline.be maintains these templates, however, researchers should always consult the funder guidelines directly for authoritative information.

Template Name	Download	Organisation Name	Last Updated	Funder Links	Sample Plans (if available)
BELSP0 DMP +	 	Belgian Federal Science Policy Office (BELSPO)	27-09-2021		
ERC DMP +	 	European Research Council (ERC)	27-09-2021		
DCC Template	 	Digital Curation Centre	27-09-2021		
Horizon 2020 FAIR DMP +	 	European Commission (Horizon)	27-09-2021		
BRAIN 2.0	 	Belgian Federal Science Policy Office (BELSPO)	13-05-2022	www.belspo.be	
FNRS DMP	 	Fonds National de la Recherche Scientifique (FNRS)	19-05-2022		
VLAIO cSBO DMP (Flemish Standard DMP)	 	Vlaams Agentschap Innoveren & Ondernemen (VLAIO)	02-09-2022		
Horizon Europe DMP +	 	European Commission (Horizon)	12-10-2022		
FWO DMP (Flemish Standard DMP)	 	Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)	21-01-2025		

Organisational Templates

Templates are provided by your organisation (UAntwerpen)

DMPonline.be

- **Web-based tool** for data management planning
 - Originally developed by the DCC – modified by DMPbelgium consortium
 - Owned by Belnet
- **DMP templates + guidance**
- Does not check or validate your answers
- **Tutorial:** DMPonline [knowledge clip](#) by Ugent

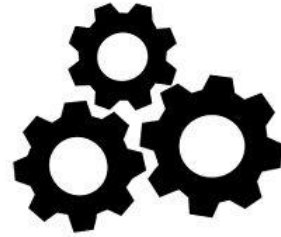
F
Findable



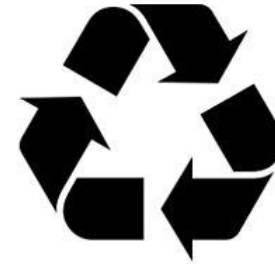
A
Accessible



I
Interoperable



R
Reusable



FAIR Data



Findable



How easily can humans (and machines) find my data?



Accessible



Once found, how easy is it to access my data?



Interoperable



How well does my data integrate with other data?

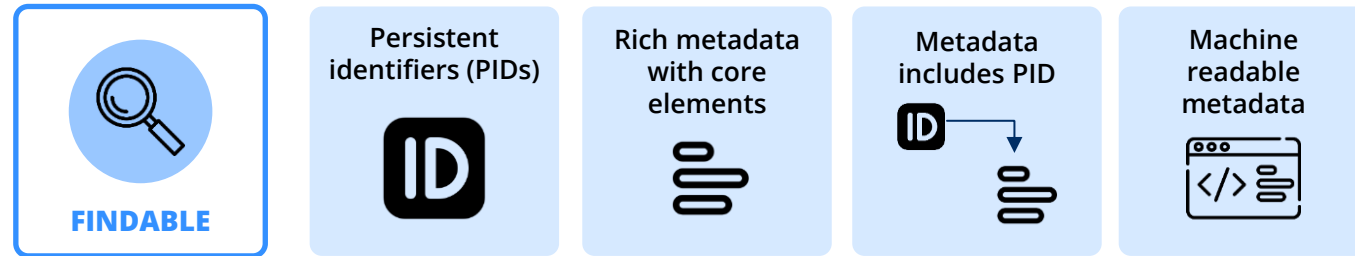


Reusable



How difficult is it to reuse my data?

F is for making data findable



- **Look for existing data in repositories**
- **Upload to and share your data via a research data repository**
- **Describe your data with as much detail as possible**
 - For search and filtering
- **Apply persistent identifiers**
 - DOI, ORCID, ROR

A is for making data accessible



Metadata has
access level &
conditions



Standard
communication
protocol

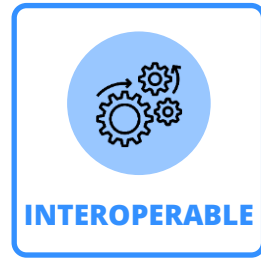


Metadata is
always available



- **Consider what can and will be shared under which conditions**
 - Determine access control
 - Access conditions clear and understandable
- **Obtain participant consent and perform risk management**
- **Share your metadata**

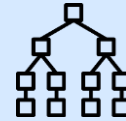
I is for making data interoperable



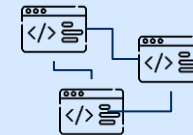
Metadata in
formal KR
language



Vocabularies &
ontologies

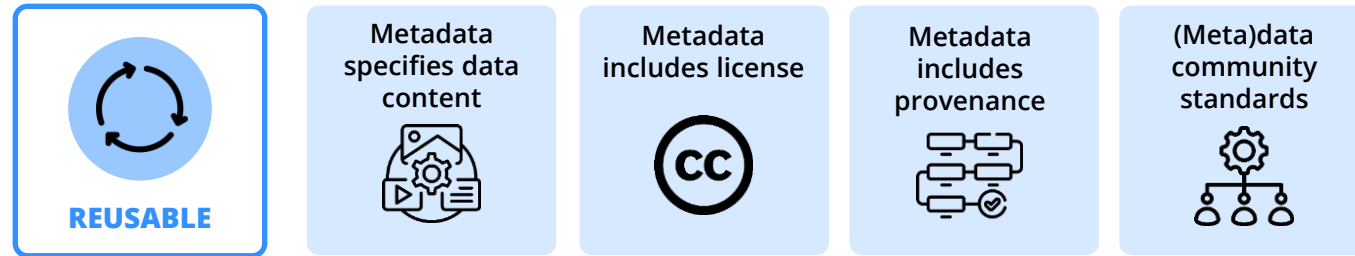


Linked
metadata



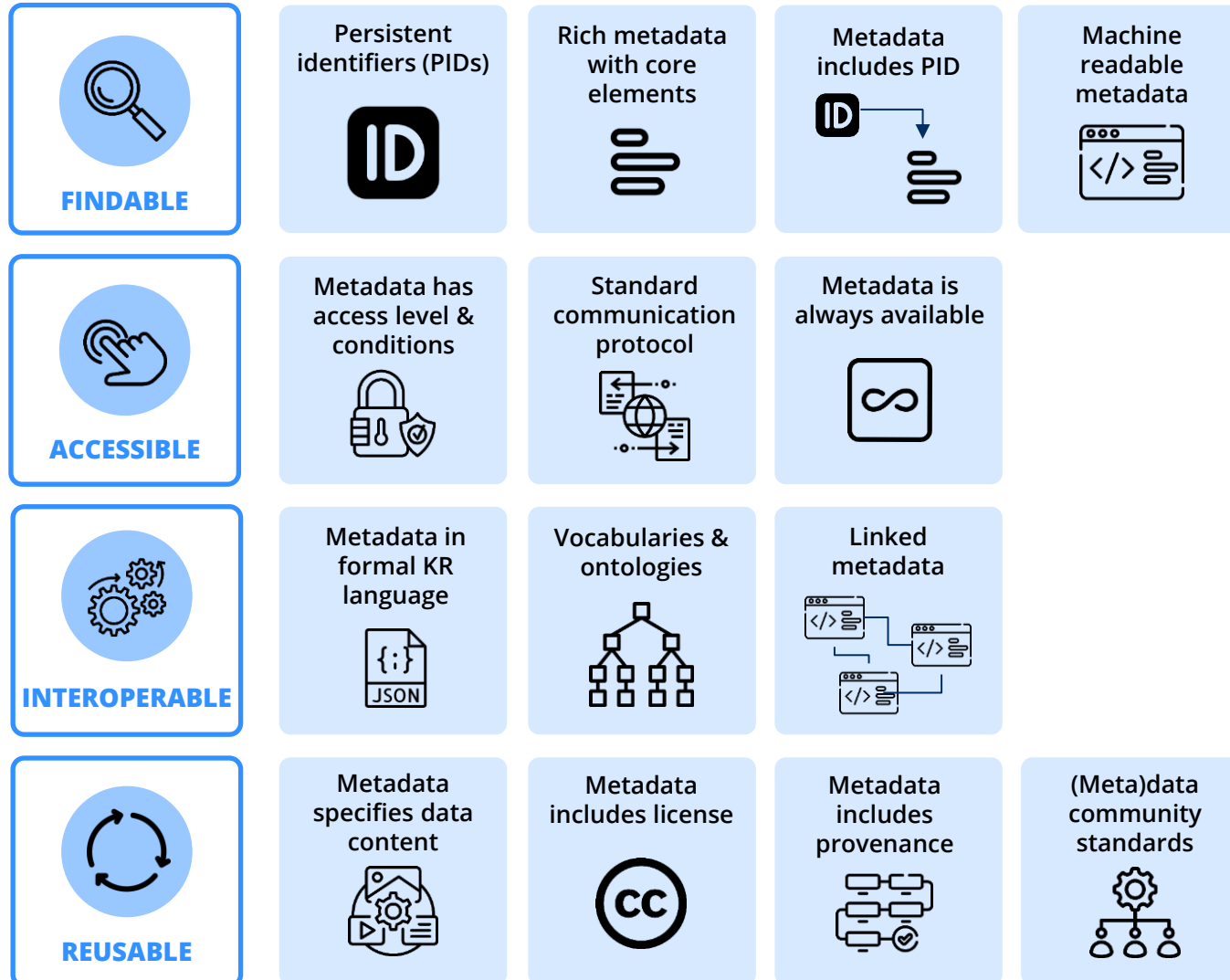
- **Use open, standardized and common formats**
- **Consistent vocabulary**
 - Preferably controlled vocabularies
- **Apply common metadata standards**
- **Linked (meta)data**

R is for making data reusable

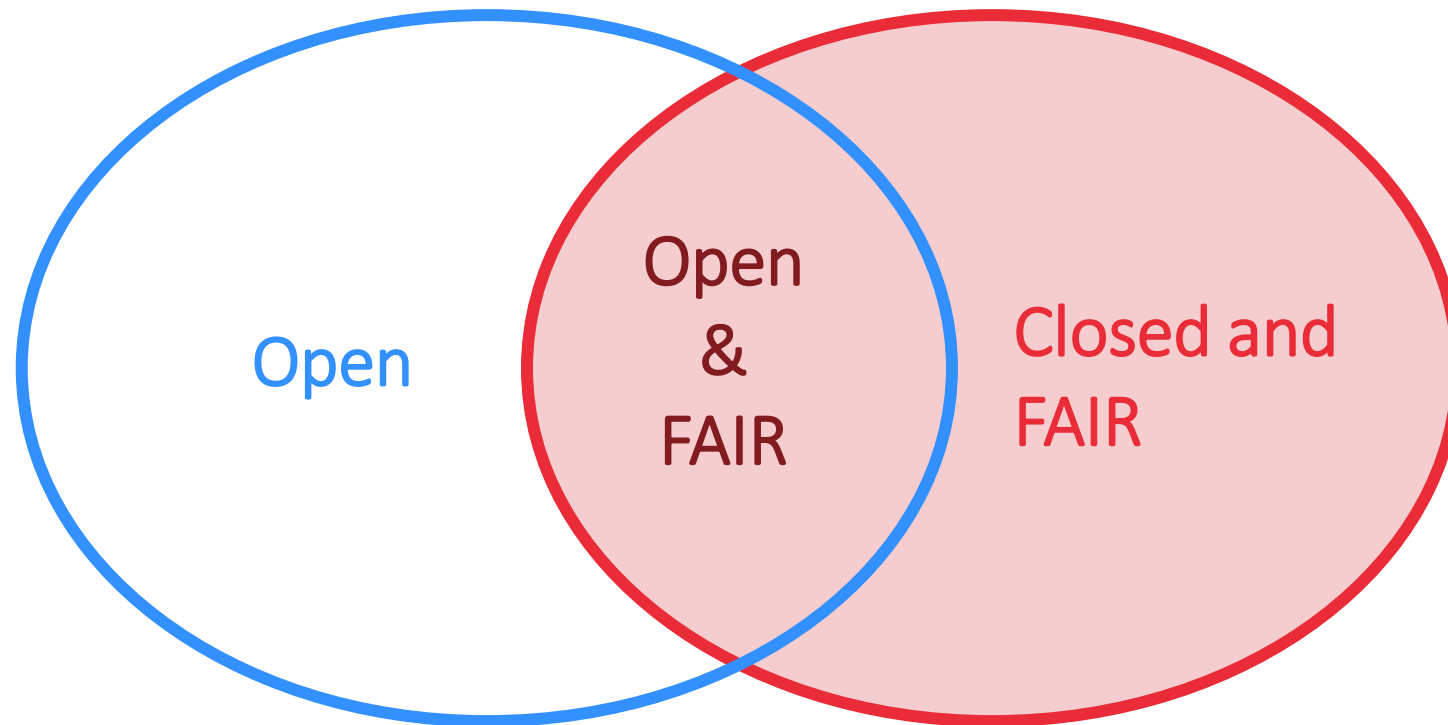


- Consider permitted use
- Apply appropriate license to specify how others can re-use your data/code
- Add sufficient documentation and provenance information
- When using data of others, give credit by data citation

FAIR Principles



FAIR ≠ Open



“As open as possible, as closed as necessary”

“As open as possible, as closed as necessary”

- FAIR data ≠ OPEN data
- There might be access controls or restrictions
- Available to whom and under what conditions?

- Reasons not to share data openly:



Personal data



Confidential data



Sensitive data



Third-party data



Copyright



Valorization

Overview of funder RDM requirements from proposal to end of project

Funder	Proposal	Initial DMP (month 6)	Final DMP (end of project)	Data preservation	Data publishing
FWO	5 RDM questions in FWO application form	Submit to rdm-support@uantwerpen.be	Submit to FWO (e-portal) with final report	Min. 5 years	Advice: data linked to publications
BOF/IOF	5 RDM questions in application form	Submit to rdm-support@uantwerpen.be	Submit to rdm-support@uantwerpen.be	Min. 10 years	Advice: data linked to publications
VLAIO	5 RDM questions	Submit to rdm-support@uantwerpen.be	Submit to rdm-support@uantwerpen.be	Min. 5 years	
BELSPO	5 RDM questions (provisional DMP) in grant application	Submit to BELSPO	Submit to BELSPO	Long-term in repository	All data and metadata in certified and trusted repository, as open as possible
Horizon Europe	RDM for FAIR data and Open Science practices	Submit to EC portal	Submit to EC portal	Long-term in repository	All data and metadata in certified and trusted repository, as open and early as possible



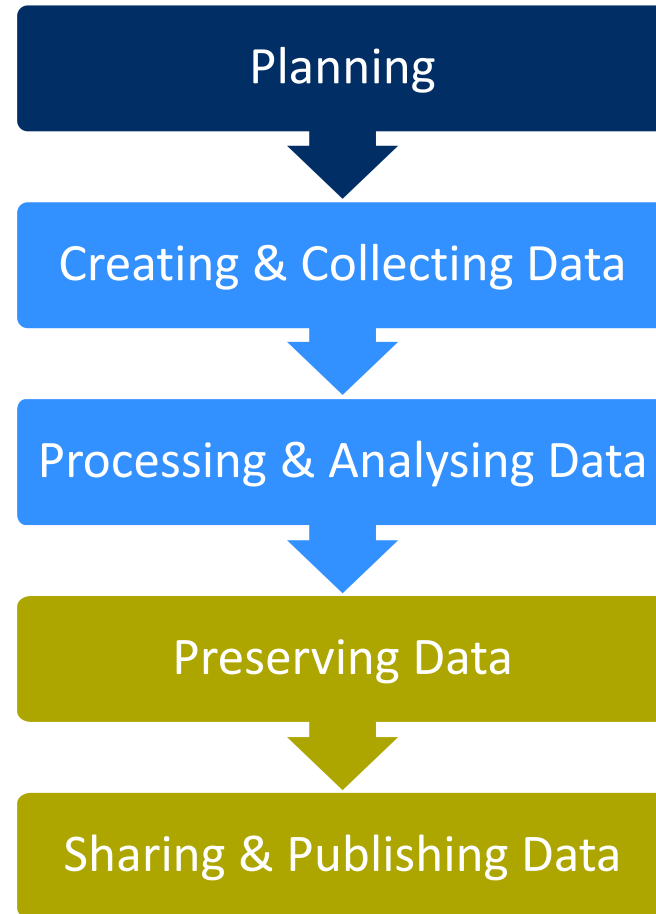
University of Antwerp RDM Policy

- **Write a Data Management Plan (DMP)**
 - Mandatory for all (partially) publicly funded research, recommended for others
 - Promotor is responsible for making sure a DMP is present
 - 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,...)**

Sidenote: Nagoya Protocol & ABS regulation

- Related to the **access and use of non-human genetic resources** and related traditional knowledge **collected abroad**
- <https://nagoya.vlir.be/>
- If there is a **chance** your research is subject to the **Nagoya Protocol**, contact us **as soon as possible** (rdm-support@uantwerpen.be)

RDM stage



RDM practices

1. Preparing a DMP

2. Organising & Documenting Data

3. Storing Data & Data Security

4. Preserving & Sharing Data

Data documentation

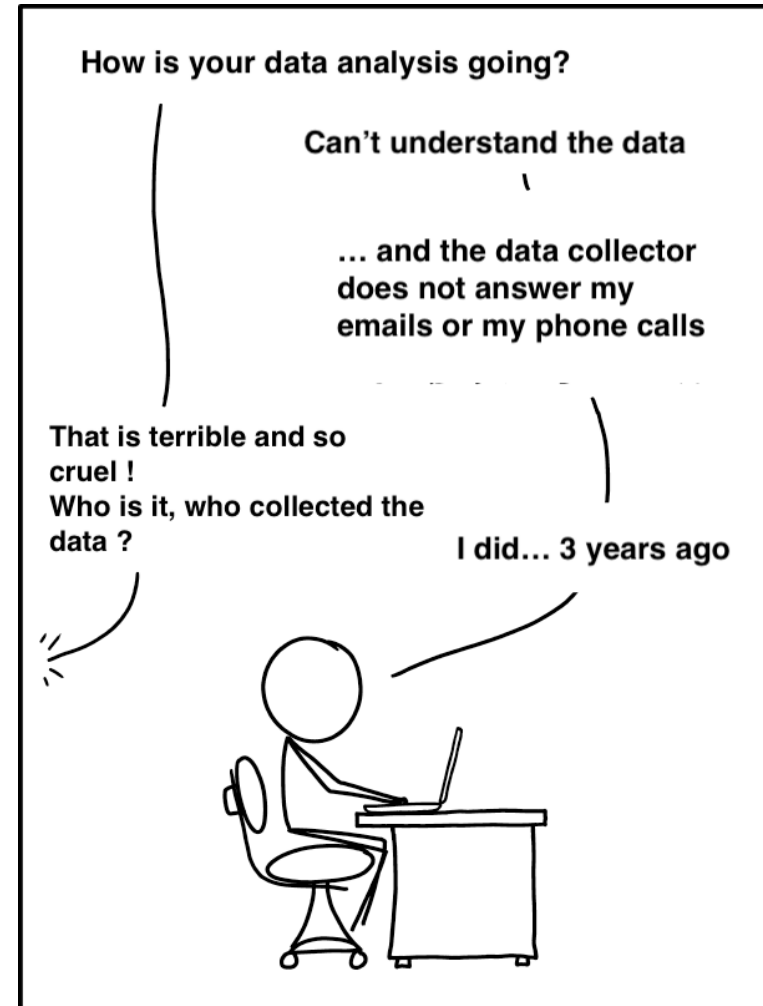
Data documentation

WHY? Needed for correct interpretation by any user (including yourself)

WHAT? Data collection methodology, analytical and procedural information, data manipulations, measurement units, data characteristics, code definitions, terminology, data quality control measures, known problems,...

Structure/organisation of data files, file naming conventions, any comments, ...

HOW? Use README.txt file, codebook, ELN or any other form of documentation. Keep these documents alongside your datasets.



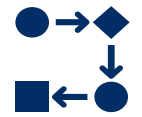


**Your first collaborators
are your future selves,
be nice to them !**

Data documentation and metadata

	Data documentation	Metadata
WHY?	Needed for correct interpretation by any user (including yourself)	Needed to find your datasets quickly in repositories, databases, project websites, ...
WHAT?	<p>Data collection methodology, analytical and procedural information, data manipulations, measurement units, data characteristics, code definitions, terminology, data quality control measures, known problems,...</p> <p>Structure/organisation of data files, file naming conventions, any comments, ...</p>	Creator, title, PID (e.g. DOI), year, publisher, date, abstract, location, file format, page number, version, language, rights, ...
HOW?	<p>Use README.txt file, codebook, ELN or any other form of documentation.</p> <p>Keep these documents alongside your datasets.</p>	Use metadata standards, that are usually determined by the tool or repository.

Levels of data documentation

DOCUMENTATION LEVEL	WHAT TO DOCUMENT	WHERE TO DOCUMENT
 Item Variable	Methods, units, possible values, synonyms, abbreviations	Codebook, data dictionary
 Sample Study object	Study object, method, instrument settings, time, place	(Register) table, protocol, notebook
 File or database	File inventory and relations, database structure	Readme, database schema, data listings (qualitative data)
 Project or study	Subject, grant, contributors, methods, process, workflow	Notebook (e-, code), document, protocol, readme, data inventory table

README

- A README file is highly recommended as an **entry point** to a published dataset
- Contains all documentation or references to other kinds of documentation
- Should be the first file created when collecting or creating data
- README = **naming convention**
- Often **.txt** file
- Template:
<https://data.research.cornell.edu/content/readme>


The screenshot shows the Zenodo interface for a dataset. At the top, there is a search bar and navigation links for 'Communities' and 'My dashboard'. Below the header, the dataset title 'Capital markets in 16th-century 's-Hertogenbosch (Bois-le-duc)' is displayed, along with the author 'Hanus, Jord' and a 'Show affiliations' button. A description follows: 'This csv-file contains a simple transcription of the so-called Bosch' Protocol, the aldermen's registers of the city of 's-Hertogenbosch, for the period October 1508 to August 1512. I transcribed this source during my master's research. This source documents real estate and annuity transactions in the wide region of 's-Hertogenbosch (the Meierij).' Below the description, a 'Files' section lists two files: 'readme.txt' (2.6 kB) and 'Transcription_BoschProtocol_15081512.csv' (642.6 kB). The 'readme.txt' file is highlighted with a red box. The interface includes buttons for 'Preview' and 'Download' for each file, and a 'Download all' button for the entire dataset.

zenodo Search records... Communities My dashboard

Universiteit Antwerpen University of Antwerp Data Repository

Published August 3, 2022 | Version v1 Dataset Open

Capital markets in 16th-century 's-Hertogenbosch (Bois-le-duc)

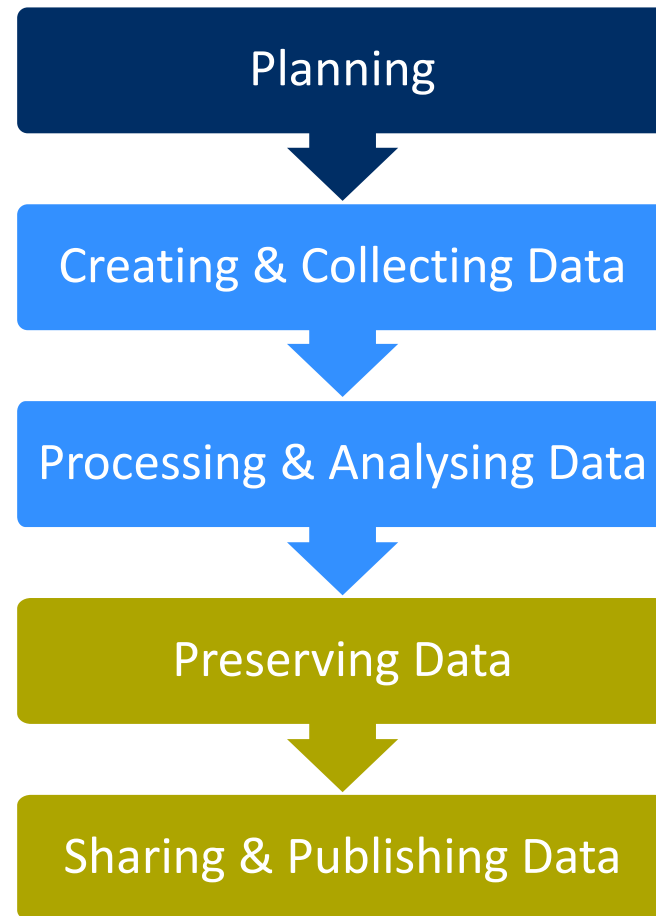
Hanus, Jord¹  [Show affiliations](#)

This csv-file contains a simple transcription of the so-called Bosch' Protocol, the aldermen's registers of the city of 's-Hertogenbosch, for the period October 1508 to August 1512. I transcribed this source during my master's research. This source documents real estate and annuity transactions in the wide region of 's-Hertogenbosch (the Meierij).

Files

Name	Size	
readme.txt	2.6 kB	Preview Download
Transcription_BoschProtocol_15081512.csv	642.6 kB	Preview Download

RDM stage



RDM practices

1. Preparing a DMP
2. Organising & Documenting Data
3. Storing Data & Data Security
4. Preserving & Sharing Data

Data sharing today?

All data in the publication

Data as supplementary material

Data made available on request

Data on project/personal websites

Through cloud storage providers

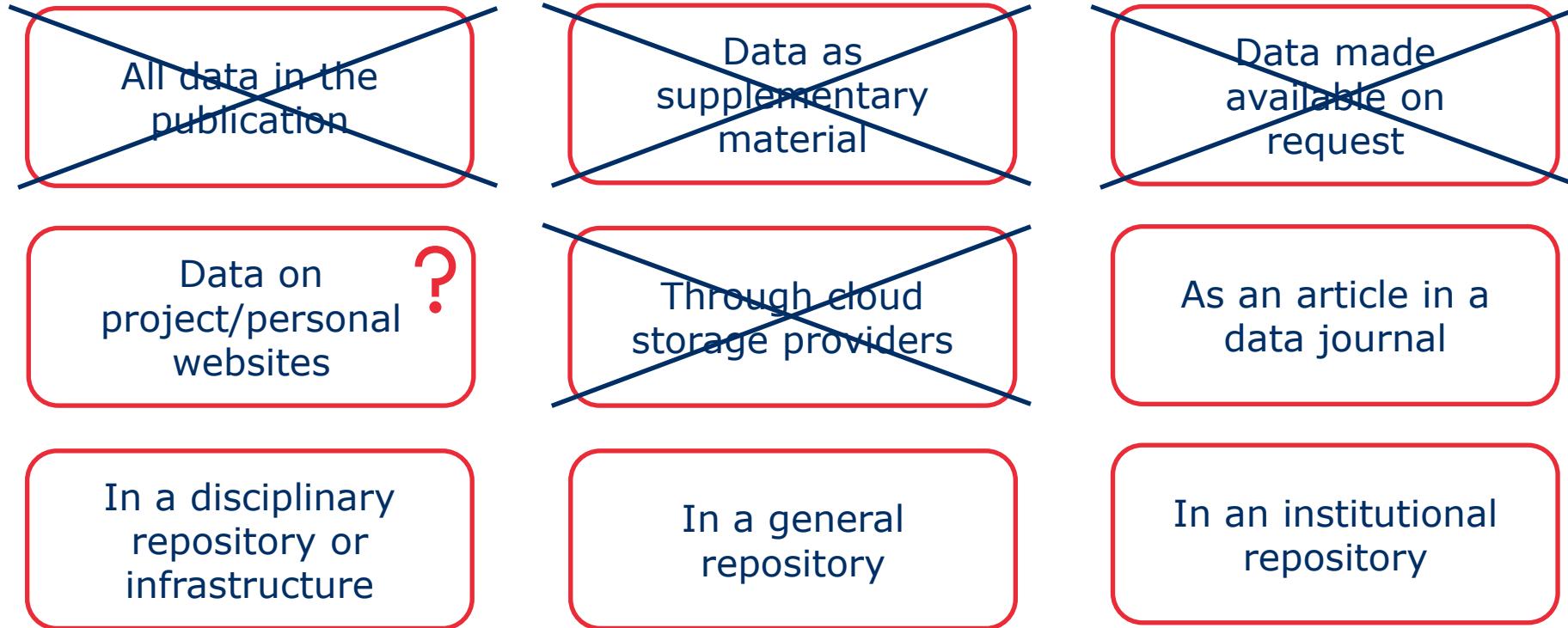
As an article in a data journal

In a disciplinary repository or infrastructure

In a general repository

In an institutional repository

Data sharing today?



Preferred method for depositing data

- 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)
- Include a **data availability statement** in your publications
- Register the dataset in **IRUA**

IRUA: Registration of metadata

- **Uantwerpen Institutional Repository (IRUA)**
- Mandatory registration of published datasets
- **Only metadata** of your published research outputs
- NOT the datasets themselves

- <https://repository.uantwerpen.be/desktop/irua>

- Guidance: [PINTRA Library](#)



Institutional Repository My bibliography Authors A-Z Faculties A-Z Research Groups A-Z Projects

Search Advanced Search Search History Save List

The UAntwerpen Institutional Repository (IRUA) is a search and Open Access publishing platform supporting the Academic Bibliography of the University of Antwerp. The Academic Bibliography provides an overview of scientific publications by researchers at the University of Antwerp.

[Read more](#)

Introduction to the Institutional Repository

Introduction to Open Science

Important? An introduction to the institution...
• Used in promotions, evaluations and project applications
• Only output in IR is listed in e-curriculum
• Updated weekly
• Data is supplied to external databases/organizations
• FRS
• VABB
• Support Open Access policy
• 2014: Open Access obligation
• 2019: ORCID obligation

An introduction to Open Science
"...the practice of science that others can contribute, where lab notes and other processes are freely available under terms that enable redistribution and its use in the research and its un..."

Wrap up

RDM Good Practice Checklist



Know what data you have.



Keep relevant documentation and metadata.



Have proper storage, backup, security.



Observe privacy and ethics principles.



Know your policies and legal obligations.



Plan for access and sharing.



Properly archive or preserve data

How to get started?

- **Know your data!**
- Discuss RDM practices with your promotor and **plan ahead!**
- 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 authoring tool** (<https://dmponline.be>)
- Reach out if you have RDM related questions (rdm-support@uantwerpen.be)
 - Check out our [support pages on RDM](#)
 - Check out our [ADS course on DMP](#)

Useful resources

- <https://fairsharing.org/>
- <https://datacite.org/>
- <https://www.dcc.ac.uk/guidance/standards>
- https://rdmkit.elixir-europe.org/metadata_management
- <https://rd-alliance.github.io/metadata-directory/>
- [Registry of Research Data Repositories \(Re3data\)](#)



Useful links for UAntwerpen guidance

- [Research Data Management | University of Antwerp](#)
- [RDM | Pintra](#)
- [ORCID | Library | University of Antwerp](#)
- [IRUA UAntwerpen Institutional Repository](#)
- [Personal data | DPO Pintra](#)
- [Storage | ICT Pintra](#)





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