



Co-funded by
the European Union

Welcome! :-)

Please turn off your microphones
and cameras.
We will start shortly.

Open for you! cycle



**Univerzita
Karlova**



**UNIVERSITÄT
HEIDELBERG**
ZUKUNFT
SEIT 1386



**PANTHÉON-ASSAS
UNIVERSITÉ**
PARIS



**SORBONNE
UNIVERSITÉ**



**KØBENHAVNS
UNIVERSITET**



**UNIVERSITÉ
DE GENÈVE**



**UNIVERSITÀ
DEGLI STUDI
DI MILANO**



**UNIwersYTET
WARSAWSKI**



Co-funded by
the European Union

Research data: Should it be open, and under what conditions?

Carolina Manfredini (University of Milan, Italy) &
Sebastian Zangerle (University of Heidelberg, Germany)

Open for you! cycle



Univerzita
Karlova



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386



PANTHÉON-ASSAS
UNIVERSITÉ
PARIS



SORBONNE
UNIVERSITÉ



KØBENHAVNS
UNIVERSITET



UNIVERSITÉ
DE GENÈVE



UNIVERSITÀ
DEGLI STUDI
DI MILANO



UNIwersYTET
WARSAWSKI

4EU+ alliance

- 4EU+ is a transnational strategic university association.
- Vision: to create one comprehensive research-intensive European University through a new quality of cooperation in
 - teaching
 - education
 - research
 - administration
- For more: <https://4euplus.eu/>

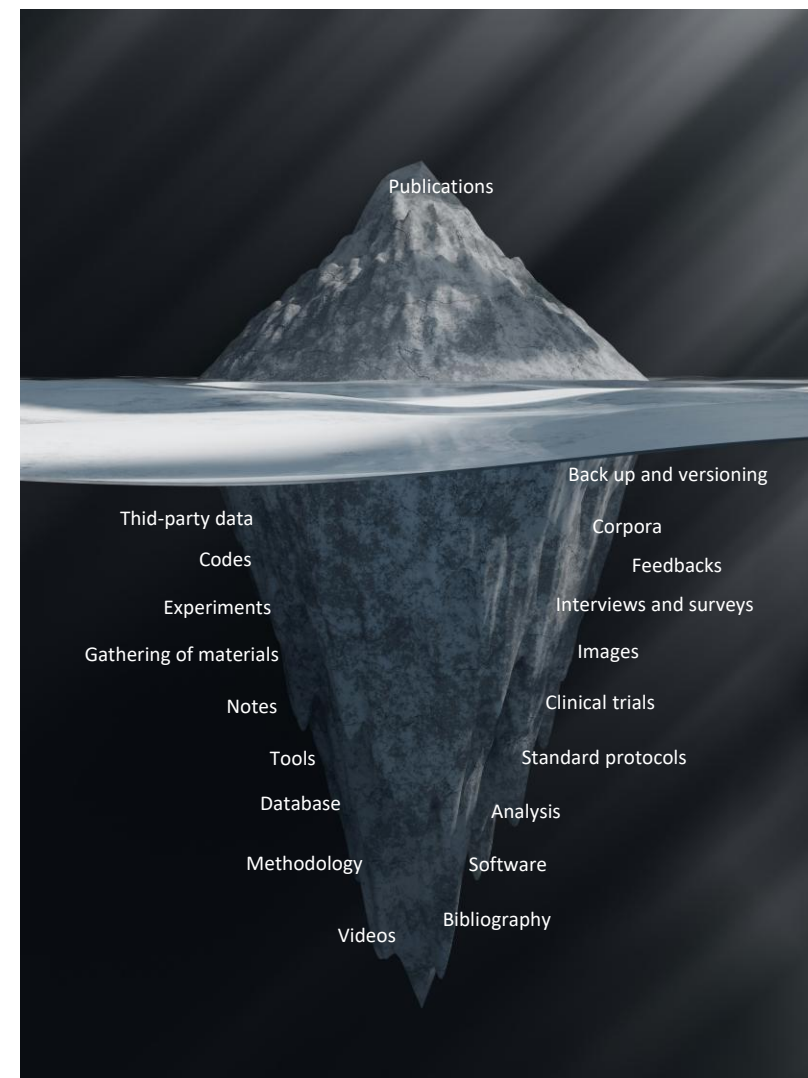


Agenda

- Identifying researchers' rights and obligations regarding the production, management and sharing of research outputs
- Introduction to the FAIR data principles
- Understanding legal tools for opening data and codes
- Data Management Plans and research funders' requirements
- Research data repositories
- Contact points for assistance with open research practices

What is research data?

Any data that is **collected, observed, generated, created, elaborated** in order to validate or reproduce your research results, as well as the documentation you need to understand and manage the data.



Why?

Why?

- **Scientific Ethics**

- Accessibility for public
- Taxpayers Argument
- Knowledge Commons
- Digital Divide

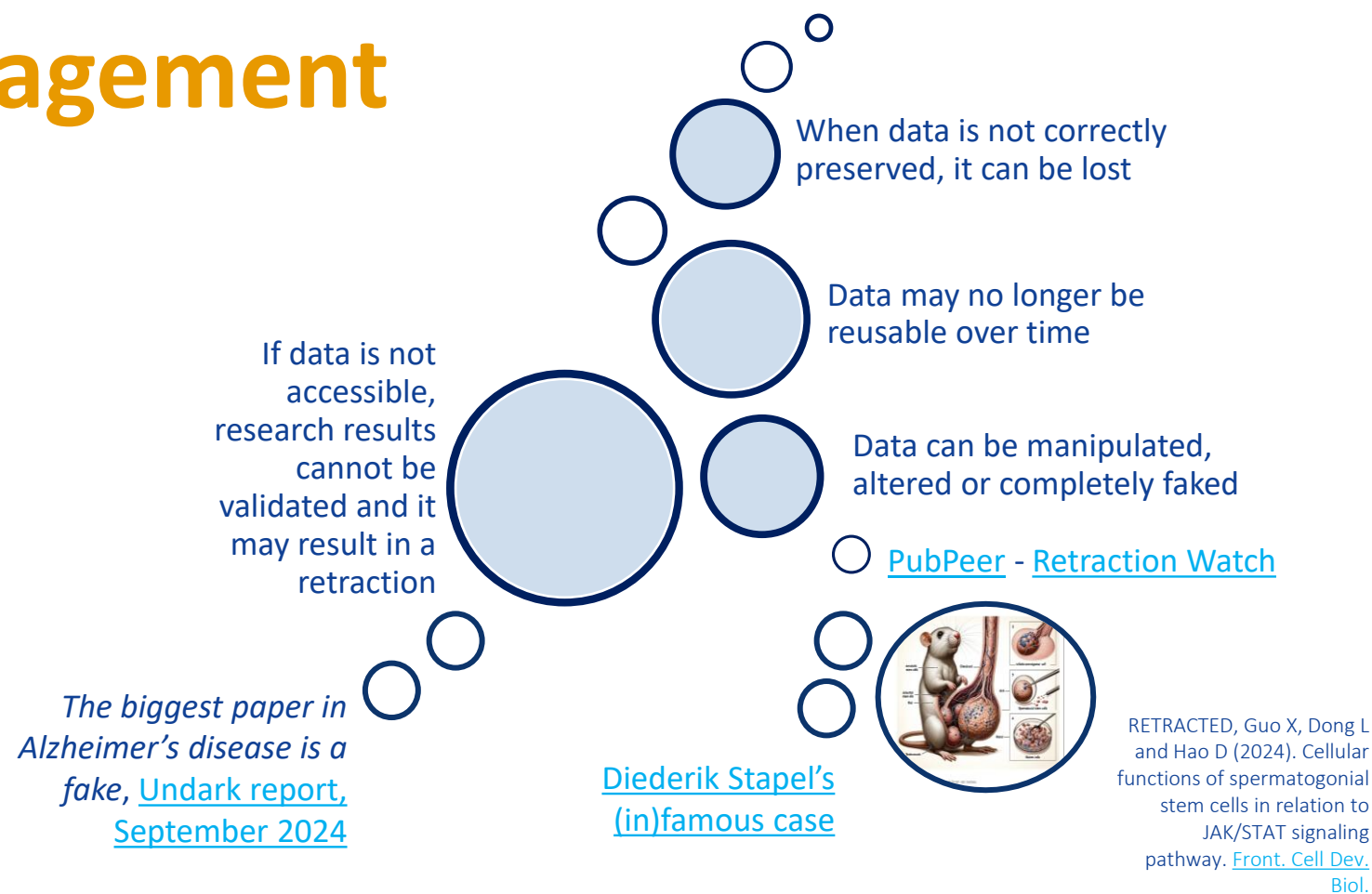
- **Visibility and Citations**

- **Evaluation of every scientific output**

- ***Better Science***

- Acceleration of Science
- Scientific Integrity and Reproducibility
- Interdisciplinarity and Cooperation
- Long-term Access, Data Safety
- Higher Data quality

The need for research data management



These problems can be partially solved through responsible research data management

«Research data management concerns the organisation of data, from its entry to the research cycle through to the dissemination and archiving of valuable results. It aims to ensure reliable verification of results, and permits new and innovative research built on existing information».

Source: Whyte, A., Tedds, J. (2011). Making the Case for Research Data Management. [DCC Briefing Papers](#). Edinburgh: Digital Curation Centre

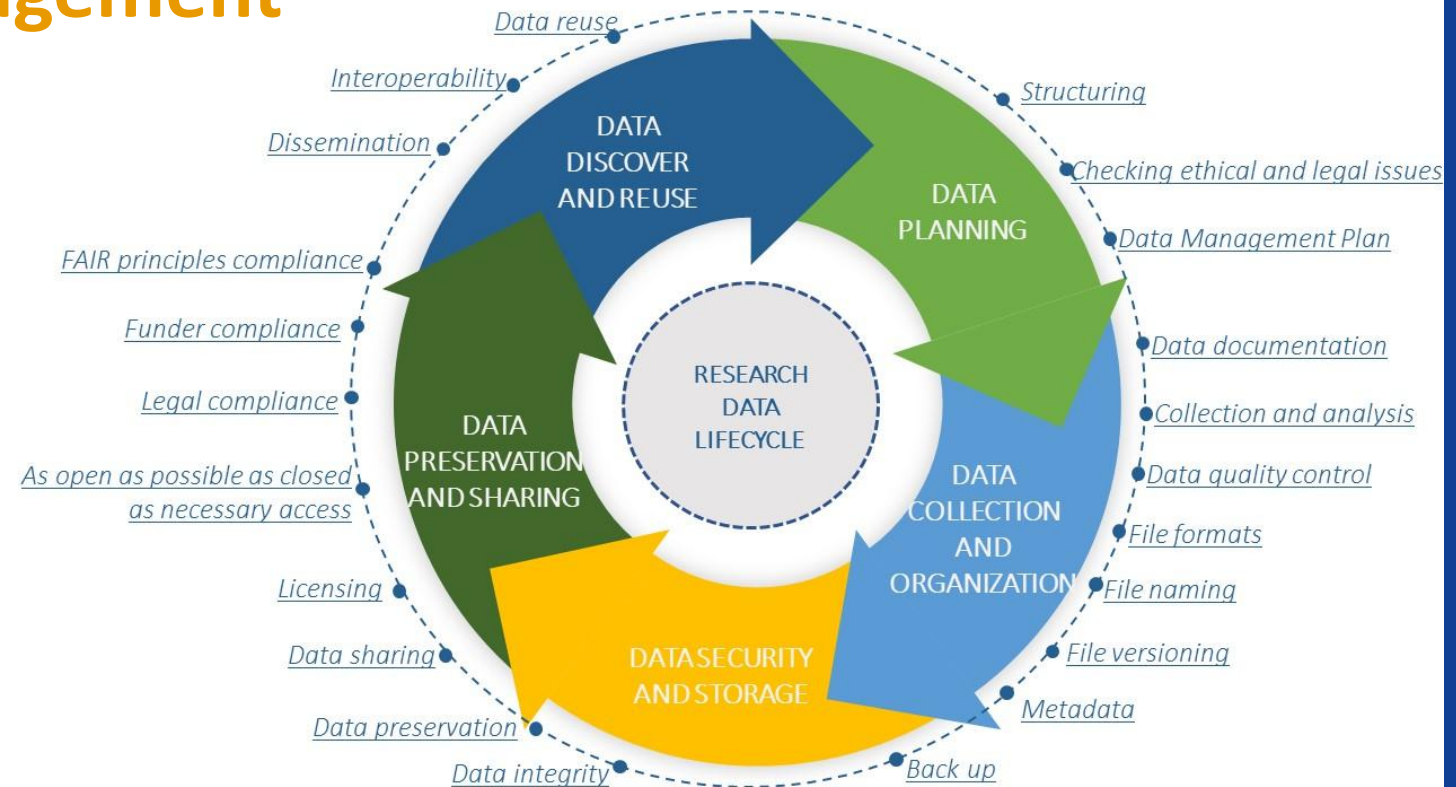


Image created by C. Manfredini for <https://rdm.unimi.it/vision-what-is-rdm/> - subjected to CC-BY-SA license

Really Why!

Institutional actions towards RDM & Open data

UNIVERSITY RESEARCH STUDY TRANSFER

PROFILE

RESEARCH DATA POLICY

PRINCIPLES FOR THE MANAGEMENT OF RESEARCH DATA

The availability of research data guarantees that research data can be reproduced as well as put to further use. This is a key aspect of good scientific practice and according to the highest standards builds on this principle and is part of the

1. Responsibility for the life cycle (*) of research data, in particular saving and supplying research data for long-term use by Investigators (PIs).

Heidelberg's policy

UNIVERSITÀ DEGLI STUDI DI MILANO

Research Data Management Policy

1. Preamble

The definitions of "research data" and of "affiliates of the University of Milan" that are specified in this Policy are specified in the attached document N.1.

The University of Milan acknowledges the fundamental importance of research data as value and the University recognizes the relevance of records management for supporting the progress of scientific research and is committed to pursuing the highest standards of data collection, storage and management.

The University of Milan adheres to the principles of Open Science and in particular to the principle of sharing of scientific publications and data produced by its staff with public funding as far as possible as closed as necessary" principle. In particular, the University of Milan adheres to the FAIR (Findable, Accessible, Interoperable, Reusable) paradigm.

UNIMI's policy

SORBONNE UNIVERSITÉ

RESEARCH COMMISSION OF 24 JUNE 2018

Open research data policy within the Sorbonne University Alliance

Since its creation, the Sorbonne University Alliance has chosen to be fully involved in the open science movement and to promote access to knowledge for all. With this in mind, it is committed to opening up the results of the research of its member institutions and promoting their dissemination, in order to share with citizens the fruits of their common research and to develop a trustworthy science.

Taking up for itself the terms of the Sorbonne declaration on research data rights of 27 January 2018, the Sorbonne University Alliance recalls that :

- access to and sharing of data enables the development of new knowledge and accelerates the benefit of society, while promoting economic development and employment;

Sorbonne's policy

Research Data Policies and Rules for good Scientific Practice!

European actions towards RDM & open data

Horizon 2020 & Horizon Europe: FAIR Data Management



- Participating projects will be required to develop a **Data Management Plan (DMP)**
- Participating projects are **required to deposit research data**, preferably into a research data repository
- “[...]as far as possible, projects must then **take measures to enable for third parties to access**, mine, exploit, reproduce and disseminate (free of charge for any user) this research data.”

National actions towards RDM & Open data



“[...] Assuming that the publication of research data from a DFG-funded project does not conflict with the rights of third parties (in particular data protection or copyright), **research data should be made available as soon as possible.**

[...]

Applicants may request funding for project-specific costs that arise in connection with a scientific project, for the preparation of research data for subsequent reuse and/or the transfer of research data to existing infrastructures as part of a proposal to the DFG. [...]”

Piano Nazionale di Ripresa e Resilienza

#NEXTGENERATIONITALIA



Coerentemente con le finalità del presente Avviso, ai risultati del Programma di ricerca e ai relativi dati (ad esempio, le pubblicazioni di risultati originali della ricerca scientifica, i dati grezzi e i metadati, le fonti, le rappresentazioni digitali grafiche e di immagini e i materiali multimediali scientifici) deve essere garantito un accesso aperto al pubblico nel minor tempo e con il minor numero di limitazioni possibile, **secondo i principi “Open science” e “FAIR Data”.**

Editorial actions

PLOS ONE

[Plos](#)

- Introduction
- Minimal Data Set Definition
- Acceptable Data Sharing

Data Availability

The following policy applies to all PLOS journals, unless otherwise noted.

Introduction

PLOS journals require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data.

When submitting a manuscript, authors must provide a Data Availability Statement describing compliance with PLOS' data policy. If the article is accepted for publication, the Data Availability Statement will be published as part of the article.

Acceptable data sharing methods are listed [below](#), accompanied by guidance for authors as to what must be included in their Data Availability Statement and how to follow [best practices in research reporting](#).

PLOS believes that sharing data fosters scientific progress. Data availability allows and facilitates:

SPRINGER NATURE

[Authors](#) [Publish an article](#) [Publish a book](#) [Publish conference proceedings](#)

[Research data](#)

Research data policy

[Research data policy](#)[Data availability statements](#)[Data repository guidance](#)[Sensitive data](#)[Data policy FAQs](#)[Research data helpdesk](#)

Data availability statements

Guidance for authors and editors

An article's data availability statement lets a reader know support the results and analysis. It may include links to p analysed or generated during the study, descriptions of w

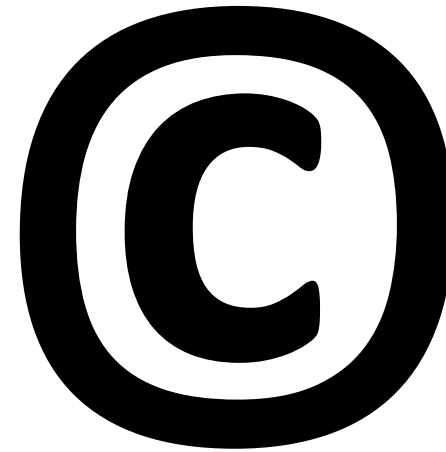
Wiley's Data Sharing Policies

[Wiley](#)

Wiley is committed to a more open research landscape, facilitating faster and more effective research discovery by enabling reproducibility and verification of data, methodology and reporting standards. We encourage authors of articles published in our journals to share their research data including, but not limited to: raw data, processed data, software, algorithms, protocols, methods, materials.

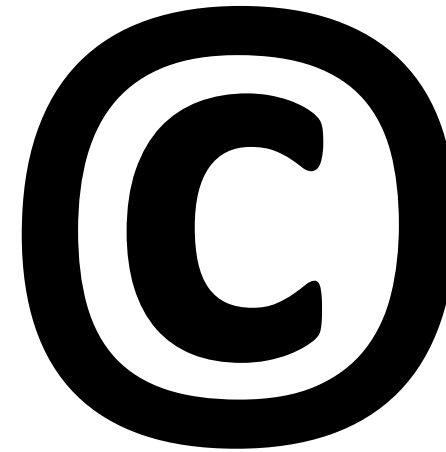
Research data and copyright

- Textual data typically are protected by copyright
- Copy right holder can grant simple or exclusive usage rights
- For publications in subscription journals: typically unlimited and irrevocable transfer of rights to the publishers



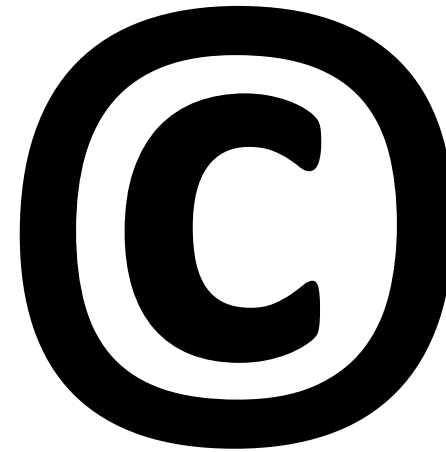
Research data and copyright

- Facts like measurements generally do not reach the **threshold of originality**, even though the data collection can be very sophisticated.
- Therefore: According to European copyright laws, research data are in many cases not copyrighted.



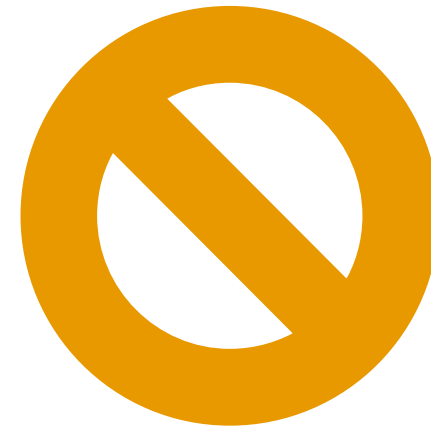
Research data and copyright

- But many data are in databases and there is some kind of protection for these (EU directive 96/09/EG).
- Virtually all data are useless without documentation. This documentation might very well be protected by copyright.
→ **threshold of originality**



(Good) Reasons not to publish ones Data

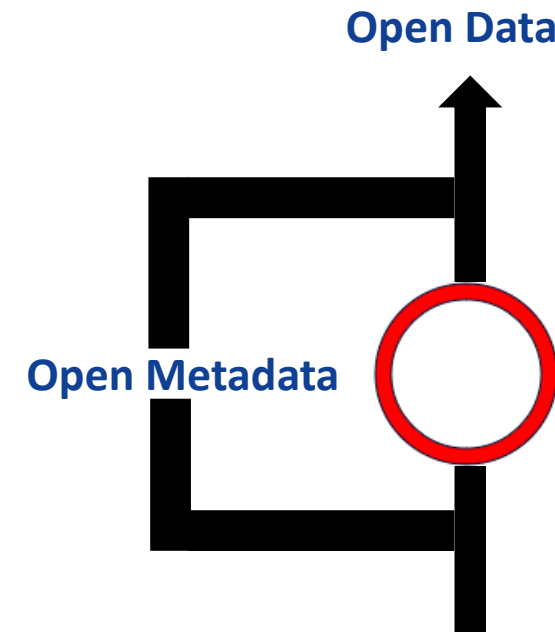
- Data Sensitivity
 - Personal data, racial or ethnic origin, political opinions, religious or philosophical beliefs, genetic data, biometric data, health-related data, data concerning a person's sex life or sexual orientation (GDPR definition)
 - Impossible to remove or anonymize
- Dual Use
- Copyright/Licences
 - Copyright protected data
 - 3rd party data



Alternatives

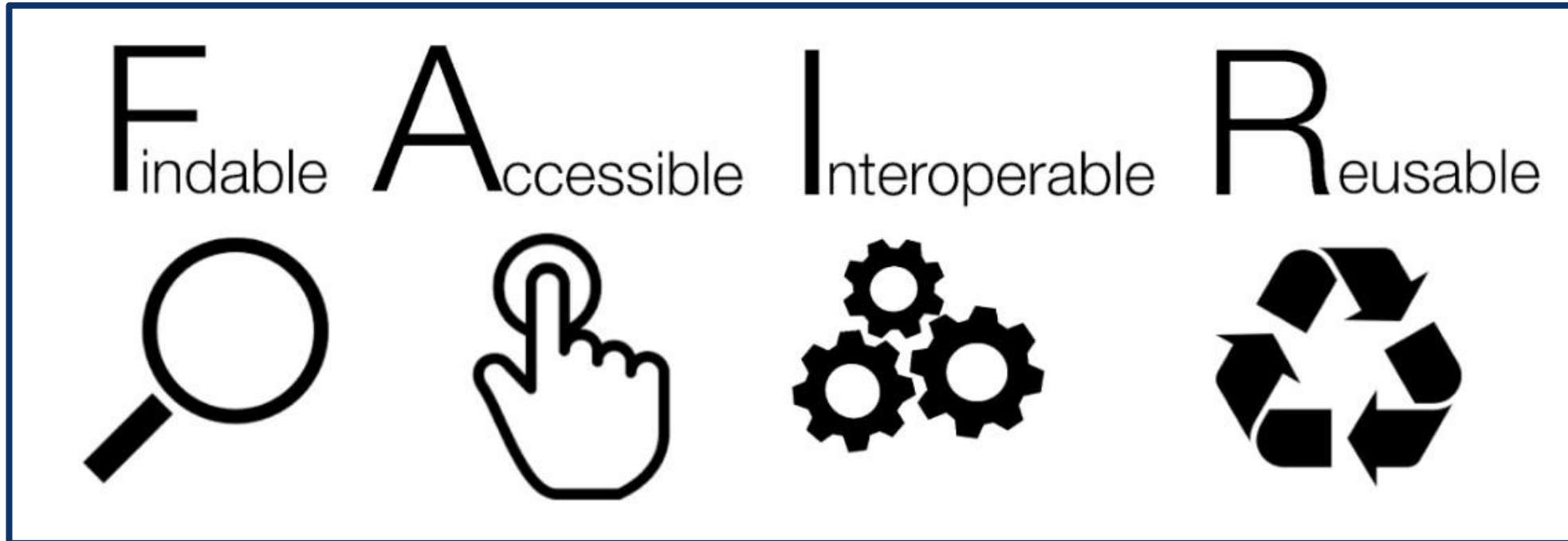
Even if you can not publish your data, you can inform about it!

- Use specialized data repositories
- Publish your metadata
- Enable restricted access
- Publish a data usage agreement



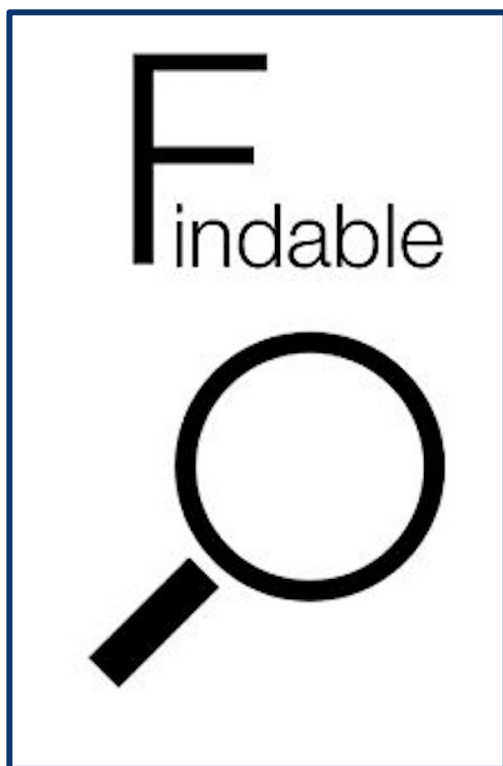
FAIR data principles

Image by SangyaPundir [CC BY-SA 4.0, from Wikimedia Commons]



Wilkinson et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship,
[Scientific Data 3](#)

F for Findable



- ❑ use globally unique persistent identifiers (**PID**)
- ❑ use rich and detailed metadata that describe your data
- ❑ promote (e.g. social networks)
- ❑ deposit in **repositories**, catalogs and databases which enable automatic harvesting of metadata



<https://doi.org/10.11588/DATA/HR9XFS>

PIDs are sequences of letters and/or numbers in specified format, managed by some service provider:

- ❑ **Unique** – different things have different identifiers (DOI, ORCID, ROR)
- ❑ **Persistent** – should work forever
- ❑ **Resolvable** – there is a way how to get representation of identified object (a Url is not a PID, but many PIDs are resolvable in Urls)

If you choose the right repository for your data half of the work is done!

A for Accessible



- ☐ share your data in a dedicated **repository**
- ☐ **openly** publish your **metadata** with a **public domain license**
- ☐ use clear statements and **instructions for closed data**

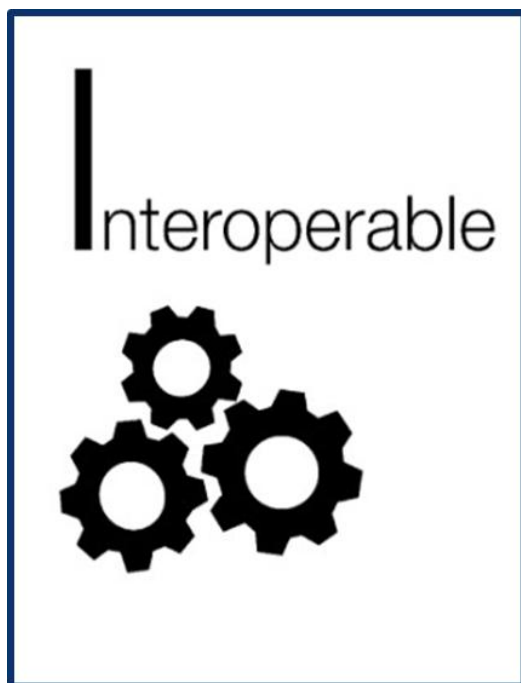
https://doi.org/10.13130/RD_UNIMI/BFU3KI

Remember:

- ☐ Do not restrict if not necessary (apply an embargo or other policies)
- ☐ Be clear on restrictions and make it easily understandable how to access your data

«as open as possible, as closed as necessary»

I for Interoperable



- ❑ describe your data in a detailed and comprehensible way in the **metadata** fields (as in this [example](#))
- ❑ add helpful and detailed documentation on data creation and processing: **compile a readme file**
- ❑ **use the right language:** well-known terminology, domain standards, English language
- ❑ **use readable, non-proprietary and [open formats](#)** (or, at least, give information on how to open the proprietary format file, as in this [example](#))

	0	1
1	0	2
2	0	3
3	0	4
4	0	5
5	0	6
6	0	7
7	0	8
8	0	9
9	0	10
10	0	11
11	1	12
12	0	13
13	0	14
14	0	15
15	0	16
16	0	17
17	0	18
18	0	19
19	0	20
20	0	21
21	1	22
22	0	23

R for Reusable



- ☐ Data and its **provenance** is trustworthy
- ☐ **Open licenses** have been used for both metadata and data to make clear what it is allowed to do with the data

https://doi.org/10.13130/RD_UNIMI/E5WJHP

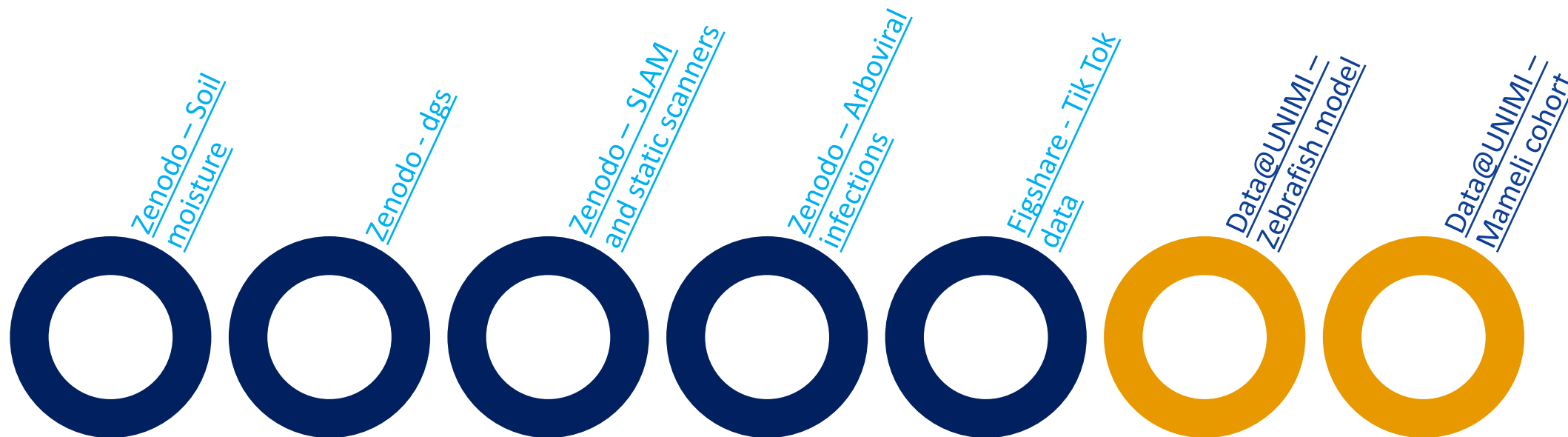
Describe in detail in the **metadata and in the readme file** the methodology for data collection and data processing, the quality control operated on data, and all the analysis elaborated to obtain the final results

Use license/s describing what can be done with your data – Free to share? Commercial applications?

FAIR data VS open data

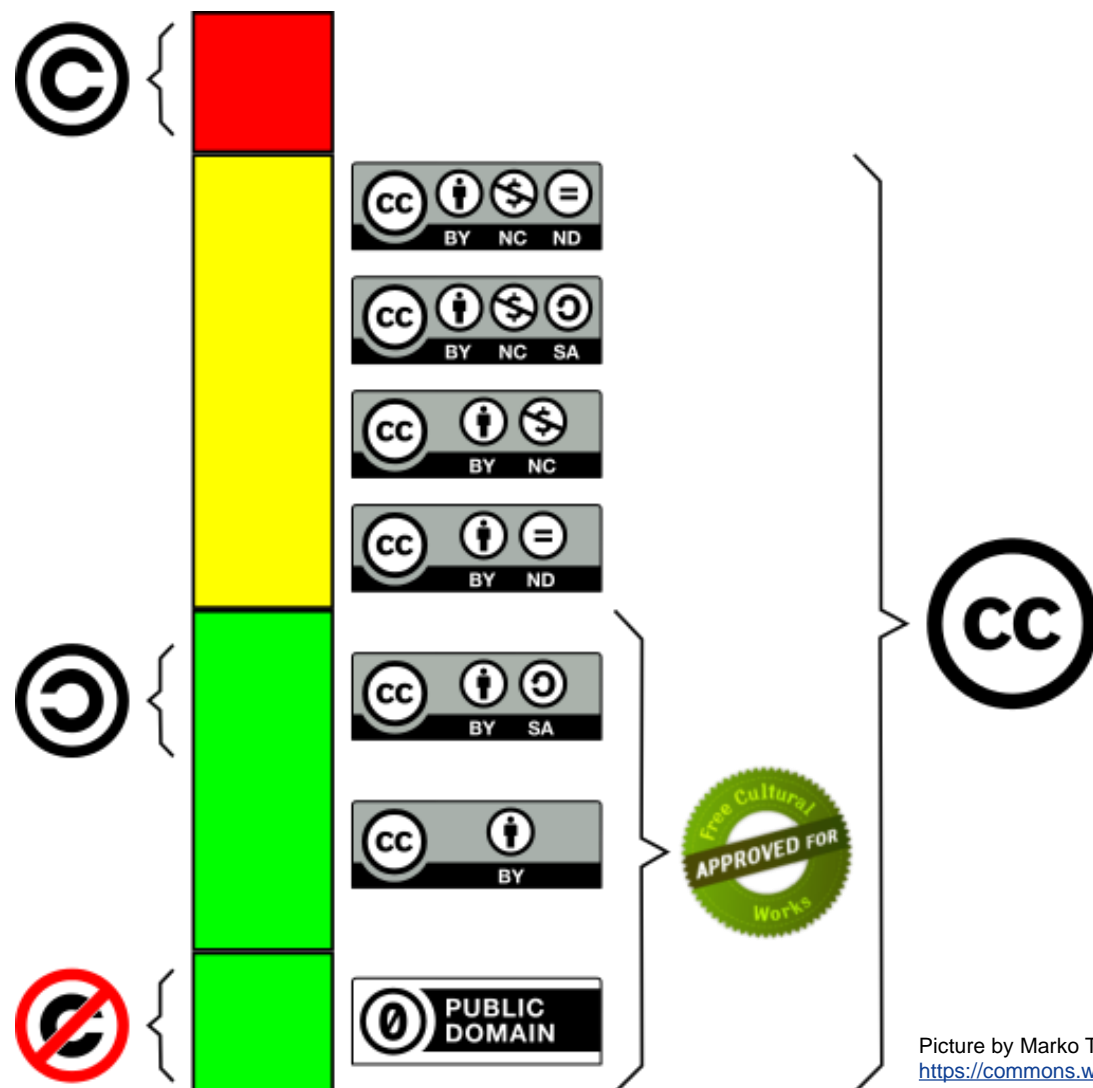
- FAIR is different from OPEN!
 - **FAIR data cannot always be open**
 - **Open data must necessarily be FAIR**
- Making data available is useless if it is not reusable. To do this, it must be accompanied by **clear, detailed, contextualized information** and a clear description of the data. And **data quality** should be guaranteed!

FAIR or non-FAIR?



Legal Tools – Creative Commons

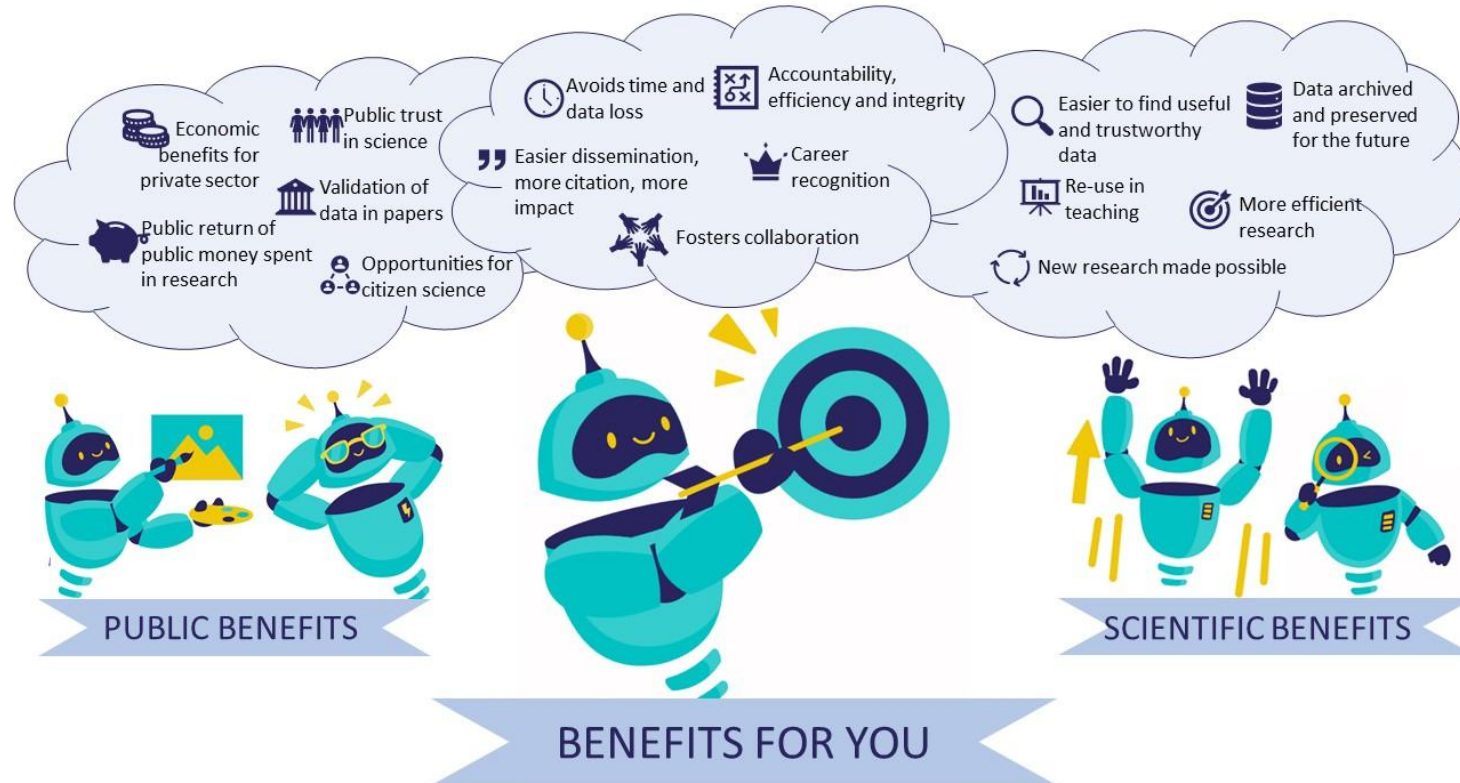
- Standard licenses that determine the scope of use of a work
- Combination of layperson-friendly formulation and a legally proper license text (adapted to the relevant national law – version 3.x)
- License content and metadata are available in machine readable form and can be added to a document.
- Modular structure with differing “degrees of freedom”



Picture by Marko Txopitea "Txopi" [CC0], via Wikimedia Commons,
https://commons.wikimedia.org/wiki/File%3ACreative_Commons_Semaforoa.svg

Legal Tools – Licenses

- CC Licenses the most common
 - KI Signals in development
- Alternatives e.g. Open Data Commons
- Specialized licenses for (Research Software)
 - MIT License: Highly permissive, simple, and popular for research.
 - Apache License 2.0: Permissive, with explicit patent rights.
 - GNU GPL (v2/v3): Strong copyleft, ensuring the software and its derivatives remain open.
- Tools (and service points) help to choose a fitting license

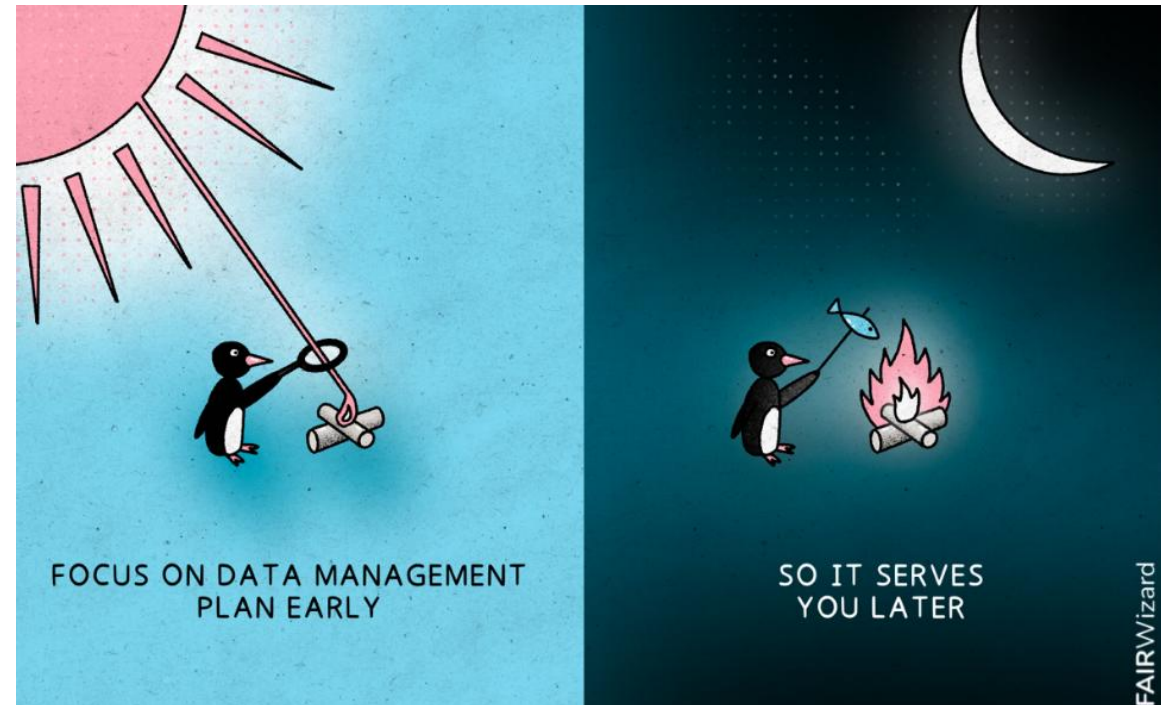


- Research data management is **an integral part of the research process**, and it should not be seen as an obligation
- If performed properly:
 - It is useful for researchers
 - It ensures that both the public and the scientific communities benefit from public research results

The Data Management Plan (DMP) is useful tool to help you properly manage your data and comply with funders and editors' requirements

What is a DMP?

- A structured document that describes what research data is created, how it will be managed and the different responsibilities. It enlists and describes the **actions that will be taken to produce, analyze, store, preserve and share the data generated by a research.**
- It is written at the **beginning of a research**, but is a **living document** which should be updated whenever the need arises



What should a DMP contain?

1. Summary

Type of data
used and reused

Aim of the data
collection/reuse

Sources of data

2. FAIR data

Findability

Accessibility

Interoperability

Reuse

3. Research results

Physical outputs

Digital outputs

4. Resources

Management
costs

Management
responsibilities

Researchers
involved in
management

5. Security

Integrity and
quality
preservation

Archiving

Long-term
preservation

6. Ethics

Legal and
ethical issues

Consent forms,
other relevant
documentation

[Check here Horizon EU's Template](#)

Workshops

These workshops are open to everyone and will be held online. They will focus on putting into practice what was covered in the webinars. **Active participation in the exercises part is expected from those who register.**

Two dates are scheduled for each workshop to encourage discussion within smaller groups. **Please register for only one date per workshop.**

Important: Please note that registration for the workshops will only open once the corresponding webinars have taken place.

Data Management Plans in practice: From principles to action

Tuesday, 28 April 2026, 10:00-12:00 AM

Tuesday, 5 May 2026, 10:00-12:00 AM

Meriç Akdoğan (Sorbonne University, France), Laura Giovinnazzi (University of Milan, Italy), Milan Janíček (Charles University, Czech Republic) and Carolina Manfredini (University of Milan, Italy)

A Data Management Plan (DMP) is a key research data management tool and a mandatory requirement for many funded projects. This hands-on workshop will allow the participants to analyse a partially completed DMP for a fictional research project and to contribute to the completion of key sections. To participate in this workshop, attendance at the webinar on open data is strongly recommended.

**Find out more
about DMPs in
the dedicated
workshop:**

[Enroll here](#)



Types of repositories

- **Institutional repositories**

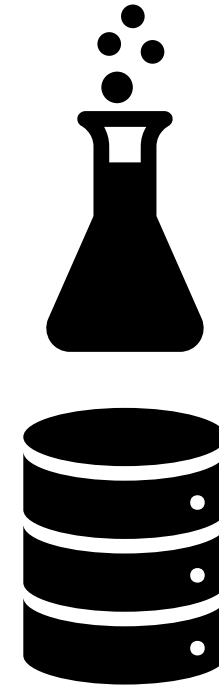
- Known contacts, part of broader cooperation
- Easy, quick, direct, flexible (personal and systems)
- Control over data security and long-term accessibility
- General metadata and services
- Medium visibility



Types of repositories

- **Subject specific repositories**

- Tailored metadata fields and services
- Specialized personal
- High visibility (in community)
- Not available for every subject/field
- High requirements

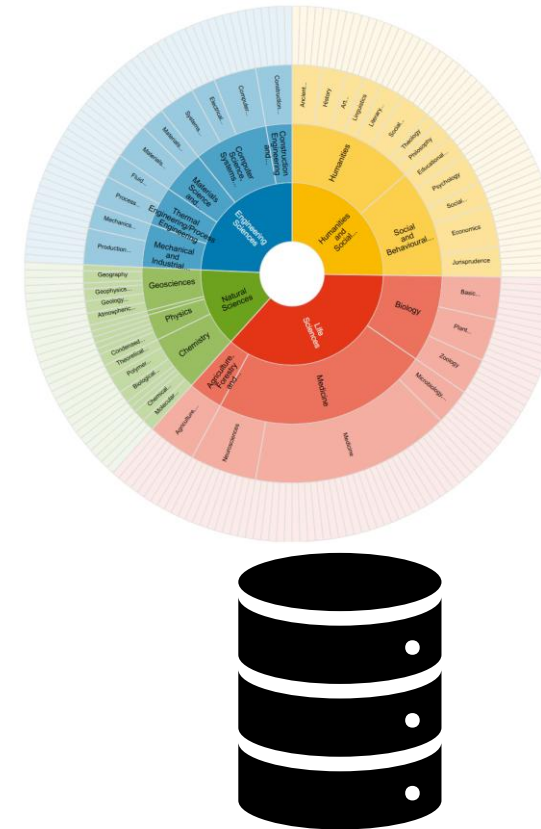


Types of repositories

- **General repositories**

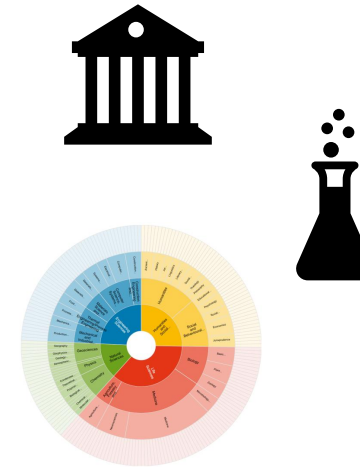
- Quick and easy publication
- Flexible licenses
- No or only general support
- Unsure (longterm) accessibility
- No curation, unsure (meta-)data quality

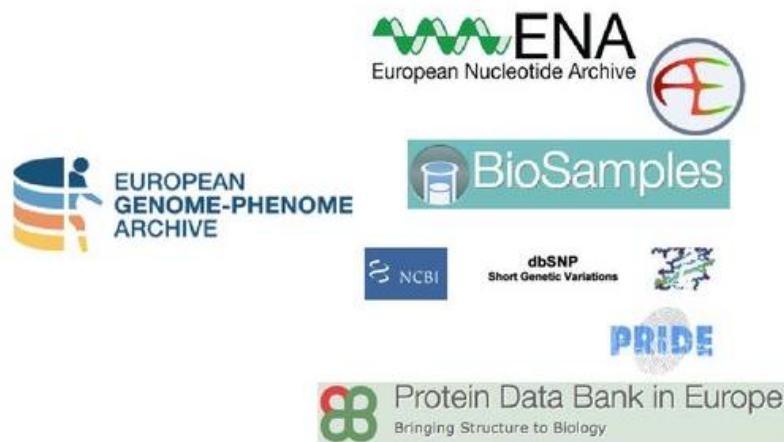
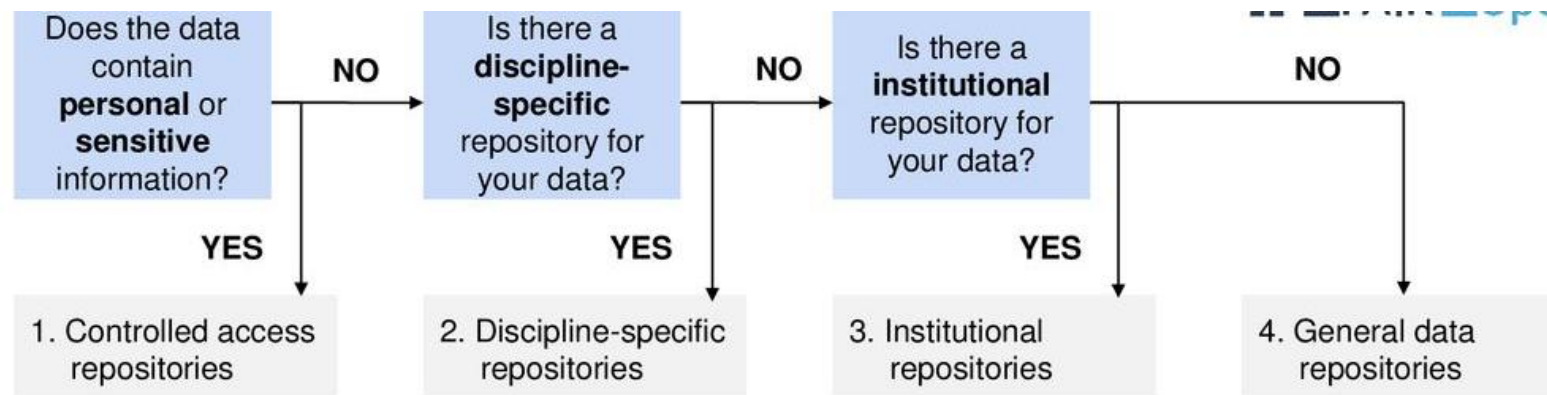
<https://www.re3data.org/browse/by-subject/>



Types of repositories

- **Types of repositories**
 - Institutional
 - Subject Specific
 - General
- All have advantages and disadvantages
- Dual deposit in most cases possible





Etc...



Slide adapted from: N. Jareborg (2019), „Data management and repositories“, <https://player.slideplayer.com/105/17629367/>.

How to find a repository for my data

- How to find
 - Where do **you** find interesting data
 - Where do **colleagues** find and publish data
 - Use databases on databases
e.g. RE3Data
 - Suggestion/requirements of **publishers**
- Ask your local service point

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

  Search

Know who you can contact at your University for assistance with data management and open research practices

Charles University: *Open Science Support Centre* – researchdata@cuni.cz

Heidelberg University: *Research Data unit* – data@uni-heidelberg.de

Paris-Panthéon-Assas University:

Sorbonne University: *Research Data and Digital Humanities Team, Direction of Archives, Libraries and Museum collections* – labrador@sorbonne-universite.fr

University of Copenhagen: *Research Support, Copenhagen University Library* – forskervservice@kb.dk

University of Geneva: *Research Data team, Division de l'information scientifique* – researchdata-info@unige.ch

University of Milan: *Research Data Management support team, Performance, Quality Assurance, Assessment, and Open Science Policy Division* – rdm@unimi.it

University of Warsaw: *ICM UW, Mrs. Agnieszka Cybulska-Phelan* – a.cybulska@icm.edu.pl

4eu+



Co-funded by
the European Union

Thank you !

Open for you! cycle



Univerzita
Karlova



UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386



PANTHÉON-ASSAS
UNIVERSITÉ
PARIS



SORBONNE
UNIVERSITÉ



KØBENHAVNS
UNIVERSITET



UNIVERSITÉ
DE GENÈVE



UNIVERSITÀ
DEGLI STUDI
DI MILANO



UNIwersYTET
WARSAWSKI