Research Data Management – AVRE Training

Sébastien Hoyas

Data Project Manager & Business Analyst

Place du Parc, 22
Office 1.12
7000 Mons
+32 65 37 34 54
sebastien.hoyas@umons.ac.be

Image created with Dall-E by OpenAI

Outline

Scope of this training session

What are (research) data?

What is data management?

Data management advices

Data management plan (DMP)

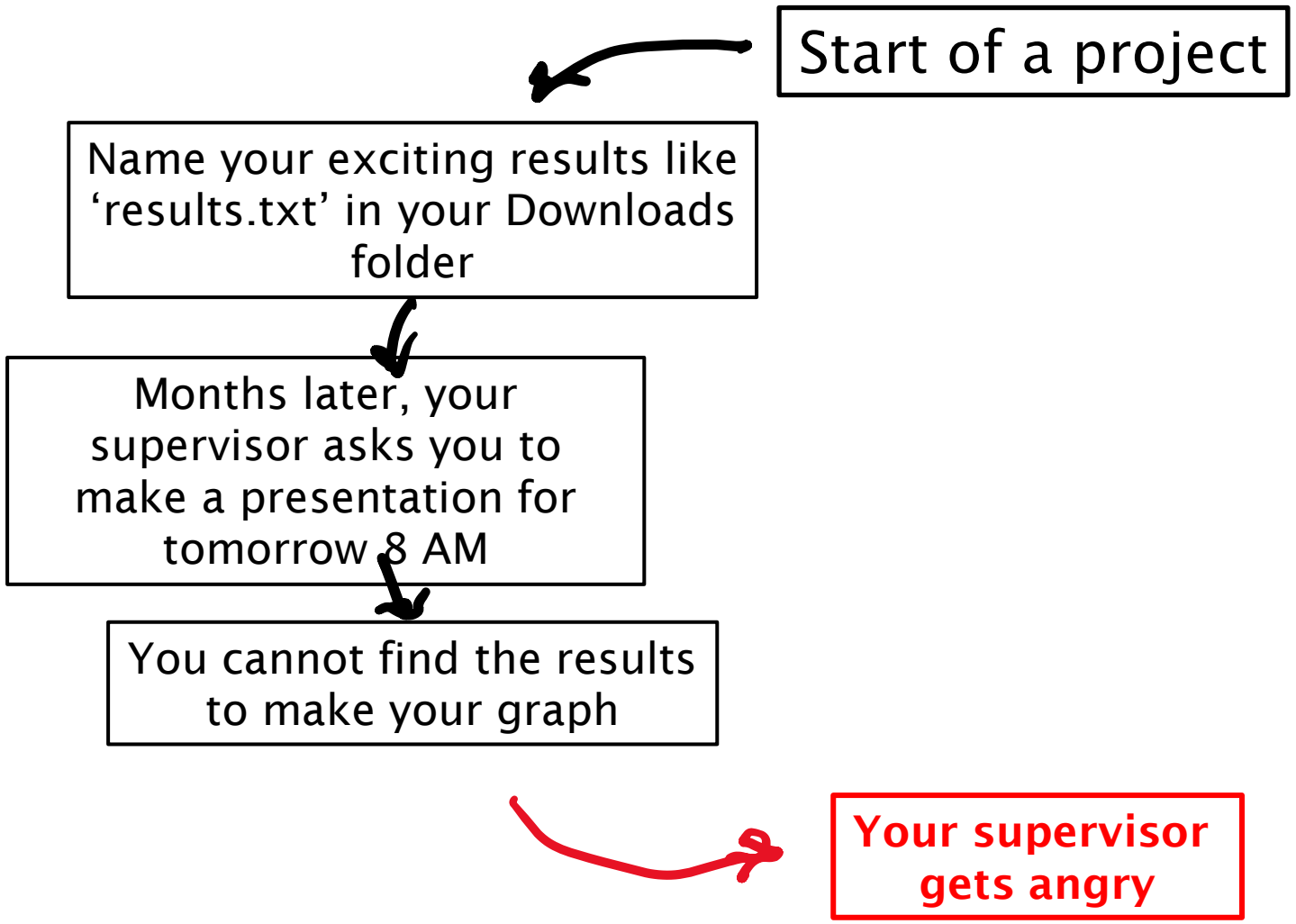
Closing remarks

Q&A

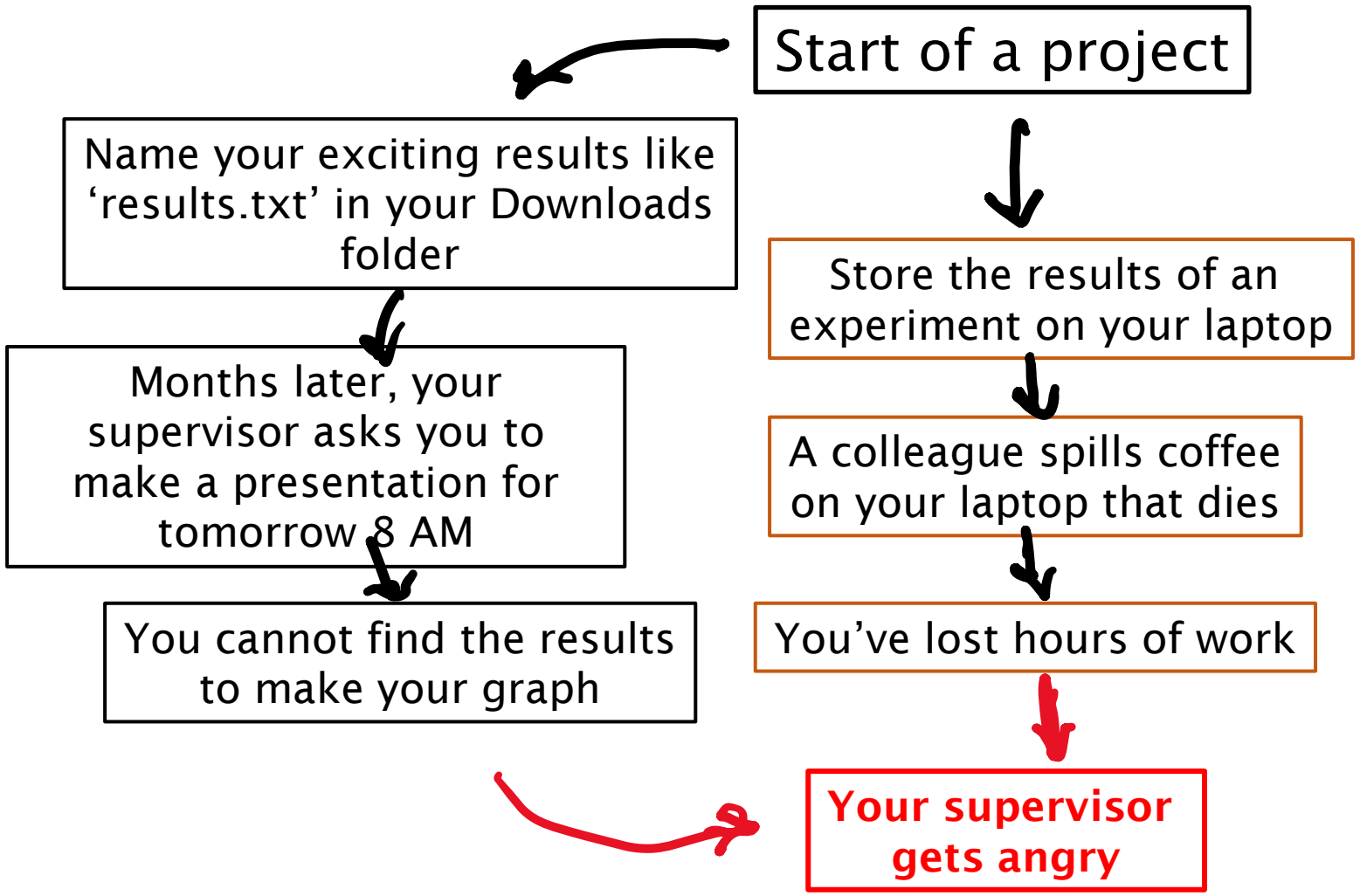
Typical journeys in research… ☺

Start of a project

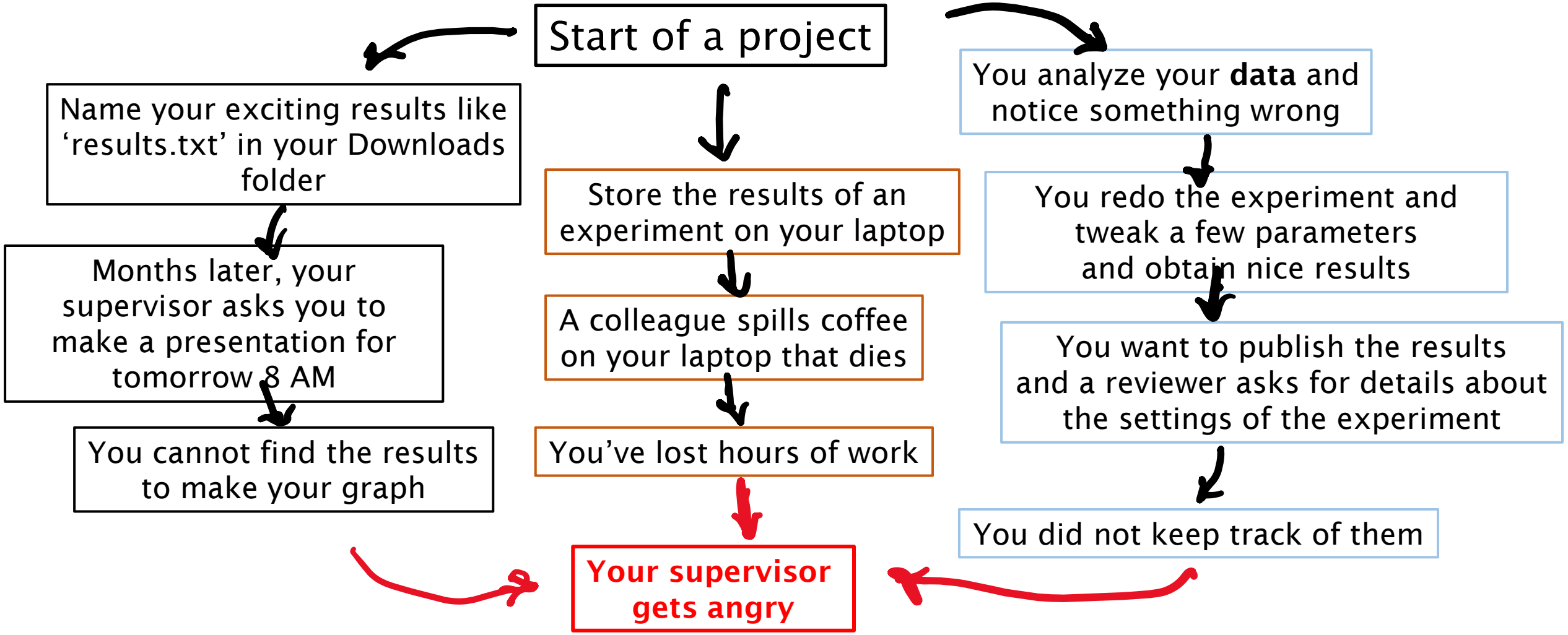
Typical journeys in research… ☺



Typical journeys in research… ☺



Typical journeys in research… ☺



Scope of this training

All these problems are related to ***data*** and their ***management***

Scope of this training:

To try help you manage your data

Make you aware about resources to help you

What I will *not* be able to do:

Provide a unique answer to manage your data

What are (research) data?

Research data: information that is collected, observed, or created for purposes of analysis to produce original research results:

- Survey answers
- Code
- Measurements
- Samples
- …

Quantitative: numerical (number of heartbeats per minute)

vs.

Qualitative: textual, visual (observations of a patient)

Different formats: text, numerical, image, audio, video, physical, etc.

What is data management?

Data management is the **handling** of research data **during *AND* after** a research activity:

- Collection
- Documentation
- Organization
- Storage
- Sharing

Good data management helps to ensure that researchers share data in a **FAIR** way (see later)

Research organizations/funders increasingly require their researchers to ***plan*** how they will manage their data to ensure that all aspects are considered from the start and they can get a return on investment (public goods)

Benefits of data management

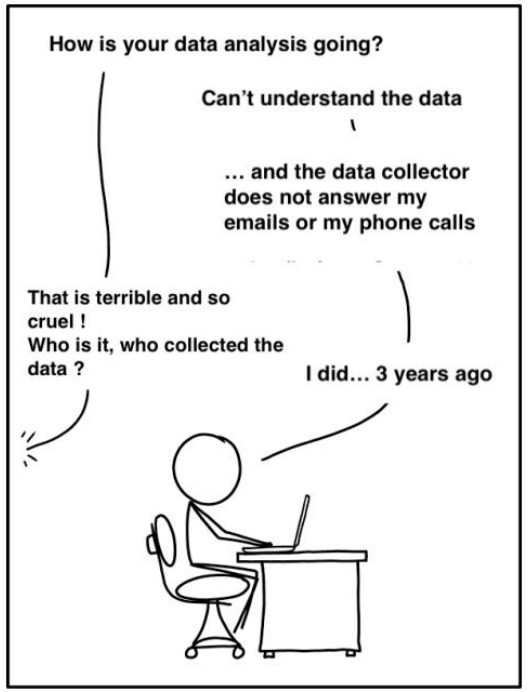
1. **Efficiency**: Proper data management can make your research process more efficient, saving you time and effort in the long run.
2. **Reproducibility**: Crucial for reproducibility of research. It ensures that you and others can understand and replicate your work in the future.
3. **Data Integrity**: It helps maintain the integrity of your data and reduces the risk of data loss.
4. **Collaboration**: It makes it easier to share and collaborate with others, both within and outside your research team by setting proper licences.
5. **Compliance**: Many funding agencies require **DMP** and data sharing (*see later*).
6. **Preservation**: It ensures your data is preserved for future use and reuse.
7. **Return on investment**: Making data discoverable, accessible, and reusable maximizes the research potential of the data and provides greater returns on public investments in research.

Why is it so important?

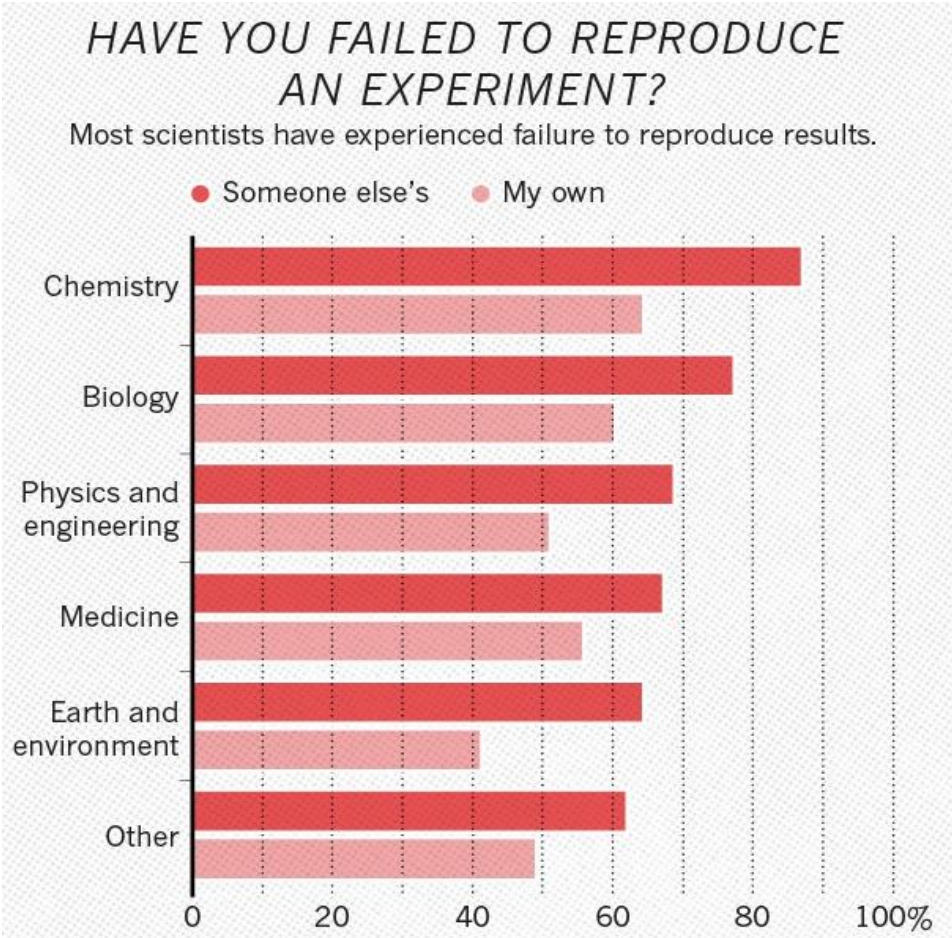
Most scientific results are difficult, even **impossible**, to reproduce and/or replicate

Research **integrity** is not favoured (*publish or perish*)

Avoid reiventing the wheel



your future self, by Julien Colomb, CC-BY-NC, derived from .NORM Normal File Format, CC-BY-NC, by Randall Munroe



Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a

Challenges in data management

Many **different actors** in a research project with **different needs**:

- **Primary Researcher or Principal Investigator**: Creates and uses data
- **Institution**: Sets internal data management policy (including Data Ambassadors, Promoters)
- **Data Repository**: Curates and provides access to data
- **User**: Uses 3rd party data
- **Funder**: Provides the resources to support a research project
- **Publisher**: Disseminates discoveries and maintains the scientific records

Data management advices

Data management is not easy (who said research is easy?) and heavily **depends on *you*, *your project* and *your peers***

To help you, general advices in:

- **Data organization**
- **Data description & documentation**
- **Data storage**
- **Data sharing**

All these « categories » are linked together

Data organization

Data organization

It is better to think in advance about ***how*** you will organize your data, but you can also change your mind during your research as long as you keep track of those changes.

Take into consideration that you may not be the only one working on those data, so **make it clear for *anyone***!

In general, most research data are digitalized, but it also applies to physical data (samples, etc.).

Tips & tricks

Existing procedures: check if there are already established ways to organize your data in your team (and check if it suits your needs!).

File organization: there is no universal answer ☹ You need to apply an organization that is compatible with you, your project and your team.

Above all, you must be ***CONSISTENT*** accross your project.

Data organization

File organization: examples

Hierarchical structure

```
ProjectName/
│
├── LiteratureReview/
│   ├── Paper1.pdf
│   ├── Paper2.pdf
│   └── ...
│
├── Data/
│   ├── RawData/
│   │   ├── Experiment1/
│   │   ├── Experiment2/
│   │   └── ...
│   │
│   ├── ProcessedData/
│   │   ├── Experiment1/
│   │   ├── Experiment2/
│   │   └── ...
│   │
│   └── Metadata/
│
├── Analysis/
│   ├── Script1.R
│   ├── Script2.py
│   └── ...
│
└── Outputs/
    ├── Figures/
    ├── Tables/
    └── ...
```

Data organization

File organization: examples

Hierarchical structure

```
ProjectName/
│
├── LiteratureReview/
│   ├── Paper1.pdf
│   ├── Paper2.pdf
│   └── ...
│
├── Data/
│   ├── RawData/
│   │   ├── Experiment1/
│   │   ├── Experiment2/
│   │   └── ...
│   │
│   ├── ProcessedData/
│   │   ├── Experiment1/
│   │   ├── Experiment2/
│   │   └── ...
│   │
│   └── Metadata/
│
├── Analysis/
│   ├── Script1.R
│   ├── Script2.py
│   └── ...
│
└── Outputs/
    ├── Figures/
    ├── Tables/
    └── ...
```

Chronological structure

```
ProjectName/
│
├── 2022/
│   ├── Q1/
│   │   ├── January/
│   │   ├── February/
│   │   └── March/
│   │
│   ├── Q2/
│   │   ├── April/
│   │   ├── May/
│   │   └── June/
│   │
│   ├── Q3/
│   │   ├── July/
│   │   ├── August/
│   │   └── September/
│   │
│   └── Q4/
│       ├── October/
│       ├── November/
│       └── December/
│
└── 2023/
    └── ...
```

Data organization

File organization: examples

Hierarchical & chronological structure
→ basically, any combination that allows you to organized AND consistent

```
ProjectName/
│
├── LiteratureReview/
│   ├── Paper1.pdf
│   ├── Paper2.pdf
│   └── ...
│
├── Data/
│   ├── RawData/
│   │   ├── 2022/
│   │   │   ├── January/
│   │   │   │   ├── Experiment1/
│   │   │   │   ├── Experiment2/
│   │   │   │   └── ...
│   │   └── 2023/
│   │       ├── January/
│   │       └── ...
│   │
│   ├── ProcessedData/
│   │   ├── Experiment1/
│   │   ├── Experiment2/
│   │   └── ...
│   │
│   └── Metadata/
│
├── Analysis/
│   ├── Script1.R
│   ├── Script2.py
│   └── ...
│
└── Outputs/
    ├── Figures/
    ├── Tables/
    └── ...
```

Data organization

Folder names: keep folder names short (max 15-20 characters) and make them descriptive of what is inside, without being redundant with the folder structure.

- Bad example

```
ProjectName/
│
├── 2022/
│   ├── ProjectName Q1 Results of Experiments/
│   │   ├── Results of January/
│   │   ├── Results of February/
│   │   └── …
```

- Good example

```
ProjectName/
│
├── 2022/
│   ├── Experiments/
│   │   ├── Spectrometry/
│   │   ├── Spectroscopy/
│   │   └── …
```

NOTES

- Avoid using spaces, dots and special characters (&, ?, etc.)
- Use hyphens (-) or underscores (_) to separate elements
- Use a minimum of two leading zeros for padding (001, 002 ,etc.) to properly sort folders by names

Data organization

File names: keep file names brief and explicit, without being redundant with the folder structure.

- Avoid using spaces, dots and special characters (&, ?, ;, etc.)
- Use **hyphens** (-) or **underscores** (_) to separate elements (easier to recover in the OS)
- Use a minimum of two **leading zeros** for padding (001, 002 ,etc.) to properly sort files by names
- Use an extension that matches the file format
- If your files cannot be integrated in a versioning tool like Git, include a **version number** at the end. Keep a logfile where you briefly state changes in each of the new versions
- Include elements such as the date (**YYYYMMDD** format, best to sort) at the beginning of the file name
- Avoid starting to name files with « draft », « final » or the version number

Example: 20230130_RDM_Training_**V001.pptx**

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/File-naming-and-folder-structure
https://library.si.edu/sites/default/files/tutorial/pdf/filenamingorganizing20180227.pdf

Data organization

Versioning: keep older versions in a separate folder, and do not delete them unless you are absolutely sure you can. Keep a logfile that briefly explains the changes in each version.

Key considerations: there are bad practices, but there are ***no unique answer*** or best method to organize data. You must be ***consistent*** throughout your project so that you and your team can work on.

https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/2.-Organise-Document/File-naming-and-folder-structure
https://library.si.edu/sites/default/files/tutorial/pdf/filenamingorganizing20180227.pdf

Data description & documentation

Data description and documentation

Once you know how you will organize your data, you can start collecting them.

You should also start describing them using **metadata** and **documentation**.

- **Metadata** are data that describe data… ☺ (make data findable)
- **Documentation** of your data should include the method(s) you used to obtain them, how they were analyzed, processed, where you can find them, etc.

Those information should be stored in distinct files in the relevant folders.

Data description

Data description and documentation

Standard metadata → depends on your discipline/data type/purpose

Metadata are important to **find** the data, ensure **reproducibility** and **reuse**

Common elements

- **Title**: The name given to the dataset.
- **Author**: The main researchers involved in producing the data.
- **Date of creation**
- **Identifier**: A unique code assigned to the dataset. May be added later, when *sharing* the dataset.

Standards commonly used:

- Dublin Core: 15 properties for describing a wide range of resources (general purpose)
  Check this website to generate your metadata file !
- Digital Curation Center: social sciences
- Biology, earth sciences, physical sciences

Data documentation

Data description and documentation

Data documentation: document that explains your data. It will help others that would use your data, but also yourself to remember how you obtained and processed those data (*try to do it on the fly*). It also helps you and others to reproduce your results.

- **Folder organization:** you should explain how data are organized so that anyone starting to collaborate with you can understand what is going on.
- **Data collection:** explain how you obtained the data (from known datasets or from an experiment, survey, simulations, etc.).
- **Data cleaning**: explain your investigation and why you removed part of them (errors, inaccuracies, etc.) .
- **Data analysis:** how you analyzed the cleaned data (which software, parameters, results of the analysis, etc.).
- **Plan for Change**: keep in mind that your data may change over the course of your project. Plan for how you will document and manage these changes.
- **Consider Your Audience**: remember that the description of your data may be read by people who are not experts in your field. Try to write in a way that is accessible to non-experts.

Data description and documentation

Key considerations: **Metadata** and **data documentation** files should ***always accompany your data*** (same folder or dedicated subfolder and the dataset). The more **descriptive** the better for future reuse and reproducibility.

Data storage & sharing

Data storage and sharing

Data Storage

- **Reliability**: Use reliable storage solutions that ensure data integrity and availability **NOT YOUR LOCAL LAPTOP/DESKTOP!**
- **Accessibility**: Ensure that data is easily accessible to authorized users.
- **Scalability**: Choose storage solutions that can grow with the size of your data.

NOTE: Backup is different than preservation

- **Backup** = periodic snapshots in case current version is lost or destroyed (***cloud, NAS, etc.***)
- **Preservation** = archival, usually the final version of a dataset, stored for long-term and further use (***data repositories***)

Data storage and sharing

Backup:

- **Frequency**: Regularly back up data according to the importance and frequency of change.
- **3-2-1 Rule**: Keep at least three copies of your data, on two different media, with one backup offsite[1].
- **Verification**: Regularly verify the integrity of backup copies.
- **Disaster Recovery**: Have a disaster recovery plan in place to restore data if needed.

Nice **free open source** software for Linux, Windows and MacOS: FreeFileSync: Open Source File Synchronization & Backup Software

https://library.si.edu/research/best-practices-storing-archiving-and-preserving-data

Data storage and sharing

Backup: *where?*

- **Cloud**: UMONS provides 1To/user or team (see sharepoint) for free
- **Cloud**: if you need extra space or you want to backup somewhere else, you must pay for another service (not offered by UMONS)
- **Local server**: you can setup a Network Attached Server (NAS) for you and your team with backup services (not offered by UMONS)

https://library.si.edu/research/best-practices-storing-archiving-and-preserving-data

Data storage and sharing

Long-term preservation: ***data repositories***

A trusted digital repository provides reliable long-term access to managed digital resources to its designated community, now and in the future!

Only completed datasets with the purpose to ***publish***, ***share*** and/or ***preserve*** them should be uploaded (**not all research data**).

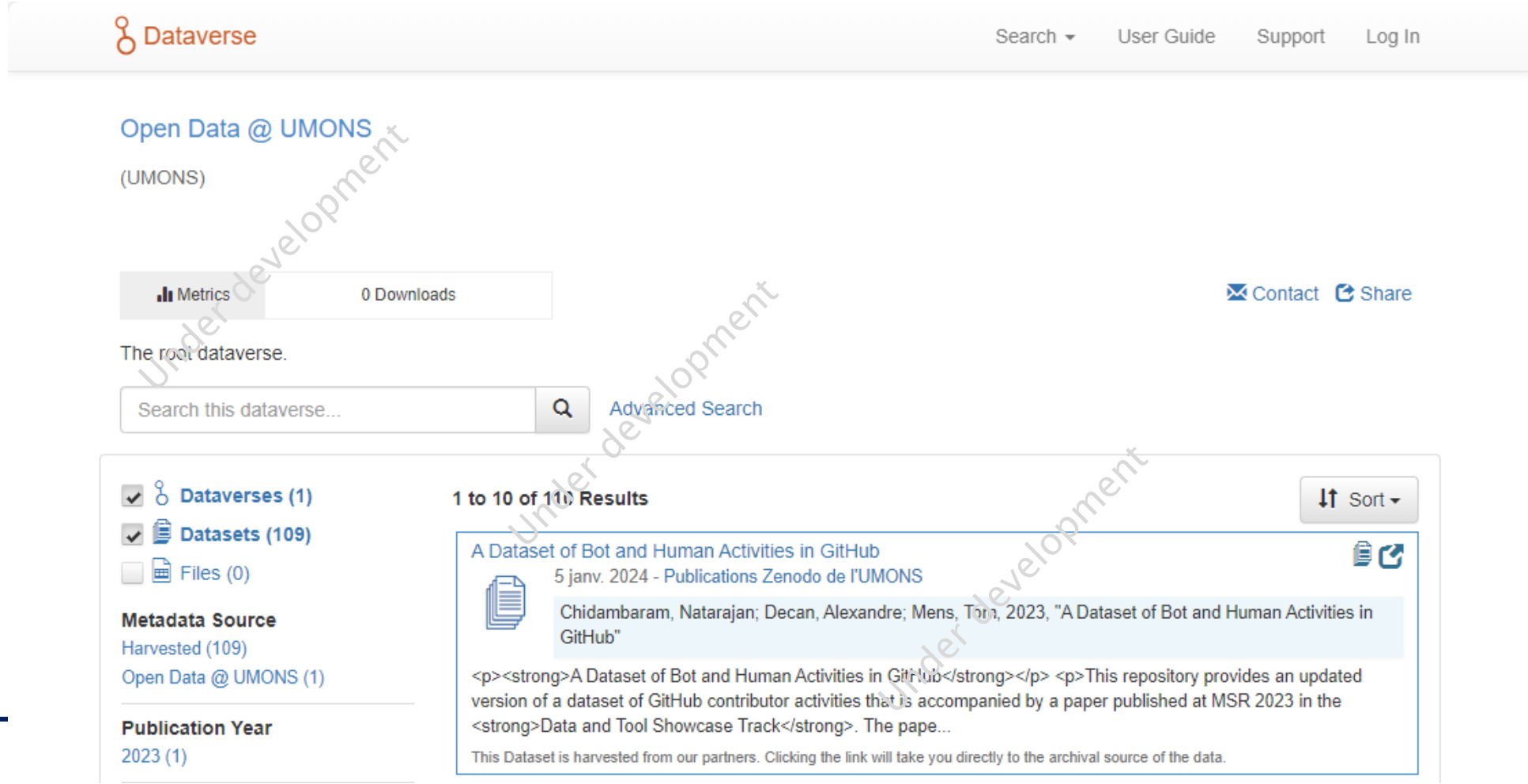
Typical files included in a dataset repository: inputs, outputs, method and metadata, not all intermediary data! With only those information, anyone should be able to obtain the same outputs as you.

Examples:

- Zenodo: general purpose repository
- SODHA: the federal Belgian data archive for social sciences and the digital humanities

Data storage and sharing

Dataverse: our institutional repository under development to host your clean datasets



Data storage and sharing

Data repositories also allow you to ***share*** your data.

Your data should meet the **FAIR data principles**:

Findable → **Metadata**

- Data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services.

Accessible → **Data repository and unique identifier (DOI**

- Once the user finds the required data, they need to know how they can obtain them. The data might be publicly accessible, or access may be restricted but metadata should remain accessible.

Interoperable → **Metadata, documentation**

- The data usually need to be integrated with other data regardless of the systems or tools being used → non-proprietary data formats (.txt, .csv, .md, .pdf, etc.)

Reusable → **Documentation**

- The ultimate goal of FAIR is to optimize the reuse of data. To achieve this, metadata and data should be well-described so that they can be replicated and/or combined in different settings.

Data licensing

Data sharing and licenses

Key take-away: *To effectively share data, resolve any data ownership or intellectual property rights issues early. Consult your institution (AVRE) to determine what policies might affect data ownership and sharing.*

- **Understand Your Goals**: What do you want others to be able to do with your data? This can help guide your choice of license.
- **Public Domain**: If you want to give people the most freedom, consider a public domain license like CC0.
- **Attribution**: If you want to allow free use but also want to be credited, consider a license that requires attribution, like CC-BY.
- **Share-Alike**: If you want any derivatives of your work to be licensed under the same terms, consider a share-alike license, like CC-BY-SA.
- **Non-Commercial**: If you want to restrict the commercial use of your data, consider a non-commercial license, like CC-BY-NC.
- **Understand the Implications**: Make sure you understand the implications of the license you choose. Some licenses may have implications for how your data can be used or shared.
- **Seek Legal Advice**: If you're unsure, consider seeking legal advice. Licensing can be complex, and it's important to get it right.

Data sharing and licenses

Common licenses for datasets

- CC0 Public Domain Dedication
- Open Database License
- Open Licence Etalab
- Open Data Commons Public Domain Dedication and License
- Open Data Commons Attribution License

In more details: SPDX License List | Software Package Data Exchange (SPDX)

Budget for storage

Budget for data storage

Data production can be costly (equipment, products, time).

In addition to the costs of data collection, data management, curation, documentation, **storage can be expensive**.

→ Need to think about all costs (IT, server, etc.).

At the moment, **1 TB/person for free** with Microsoft OneDrive, but may not be enough or practical for your project.

These costs may be eligible in some calls for projects.

Examples:

- Cloud storage on Microsoft Azure: 250 GB/month → min. 50€ → 4 years = 2400€ only for your research
- Network attached storage: fixed cost NAS 500€ + UPS 200€ + 4TB disks 120€ * 3 = 1060€ → multiple users, research, etc. (*no support from the university yet…*)

* storage, backup, second backup → 3-2-1 rule

Data Management

Data management

Common element in previous advices about data management: ***planning***

To help you managing your data, it is encourage (and even more often than before required) to create a ***Data Management Plan*** (**DMP**).

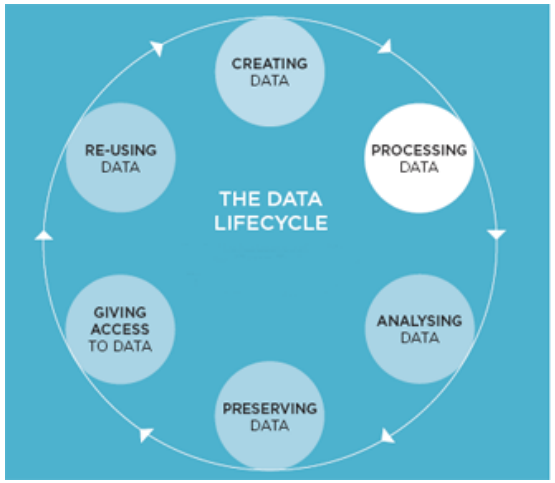
DMP: Document that describes the data produced/used in all stages of a research project and outlines the strategy to manage your data, before and during the project.

This document is ***not static*** and can ***evolve*** during your research (research can be unpredictive 😉).

Data lifecycle in research

To create a DMP, you need to understand the various stages data goes through during a research project == "***data lifecycle***".

Many different models:



Common elements:

- Planning
- Collecting
- Analyzing
- Organizing
- Preserving data for future use and share

[1] RDM explained — Ghent University (ugent.be)
[2] Data Lifecycle - Data Module #2: Planning for Your Data Use - All Guides at Macalester College
[3] About – European DDI User Conference (eddi-conferences.eu)

Data management plan

More and more funding agencies require to submit a **DMP** to ensure that data will be accessible and usable on the long run → *return on investment*

Better to write a DMP **BEFORE** starting a project, but also beneficial to do it **NOW**.

Two helpful and free DMP tools:

- DMPTool, created by the University of California Curation Center of the California Digital Library
- DMPOnline, developed by the UK Digital Curation Centre → **Mandatory for European Projects!**

Both of these tools provide guidances and templates for creating DMPs in compliance with institutional and funder requirements.

Data management plan

Typical questions in a DMP:

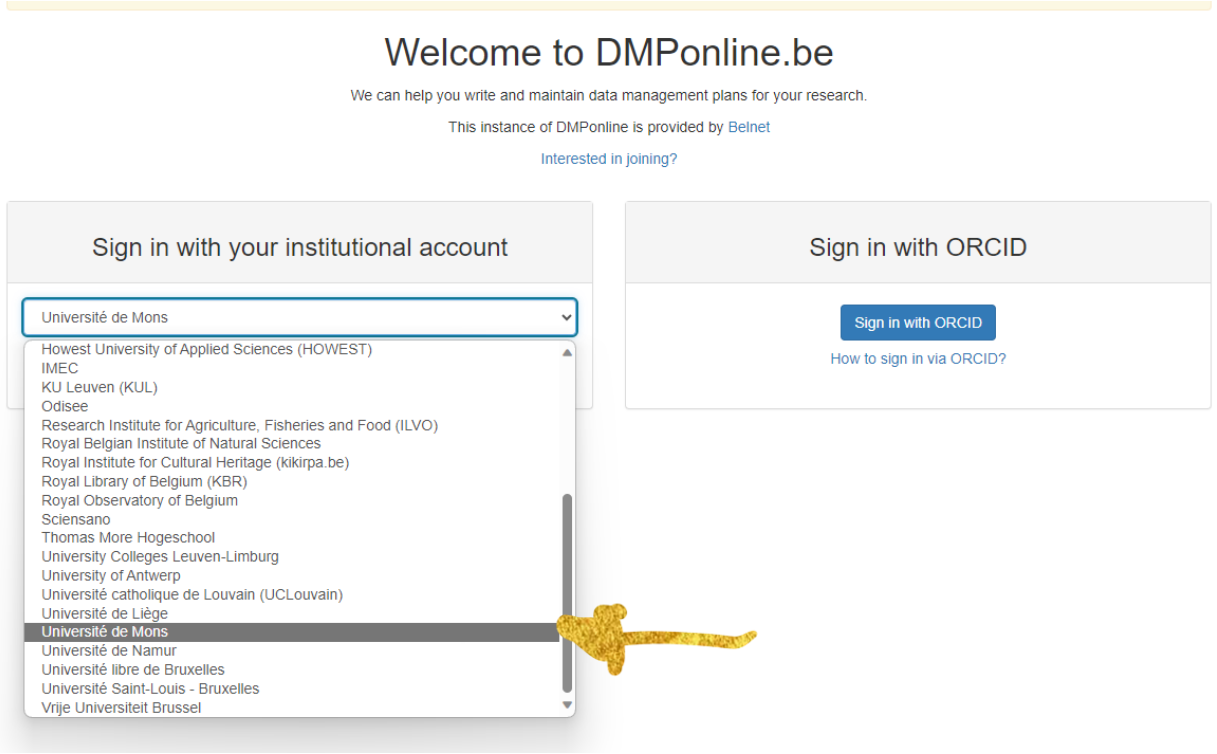
- What data will you collect or create?
- How will the data be collected or created?
- What documentation and metadata will accompany the data?
- How will you manage any ethical issues?
- How will you manage copyright and intellectual property rights issues?
- How will the data be stored and backed up during research?
- How will you manage access and security?
- Which data should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?
- How will you share the data?
- Are any restrictions on data sharing required?
- Who will be responsible for data management?
- What resources will you require to implement your plan?

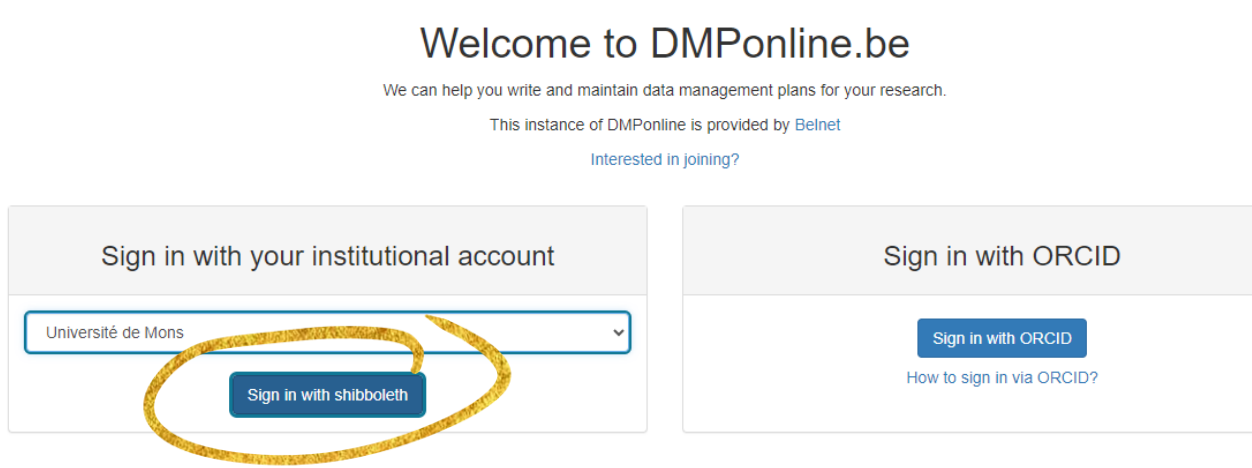
DMP Online

Live demo: Welcome to DMPonline.be

DMP Online

Log in using your institutional credentials





DMP Online

Welcome to DMPonline.be

We can help you write and maintain data management plans for your research.

This instance of DMPonline is provided by the DMPbelgium Consortium, which was founded in 2017 by:

- Instituut voor Natuur- en Bosonderzoek
- Université Libre de Bruxelles
- Universiteit Antwerpen
- Universiteit Gent
- Universiteit Hasselt
- Vrije Universiteit Brussel
- Wetenschappelijk Instituut Volksgezondheid – Institut Scientifique de Santé Publique (Sciensano)

In 2018 they were joined by:

- Université Catholique de Louvain
- Université de Liège
- Université de Mons
- Université de Namur
- Vlaamse Instelling voor Technologisch Onderzoek

Since then, the Consortium has been joined by:

- Arteveldehogeschool
- Instituut voor Landbouw-, Visserij- en Voedingsonderzoek
- Universitair Ziekenhuis Gent
- Vlaams Instituut voor de Zee
- Vlerick Business School
- Hogeschool Gent

Interested in joining the Consortium?

Sign in with your institutional account

Flanders Make (Belgium) (flandersmake.be)

Flanders Marine Institute

Flemish Institute for Technological Research

Ghent University

Ghent University (UZGent)

Hasselt University

Hogeschool Gent (HOGENT)

Hogeschool VIVES (VIVES)

IMEC

KU Leuven (KUL)

Research Institute for Agriculture, Fisheries and Food (ILVO)

Royal Institute for Cultural Heritage (kikirpa.be)

Royal Library of Belgium (KBR)

Royal Observatory of Belgium

Sciensano

Thomas More Hogeschool

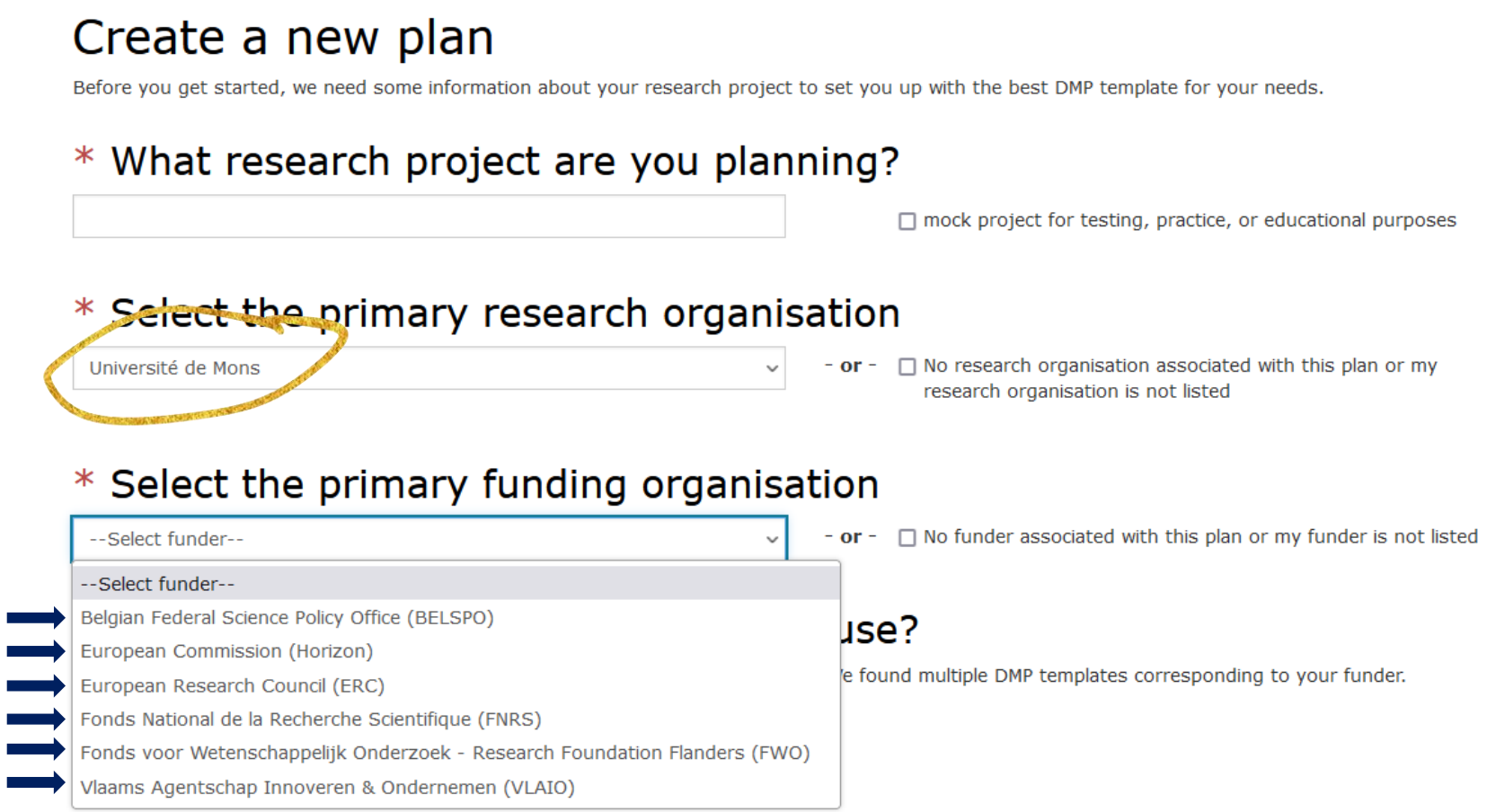
Université catholique de Louvain (UCLouvain)

Université de Liège

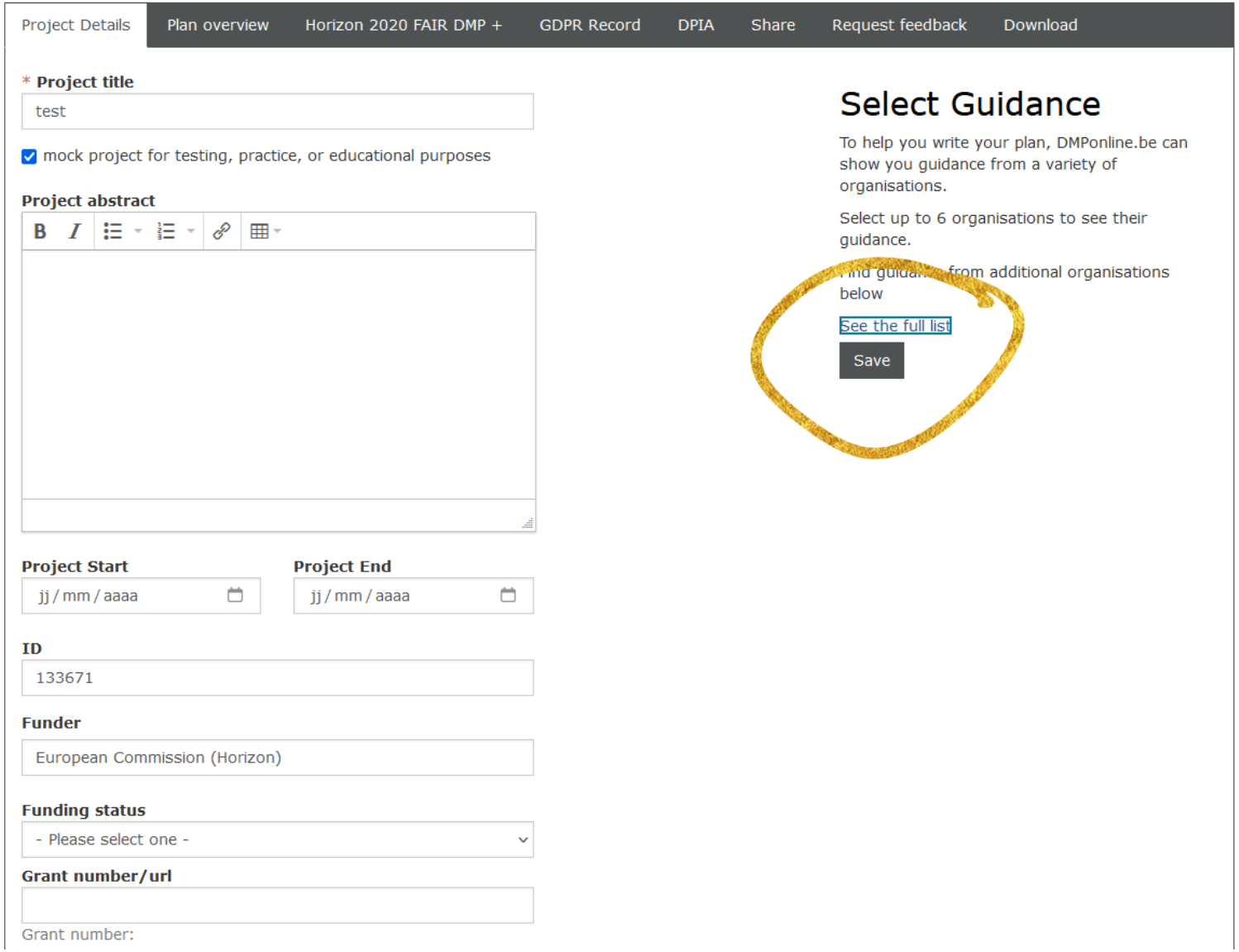
Université de Mons

Université de Namur

DMP Online


Create a new plan
Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.
* What research project are you planning?
mock project for testing, practice, or educational purposes
* Select the primary research organisation
Université de Mons
- or -
No research organisation associated with this plan or my research organisation is not listed
* Select the primary funding organisation
--Select funder--
- or -
No funder associated with this plan or my funder is not listed
--Select funder--
Belgian Federal Science Policy Office (BELSPO)
European Commission (Horizon)
European Research Council (ERC)
Fonds National de la Recherche Scientifique (FNRS)
Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)
Vlaams Agentschap Innoveren & Ondernemen (VLAIO)
use?
e found multiple DMP templates corresponding to your funder.

DMP Online


Project Details
Plan overview
Horizon 2020 FAIR DMP +
GDPR Record
DPIA
Share
Request feedback
Download
* Project title
test
mock project for testing, practice, or educational purposes
Project abstract
Project Start
jj / mm / aaaa
Project End
jj / mm / aaaa
ID
133671
Funder
European Commission (Horizon)
Funding status
- Please select one -
Grant number/url
Grant number:
Select Guidance
To help you write your plan, DMPonline.be can show you guidance from a variety of organisations.
Select up to 6 organisations to see their guidance.
from additional organisations below
See the full list
Save

DMP Online

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)
BELSPO 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)	24-10-2022		

Resources from the Data Ambassadors Network

You may need help processing those information or have questions about RDM:

The Data Ambassadors Network is there for you!

UMONS Data Ambassadors:

VISEUR	Robert	Business and Economics	Robert.VISEUR@umons.ac.be
GALLAS	Mohamed-Anis	Architecture	Mohamed-Anis.GALLAS@umons.ac.be
COPPEE	Frédérique	Medicine	Frederique.COPPEE@umons.ac.be
GROSJEAN	Philippe	Sciences	Philippe.GROSJEAN@umons.ac.be
DUPONT	Nicolas	Applied Sciences	Nicolas.DUPONT@umons.ac.be
MEYERS	Charlène	Languages	Charlene.MEYERS@umons.ac.be
RIVIERE LORPHEVRE	Edouard	Applied Sciences	Edouard.RIVIERELORPHEVRE@umons.ac.be
SIMOES LOUREIRO	Isabelle	Psychology	Isabelle.SIMOESLOUREIRO@umons.ac.be

+ webinars (data anonymization, how to archive data, etc.)

Closing remarks

- Research data management is not an easy task but is not impossible either.
- There is no unique way to properly manage your data, as long as your are ***consistent***, ***descriptive*** and organize in a way that ***anyone can understand***.
- Different tools are available (**DMPTool**, **DMP online**) to help you to plan how to manage your data.
- **Data Ambassadors** and **AVRE** people are also here to help you.

Q&A

If I was not able to answer your question, please fill in this Microsoft Forms for Q&A and I will contact you later

UMONS Université de Mons

UMONS Université de Mons — DIRECTION AVRE Administration et Valorisation de la Recherche

Thank you for your attention

Credits to Céline Thillou, Judith Biernaux, David Lhoir & Edouard Rivière-Lorphevre

UMONS Université de Mons · LIÈGE université · UMONS Université de Mons · UMONS Université de Mons

Resources for data management

Share personal data through a repository

Facilitating FAIR practices in Research Methods, Data, And Software in Natural and Engineering Sciences

Complete training about RDM by Macalester College Library

DocFetcher - Fast Document Search (sourceforge.io) to index your files and quickly find their content

Understanding Research Data Management – University of Pittsburgh

https://libereurope.eu/event/data-management-plans-use-and-reuse-webinar/ ; coming webinar

Challenges in RDM

Completed DMP for « PURE » European Union's Horizon 2020 project Pure Project Data Management Plan (zenodo.org)

Other DMPs for different disciplines: Example DMPs and guidance | DCC

Do's and Don'ts of DMP

Resources to find data for your research

EOSC (Europen Open Science Cloud): https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud

Mendeley Data website: https://data.mendeley.com/datasets

OpenAire : https://explore.openaire.eu/search/find

re3data.org: https://www.re3data.org/: Harvesting several data repositories

Google: https://toolbox.google.com/datasetsearch

FigShare: figshare - credit for all your research

Zenodo: Zenodo - Research. Shared.