



DataHub

Onboarding Training for Data Professionals

Maria Vivas-Romero

Data-Hub

Generic Data Steward/Training Capacity Building

m.vivasromero@maastrichtuniversity.nl

Visit us at
<https://datahubmaastricht.nl>



About the workshop

- This **onboarding workshop** touches upon key aspects of data management in the health and life sciences at Maastricht University
- Participants are **Data Professionals** with different professional backgrounds and levels of expertise in different tools regarding RDM. This means some topics might be too specific or general for some.
- This workshop allows to **inhale but also exhale information** in an interactive fashion. You can apply all concepts learned here in your daily data management practices
- This is a pilot version so all of your **comments and feedbacks** are welcome
- Finally, we hope you feel **your time was well spent** ☐ and the information taken home is of benefit.

What do you expect to get out of this training?

Menti

code **1270 5395**



Click to edit Master text styles

What's your data super power?

- Menti
- The code **1270 5395**



Schedule

Schedule	Topic	Trainers	Material Type
9.00 9.05	Introduction to the workshop	Maria Vivas-Romero	Presentation
9.05 9.15	Ice-Breaker: What's your data super power	Maria Vivas-Romero	Group Activity
9.15- 10.15	Why FAIR and Open Science at MUMC+? Road Map of Tools and Services	Dennie Hebels (Open Science Officer-UB) (?) Maria Vivas-Romero Anouk Cuijpers	Presentation
10.15 – 10.20	Coffee Break	Participants	Group Activity
10.20 11.00	Ethical Requirements and GDPR	Helena Bossini-Castillo Alfons S	Computers and tools services
11.00 12.00	Hands on exercise: Part I MDR or DMP	Maria Vivas-Romero Olav Palmen (Datahub)	
12.00 13.00	Lunch		
13.00 – 13.30	Hands on exercise Part II GDPR form	Maria Vivas-Romero	Computers and tools and services
13.30 – 14.00	Plenary discussion : Take home message	Maria Vivas-Romero	

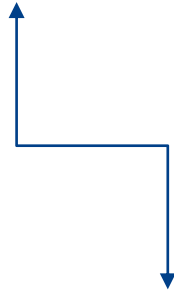
What do we mean we data?

From Information Theory:

- ***Data are those bits of observable information that are recorded***

Two Pillars of Research Data

Data Management Procedures



Information organized and ready to be FAIR



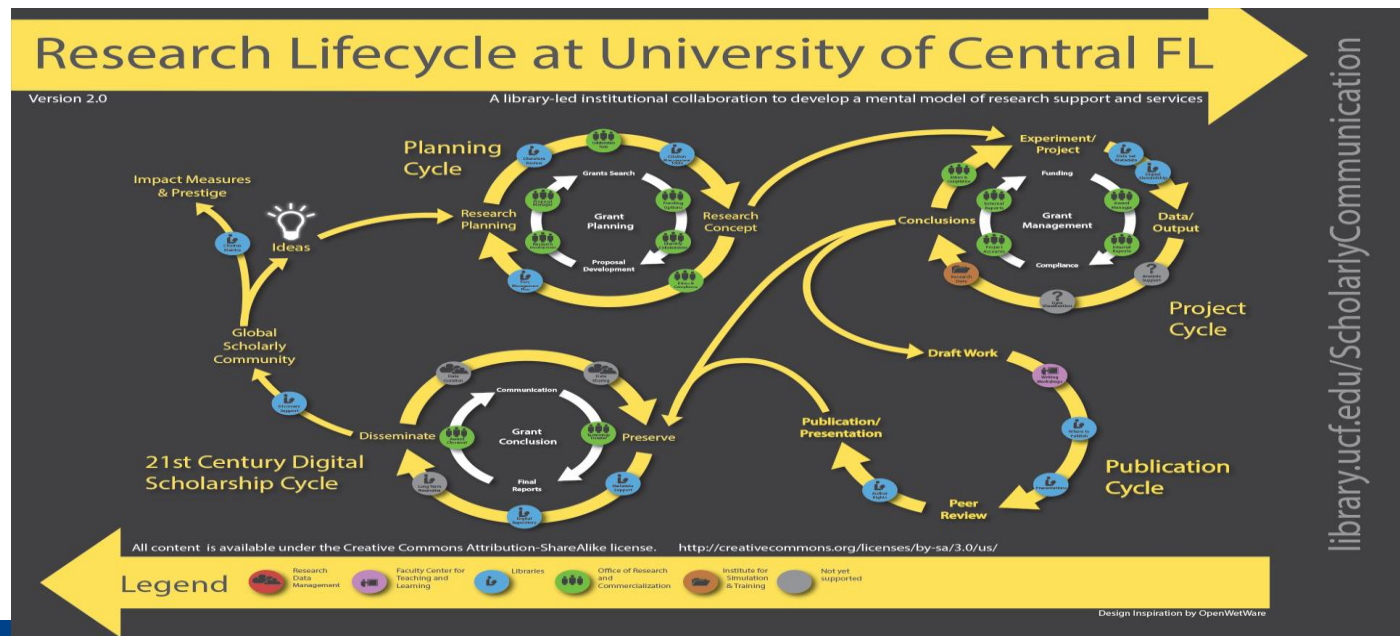
Research Data Management

Research Data Management & and Data Life-cycle

Sustainability of Science and Data: Good Data Management assures that data just like any other resources lives on the future.

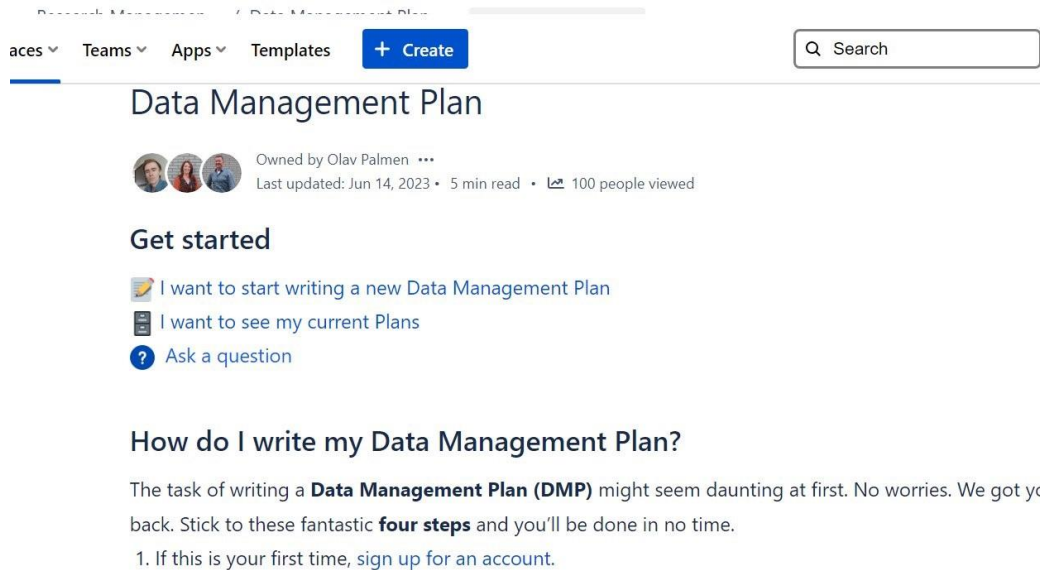
FAIR principles are one way to assure this with research data, so that reproducibility is a possibility

FAIR is not the only way to go, but it's the recipe for RDM we work with at UM



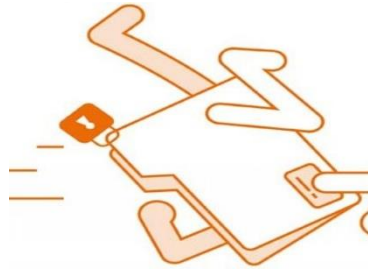
Road Map Data Management Through the Research Life Cycle : Tools and Tricks

1- Plan for their data collection: DMP (Data Management Plan)



The screenshot shows the 'Data Management Plan' page on the Research Management website. At the top, there are navigation links for 'aces', 'Teams', 'Apps', 'Templates', and a '+ Create' button. A search bar is also present. The main heading is 'Data Management Plan'. Below it, it says 'Owned by Olav Palmen' and 'Last updated: Jun 14, 2023 • 5 min read • 100 people viewed'. Under the heading 'Get started', there are three options: 'I want to start writing a new Data Management Plan', 'I want to see my current Plans', and 'Ask a question'. Below this, the section 'How do I write my Data Management Plan?' is shown, followed by an introductory paragraph and a list of steps starting with '1. If this is your first time, sign up for an account.'

2 - How will they collect their data and where will they store it : Local Servers, Surf, MDR,



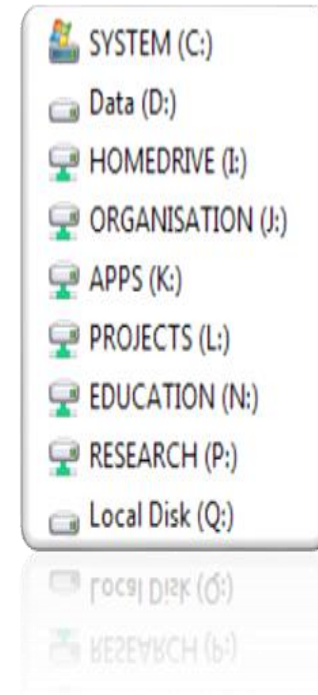
Storage Facilitated by UM

- Faculty Systems
- Research network drive (P:) (provided by ICTS)



- Sharepoint MS

- Current Research Information System
Research publications portal (Pure)



Data Science Research Infrastructure

External Storage and data sharing

SURFdrive

SURFfilesender

SURF research cloud



Welcome to SURF research cloud

 LOG IN







[SURF.nl](https://surf.nl)

Collaborative organisation for ICT
in Dutch education and research



Motivation for RDM

GDPR enforcement tracker

	ETid-569	 HUNGARY	2020-12-10	22,200	Budapesti Műszaki és Gazdaságtudományi Egyetem (Budapest University of Technology and Economics)	Art. 5 (1) a), b), c) GDPR, Art. 6 (1) GDPR, Art. 9 (2) GDPR, Art. 12 GDPR, Art. 13 GDPR	Insufficient legal basis for data processing	link
	ETid-552	 IRELAND	2020-12-17	70,000	University College Dublin	Art. 5 (1) e), f) GDPR, Art. 32 (1) GDPR, Art. 33 (1) GDPR	Insufficient technical and organisational measures to ensure information security	link
	ETid-527	 POLAND	2021-01-05	5,500	Śląski Uniwersytet Medyczny (Medical University of Silesia)	Art. 33 (1) GDPR, Art. 34 (1) GDPR	Insufficient fulfilment of data breach notification obligations	link
	ETid-482	 SWEDEN	2020-12-11	54,000	Umeå University	Art. 5 (1) f) GDPR, Art. 32 (1), (2) GDPR	Insufficient technical and organisational measures to ensure information security	link
	ETid-472	 SWEDEN	2020-12-03	390,100	Karolinska University Hospital of Solna	Art. 5 (1) f) GDPR, Art. 5 (2) GDPR, Art. 32 (1) GDPR, Art. 32 (2) GDPR	Insufficient technical and organisational measures to ensure information security	link
	ETid-471	 SWEDEN	2020-12-03	341,300	Sahlgrenska University Hospital	Art. 5 (1) f) GDPR, Art. 5 (2) GDPR, Art. 32 (1) GDPR, Art. 32 (2) GDPR	Insufficient technical and organisational measures to ensure information security	link
	ETid-449	 IRELAND	2020-08-18	65,000	Cork University Maternity Hospital	Art. 5 GDPR, Art. 32 GDPR	Insufficient technical and organisational measures to ensure information security	link

Why FAIR and Open Science at MUMC+?

Dennie Hebels

16 February 2024



Background

Intro

MERLN Institute for Regenerative Medicine (0.6 FTE)

- Project manager
- Embedded data steward



Open Science Community Maastricht

- Ambassador for Faculty of Health Medicine and Life sciences



Maastricht University Library (0.4 FTE)

- Open Science Officer
- Implementation of the UM Open Science policy



What is Open Science?

What is Open Science?

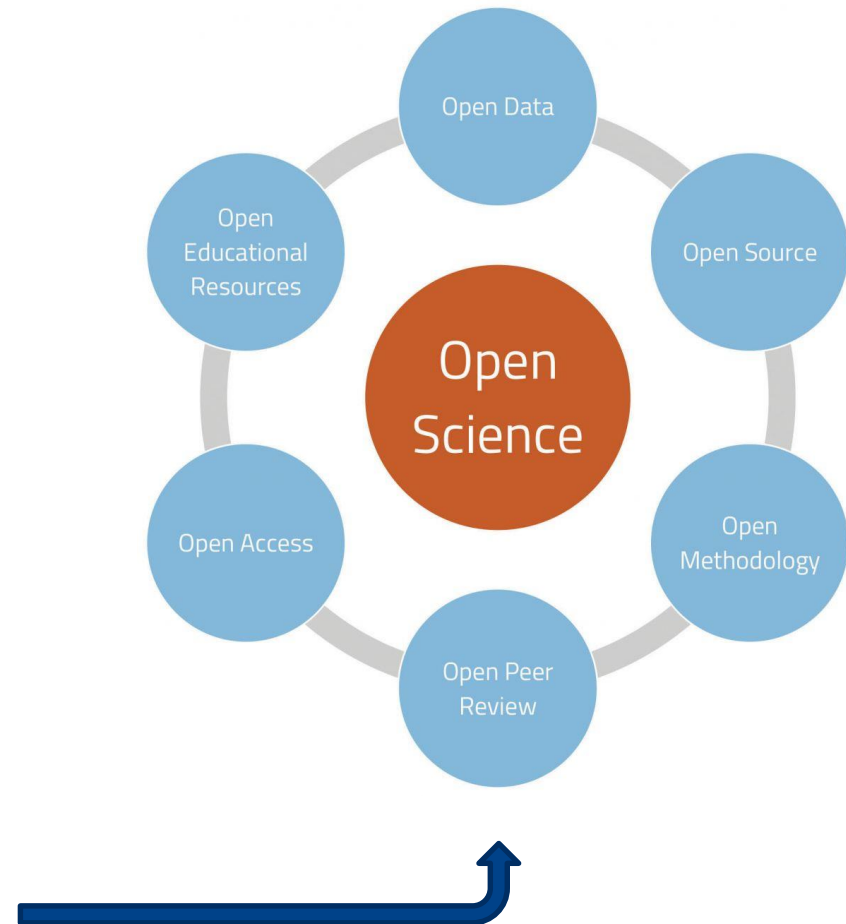
wooclap

Open Science

Definition in two parts:

- Open Science aims to make the primary outputs of publicly funded research results – publications and the research data – publicly accessible in digital format with no or minimal restriction to all levels of society (both amateur or professional) to maximize its **(societal) impact**
- Open Science is **transparent** and **accessible** knowledge that is shared and developed through collaborative networks

Focus is often on Open Research Data and Open Access to publications, but it extends to many other resources

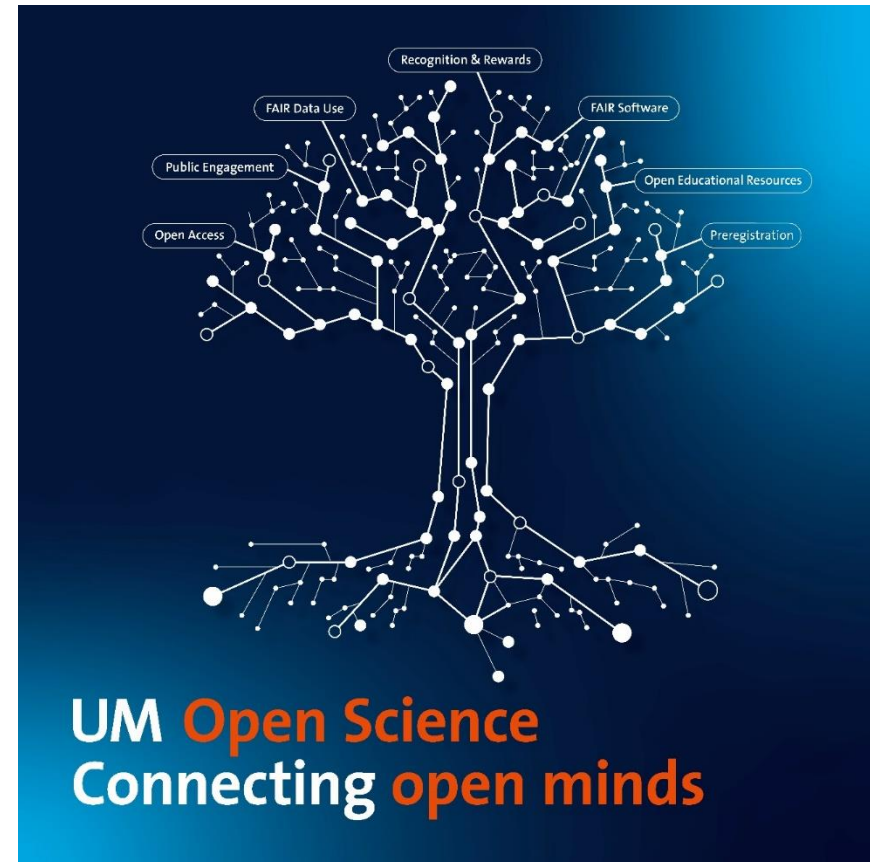


UM Open Science policy 2022-2026

Seven themes play a central role in UM's policy:

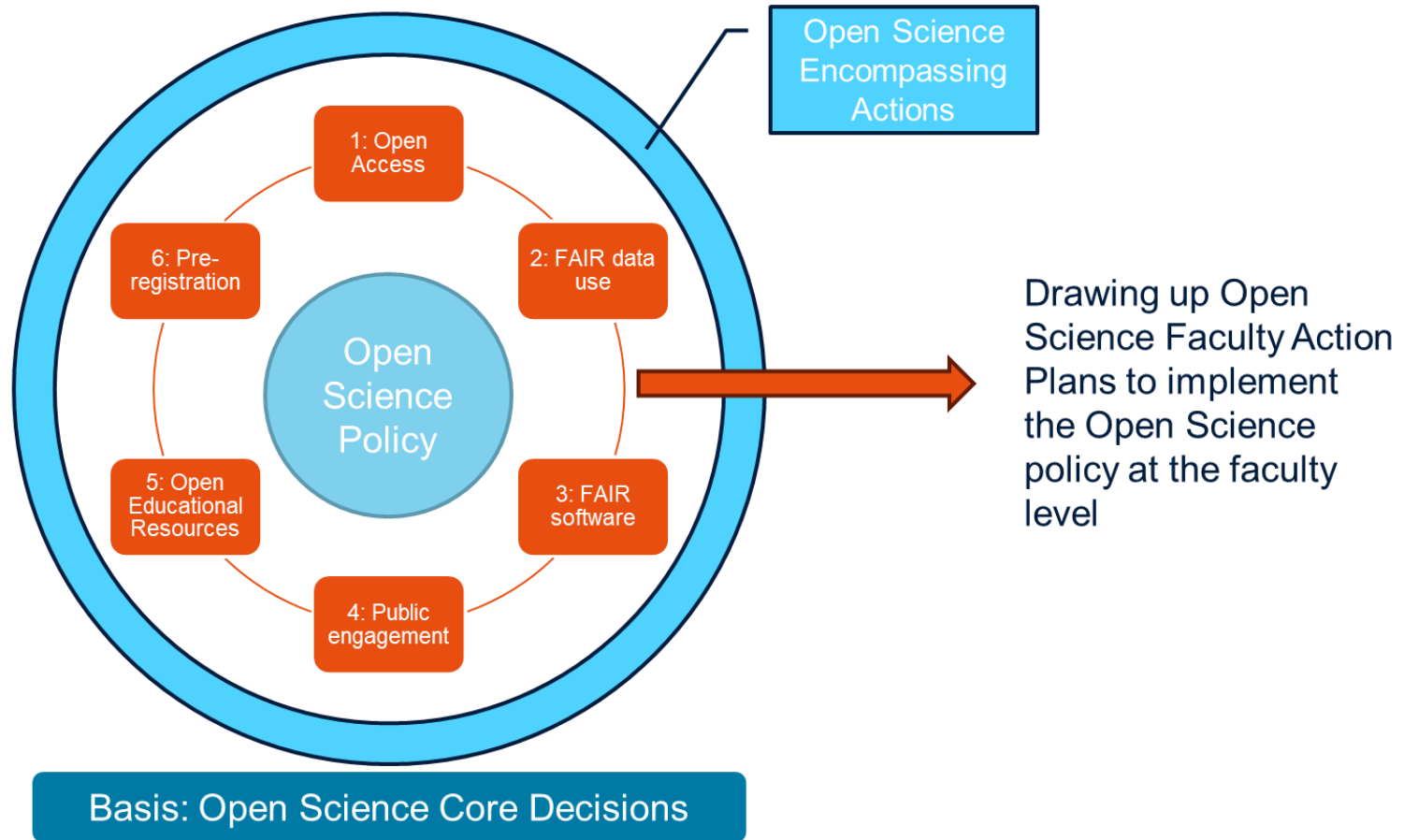
- Recognition & Rewards (separate UM strategy)
- Open Access
- FAIR data use
- FAIR software
- Public engagement
- Open Educational Resources (OER)
- Preregistration

<https://www.maastrichtuniversity.nl/file/um-open-science-policy-update-2022-v13pdf>



UM Open Science policy 2022-2026

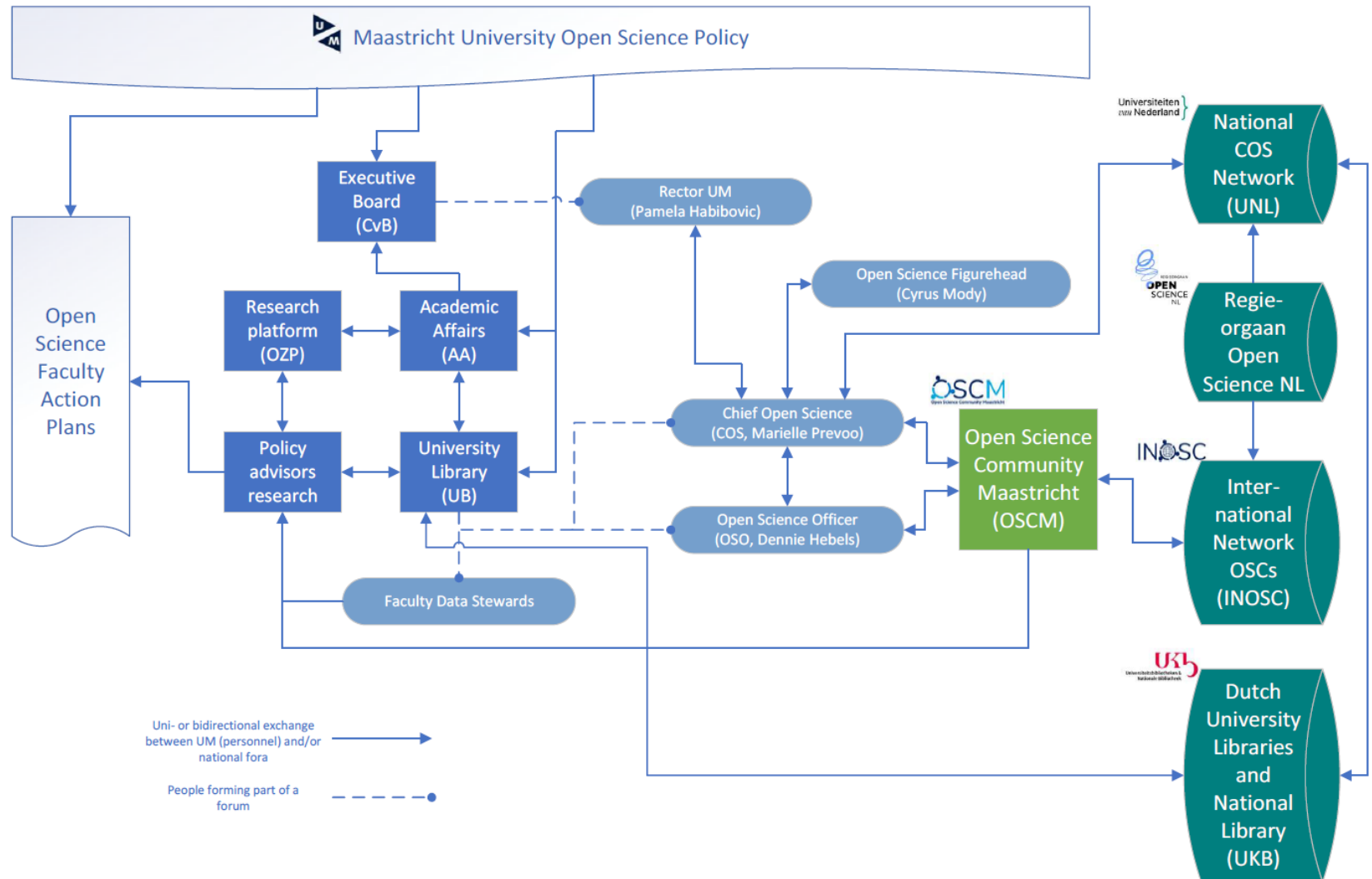
Written by Academic
Affairs and UM
Library, approved by
faculties



The FHML Open Science Faculty Action Plan

- **FHML started with a survey to get an overview of the current state of affairs on the six Open Science themes**
 - Some topics are quite widespread while others are not familiar territory
- **Early 2023, the first Action Plan was presented to the Executive Board, detailing where we are now and where we want to go**
 - Living document → will be updated yearly
- **2024: inclusion of MUMC in Open Science team to jointly implement the Open Science policy**
 - Take into account the close collaborations between FHML and MUMC, shared infrastructure, shared research and education, etc.

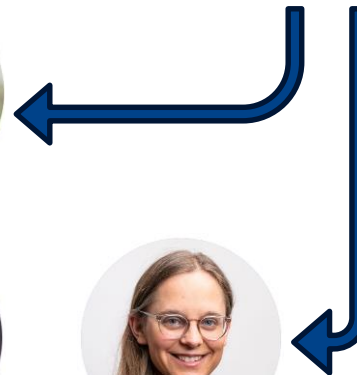
Open Science governance at UM



Open Science Community Maastricht



Founded in 2020
Representation across all
faculties (ambassadors)



FAIR Coffee
Lectures



FAIR Essentials
Workshop



ReproducibiliTea
Journal Club



Join OSCM to stay up to
date on Open Science
news and activities!



Rianne Fijten
Founder



Egon Willighagen
Founder



Matthijs Sloep
Founder



Gijs van Dijk
Open Science Ambassador LAW



Ewout Meijer
Open Science Ambassador FPN



Roman Briker
Open Science Ambassador SBE



Dennie Hebels
Open Science Ambassador FHML



Emmy van den Heuvel
Open Science Ambassador UM Campus



Gerasimos (Jerry) Spanakis
Open Science Ambassador FSE

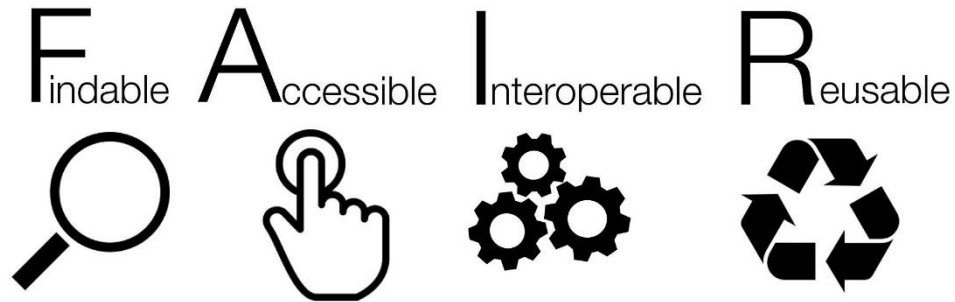


Monika Barget
Open Science Ambassador FASoS

What is FAIR? (Findable, Accessible, Interoperable, and Reusable)

wooclap

The FAIR principles



FAIR is a set of principles to define best practices for research data management to facilitate discovery and reuse by humans and machines. It involves a community agreement of creating open knowledge to serve humanity.

The FAIR principles



Findable

- (Meta)data are assigned a globally unique and persistent identifier
- Data are described with rich metadata
- Metadata clearly and explicitly include in the identifier of the data it describes
- (Meta)data are registered or indexed in a searchable resource



Accessible

- (Meta)data are retrievable by their identifier using a standardized protocol
- The protocol is open, free and universal
- The protocol allows for authentication and authorization, as needed
- Metadata are accessible, even when the data are no longer available



Interoperable

- (Meta)data use a formal, accessible, shared and broadly applicable language
- (Meta)data use vocabularies that follow FAIR principles
- (Meta)data include qualified references to other (meta)data



Reusable

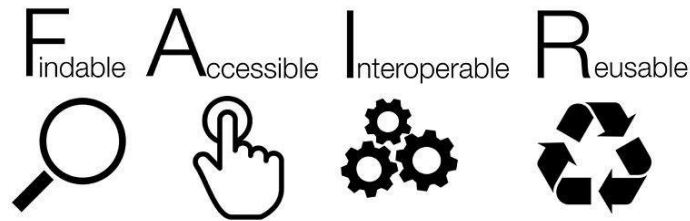
- (Meta)data are richly described with a plurality of accurate and relevant attributes
- (Meta)data are released with a clear and accessible data usage licence
- (Meta)data are associated with a detailed provenance
- (Meta)data meet domain-relevant community standards



**When you do
Research Data
Management (RDM)...**



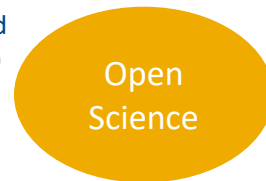
**...following the
FAIR Principles...**



**...you get closer to
Open Science**



Effort towards making
science more accessible and
reproducible (not just data)



How to describe your
(meta) data



Integral part
of

Offering principles for

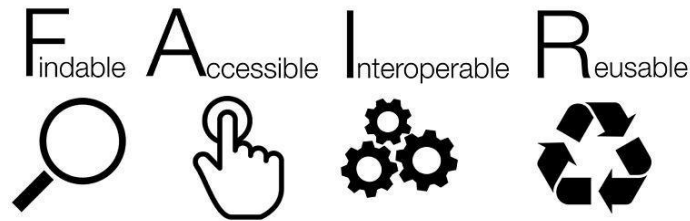


How to manage
your data

When you do
Research Data
Management (RDM)...



...following the
FAIR Principles...



...you get closer to
Open Science



Open not in
the sense of

~~Open
Datasets~~

But *Open* in the
sense of

Transparent
Science



Transparency also contributes to Research Integrity

Research Integrity: Conducting research in such a way that allows others to have confidence and trust in the methods and the findings of the research

4 FUNDAMENTAL PRINCIPLES OF RESEARCH INTEGRITY

The European Code of Conduct for Research Integrity, 2017



Combining FAIR with Open Science

(while adhering to Research Integrity code of conduct)

An Open dataset is not necessarily FAIR, for example:

John deposited his dataset in a public repository, but provided no metadata, no codebook, and did not include information on the terms of use → the data are not FAIR

A FAIR dataset is not necessarily (fully) Open, for example:

Julia implemented the FAIR principles on her dataset, but since her research involved privacy-sensitive data, she cannot make the data available → the research data are not (fully) Open

Golden combination
is Open data that are
FAIR → this is the
most impactful



European Commission requires beneficiaries of research and innovation funding to make their publications available in open access and make their data as open as possible and as closed as necessary.

In this case, Julia *did* make the metadata openly available (part of FAIR principles) → others are aware of the dataset's existence → transparency → this is perfectly acceptable!

Strict terms of use can be defined here → e.g. researchers interested in using the data should contact the data owners to see what is possible in terms of data reuse

Reproducibility/Replication crisis

Role of Open Science, FAIR, Research Integrity

Open access, freely available online

Essay

Why Most Published Research Findings Are False

John P. A. Ioannidis

Summary

There is increasing concern that most current published research findings are false. The probability that a research claim is true may depend on study power and bias, the number of other studies on the same question, and, importantly, the ratio of true to no relationships among the relationships probed in each scientific field. In this framework, a research finding is less likely to be true when the studies conducted in a field are smaller; when effect sizes are smaller; when there is a greater number and lesser preselection of tested relationships; where there is greater flexibility in designs, definitions, outcomes, and analytical modes; when there is greater financial and other interest and prejudice; and when more teams are involved in a scientific field in chase of statistical significance. Simulations show that for most study designs and settings, it is more likely for a research claim to be false than true. Moreover, for many current scientific fields, claimed research findings may often be simply accurate measures of the prevailing bias. In this essay, I discuss the implications of these problems for the conduct and interpretation of research.

Some causes:

- Lack of information
- Bad study design
- Publication bias
- P-hacking
- Fraud
- *Publish or perish* mentality in science

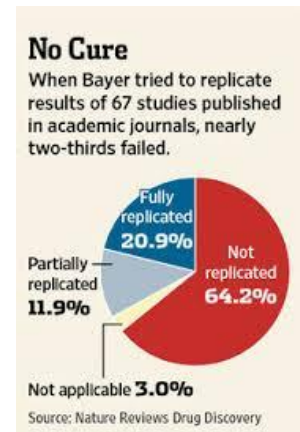


NATURE | NEWS

Over half of psychology studies fail reproducibility test

Largest replication study to date casts doubt on many published positive results.

Monya Baker



The reproducibility crisis can be remedied by applying (among others) Open Science, the FAIR principles and a Research Integrity policy.

<https://doi.org/10.1371/journal.pmed.0020124>

Reproducibility/Replication crisis

NUMBERCRUNCHER

Inadequate reporting, variability in reagents and the pressure to publish all contribute to the reproducibility crisis that is affecting life sciences...

\$23.2 MILLION 3% of 1,576 researchers said there wasn't a reproducibility crisis
Baker, Nature 533 (2016)

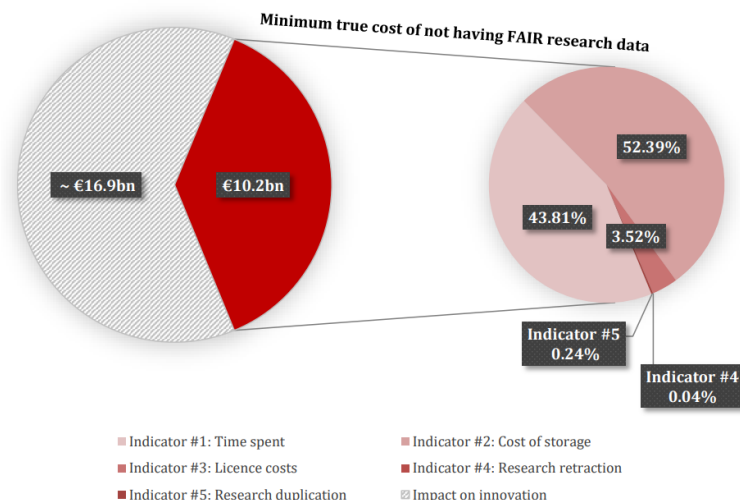
has funded the US Center for Open Science, dedicated to reproducibility in the sciences

50% of scientific resources used in previously published articles were unidentifiable
Vasilevsky et al. (2013), PeerJ

\$28 BILLION annually in the US alone wasteful spent on research that cannot be replicated
Freedman, L. P et al. PLOS Biol (2015)

70% of researchers said they'd tried and failed to reproduce another group's experiments
Baker, Nature 533 (2016)

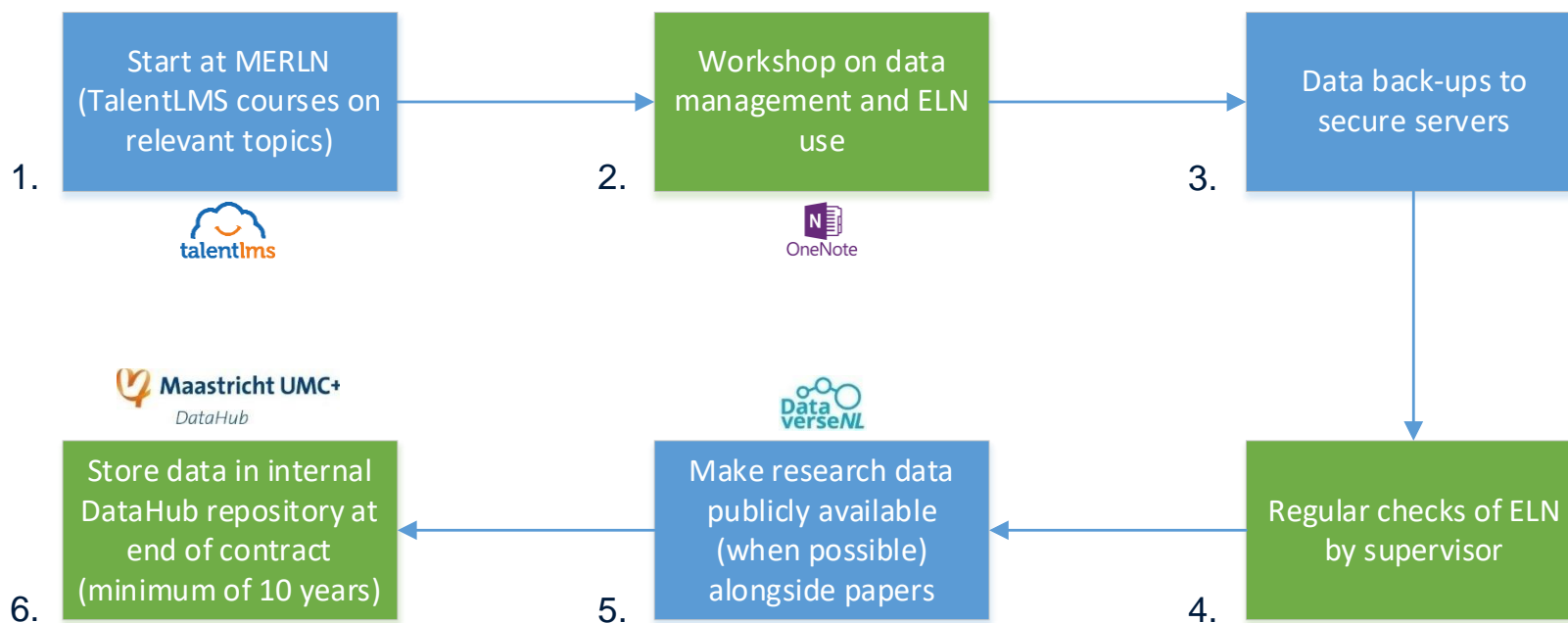
Likely cost of not having FAIR research data



- 50% of scientific resources used in previously published articles were unidentifiable. Vasilevsky et al. (2013), PeerJ
- \$28 billion annually in the US alone wasteful spent on research that cannot be replicated. Freedman, L. P et al. PLOS Biol (2015)
- 70% of researchers said they'd tried and failed to reproduce another group's experiments. Baker, Nature 533 (2016)
- Not having FAIR research data costs the European economy at least €10.2 billion every year. PwC EU services: <https://tinyurl.com/mwcjp48z> (2018)

An example of Open Science and FAIR in practice

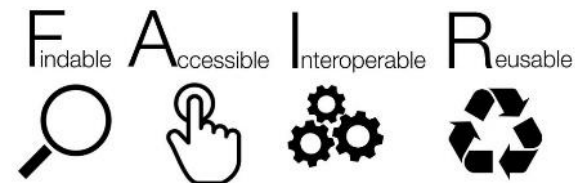
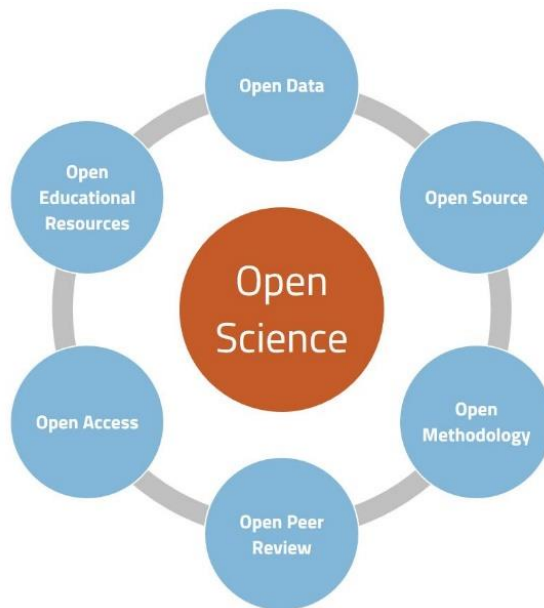
Data management in MERLN: From start to finish



ELN = Electronic Lab Notebook

1: TalentLMS intro for employees

- Online learning environment with many different courses relevant to MERLNers
- Includes a course on data management and its embedment in Open Science and the FAIR principles



Other things to consider at (or prior to) this stage

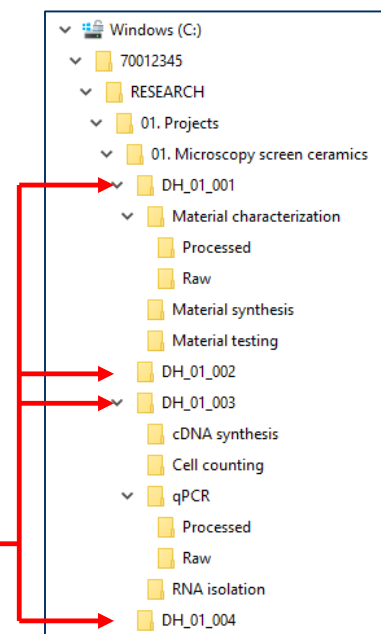
- Depending on the person, researchers will either start working on a grant proposal (usually the more senior researchers) or a grant was already acquired by the supervisor (usually the case for PhD students and postdocs)
- In either case, a Data Management Plan (DMP) needs to have been/be prepared → assistance from the data steward → make sure the FAIR principles are covered
- In rare cases (for MERLN), a GDPR form needs to be completed
- Ethics committee review (for animal or human studies) may also be needed

2: Data management and ELN workshop

- **Well organized data** (logical folder structuring and naming)
- **Clear file names** (or file names explained)
- **An electronic lab journal**
 - Logical structure
 - Shareable with colleagues
 - Flexible to all content (text/tables/images/video/audio/files)
 - Re-arranging or adding/editing data should be easy and clean
 - No cryptic handwriting issues
 - Experiment coding → unique identifiers for each experiment
 - Experimental protocols should be part of lab journal

2: Data management and ELN workshop

- There are many ELNs available. MERLIN uses Microsoft OneNote
 - Very easy to use
 - Part of Office package
 - Can be easily shared between researchers (also outside University)
 - PDF exports possible



An example of an ELN in use

HF_01_001_ZA22 - OneNote

File Home Insert Draw History Review View

Clipboard Basic Text Styles Tags

ELN Héla Fernandes

01 - Platform to develop 3D spheroids

Table of Contents Project info Experiments Work Discussions

HF_01_001_ZA22

01 May 2022 17:33

Title
Training on stamp development with Rubber ZA22 blue
First trials alone, clean acrylic casts and solve leakage issues
Rubber ZA22 blue

Date
2022-03-08 to 2022-03-15

Researcher(s)
Héla Fernandes
Carlos Mota (only on 2022-03-08)

SOP
P01_Rubber stamps_v01

1. Introduction/background

The Silicon Rubber ZA22 blue was used to get acquainted with the protocol to produce the stamps. These will later on be used to replicate the desired topography of the acrylic cast designed and developed by Carlos Mota.

This product consists of two bottles, to be used in a 50-50% ratio (one is blue and the other white). The advantage of this rubber is that it facilitates the initial experiments due to the visual aid during the mixing (the solution becomes light blue and it is easier to understand when it becomes homogeneous).

2. Objective(s) of the experiment

With these experiments, we want to successfully replicate the acrylic cast made by Carlos Mota. These will later on be used to develop 3D Spheroids on a large scale, with a more homogeneous size. Each replicative base will fit a well in a 6-well plate and will allow the production of around 15 000 3D spheroids, as according to the plan.

3. Conclusions

Some initial tests were done to develop the stamps and try to replicate the topography of the acrylic casts. Some additional tests need to be done and compare different curing times and check stamp resistance and roughness over time. Besides, plasma + ethanol 100% treatment will be necessary to allow the replication of the casts and their effect needs to be assessed.

4. Data location

Data are located here:
[C:\Users\Héla\Documents\Research\01 - Project\01 - Platform to develop 3D spheroids\HF_01_001_ZA22](#)

5. Materials and Methods

The following SOP was optimized and used during these experiments:
• [SOP_Rubber Stamps_v01](#) (for more information, consult the SOP in the dedicated SOP's separator, outside the Project section).

6. Results



 

Figure 1. These are the acrylic casts that were made with the topography that we want to replicate.

Figure 2. The casts are attached to the bottom of a 6-well plate, as can be seen in this image.

This is a representative image of the stamps and the bases and how they look after production.

Figure 1:
→ The surface with the desired topography is present in the bottom of the well.
→ The orifice present in the top of the stamp is done by placed a 15ml tube lid in the top of the solution before curing. This facilitates the work with the stamps, including extraction, cleaning (with air pressure gun) and even posterior use to replicate the cast.

Figure 2:
→ These bases will have the same topography as the acrylic casts made by Carlos Mota and will later one be used to culture cells in the microcavities made.



Figure 1. These are the stamps produced.

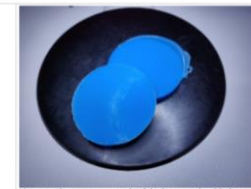


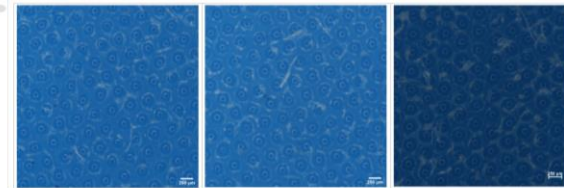
Figure 2. These are an example of the bases made with the stamps, to replicate the surface of the acrylic casts.

2022-03-08 - First Training

The protocol was done by Carlos Mota and observed by me. I was able to take some notes and understand how to develop the stamps and replicate the cast developed previously by Carlos.

→ Curing: 30min at 80°C

A first attempt was done and after curing, the stamps were observed under the microscope. It was noticed that a lot of residues were present between the pillars. This can be observed in the following images. It was hypothesized that the problem might be the reduced curing time. Therefore, higher curing times must be tested to try and solve this issue.

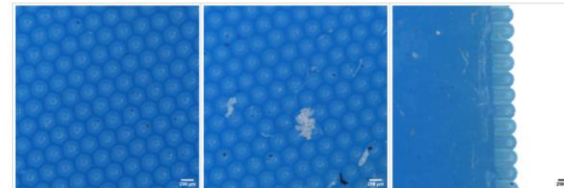


2022-03-09 - First trial alone

The protocol was done by myself, alone for the first time. The protocol done by Carlos Mota was replicated to develop more stamps. A new curing condition was tested.

→ Curing: 3h at 80°C

A first attempt was done and after curing, the stamps were observed under the microscope. It was noticed that a lot of residues were present between the pillars. This can be observed in the following images. It was hypothesized that the problem might be the reduced curing time. Therefore, higher curing times must be tested to try and solve this issue.




The increasing of the curing time permitted a better stamp surface. It is possible to observe that there are not residues present in between the pillars. The pillars are cleared and well-defined. However, as the protocol is done in non-sterile conditions, especially after removal of the stamp from the plate, it is normal to observe some residues that are not affecting the structure but are present in a non-sterile environment.

In the normal images of the stamps, we can clearly determine the width of the pillars and with a side-view image of the stamp, we can determine their depth.

NOTE: after the protocol, the acrylic casts are observed under the microscope to try and understand if all the solution poured got cured and removed. Sometimes, some pillars break inside the microcavities of the acrylic casts.

3: Data backups to secure UM servers

- Daily backups of all research data from employee laptops to UM server with multiple server backups
- Use of automated software (Iperius Backup)  IPERIUS[®] BACKUP
- Prevents loss of critical data in case of accidents



4: Regular checks of ELNs

- Follow-up by MERLN data steward a few weeks after workshop
- Afterwards, supervisors are responsible for checking compliance of employees with data management strategy



5: Make research data publicly available

- Stimulate the **public sharing of data** alongside (**Open Access**) publications
- Use of public repositories such as **DataverseNL**
- **Creative Commons Zero** license (default)
- Increasingly **demanded by journals**
- Requires researchers to make their data **reusable** → e.g., include a readme file
- In case **Intellectual Property** and/or **GDPR** legislation prevents the sharing of (part of) the research data, the **metadata** should always be shared
- **Fully anonymized** data can be shared
- DataverseNL uniquely offers the possibility to deposit **pseudonymized** data under **access request**



Benefits of Open Access



6: Storage in Maastricht Data Repository (MDR) at end of contract

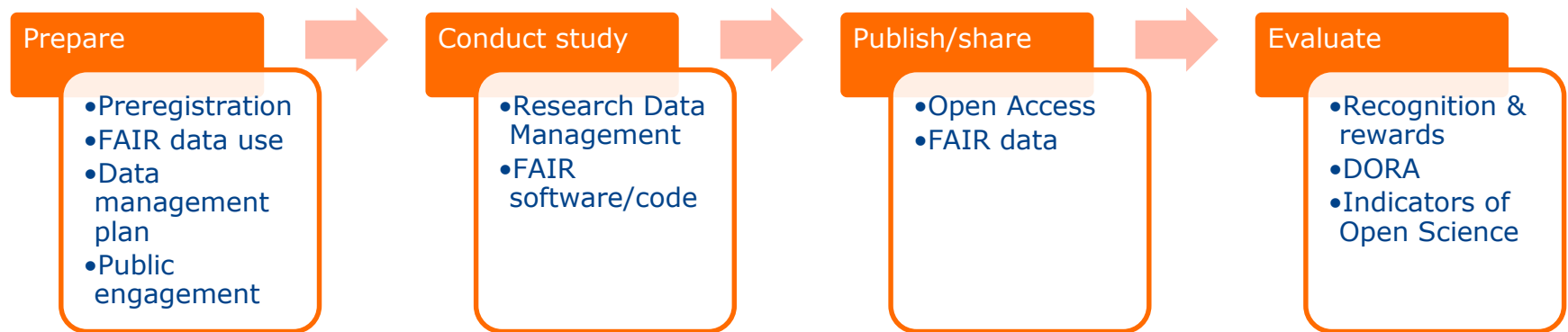
- MDR offers long-term archival private storage of data on our own servers (not publicly accessible, in contrast to DataverseNL)
- Advantages:
 - Free up space from your laptops/PCs and/or servers
 - Data are secure and backed up (2 locations)
 - Large capacity
 - Easy to use
 - Easy to access
 - Relatively cheap (€ 15-62/TB/year)
 - Also suitable for sensitive data (e.g., IP-protected, GDPR)
 - Access permission can be given to anyone within UM/MUMC+
 - Meet UNL criterion to store data for a minimum of 10 years

<https://mdr.datahubmaastricht.nl/>

Universiteiten
van Nederland }

Take home messages

- Open Science and FAIR (and of course Research Integrity) are concepts that need to be on researchers' radars constantly, from the start of a research project till the end



- Local Data Professionals (embedded data stewards) are in a unique position to stimulate researchers, point them in the right direction, and keep them up to date on new developments

Take home messages

- Join the OSCM and UM Library Umployee groups to stay up to date on Open Science events, trainings, and news



<https://umemployee.maastrichtuniversity.nl/en/groups/open-science-community-maastricht/101>



<https://umemployee.maastrichtuniversity.nl/en/groups/maastricht-university-library/144>

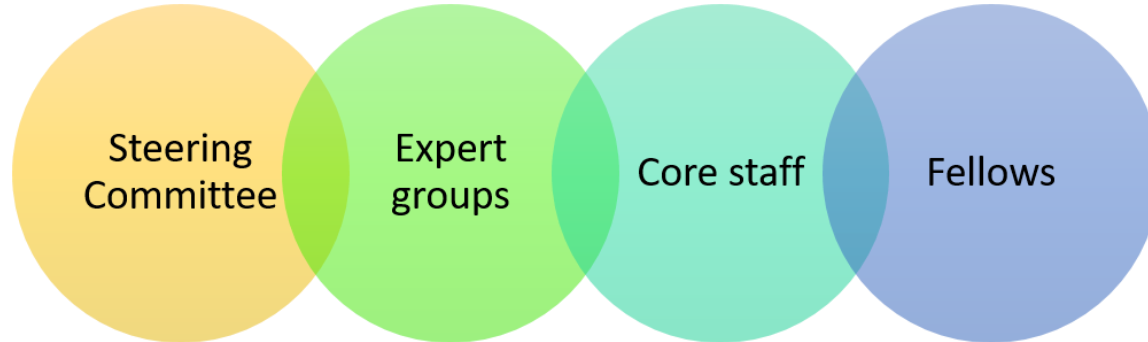
MUMC+

Data Expertise Community

Maastricht

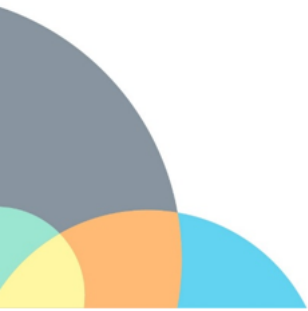
FHML

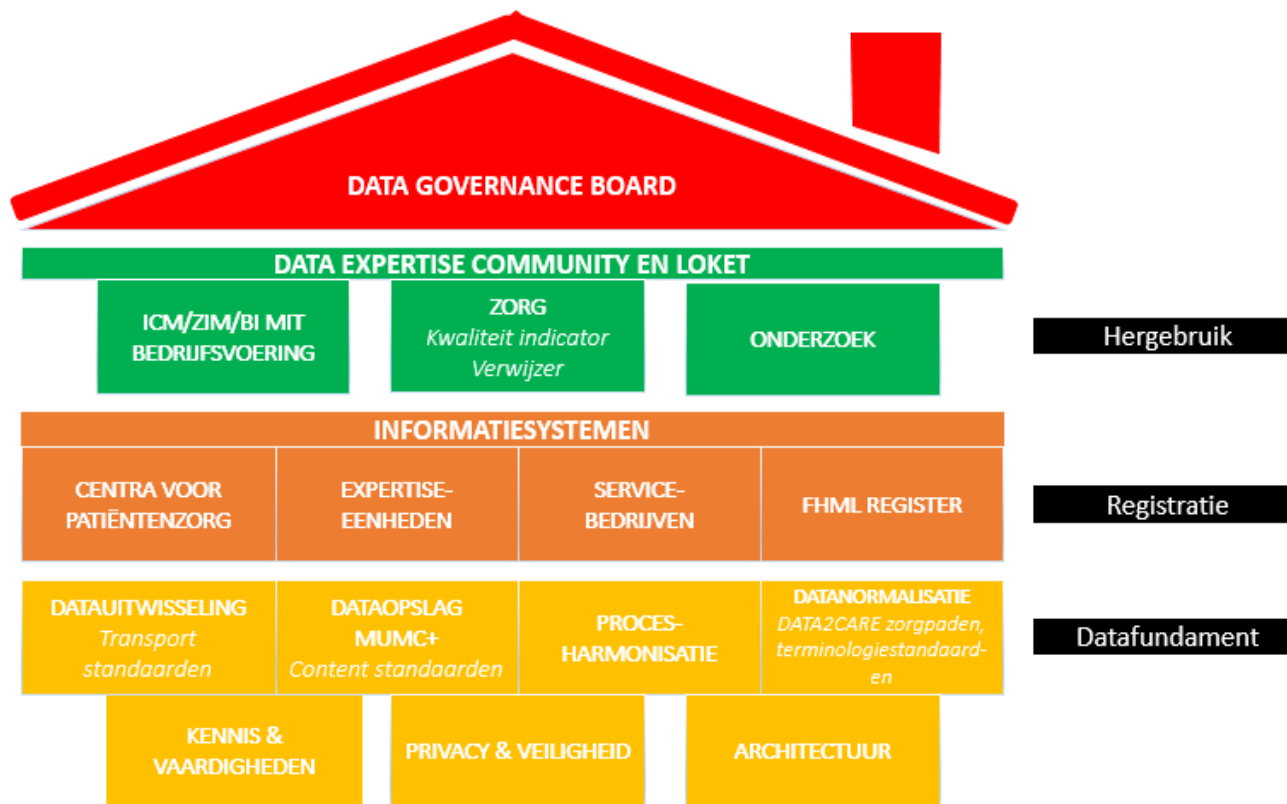
Who are we?



What are our goals?

- Structural **support** for cohorts and biobanks
- Sustaining input data **quality assurance**
- Facilitating **collaboration** in data infrastructure
- Simplifying **access** to patient and lab data for research
- Expanding **knowledge** and **expertise**





Where to go to?

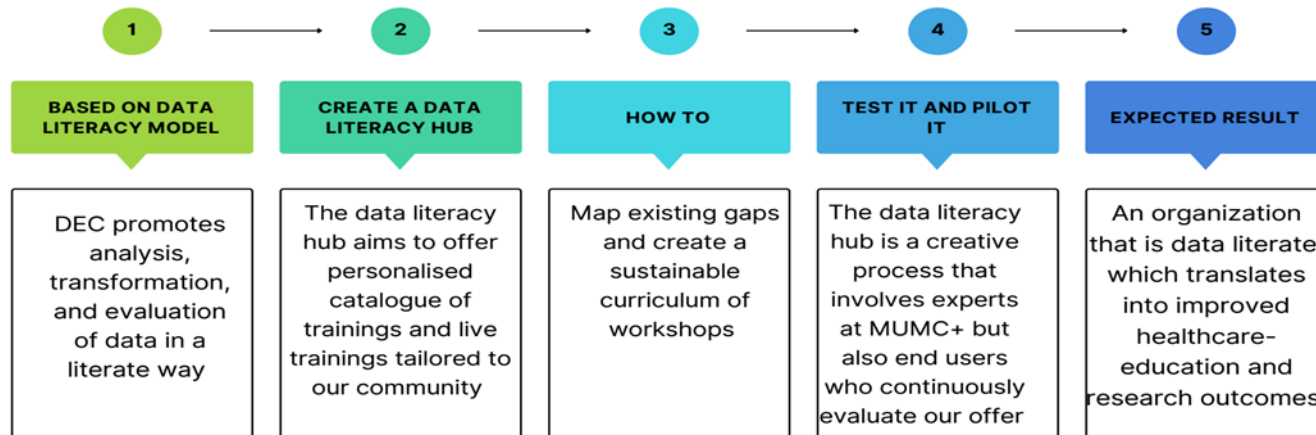


How to implement – DEC Domain Health Care

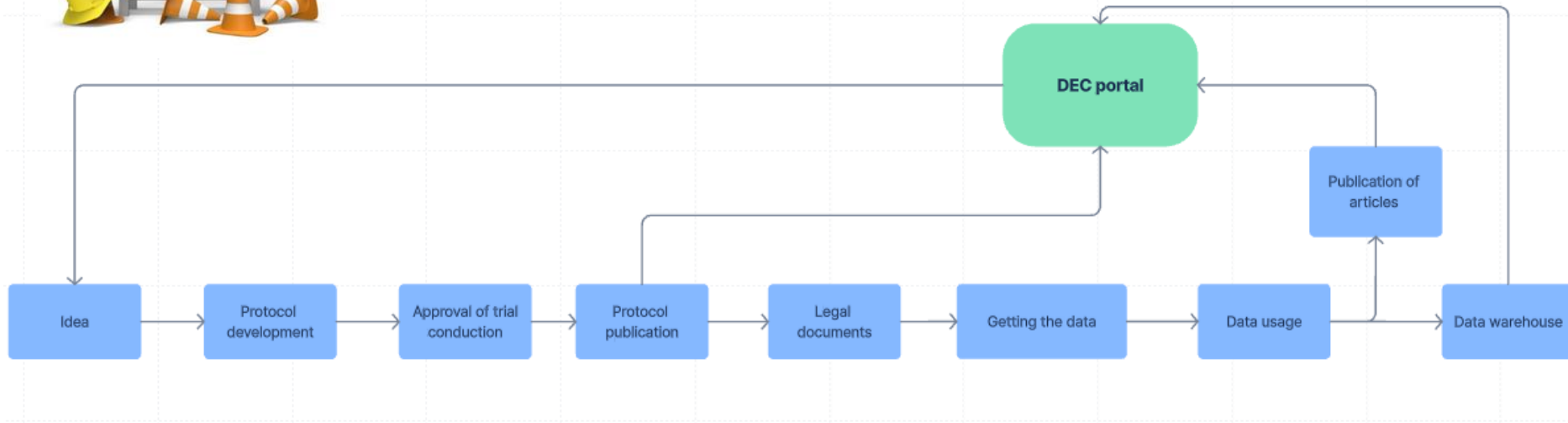
Health Care professionals and other users make interactive decisions



How to implement – DEC Domain Education



How to implement – DEC Domain Research



How to implement – DEC Domain Research

Research flow

- Retrospective
- Prospective
- Beyond the Maastricht UMC+

Zorgpaden

Procedures

- GDPR
- Data Management Plan
- nWMO/WMO
- Legal documents

Data repository

- Metadata overview
- Research data & Health Care data (invisible to the user)
- Publications
- Conducted clinical trials

Data requests

Questions & Contact Information



dec@mumc.atlassian.net

Data Protection Training

13th February-2023

Helena Bossini Castillo

Alfons Schroten

Vasco Trindade Veiga

Program

- Why GDPR?
- History
- Definitions
- GDPR and scientific research
- Consequences of non-compliance

GDPR

The GDPR or **General Data Protection Regulation**

What is the GDPR?

- The General Data Protection Regulation (GDPR) is the European privacy legislation, which has been in force throughout the European Union (EU) since 25 May 2018. The GDPR is about the protection of **personal data**. **Citizens** have more **(privacy) rights**, **organizations** have more **responsibilities** and the National **Authorities** have more **powers**.

Where does the GDPR apply?

EU countries (27)
+
Norway, Iceland and Liechtenstein
=
European Economic Area (EEA)

Personal Data

- **Any information** relating to an **identified or identifiable natural person** ('data subject'); an identifiable natural person is one who can be identified, **directly or indirectly**.
- For example: date of birth, telephone number, IP address, IBAN, license plate number, photos, images, mental or physical characteristics, location data, patient numbers, information about income, medical data, personnel file, answers to an exam, etc.
- **Data that does not fall within the definition** of personal data is data of **legal persons, anonymous data** and, **in some European countries**, such as the Netherlands, **data of deceased persons**.

Special categories of Personal Data

- Certain personal data with **extra sensitive information** such as:
 - Data on race, religion, sexual orientation, trade union membership, political preference, medical and biometric data, genetic data...

Have special protection in the GDPR

- Criminal conviction and offense data is not, by definition, a special category of personal data, however it does have special protection.

Special categories of Personal Data

- The **processing of special categories** of Personal Data **is prohibited**
- **Unless** there is a legal **exception**

Pseudonymization

- **Process in which the Personal Data is adjusted so that it can no longer be traced directly to the person without additional information/data**
- Certain information is replaced with a code or key
- Keep key separately and take measures to prevent re-identification
- Security measure must be implemented
- **Pseudonymized data is still personal data and must therefore comply with the GDPR**

Anonymization

- With anonymization, there is no traceability of the person, not even through a key
- Anonymous data is not personal data and therefore the GDPR does not apply
- Anonymous data is not always anonymous
 - Rare diseases/occupations
 - Combinations (Occupation: mayor, residence: Maastricht)
 - With the help of new techniques it will be easier and easier to know who the person is
- Anonymization of data is also processing personal data

What is processing?

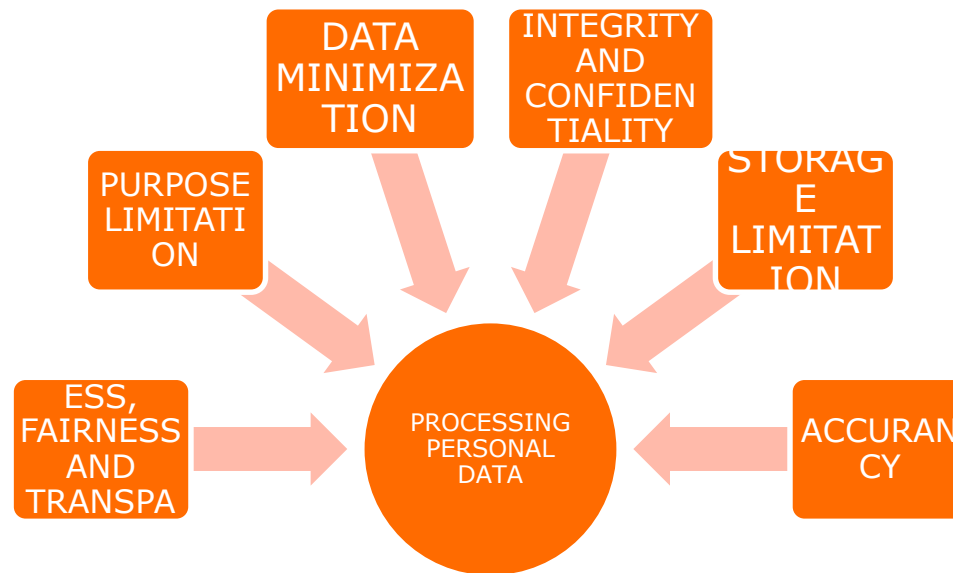
- Any action with or on Personal Data
- For example: collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.



"Before I write my name on the board, I'll need to know how you're planning to use that data."

Principles of data processing

Check the accuracy of your research in 6 steps



Legal base to process Personal Data

1. Consent
2. Execution of an agreement
3. Legal obligation
4. Protect vital interests
5. Public interest or Public Authority
6. Legitimate interest

Plenary session

- What can we offer each other?
- Formulate future points of collaboration between both organizations as well as outside?

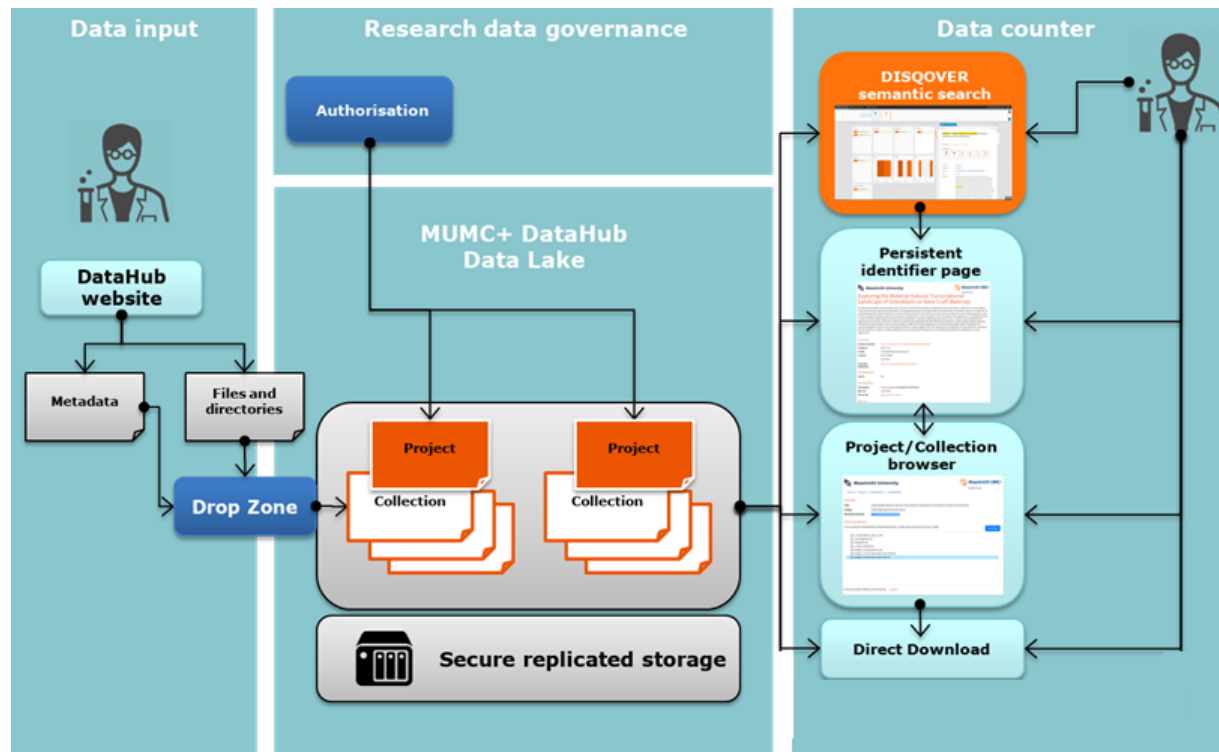
Let's Play with the tool





**A self-service FAIR data station for responsible
research data management**

Maastricht Data Repository Infrastructure



Storage is not for free !

Projects

View as:
Default **Principal Investigator**

Show 50 entries Search:

Project	Title	Size	Size (GB)	€ / GB	€ / month*	Budget no.
P000000022	Imaging of the lung organoids	13 TiB	14681	0.0267	391.49	UM-30991022N
P000000071	Single Particle Analysis on ESX1	10 TiB	10981	0.0267	292.83	UM-30991022N
P000000039	Tools for near-atomic resolution in single-particle cryogenic electron microscopy (Pavel Afanasyev)	6 TiB	6240	0.0267	166.4	UM-30991022N
P000000018	Cryo FIB/SEM lamellae tomography workflow of infected cells	5 TiB	5852	0.0267	156.05	UM-30991022N
P000000038	Archive of data Peter Peters at NKI	5 TiB	5780	0.0267	154.13	UM-30991022N

**High
Performanc
e Spinning
Disk
Storage
(NAS)**

€ 132 / TB /
year

**Medium
Performanc
e Spinning
Disk
Storage
(Ceph)**

€ 62 / TB /
year

**Tape
Archive**

€ 15 / TB /
year

<https://portal.datahubmaastricht.nl/pages/viewpage.action?pageId=44269570>

MDR exercise

- Log in to VPN- Horizon Client
- Login to MDR
- Explain the request project step- demo by Olav
- create a drop-zone, enter meta-data, add data/files, ingest, and edit meta-data, report persistent identifier

DMP exercise: Different DMP templates

ANNEX 1: Horizon 2020 FAIR Data Management Plan (DMP) Template

INTRODUCTION

This Horizon 2020 FAIR DMP template has been designed to be applicable to any Horizon 2020 project that produces, collects or processes research data. You should **develop a single DMP for your project** to cover its overall approach. However, where there are specific issues for individual datasets (e.g. regarding openness), you should clearly spell this out.

FAIR data management

In general terms, your research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation solution.

This template is not intended as a strict technical implementation of the FAIR principles, it is rather inspired by FAIR as a general concept.

More information about FAIR:

[FAIR data principles \(FORCE11 discussion forum\)](#)

[FAIR principles \(article in Nature\)](#)

Structure of the template

The template is a **set of questions** that you should answer with a level of detail appropriate to the project.

It is not required to provide detailed answers to all the questions in the first version of the DMP that needs to be submitted by month 6 of the project. Rather, the DMP is intended to be a **living document** in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, DMPs should have a clear version number and include a timetable for updates. As a minimum, the DMP should be updated in the context of the periodic evaluation/assessment of the project. If there are no other periodic reviews envisaged within the grant agreement, an update needs to be made in time for the final review at the latest.

The **main sections** to be covered by the DMP are outlined.

At the end of the document, Table 1 contains a summary of these elements in bullet form.

This template itself may be updated as the policy evolves.

1. DATA SUMMARY

What is the purpose of the data collection/generation and its relation to the objectives of the project?

What types and formats of data will the project generate/collect?

Will you re-use any existing data and how?

What is the origin of the data?

What is the expected size of the data?

To whom might it be useful ('data utility')?

2. FAIR DATA

Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

What naming conventions do you follow?

Will search keywords be provided that optimize possibilities for re-use?

Do you provide clear version numbers?

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

How will the data be made accessible (e.g. by deposition in a repository)?

What methods or software tools are needed to access the data?

Is documentation about the software needed to access the data included?

Is it possible to include the relevant software (e.g. in open source code)?



Where will the data and associated metadata, documentation and code be deposited? Preference should be given to ...




Data Management Plan Template











This template is intended for creating a data management plan, based on the data management section that was part of your research proposal. NWO expects you to incorporate any comments received from the referees and/or the committee about the data management section in this data management plan.





The UM template

  Home Recent ▾ Spaces ▾ Teams ▾ Apps ▾ Templates [+ Create](#)




 6  

Research Managemen... / Data Management Plan **UNPUBLISHED CHANGES** Ready for review |     |  Summarize |   [Share](#) 

Data Management Plan

   Owned by Olav Palmen ...
Last updated: Jun 14, 2023 • 5 min read •  114 people viewed

Get started

-  [I want to start writing a new Data Management Plan](#)
-  [I want to see my current Plans](#)
-  [Ask a question](#)

How do I write my Data Management Plan?

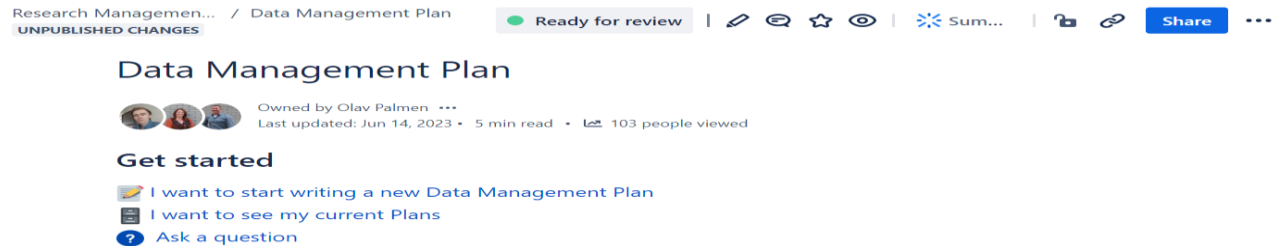
The task of writing a **Data Management Plan (DMP)** might seem daunting at first. No worries. We got your

Part I:

1- Go to :

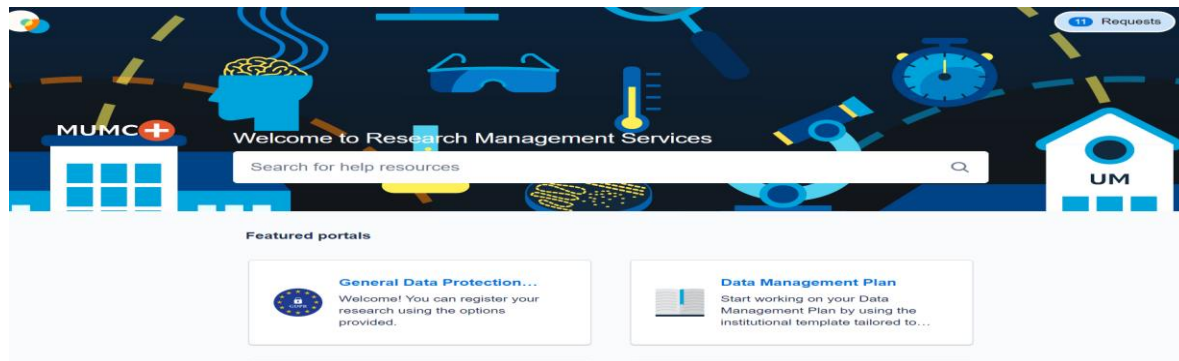
[Data Management Plan - Research Management Services - MUMC+ \(atlassian.net\)](https://mumc.atlassian.net/servicedesk/customer/portal/4)

2- If you have an account log in and if you don't sign up and create an account



3- Select data management plan module and click on it to load the form

<https://mumc.atlassian.net/servicedesk/customer/portal/4>



Part II:

4- Fill out the mandatory fields, select data steward (Maria Vivas-Romero), write down case study with data descriptions/ Choose a chapter to work on

The screenshot shows the 'Data Management Plan' form. It includes a header with 'Research Management Services / Data Management Plan' and a sub-header 'Data Management Plan'. Below this, there is a brief introduction and a section titled 'What can we help you with?'. The form contains several input fields: 'Project title', 'Project purpose', and 'Data Steward'. The 'Data Steward' dropdown menu is open, showing 'Maria Vivas-Romero (m.vivasromero@maastrichtuniversity.nl)' as the selected option. There are also checkboxes for 'Project title' and 'Project purpose'.

1. General Information

1.1a Please select your institute

The address details of your institute will be visible in the export of your data management plan

Select...

1.1a Comment

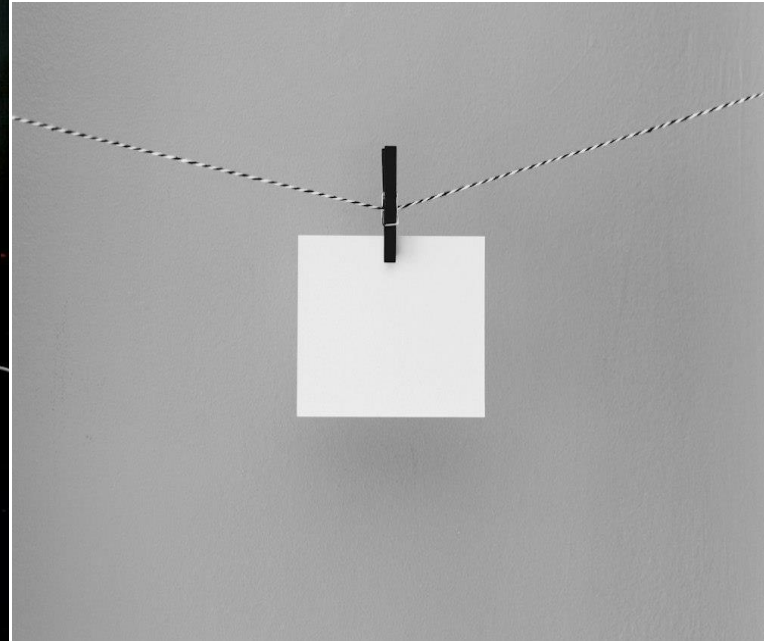
Add comments to share with collaborators

5- Send it, make an edit, leave a comment on the right hand side of the form

6- Submit the plan / approval

The screenshot shows the bottom section of the form. It includes a warning message: 'When clicking the "Send" button everything is saved (but not submitted!). After that you can continue at any time via the "Edit" button and "Submit" the final version.' Below this is an 'Attachment' section with a text input field and a 'Browse' button. At the bottom, there are 'Send' and 'Cancel' buttons.

QUESTIONS AND COMMENTS





Visit us at <https://datahubmaastricht.nl>



Maastricht UMC+

DataHub



Pascal Suppers
Managing Director

DataHub Maastricht
Paul-Henri Spaaklaan 1
6229 EN Maastricht **T** +31 6 27 07 16 54
The Netherlands **E**
p.suppers@maastrichtuniversity.nl