

iEntrance FAIR Data Day, 29th November 2023

FAIR RESEARCH DATA MANAGEMENT

Advantages and practices

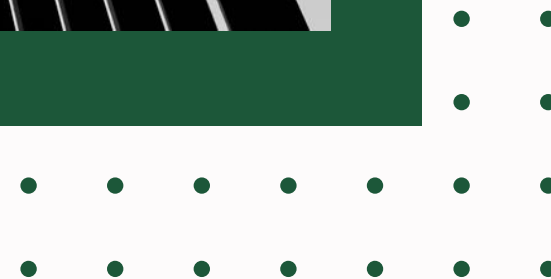
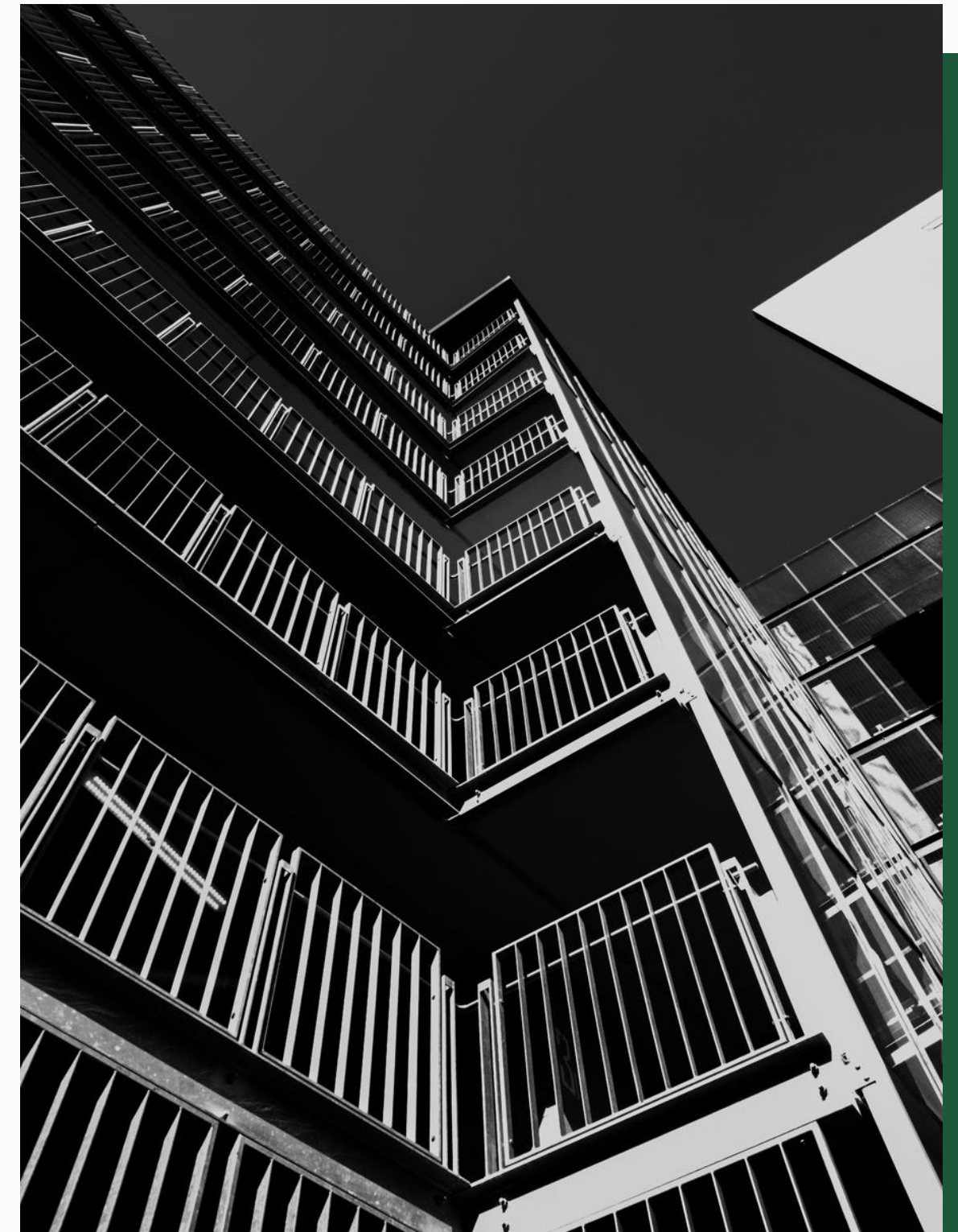
Giulia Caldoni, Data Steward
<https://orcid.org/0000-0001-6105-721X>



29th November 2023

Content

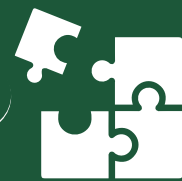
- 01 The 5 Ws of FAIR
- 02 FAIR in practice:
Data Management Plan
- 03 To summarize



The 5 Ws of FAIR

FAIR DATA MANAGEMENT

01



WHAT

WHAT DO WE MEAN BY DATA?

DATA



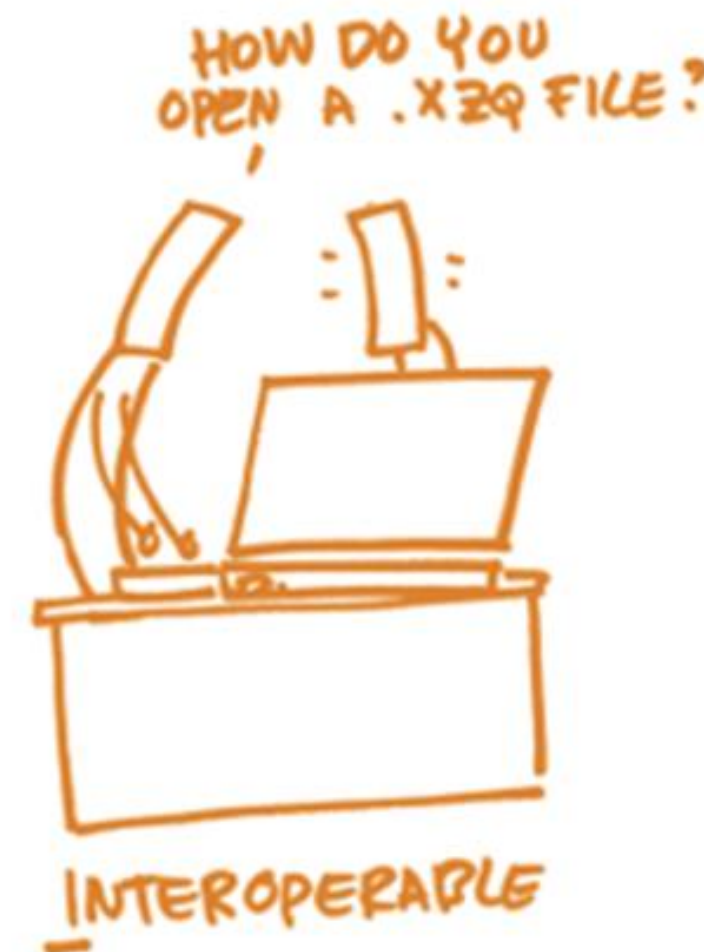
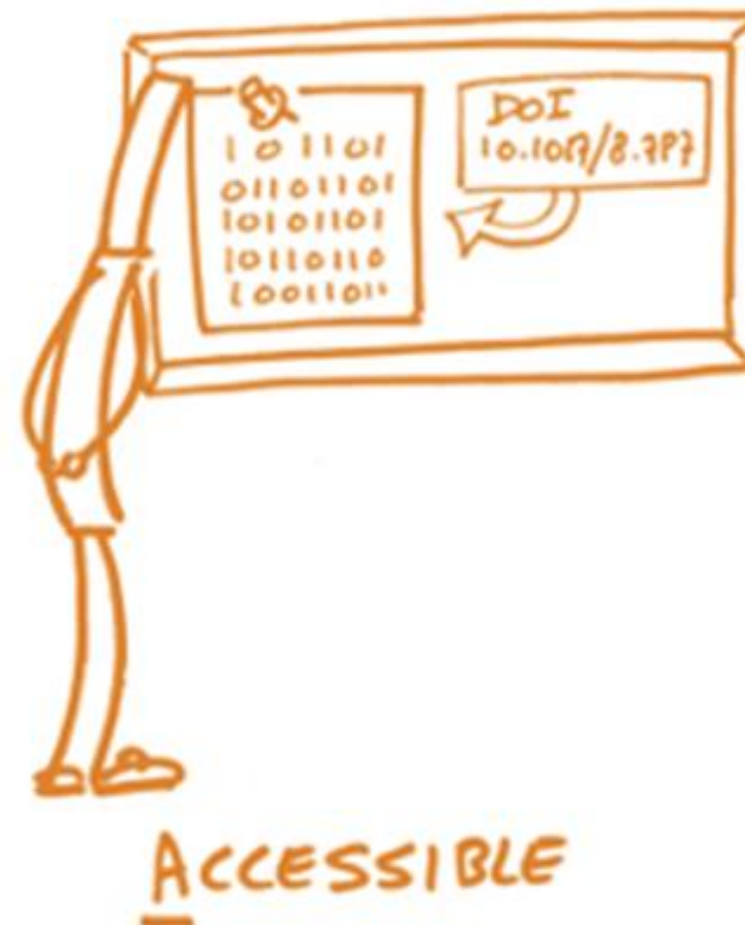
- **Information** collected to be examined and used as a **basis for reasoning, discussion or calculation.**
- Any research material different from publications.

DATASET



- **Collection of data** assembled with a common purpose, and **organized** to reflect the results of a research activity.

FAIR DATA PRINCIPLES



(Meta)data can be reliably found online

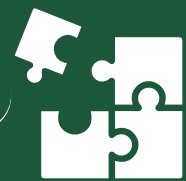
(Meta)data are always accessible

(Meta)data are created and managed using recognised standards and tools

(Meta)data are accompanied by clear rules on access and reuse

FAIR DATA MANAGEMENT

01



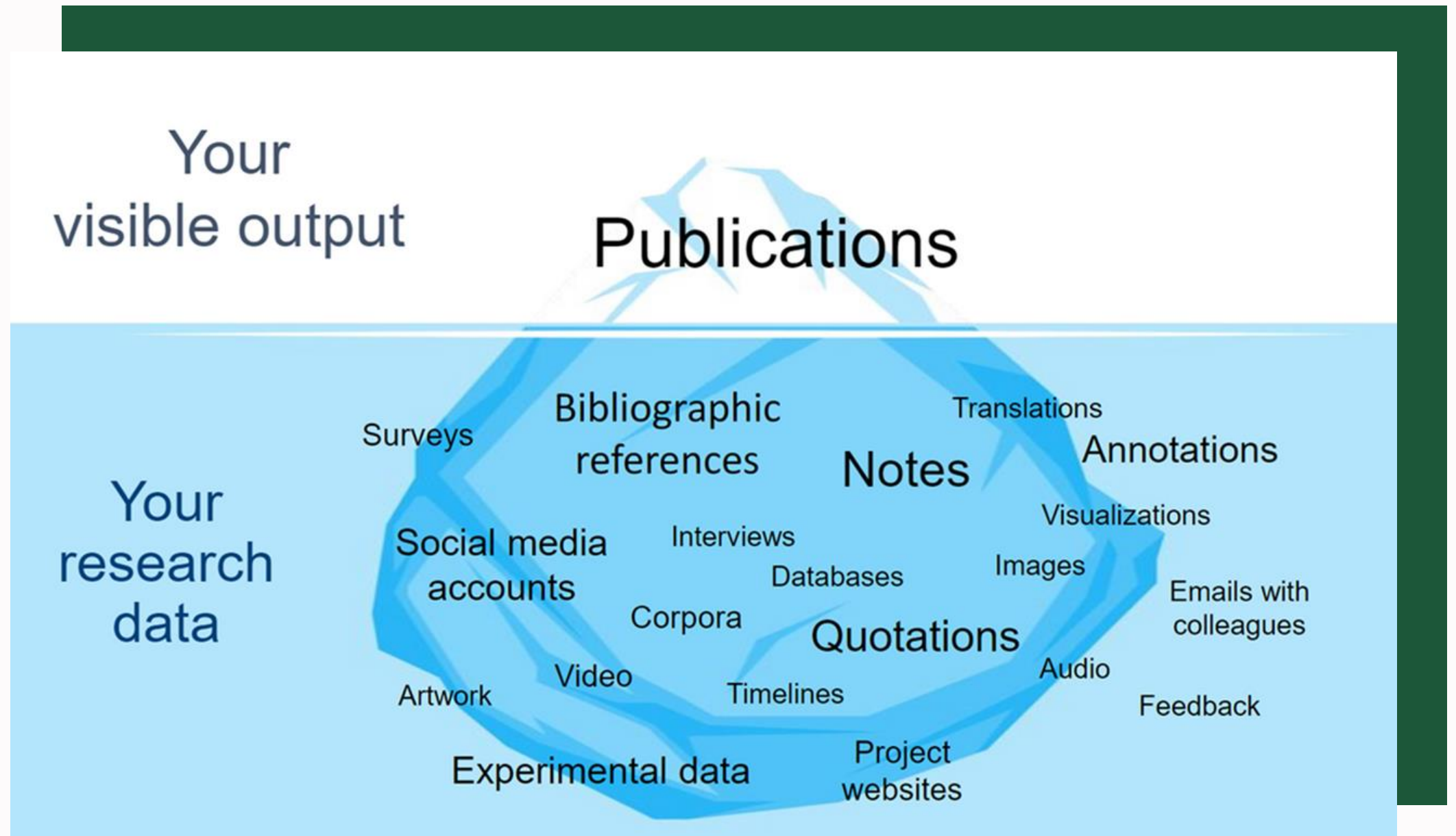
WHAT

WHY



02

REPRODUCIBILITY



The main objective is to provide the scientific community with the necessary data, methodologies, and research tools to validate (and in some cases replicate) the conclusions of a research.

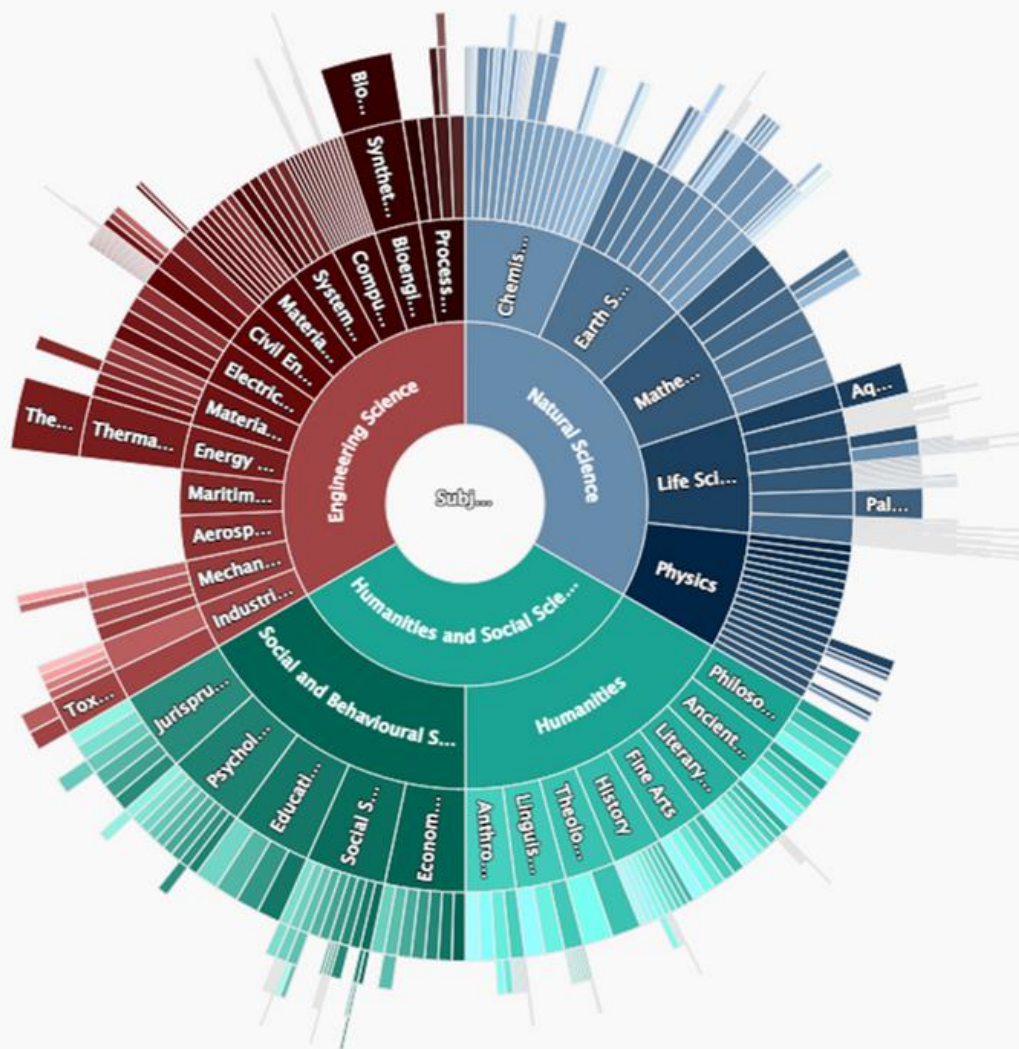


Having data available is of no use if it cannot be understood.

The use of community-defined standards for metadata is vital for reproducible research and allows for the comparison of heterogeneous data from multiple sources, domains and disciplines.

Metadata standards are also **discipline-specific**.

<https://the-turing-way.netlify.app/reproducible-research/rdm/rdm-metadata>
<https://fairsharing.org/browse/subject>



FAIR DATA MANAGEMENT

01



WHAT

WHY



02

03



WHO



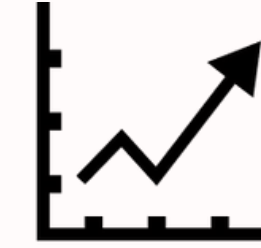
For researchers

- The chances of being cited grow, as do opportunities for collaboration
- Future research assessment is likely to include data
- Analyses and conclusions are more robust
- It is required by more and more funding bodies



For research

- Increases efficiency: less time is spent reinventing the wheel
- Institutions can measure their impact beyond scientific publications
- Increases the integrity and transparency of research



For society

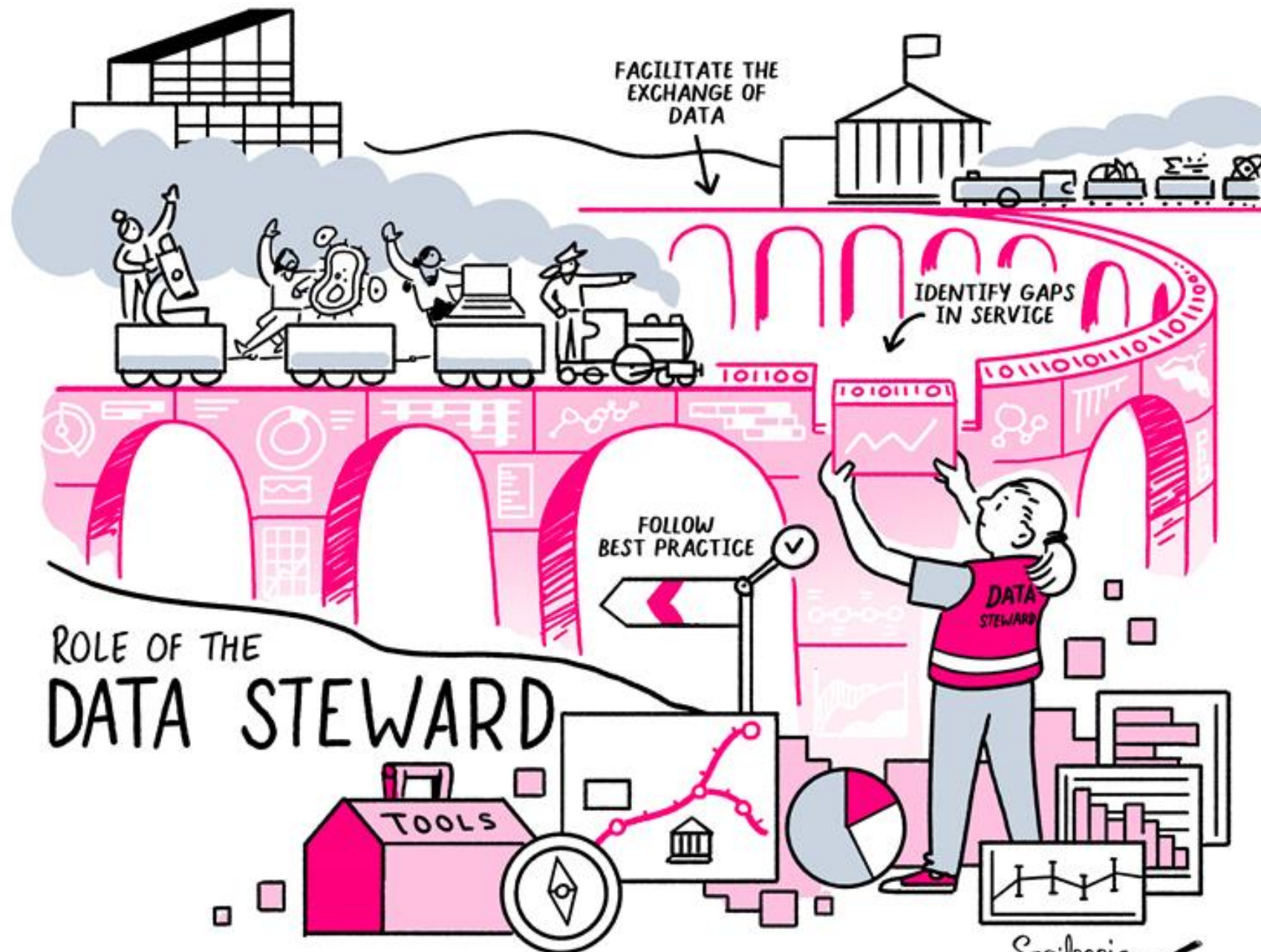
- Public funds are used more efficiently
- Better research leads to better decision-making
- Integrity and transparency help promote the public's trust in science

DATA STEWARDSHIP

Researchers are supported in data handling.

(Some of) The Roles of a Data Steward:

- Act as a knowledge and communication hub;
- Advise and support on policy, guidelines, infrastructure, services and tools;
- Align support functions, tool landscape and RDM policy to the FAIR (Findable, Accessible, Interoperable, Reusable) data principles and the principles of Open Science;
- Develop, implement, monitor and evaluate policies regarding research data;
- Translate RDM policy and legislation to RDM guidelines and tool selection that researchers can easily use.



FAIR DATA MANAGEMENT

01



WHAT

WHY



02

03



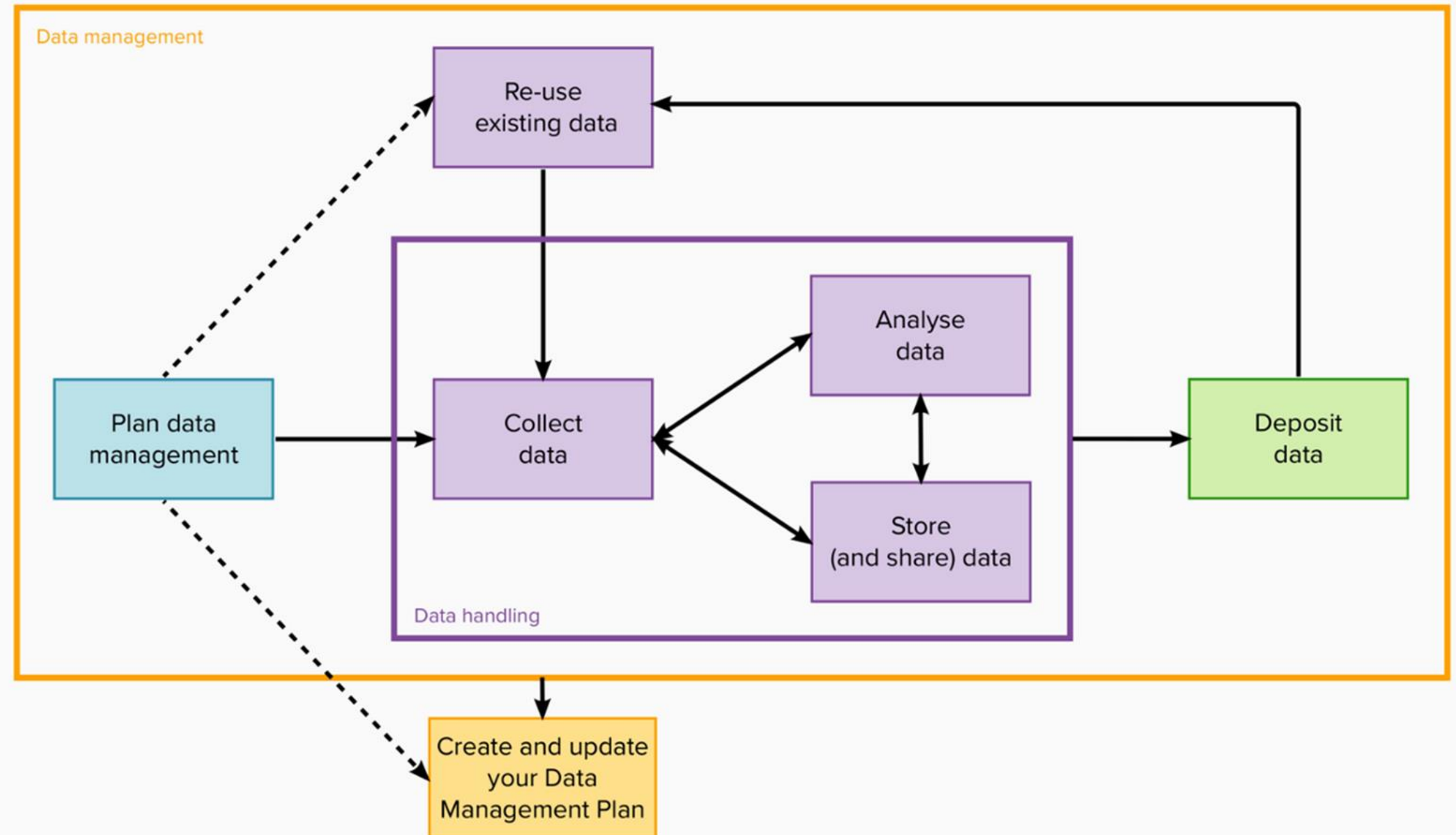
WHO

WHEN



04

THE RESEARCH DATA LIFECYCLE



PLANNING

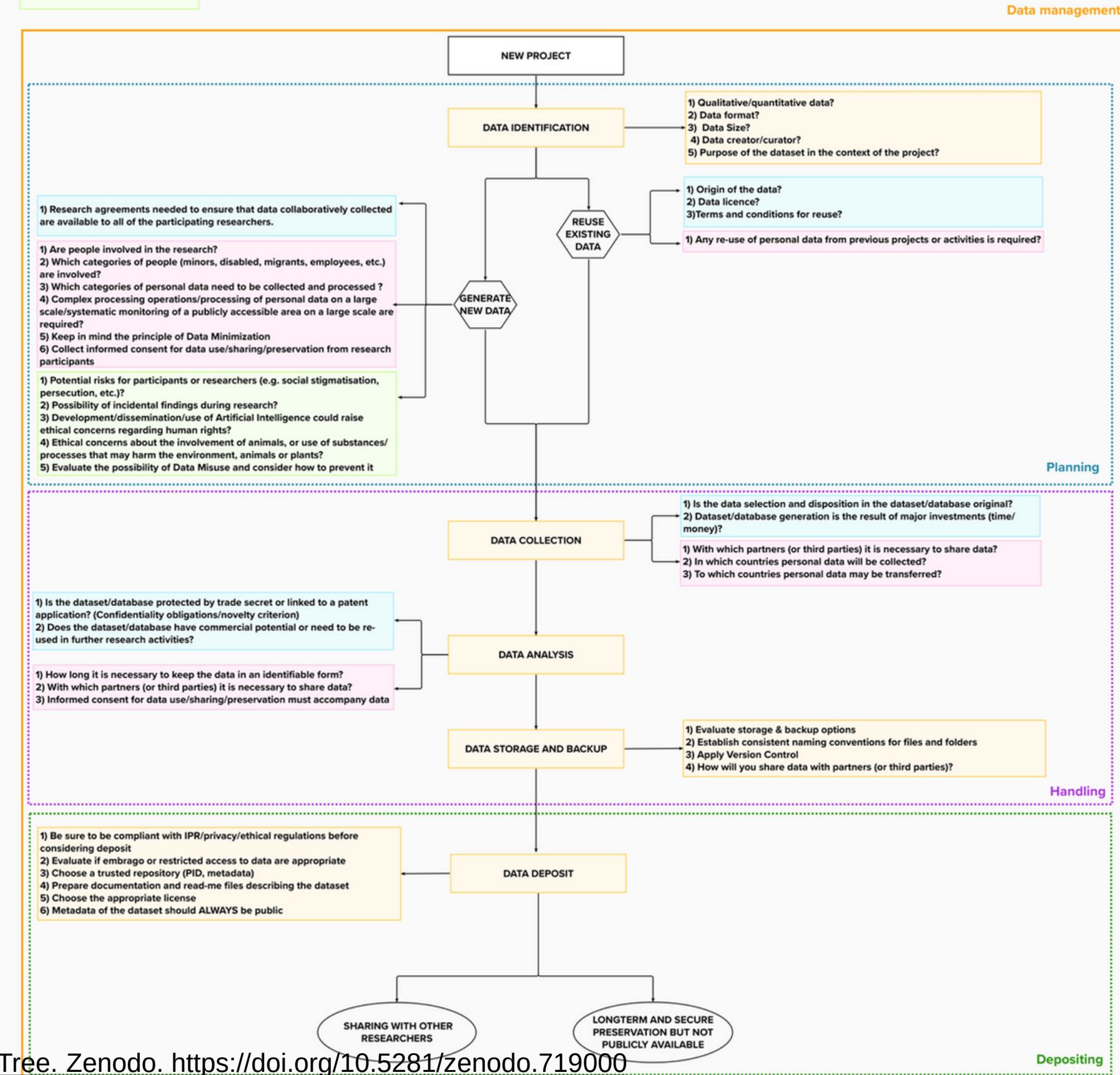
Planning data management is key, especially when many researchers come together in a big consortium.

Decisions need to be made early on, according to the nature of the research project, the type of data, and legal and ethical matters that may force or prevent certain decisions.

Legend:



DECISION TREE FOR DATA MANAGEMENT



FAIR DATA MANAGEMENT

01



WHAT

WHY



02

03



WHO

WHEN



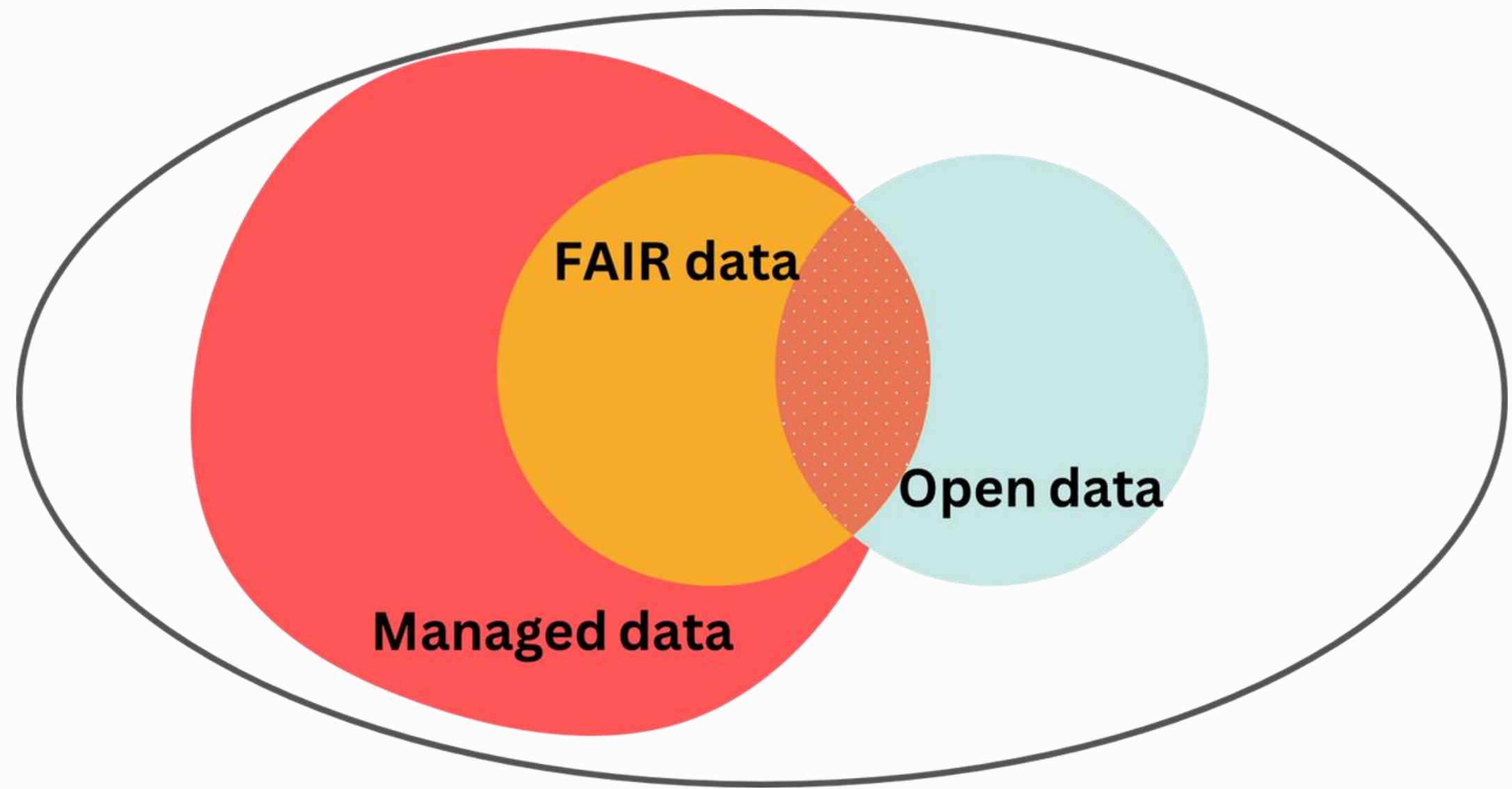
04

05



WHERE

**"AS OPEN AS
POSSIBLE, AS
CLOSED AS
NECESSARY"**



ACCESSIBLE DATA ARE NOT ALWAYS OPEN DATA

Open data are FAIR data released under an open license.
This is often Creative Commons Zero (CC 0), Creative Commons Attribution (CC BY)
or an equivalent license. See <https://opendefinition.org/licenses/>

DATA REPOSITORY

A data repository doesn't do all the work of making your data as FAIR as it can be.

However, it offers a **great framework** for getting the basics right.

- ✓ Persistent Identifiers (PIDs)
- ✓ Metadata schemas
- ✓ Licenses
- ✓ Clear Access conditions


Recent uploads

November 27, 2023 (1.0.0)

Dataset

Open

Dataset for manuscript titled "Medicinal chemistry lens o

Gadiya, Yojana 

This is the data directory for running the code available on the GitHub repository "patent data in SureChEMBL". The GitHub repository is available at <https://github.com/yogadiya/SureChEMBL> characteristics.

Uploaded on November 27, 2023

November 27, 2023 (v1)

Dataset

Open

Data from: Parentage of 920 gray-sided voles (*Myodes rufocanus*) enclosure between September 1992 and May 1994

Ishibashi, Yasuyuki ; Saitoh, Takashi

This dataset provides the estimated birth location, sex, assigned parents, and recapture data for 30 trapping sessions. We introduced 22 males and 25 females (*Myodes rufocanus*) born in a 3-ha outdoor enclosure in Sapporo, Japan between September 1992 and May 1994.

Uploaded on November 27, 2023



FAIR in practice: Data Management Plan

THE DATA MANAGEMENT PLAN



Describe the processes of data acquisition, analysis, management and storage during the research.

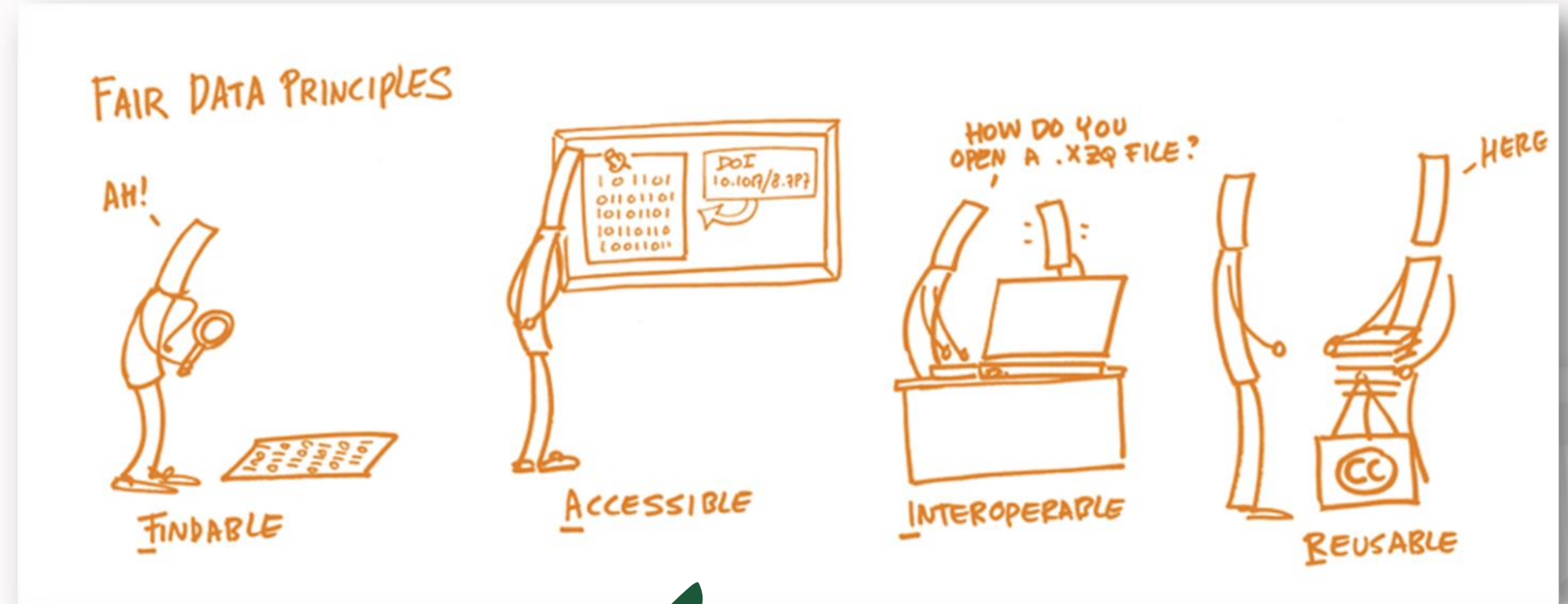
Describe who owns and is responsible for the data.

Describe how the data will be preserved and shared during and after the completion of the research.

THE DATA MANAGEMENT PLAN

A DMP is a structured document, a roadmap, consisting of a series of questions to be answered to:

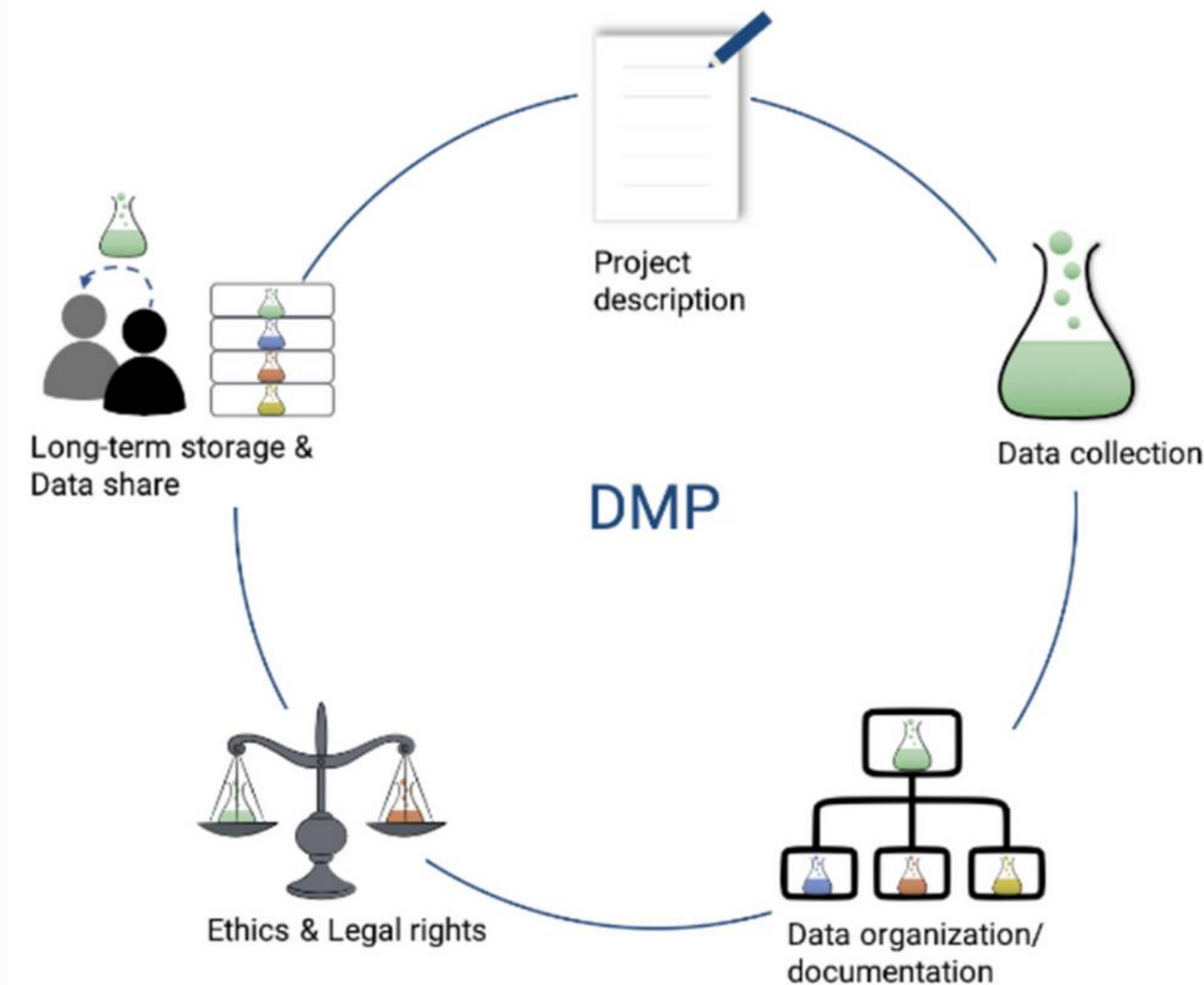
- Manage data while avoiding problems such as duplication, data loss, security breaches.
- Develop efficient and consistent procedures for data management, saving time later.
- Be aware of how you are fulfilling all your legal and ethical obligations.



THE DATA MANAGEMENT PLAN

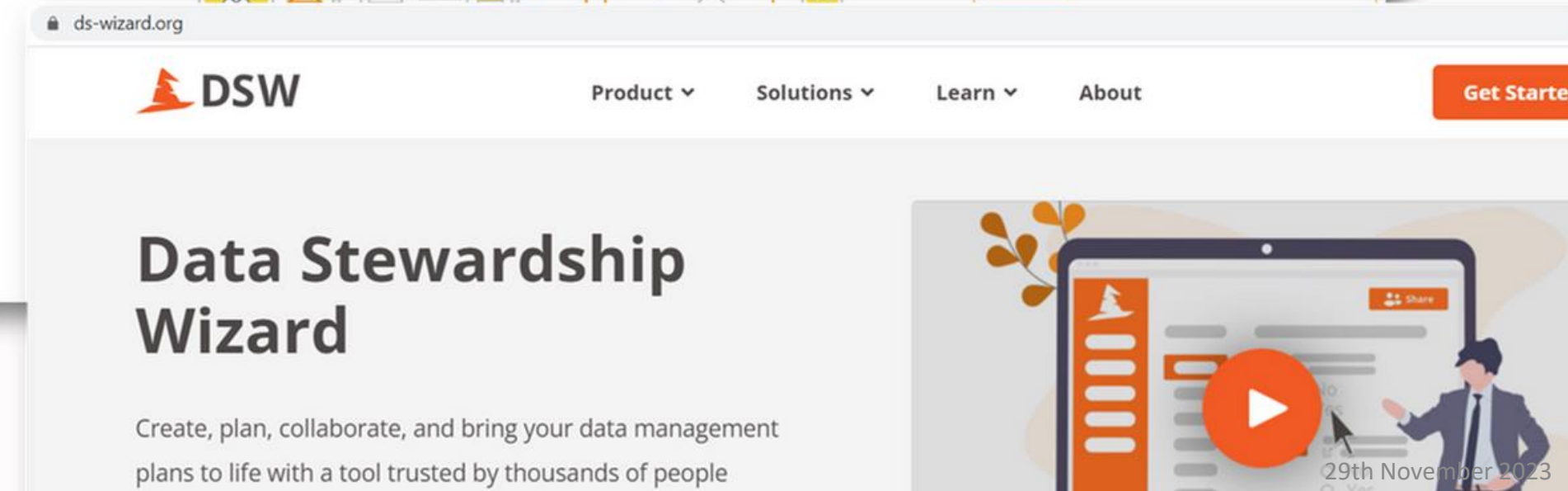
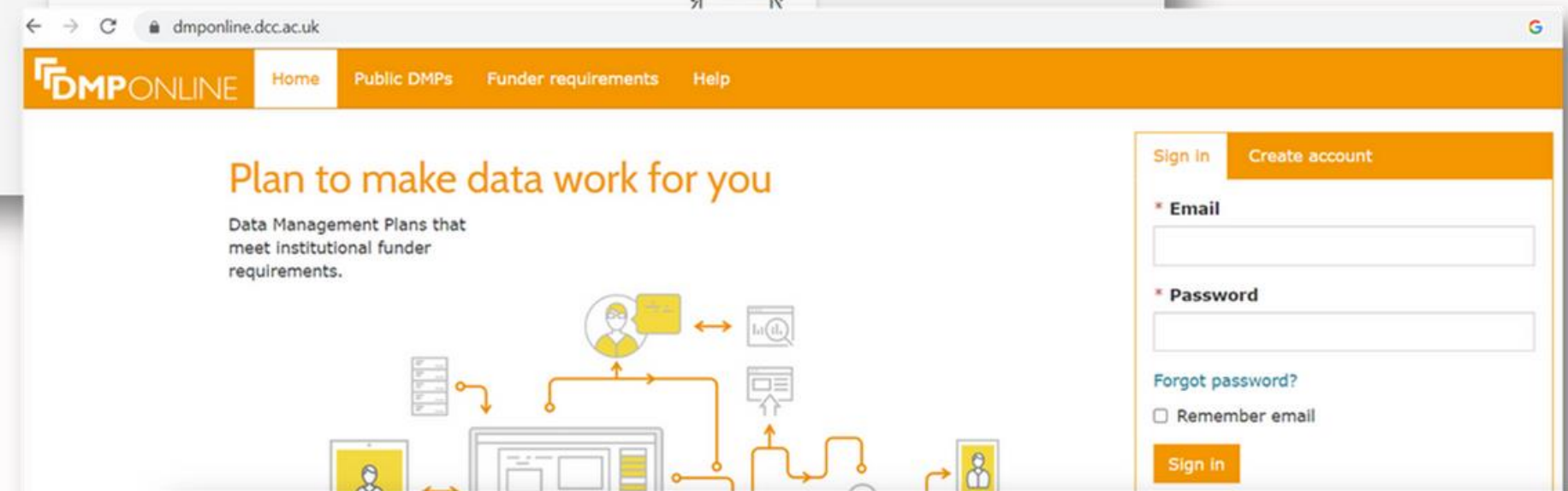
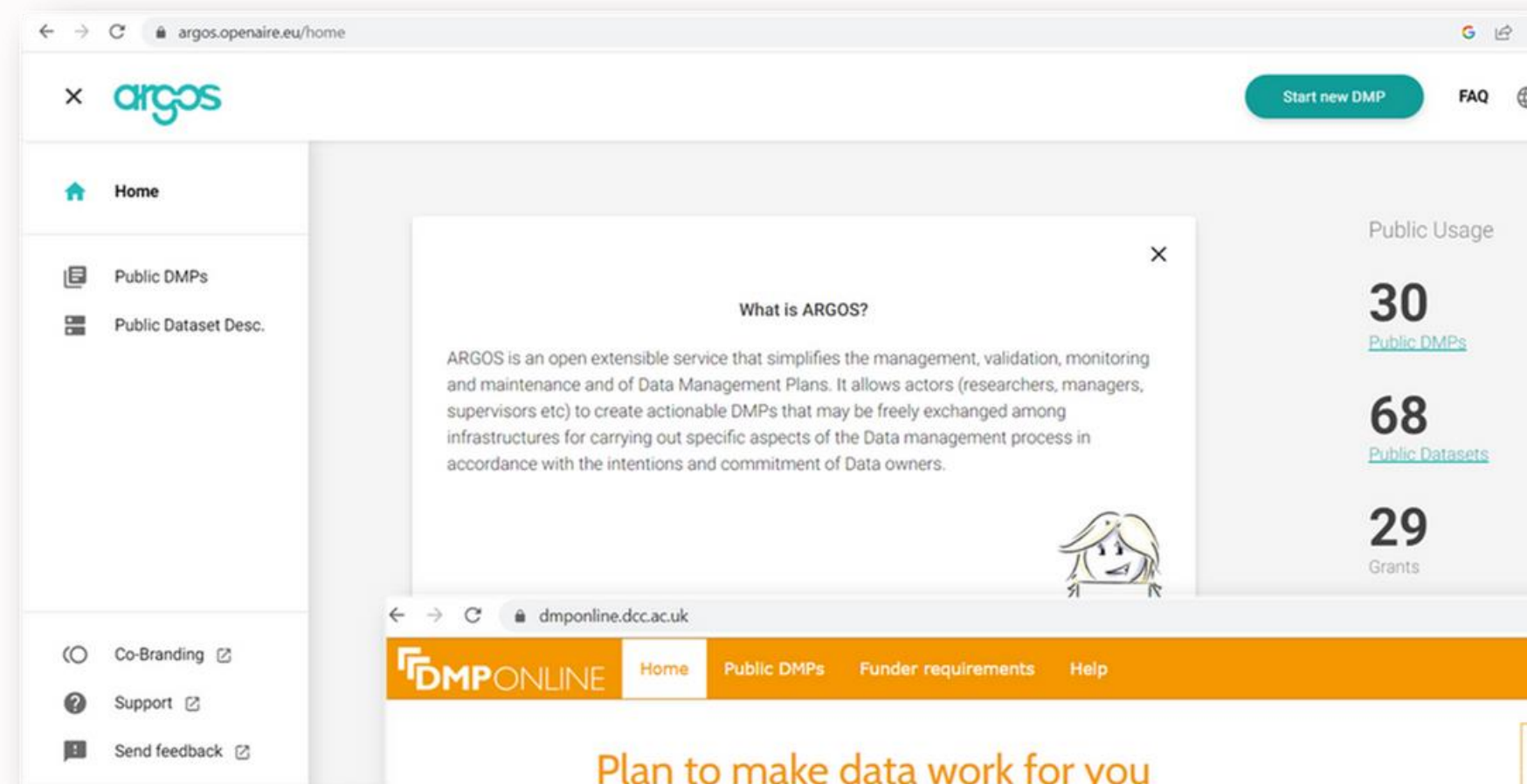
The DMP should include information on:

- The management of research data during and after the end of the activities.
- What data will be collected, processed and/or generated.
- Which methodologies and standards will be applied.
- Whether data will be shared/open access.
- How the data will be curated and stored (also after the end of the research activity).



THE DATA MANAGEMENT PLAN

Online tools were developed to help drafting a DMP without starting from scratch.



<https://ds-wizard.org/>; <https://dmponline.dcc.ac.uk/>; <https://argos.openaire.eu/home>

A photograph of a modern office interior, heavily overlaid with a dark green semi-transparent filter. The office features wooden desks, ergonomic chairs, and numerous potted plants, including large indoor trees and hanging greenery. The space is bright and airy, with large windows in the background. The text "To summarize" is centered in a large, white, sans-serif font.

To summarize



Research data must be managed

They are an asset for researchers, institutions and societies



FAIR principles provide a framework

To increase the scientific value of your data



Data needs to be preserved after the project end

Through deposit in a trusted data repository



FAIR does not necessarily mean open

But open data are a force for good

**THANK
YOU**

