

Data management plans

Elena Giglia

Soft Skills courses

Sevilla 3rd conference – November 20th/21st, 2025

Data management plans

Innovation  Acta



Seville, Nov. 21, 2025

Elena Giglia

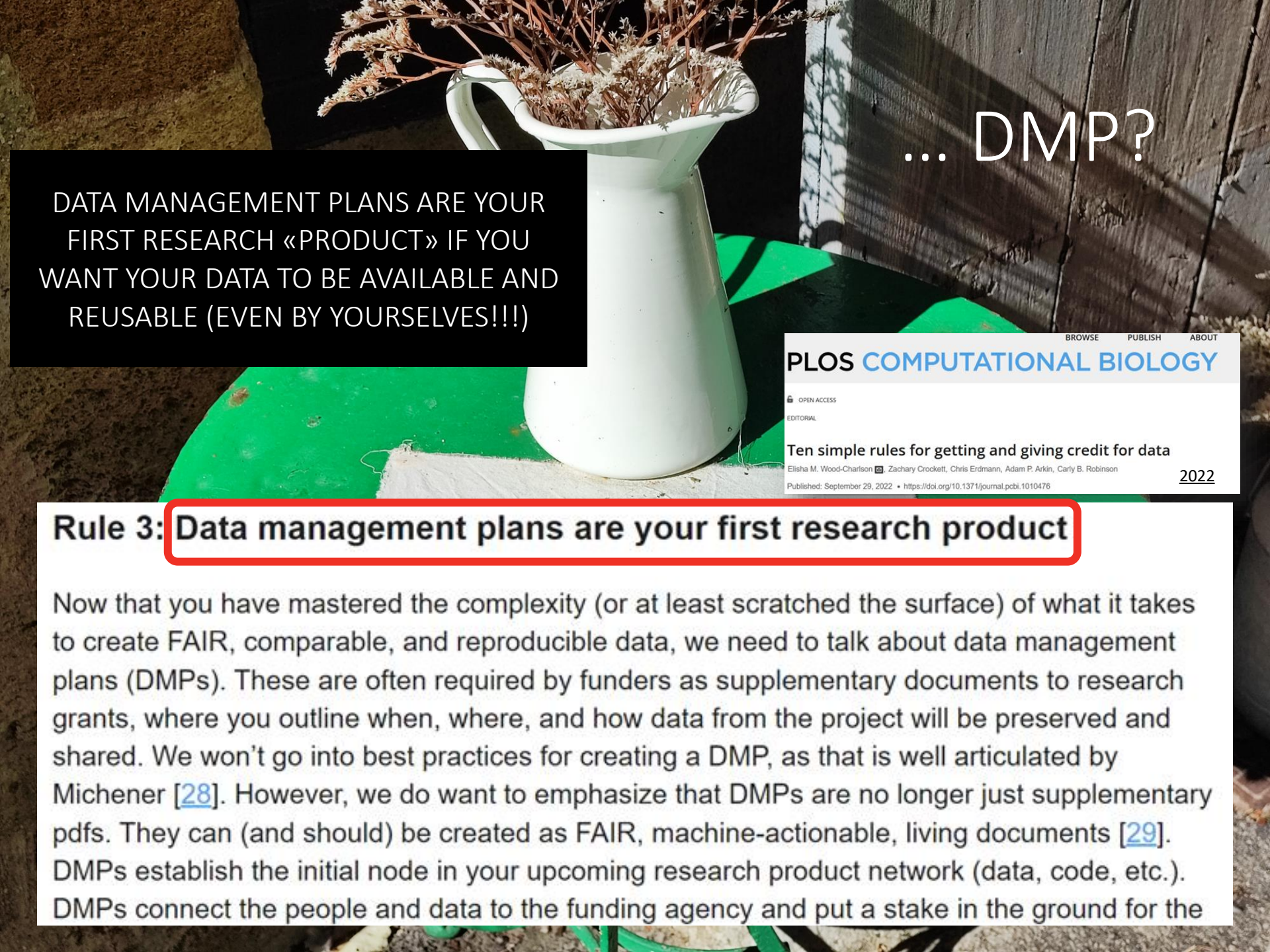
elena.giglia@unito.it



[@elegiglia.bsky.social](https://bsky.social/@elegiglia)



[ORCID](#)



DATA MANAGEMENT PLANS ARE YOUR
FIRST RESEARCH «PRODUCT» IF YOU
WANT YOUR DATA TO BE AVAILABLE AND
REUSABLE (EVEN BY YOURSELVES!!!)

... DMP?



Rule 3: Data management plans are your first research product

Now that you have mastered the complexity (or at least scratched the surface) of what it takes to create FAIR, comparable, and reproducible data, we need to talk about data management plans (DMPs). These are often required by funders as supplementary documents to research grants, where you outline when, where, and how data from the project will be preserved and shared. We won't go into best practices for creating a DMP, as that is well articulated by Michener [28]. However, we do want to emphasize that DMPs are no longer just supplementary pdfs. They can (and should) be created as FAIR, machine-actionable, living documents [29]. DMPs establish the initial node in your upcoming research product network (data, code, etc.). DMPs connect the people and data to the funding agency and put a stake in the ground for the



IT IS A STRUCTURED WAY
TO THINK OF YOUR DATA

IT IS A FORMAL DOCUMENT ABOUT
HOW YOU MANAGED YOUR DATA
(ENSURING INTEGRITY)

CLEAR RULES = LESS
MISTAKES FROM THE
BEGINNING

A NEW WAY OF THINKING TO YOUR
RESEARCH, FROM THE PERSPECTIVE
OF YOUR DATA

IT IS A «LIVING
DOCUMENT»,
IT GROWS WITH THE
PROJECT

IT IS THE RIGHT VENUE TO
JUSTIFY OPEN/CLOSED
CHOICES

IT IS THE VENUE TO
EXSTIMATE COSTS

A Data Management Plan...

DMP?

...LET'S BE CLEAR:
THE ISSUE HERE IS NOT «LEARNING»
HOW TO DRAFT A DMP
BUT **LEARNING HOW TO RESPONSIBLY
MANAGE FAIR DATA.**
DMP IS MIRRORING THAT

DMP IS WHERE YOU «MAKE
CLEAR» THE WAY YOU ARE GOING
TO MANAGE YOUR DATA

IT'S NOT BUREAUCRACY, IT'S A
RESPONSIBILITY AND A COMMITMENT
(AND A POWERFUL TOOL, LIKE A MAP TO
YOUR DATA)



Data management ABC – To start

Ask yourself this:

[DMP]

What is needed to validate the results of your research?

If you were to produce an article researching, for example, the criminal underclass in early-twentieth century New York, what data would you need to include for someone else to replicate your results? Think about it in terms your own research.

A bibliography would be the most immediate and obvious starting point, revealing to the reader all the sources that you have used to base your research. But what of the gathering mechanisms you used? Did you create a database or undertake statistical analysis? If so you need to make the database and statistics available. This doesn't just mean providing the files in a readable format, but to provide documentation and to make sure that the data is clearly identified with explicit headings, well-structured, and easily identified.

Focusing on what is needed for validation and re-use, rather than the obvious attributes of research data, is useful. It helps you to think through the process of research from a different perspective and what it is you have actually done to come to your conclusions. It also allows you to show the process you have undertaken; revealing how valuable your approach might be and making the

WHAT DO ANYONE NEED
TO REPLICATE/VALIDATE
MY RESEARCH?
THAT'S THE QUESTION
YOU HAVE TO ASK
YOURSELF



DMP – to get started

2023

News and insights
from Open Science

And how to exploit them
in your scientific work

Gina Pavone, Cnr-Isti, ORCID 0000-0003-0087-2151
How to write a Data Management Plan, with hands-on session
12 July 2023

Why?

The DMP is a structured approach to data management: instead of improvising when a need arises, thoughtful choices are made across the entire data lifecycle.

For oneself

- Save time and try to prevent problems in the future -
- Estimate costs -
- Get credit for your data and do not drown in irrelevant data -

For mandates

- It is mandatory in EC and ERC funded projects -
- Also other funders ask for it -
- RPOs may have their own policy on RDM -

For others

- Produce FAIR data, easier to find, understand and reuse -
- DMPs may also be required as part of the ethical approval process -

GDPR

- Even if a full DMP is not required, a record of processing activities is needed to comply with the GDPR when working with personal data



GOOD TIME INVESTMENT

"The time invested in setting up a good data management strategy pays off when the time comes to reproduce your analysis and results.

You will be able to easily find and understand your data, increase your data's reuse potential and comply with funder mandates at the same time."

DMP – to get started

2023

News and insights
from Open Science

And how to exploit them
in your scientific work

Gina Pavone, Cnr-Isti, ORCID 0000-0003-0087-2151
How to write a Data Management Plan, with hands-on session
12 July 2023

DMP: so many benefits



Make data FAIR

- Makes structuring and documenting of your datasets simpler, thus making it easier for others as well as your future self to find and understand the material;
- Encourages you to think about the data format which is best suited for reuse;
- Allows you to think about the reuse license you would want to apply to your data;
- Choose a proper repository etc.



Clarifies needed budget

- calculating time and resources for careful documentation as well as server space, backup solutions, hardware and software etc.
- Calculating time and resources (money and expertise) for collecting, analysing, and publishing on data.



Allows for easy project management

- An important function of a DMP is to work as a one-stop shop to find project-related information.
- Questions surrounding data management are being gathered in one place and project-related details are readily available rather than just vaguely remembered or simply forgotten.



Shows accountability

If you draw up a DMP, you are showing your affiliated institution, funders and project partners a serious approach to research data management, that includes a responsible approach towards research funds and research participants.

DMP – to get started

2023

News and insights
from Open Science

And how to exploit them
in your scientific work

Gina Pavone, Cnr-isti, ORCID 0000-0003-0087-2151
How to write a Data Management Plan, with hands-on session
12 July 2023

Stop drowning in irrelevant stuff



Time is the enemy!

- Accumulating large quantities of disorganized files
- Lack of information describing content in digital files
- Changes to hardware, software, and file formats in common use
- File corruption
- Failure of storage media
- Data leaving with collaborators

And - last but not least - tons of daily activities!



DMP – to get started

2023

News and insights
from Open Science

And how to exploit them
in your scientific work

Gina Pavone, Cnr-isti, ORCID 0000-0003-0087-2151
How to write a Data Management Plan, with hands-on session
12 July 2023

What is normally a DMP about?

Very basic aspects



Identify

the data you are working
with in your project.

- Accurately describe the types of data to be used
- Why do you need that data?
- What is the research question to be answered?



Decide

the strategy to
organise your data
and the standards
you will use.

- Make careful choices to document all steps
- In the future it will be easy to understand and retrieve all the information?

Manage

Make decisions
about daily
data management.

- What is your plan for sharing your data?
- Will you have issues sharing your data?
- Will you need more resources/budget than expected?



Data summary

Main elements to describe

- types of data
- purpose of the data
- file formats
- organization: file naming and folder structure
- Expected size of the data
- Provenance (origin/source)
- Is that data potentially useful for others?

t started

Example of data types

Text: Field or laboratory notes, survey responses

Numeric: Tables, counts, measurements

Audiovisual: Images, sound recordings, video

Code and models: Python, MATLAB, R, etc.

Discipline-specific: FITS in astronomy, CIF in chemistry

Instrument-specific: Equipment outputs

Spatial, temporal, instrument-generated, models, simulations, images, video

File naming

About:

- cues to the content
- status of a file
- uniquely identify a file
- help in classifying files

What can contain:

- project acronyms
- researchers' initials
- file type information
- a version number
- file status information
- date

Best practices:

- create meaningful but brief names
- Provide useful cues to content, status and version
- Make it unique, descriptive and informative about the content.
- Make it independent of the location of the file
- avoid using spaces and special characters
- Keep it short
- keep the file extension at the end

Adapted from: <https://dam.ukdataservice.ac.uk/media/62247/managingsharing.pdf> and https://dam.eislr-belgium.org/file_naming.html

A good example

http://www.data.cam.ac.uk/files/gdl_tildocnaming_v1_20090612.pdf

3. Version
(upper case, max 4 chars, optional)

For documents that will continue in various versions use V followed by the version number. Use an underscore to indicate a decimal point if necessary.
Eg. PMF_PRP_ZenMonkeyProject_V2_20090607.docx

New versions should not be created for each iteration of the document, but rather at significant changes or when it has been reviewed or changed by another author.

Document naming for the TILS Division should follow this convention:

GDL TILSDocNaming V1 20090612.docx

A prefix shows document type

The document title

The version number

The date in the

Prefix	Meaning
AGD	Agenda
AGR	Agreement
GDL	Guideline
MEM	Memorandum
MIN	Minutes and Notes
PRE	Presentation
PRO	Procedure
PRP	Proposal
REP	Report
TEM	Template

2. Document title/ Description
(mixed case, max 30 chars, no spaces)

- Describes the purpose or "business" of the document. Acronyms, capitalisations, abbreviations can be used, keep in mind that descriptions should be **meaningful** to anyone reading the file name
- In the case of project documentation use the **project name** or its usual abbreviation
- If possible Department Branch and/or Section should be integrated into this field to indicate origin / ownership of document.
- Use only alpha-numeric characters, plus the hyphen and underscore.
- **Do not use spaces.**

Folder structure

Do's

- List all the type of files required for your work and try to group them according to logic criteria.
- Consider the best hierarchy for files
- Go from a general, high-level folder (e.g., a single folder for the project, using its name or acronym) to more specific lower-level folders
- Make it intuitive, clear and understandable for everyone.
- Apply file naming guidelines for naming files.
- Include documentation to explain a complex folder structure (such as a README.txt) at the root of your folders.
- Separate raw and processed data

Dont's

- Make your structure too deep nor too shallow (the number of levels depends on the project).
- Use a generic "current stuff" or "my stuff" folder.
- Create researcher-specific folders ("Name_Surname" folder) within a project: folders are about the contents, not the authors.
- Create similar folders in different places (overlapping categories or folder redundancy).
- Create copies of files in different folders.

For example: Projects (README.txt); Administration (Budget, Approval, Travels); Planning (DMP, Ethics,); Literature; Experiments (ExperimentA, ExperimentB...); Dissemination (Posters, Presentation, Publications).

Adapted from: https://rdm.elixir-belgium.org/folder_structure#sources-and-further-reading

News and insights
from Open Science

And how to exploit them
in your scientific work

ne, Cnr-Isti, ORCID 0000-0003-0087-2151
ite a Data Management Plan, with hands-on session
!3



Findability

Elements to include in the DMP

Some "F" aspects to consider

- Use metadata and specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how.
- Use search keywords
- Persistent and unique identifiers such as DOI
- File and folder naming conventions
- Versioning of the datasets and clear version numbers

2023

News and insights
from Open Science

And how to exploit them
in your scientific work

Gina Pavone, Cnr-isti, ORCID 0000-0003-0087-2151
How to write a Data Management Plan, with hands-on session
12 July 2023

ACCESSIBILITY

Elements to include in the DMP

- Mention the repository you intend to use
- Info on the repository features and policies (eg. authentication procedure, long term preservation, PID/DOI...)
- Availability of metadata
- Access rights

Some "A" aspects to consider

- Explain which data can't be shared openly, if any
- Specify how access will be provided in case of restrictions, e.g. through a data committee, a license, or arranged with the repository
- Will methods or software tools needed to access the data (if any) be included or documented?
- Deposit the data and associated metadata, documentation and code preferably in certified repositories which support Open Access

Interoperability

Main aspects to include

- Vocabularies
- Reference to other metadata

	A	B	C
1	Attribute	Attribute Description	Example
2	Study Title	Title of your experiment	Genes in response to drug
3	Study Description	What your experiment is about (hypothesis, methods etc...)	Genes expression in liver of mice treated with DrugY
4	Person Name	Name Surname	My Name
5	Person Role	Performer, Supervisor etc...	Performer
6	Experimental factor(s)	Independent variable manipulated by the experimentalist	Drug
7	Experimental factor levels	Groups you want to compare	PBS DrugY
9	Replication type(s)	Experimental Unit, Observational Unit, Independent replicates, pseudoreplicates etc...	Experimental Unit
10	Replication description	What your replicates are	Mouse
11	Organism	Species according to one taxonomy	Mm
12	Genotype	Genetically modified or not	geneX_WT
13	Date	Format as YYYYMMDD	YYYYMMDD
14	Extract_Material	Extracted material type	RNA
15	LibraryLayout	Single end or paired end (SE, PE)	SE
16	MillionReads_Asked	How many reads are needed	10
17	ReadLength_bp	Reads length in bp	50

Reusability

Main elements to include:

- documentation
- licences
- quality assurance

Some hints for reusability

- Describe the **scope of your data**: for what purpose was it generated/collected?
- Mention any **particularities** or **limitations** about the data that other users should be aware of.
- Specify the date of **generation/collection** of the data, the lab conditions, who prepared the data, the **parameter settings**, the **name and version of the software** used.
- Is it **raw** or **processed** data?
- Ensure that all **variable names** are explained or self-explanatory (i.e., defined in the research field's controlled **vocabulary**).
- Clearly specify and document the **version** of the archived and/or reused data.

To be operational

IN UNITO WE NEED TO SET A CLEAR DMP WORKFLOW.

IT HAS TO BE CLEAR WHO DOES WHAT AND AT WHAT STAGE (TO AVOID DELAYS)

WHAT TO ASK

- TYPE OF DATA (AND VOLUME)
- GDPR ISSUES? ETHICAL ISSUES?
- EXPLAIN «AS OPEN AS POSSIBLE»
- GIVE REFERENCES ON WHERE TO FIND SPECIFIC TOOLS

1) PRELIMINAR INTERVIEW

(IT'S POINTLESS TO «REVIEW» AFTER THEY DRAFT – USUALLY A MEANINGLESS DOC AS THEY HAVE NO CLUE ON WHT TO PUT IN A DMP)

2) SHARED SET UP OF THE ONLINE DOC

3) ONLY THEN, A FINAL REVISION

Tips and tricks

Top tip - keep it short and specific!

This very short extract from a presentation by Peter Dukes, Medical Research Council, gives really useful advice on writing a DMP from the funding body perspective. The advice applies to all disciplines. The quality of the video isn't great, but the advice is definitely good!



BE SINTETIC AND
SPECIFIC

DO NOT
COPY/PASTE

BEING GENERIC IS USELESS
[«we expect a huge size of
data»; «data will be available»]

EVERY DATASET IS
UNIQUE, EVERY
INFRASTRUCTURE IS
DIFFERENT, EVERY
RESEARCH HAS
DIFFERENT
PARTNERS/POLICIES

- LET'S USE TABLES AND BULLET POINTS
- BE CLEAR, SHORT SENTENCES. IT'S NOT A DISSERTATION

- IF YOU DON'T KNOW ABOUT A SINGLE ASPECT, JUST TELL IT
- IF NOT, IT LOOKS LIKE YOU ARE NOT AWARE OF THAT SPECIFIC ASPECT [SAME DIFFERENCE AMONG A BLANK CELL AND A CELL WITH "N.A" OR ZERO]

WHAT YOU STATE IN THE
DMP THEN HAS TO BE
DONE...

DON'T SHOW OFF
DON'T DECLARE
SOMETHING YOU CAN'T
GET

e.g. PSEUDONIMYZED
DATA, not ANONIMYZED



10 TIPS FOR WRITING A DATA MANAGEMENT PLAN

Tips&tricks / 2

1

START EARLY

Read the guidance and ask for advice early on in the process, as writing a DMP may take some time

2

CONSIDER REUSE

Think about reusing existing data. Describe what you will need to know about your data five years from now

3

CHECK POLICIES

Talk to your supervisor or lab members about existing data management policies and standards

4

MAKE USE OF SUPPORT

Use your in-house support services like RDM Support, the Library, IT department or legal desk

5

THINK BROAD

Also address software code, algorithms and any other valuable research assets in your DMP

6

COPY WHERE YOU CAN

Look at other (submitted) plans and copy when appropriate

7

BE UNIQUE WHERE NEEDED

Since every research project is unique, so are the data it generates. Copying from sample DMPs is not sufficient

8

BE CONCRETE

Make your answers as concrete as possible. Show that you have consulted RDM experts

9

SAY SO IF YOU DON'T KNOW

Indicate what you do not yet know and how you will resolve these questions later

10

UPDATE

DMPs add to the planning of your research methods. Therefore define, carry out and update your DMP just as you would any method

A data planning process ensures that all aspects of data management are holistically explored at the start of a project. Short-term and long-term aims can be balanced, so that decisions made early in a project do not negatively impact on the ability to find and use the research data in future.

Effective management of data provides researchers with many benefits, including

- time saved through reduced duplication of effort
 - decreased risk of loss, theft or inappropriate use of data
 - good research practice ensures the integrity and quality of data
 - data can be understood and used now and in the future
- helps researchers find and gain access to data management – expertise and infrastructure offered at the University
 - increased researcher profile through data dissemination and re-use.

A data planning process is particularly important in the context of collaborative research projects. Researchers may identify areas of potential difficulty or conflict, and these can be resolved with colleagues and collaborators before they escalate into issues. Clarifying ownership of data, and ensuring early agreement on technical standards and frameworks across institutions, are an important part of establishing trust and ensuring that a project runs smoothly.

DMP USEFUL IN
COLLABORATIVE
PROJECTS TO
AVOID FUTURE
CONFLICTS

Benefits of a DMP

CESSDA Guide



⊖ Benefit 3. Clarifies needed budget

Data management is not free. You do not want to find yourself running out of funding before the end of the project because you have ignored or underestimated the cost of structured, detailed, and safe data management. Therefore, an important aspect of a DMP is its use in calculating how much money will be required for managing your research data during your research project.

A DMP can be useful in the process of applying for funding. Grant applications should not only include time and resources for collecting, analysing, and publishing on data in their budget, time and resources for careful documentation as well as server space, backup solutions, and documentation software need to be included as well. A DMP is also useful once funding is granted to plan and manage your expenses. Many research funders require a DMP as part of the application and decision-making process. The arguments for making data available are several, the most popular being that the data produced by public funds should be used to the greatest extent possible and available to the public. Unless there are legal, ethical or commercial barriers, data should also be openly available so that research results can be verified, replicated and reused.

Examples of Data Management cost assessments are given by the [University of Utrecht](#) (n.d.) and the Dutch Landelijk Coördinatiepunt Research Data Management ([LCRDM](#), n.d.) inspired by the '[Data management costing tool](#)' by UK Data Service, 2013.

CRUCIAL TO EXSTIMATE THE
COSTS [ALSO TO HIRE A DATA
STEWARD...5%]
FROM THE BEGINNING
[PLANNING!!!]

DMP Core Requirements

CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS



When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:



1. Data description and collection or re-use of existing data

- How will new data be collected or produced and/or how will existing data be re-used?
- What data (for example the kinds, formats, and volumes) will be collected or produced?



2. Documentation and data quality

- What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
- What data quality control measures will be used?



3. Storage and backup during the research process

- How will data and metadata be stored and backed up during the research process?
- How will data security and protection of sensitive data be taken care of during the research?



4. Legal and ethical requirements, codes of conduct

- If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
- How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
- How will possible ethical issues be taken into account, and codes of conduct followed?



5. Data sharing and long-term preservation

- How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
- How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
- What methods or software tools will be needed to access and use the data?
- How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?



6. Data management responsibilities and resources

- Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
- What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?



SCIENCE EUROPE
PRACTICAL GUIDE TO
THE INTERNATIONAL ALIGNMENT OF
RESEARCH DATA MANAGEMENT

Extended Edition
with DMP Evaluation Public

Jan. 27, 2021



SEZIONI MINIME
IN UN DMP

CORE REQUIREMENTS
FOR DATA MANAGEMENT PLANS



DMP – Rubric for [self-]evaluation



3 STORAGE AND BACKUP DURING THE RESEARCH PROCESS

Guidance for Researchers

3a

How will data and metadata be stored and backed up during the research?

- Describe where the data will be stored and backed up during research activities and how often the backup will be performed. It is recommended to store data in least at two separate locations.
- Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution. Storing data on laptops, stand-alone hard drives, or external storage devices such as USB sticks is not recommended.

Sufficiently Addressed The DMP...

- Clearly (even if briefly) describes:
 - › The location where the data and backups will be stored during the research activities.
 - › How often backups will be performed.
 - › The use of robust, managed storage with automatic backup (for example storage provided by the home institution).
- or
- Explains why institutional storage will not be used (and for what part of the data) and describes the (additional) locations, storage media, and procedures that will be used for storing and backing up data during the project.

Insufficiently Addressed The DMP...

- Provides no information or very vague reference to how data will be stored and backed up during the project.

SEE WHAT YOU ARE NOT SUPPOSED TO DO!

Guidance for Researchers

1a

How will new data be collected or produced and/or how will existing data be re-used?

- Explain which methodologies or software will be used if new data are collected or produced.
- State any constraints on re-use of existing data if there are any.
- Explain how data provenance will be documented.
- Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.

Sufficiently Addressed The DMP...

- Gives clear details of where the existing data come from and how new data will be collected or produced. It clearly explains methods and software used.
- Explains, if existing data are re-used, how these data will be accessed and any constraints on their re-use.

Insufficiently Addressed The DMP...

- Provides little or no details on where the data come from and what data will be collected or re-used.
- Does not, if applicable, provide sufficient rationale for generating new data.

Basic Information.

- State the purpose of the data collection/generation.
- Explain the relation to the objectives of the project
- Consider what data will be collected or created as part of the study (RAW data).
- Consider what data will be produced by processing the RAW data (Secondary, processed data).
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- Specify the types and formats you plan to use for the data generated/collected (raw, processed, published).
- Consider what data will be published as the result of your study (Published data).

Volume and Life Cycle of the Data.

If you are using FAIRDOM, we will look after data that will be retained and potentially exchanged by your projects. It will help with local storage for temporarily-held local data prior to processing.

For RAW data, please consider the following:

- How much RAW data you think will be produced (Estimates, per month, year, full project duration)?
- Will all of the RAW data be kept for the duration of the study or will the RAW data be deleted once it is processed?
- For large scale RAW data (images, sequence) have you planned the local storage capacity necessary for processing?
- Do you require help to organise a suitable local management system for RAW data?
- Do you have policies that govern the management and usage of RAW data?
- How long will RAW data be kept? Will there be a long-term archive?

For Secondary and Published data, please consider the following:

- What data processing is foreseen in the project?
- How much processed data will be produced, and stored (can you make estimates per month, year, full project)?
- How much of this data will be published? (Estimates per month, year, full project)?
- Does your institution, or the project funders, have policies governing the access and usage of processed data?

Additional for personally sensitive data (e.g medical data)

- When looking at the data flow through the project, define what data is:
 - aggregated (typically safe to share, if names cannot be recovered)
 - anonymized (name cannot be recovered from the data)
 - pseudonymized (name can be recovered by some)
 - non-anonymized (name linked to data)
- Determine which organisational boundaries have to be traversed by which data.
- Make sure with your "local" data protection officer and ethics commission that the data can be shared with your partners along the flow described with the anonymisation levels as described. Why local? Some laws change across surprising boundaries. E.g. in Germany Universities and other public organisations are subject to another data protection law than enterprises. Why seek advice? In some cases you may be required to be able to recover the name-data-relation, e.g. to enable study participants to "leave" a study.

Data Management Checklist

<https://fair-dom.org/knowledgehub/data-management-checklist/>

JUST AS
«SERVING
SUGGESTIONS»
FOR THE FIRST
TIMES

DMP content

Open Science and Research Data Management

...

IIT-UNIGE training module for PhD fellows
Feb 5-9, 2024

Valentina Pasquale, Research Organization Directorate, Istituto Italiano di Tecnologia
Anna Maria Pastorini, Servizio Sistema Bibliotecario di Ateneo, Ufficio Biblioteca Digitale e Open
Access, Università di Genova

Data summary

CCBY Valentina Pasquale

Dataset name	A descriptive title of your dataset
Dataset goal/utility	Define the main purpose: which need will this dataset meet? Which categories of users and/or researchers may use this dataset in the future?
Data origin	Generated within the project/Already existing
Type/formats	Indicate data type and formats, with preference for interoperable formats (UK Data Service Recommended file formats)
Expected size	MB / GB / TB ?
Expected time of delivery	Estimation of when the dataset will be completed



DMP content

CCBY Valentina Pasquale

Documentation, metadata, and data quality

Documentation

How will documentation about data collection be captured? e.g. 'readme' text files, file headers, code books, or lab notebooks. Which information will be included? e.g. methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, etc.

Metadata standards

Which metadata standards (for example DDI, DataCite, etc.) will be used? Use community metadata standards where these are in place. Find them at [DCC standards webpage](#) or [FAIRsharing.org](#).

Data quality assessment

How will consistency and quality of data collection be controlled and documented?



DMP content

Storage & backup

CCBY Valentina Pasquale

Storage & backup strategies

Where will data be stored and backed up during research? How often will backup be performed? Give preference to the use of robust, managed storage with automatic backup, e.g. provided by institutional IT support service.

Data organization

How will data be organised during the project? e.g. file naming conventions, version control, folder structures, etc. Consistent, well-ordered research data will be easier to find, understand, and re-use.

Data security & protection

Who will have access to the data during research? How will access be controlled? Consider data protection risks, particularly if your data is sensitive, e.g. personal data, trade secrets.

An abstract painting featuring various colorful, stylized buildings or structures in shades of blue, orange, green, and white, set against a dark background.

DMP content

CCBY Valentina Pasquale

Legal & ethics

Protection of personal/sensitive data	Ensure that you are compliant with personal data protection laws (e.g. GDPR) <ul style="list-style-type: none">› Gain informed consent for preservation and/or sharing of personal data› Consider anonymisation for preservation and/or public sharing› Consider pseudonymisation/encryption of personal data
Intellectual property rights	Who will be the owner of the data? Who will have the rights to control access? Make sure to cover these matters for multi-partner projects in the consortium agreement.
Ethical issues	What ethical issues and codes of conduct are there, and how will they be taken into account?



DMP content

Data sharing & preservation

CCBY Valentina Pasquale

Long-term data sharing	How will data be findable and shared (e.g. trustworthy data repository, indexed in a catalogue, use of a secure data service, etc.)
Data repository	Specify where the data will be deposited and demonstrate that the data can be curated effectively beyond the lifetime of the grant
Restrictions on sharing	Are there possible restrictions to data sharing or embargo reasons? Explain any constraint on sharing, such as embargo period or restricted access. Please make sure to justify any potential restriction.
Data curation	How will data for preservation be selected?
Licensing	Which licenses will be applied? E.g. <u>CC licenses</u>

An abstract painting featuring a row of stylized buildings in various colors including brown, blue, green, and orange, set against a dark blue background. The buildings have textured, layered appearances.

DMP content

Costing & resources

CCBY Valentina Pasquale

Roles and responsibilities	Outline the roles and responsibilities for data management/stewardship activities. For collaborative projects, explain the co-ordination of data management responsibilities across partners: indicate who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised.
Costing	What resources (for example financial and time) will be dedicated to data management? <u>Estimate the costs</u> in terms of personnel (e.g. PM needed to collect and maintain data) and other costs (storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges). Describe how you plan to cover these costs both during and beyond project duration.

https://rdmkit.elixir-europe.org/data_management_plan

Data management

Data life cycle ▾

Your role ▾

Your domain ▾

Your tasks ▴

Compliance
monitoring

Costs of data
management

Data analysis

Data brokering

Data management

Your tasks

Data management plan

What should you write in a Data Management Plan (DMP)?

Description

A DMP should address a broad range of data management aspects, regardless of funder or institution specific templates. It is important to be aware of the current best practices in DMPs before starting one. For more generic information about data management planning, see also our [Planning page](#).

Considerations

Common topics of a DMP are:

Tools

Search Type here...

https://rdmkit.elixir-europe.org/your_domain

⌵ Your domain

Bioimaging data

Data management solutions for bioimaging data.

⌵ Your domain

Biomolecular simulation data

Data management solutions for biomolecular simulation data.

⌵ Your domain

Epitranscriptome data

Data management solutions for epitranscriptome data.

⌵ Your domain

Health data

Data management solutions for human health data.

⌵ Your domain

Human data

Data management solutions for human data.

⌵ Your domain

Human pathogen genomics

Data management solutions for Human pathogen genomics

⌵ Your domain

Intrinsically disordered proteins

Data management solutions for

⌵ Your domain

Machine learning

This page is about setting best practices for data towards enabling FAIR ML

⌵ Your domain

Marine metagenomics

Data management solutions for marine metagenomics data.

⌵ Your domain

Microbial biotechnology

Data management solutions for microbial biotechnology data.

⌵ Your domain

Plant sciences

Data management solutions for plant sciences data.

⌵ Your domain

Proteomics

Data management solutions for proteomics data.

⌵ Your domain

Rare disease data

Data management solutions for rare disease data.

⌵ Your domain

Single-cell sequencing

Managing data generated from single-cell sequencing exper

⌵ Your domain

Structural Bioinformatics

Data management solutions for

⌵ Your domain

Toxicology data

Data management solutions for toxicology data.

martedì 4 giugn

RDMkit

Data management

About

Contribute

Data management

Data life cycle



Your role



Your domain



Your domain

In this section, information is organised based on different domains in life science their data. You will find:

- Domain-specific best practices and guidelines for data management.

Tools

Your role


https://rdmkit.elixir-europe.org/your_role

In this section, information is organised based on the different roles a professional can have in research data management. You will find:

- A description of the main tasks usually handled by each role.
- A collection of research data management responsibilities for each role.
- Links to RDMkit guidelines and advice on useful information for getting started with data management specific to each role.


Search Type here...



 Your role


Data Steward

Data management guidance for data stewards.

 Your role


Policy maker

Data management guidance for policy makers.

 Your role


Principal Investigator (PI)

Data management guidance for Principal Investigator (PI).

 Your role

Research Software Engineer (RSE)

Data management guidance for Research Software Engineers (RSEs).

 Your role

Researcher

Data management guidance for researchers, including how to make your data FAIR.

 Your role

Trainer

Data management guidance for trainers.

Tools

https://rdmkit.elixir-europe.org/your_tasks

Data management

Data life cycle



Your role



Your domain



Your tasks



Compliance
monitoring

Costs of data

Your tasks


In this section, information is organised around regular research data management tasks or challenges. You will find

- Best practices and guidelines for each data management task.
- A list of all the considerations to be taken into account when performing a specific data management task.
- Links to task-specific training materials.
- Links to tool assemblies implemented by others to address specific data management challenges.
- Links to a Data Stewardship Wizard for your DMP and to step-by-step instructions to make your data FAIR.
- A summary table of tools and resources relevant for the specific task and recommended by communities.

 Your tasks


Compliance monitoring & measurement

How to measure compliance to data management regulations and standards.

 Your tasks


Costs of data management

Budgeting and costing for data management

 Your tasks


Data analysis

How to make data analysis FAIR.

 Your tasks


Data brokering

Information on brokering data to data repositories on behalf of data producers.

 Your tasks


Data management plan

How to write a Data Management Plan (DMP).

 Your tasks


Data organisation

Best practices to name and organise research data.

 Your tasks


Data provenance

How to record information about data provenance.

 Your tasks

Data publication

How to prepare data and find repositories for publication.

 Your tasks

Data quality

How to ensure high quality of research data.

Data management

Data life cycle

Your role

Your domain

Your tasks

Tool assembly

COVID-19 Data Portal

CSC

FAIRtracks

Galaxy

IFB

https://rdmkit.elixir-europe.org/tool_assembly

Tool assembly

Tool Assemblies are examples of combining tools to cover data management tasks across several stages of the data life cycle. These can be tools that one or several communities combine to support RDM that can be picked up or accessed and used by others. The assemblies are aimed for users in a specific location and/or for users within a specific domain.

Filter by affiliation

Choose...

Search

Type here...

Tool assembly

COVID-19 Data Portal

The COVID-19 Data Portal brings together relevant datasets for sharing and analysis to accelerate coronavirus research.

Tool assembly

Affiliations:

Affiliations:



Tool assembly

CSC

The Center of Science (CSC) provides high-quality ICT expert services for researchers in Finland and their collaborators.

Tool assembly

Affiliations:

The French Bioinformatics Institute (IFB) offers IT infrastructure and bioinformatics expertise to support researchers in Life Sciences.

Affiliations:



Tool assembly

FAIRtracks

The FAIRtracks ecosystem provides technical solutions for the FAIRification of genome browser track files.

Tool assembly

Affiliations:

Molgenis is a modular web application for scientific data. Flexible data integration platform to find, capture, exchange, manage and analyse scientific data.

Affiliations:



Tool assembly

Marine Metagenomics

The Marine Metagenomics tool assembly aims to provide a comprehensive data management toolkit of marine genomics researchers in Norway.

Affiliations:



Tool assembly

NeLS

NeLS provides the necessary tools for data management to researchers in Norway and their collaborators.

Affiliations:



Tool assembly

OMERO

OMERO is a software platform for managing, sharing and analysing images data.

Affiliations:



Tool assembly

Plant Genomics

Tool assembly for managing plant genomic data.

Tool assembly

Plant Phenomics

Tool assembly for managing plant phenomic data.

Tool assembly

TSD

The Sensitive Data Service (TSD) provides a platform to store, compute and analyse



DMP in Horizon Europe

- IN HORIZON EUROPE
- **IN THE PROPOSAL**: 1 PAGE ON RESEARCH OUTPUT MANAGEMENT
- DMP DUE AS DELIVERABLE M6



V.1 Feb. 2021

Horizon Europe (HORIZON)
Euratom Research and Training Programme
(EURATOM)

General Model Grant Agreement
EIC Accelerator Contract

(HE MGA — Multi & Mono)



V.2 April 2021



Horizon Europe Programme
Standard Application Form (RIA, IA)

Application form (Part A)
Project proposal – Technical description (Part B)
Version 2.0
22 April 2021

⚠ *Proposals selected for funding under Horizon Europe will need to develop a detailed data management plan (DMP) for making their data/research outputs findable, accessible, interoperable and reusable (FAIR) as a deliverable by month 6 and revised towards the end of a project's lifetime.*

⚠ *For guidance on open science practices and research data management, please refer to the relevant section of the [HE Programme Guide](#) on the Funding & Tenders Portal.*

Open science: research data management

The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- establish a data management plan ('DMP') (and regularly update it)
- as soon as possible and within the deadlines set out in the DMP, deposit the data in a trusted repository; if required in the call conditions, this repository must be federated in the EOSC in compliance with EOSC requirements

DMP in Horizon Europe



HEU DMP

Horizon Europe

Data Management Plan Template

Version 1.0
05 May 2021

Data Summary

- Will you re-use any existing data and what will you re-use it for?
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

FAIR data

- 2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?
- 2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- 2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- 2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that it can be harvested and indexed?
- 2.2. Making data accessible - Repository: Will the data be deposited in a trusted repository?
- 2.2. Making data accessible - Repository: Have you explored appropriate arrangements with the identified repository where your data will be deposited?

- 2.2. Making data accessible - Repository: Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?
- 2.2. Making data accessible - Data:
Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
- 2.2. Making data accessible - Data:
If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- 2.2. Making data accessible - Data:
Will the data be accessible through a free and standardized access protocol?
- 2.2. Making data accessible - Data:
If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
- 2.2. Making data accessible - Data:
How will the identity of the person accessing the data be ascertained?
- 2.2. Making data accessible - Data:
Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?
- 2.2. Making data accessible - Metadata:
Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why.
Will metadata contain information to enable the user to access the data?

**TEMPLATE AVAILABLE
IN DMPONLINE,
ARGOS AND DSW
DUE BY MONTH 6**

Let's go working!



AGATHOCLES DMP online

Project Details Contributors Plan overview Initial DMP Detailed DMP Final review DMP Share Download

expand all | collapse all

8/9 answered

FREE TEXT. YOU NEED TO KNOW WHAT TO PUT IN

1. Data summary (1 / 1)

2. FAIR data (3 / 4)

3. Allocation of resources (1 / 1)

4. Data security (1 / 1)

5. Ethical aspects (1 / 1)

DS Wizard

Knowledge Models

Projects

WIZARD TO FILL IN THE FORM. IN THE END, YOU WRITE NOTHING AND THE SYSTEM EXTRACTS THE RELVANT INFORMATION

Leiden Booksellers - Giglia IFDS homework week 5

Questionnaire Metrics Preview Documents Settings

View

Comments

TODOS

Version history

Current Phase

Before Submitting the Proposal

III. Creating and collecting data

We will make sure that we know what data will be coming together in the project, when it will be coming. We also need to make sure that we have adequate storage space to deal with it, and that all the responsibilities have been taken care of.

Chapters

I. Administrative information

II. Re-using data

III. Creating and collecting data

IV. Processing data

V. Interpreting data

VI. Preserving data

1 What existing data formats/types will you be using?

Horizon 2020 DMP Science Europe DMP

Have you identified types of data that you will use that are used by others too? Some types of data (for example "images" or "tables") are used by many different projects. For such data, often common standards exist (in our example "JPG" and "CSV" [comma separated values]) that help to make these data reusable. Are you using such common data formats?

Please make sure you list all the data types that are important for your project. You should make sure also to list the formats used in any data sets that you are re-using.

☒ Desirable: Before Submitting the Proposal

Science: [n/a](#)

ABOUT

RESOURCES

CONTACT

LOG IN



Argos

Plan and follow your data

Create machine actionable DMPs.

Configure to best fit your discipline.

Link to EOSC components out of the box.

Share easily in your repository.

Bring your Data Management Plans closer to where data are generated, analysed and stored.

Start your DMP

WIZARD+FREE TEXT
[SOMETIMES LIMITED OPTIONS e.g.
«Research data» as type]

Comparative table



The logo for DMP ONLINE, with 'DMP' in white and 'ONLINE' in orange, next to a stylized orange 'F' icon.

- TEMPLATE HEU AND SCIENCE EUROPE
- CUSTOMIZABLE TEMPLATES

- FREE TEXT WITH FORMATTING OPTIONS (TABLES, BULLET POINTS...)

- CONTEXT-SENSITIVE GUIDANCE

- COLLABORATIVE WRITING (ASYNCRONUS!!!)

- TEMPLATE HEU AND SCIENCE EUROPE
- KNOWLEDGE MODEL (ELIXIR)

- WIZARD – VERY DETAILED (DIFFICULT TO FORGET ANYTHING)

- LINK TO BAREND MONS' BOOKS CHAPTERS (DO/DON'T)

- COLLABORATIVE WRITING

- CREATE A TO DO LIST

- MACHINE ACTIONABLE



- TEMPLATE HORIZON EUROPE AND SCIENCE EUROPE

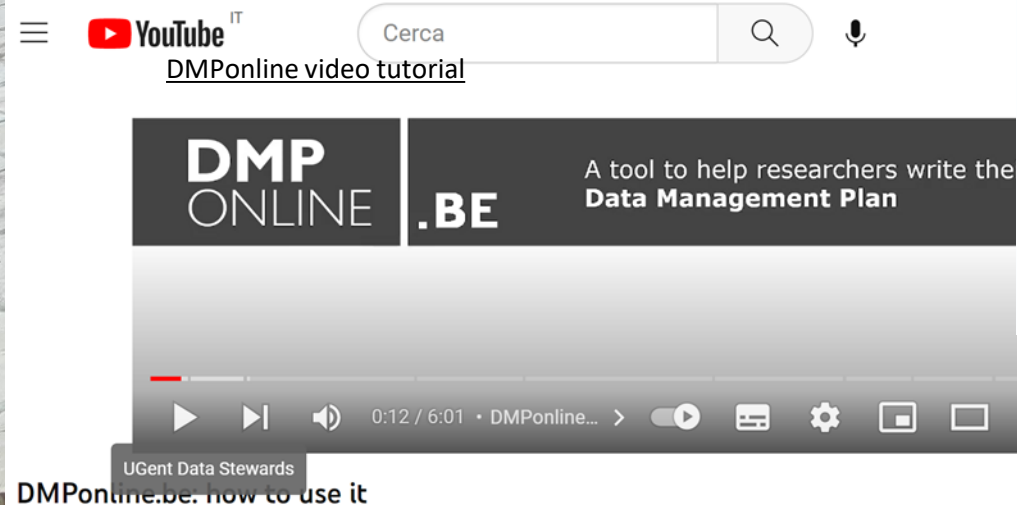
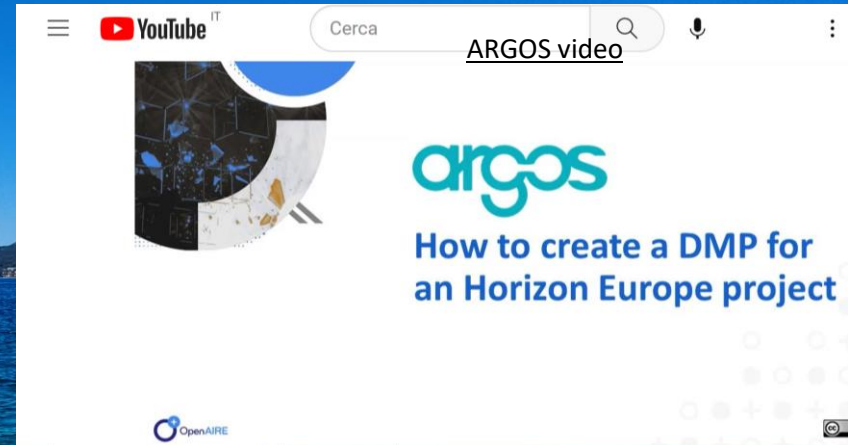
- POSSIBLE ASSOCIATION OF MANY DIFFERENT DATASETS TO THE SAME DMP

- MACHINE ACTIONABLE

- DIRECT CONNECTION TO OPENAIRE RESEARCH GRAPH (LINK DATA+PUBLICATIONS)

...you are not alone...

VIDEO TUTORIALS



Video Tutorials



Get Started

This video shows how to easily and quickly get started doing Data Management Planning with the FAIR Wizard.

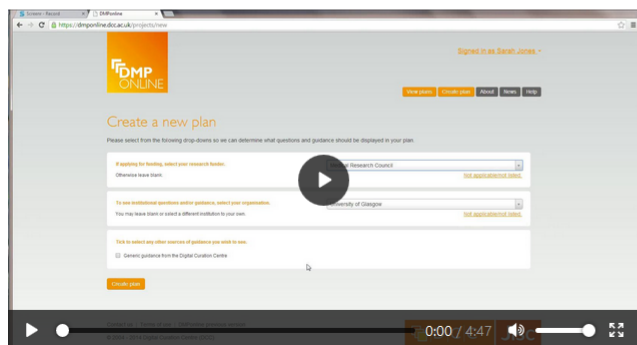
DMP online

[Home](#)[About](#)[Future plans](#)[Help](#)[Change language](#)

Welcome.

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It has been jointly developed by the Digital Curation Centre (DCC) and the University of California Curation Center (UC3).

Screencast on how to use DMPonline



FREE, YOU JUST HAVE TO REGISTER. THEN, ONCE SIGNED IN, ALL YOUR DMPS ARE DISPLAYED IN YOUR DASHBOARD

Sign in

Veteran tapes

[Project Details](#)[Plan overview](#)[Write Plan](#)[Share](#)[Download](#)[expand all](#) | [collapse all](#)[13/13 answered](#)

Data Collection (2 / 2)

What data will you collect or create?

B *I*

The "Veteran tape " project will collect and generate different types of datasets:

Type of data	Volume	Format	Storage format
Video recordings	600 x 1Gb	.mkv	.mkv
Transcriptions	600 x 1500Kb	MS Word	.txt
Structured interview text	1 x 500Kb	MS word	.txt

For the video recordings the selected format is .mkv; the same .mkv format will be used for the long-term preservation .

Transcriptions will be written in MS Word and then stored as .txt files.

We checked the format compatibility against EASY File format
<https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats>

As the total volume of data is greater than 50Gb, DANS requires a fee for the storage. We are currently in touch with EASY to determine the costs of archiving.

[Save](#)[Guidance](#)[Comments \(1\)](#)[DCC](#)[DCC guidance](#)

Guidance

Questions to consider:

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can reuse?

Guidance:

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

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

- or - ☐ No research organisation associated with this plan or my research organisation is not listed

* Select the primary funding organisation

☒ European Commission (Horizon 2020)☐ European Research Council (ERC)

- or - ☐ No funder associated with this plan or my funder is not listed

FUNDERS**DOWNLOAD THE
FINAL PDF**

Prova

[Project Details](#) [Plan overview](#) [Initial DMP](#) [Detailed DMP](#) [Final review DMP](#) [Share](#) [Download](#)

* Project title

☒ mock project for testing, practice, or educational purposes

Funder

Grant number

Project abstract

Briefly summarise your research project to help others understand the purposes for which the data are being collected or created.

**SHARE TO
COLLABORATIVE
WRITING**[See the full list](#)[Save](#)**GUIDE****IDENTIFIERS**

ID

Principal Investigator

Name

ORCID iD

Email

Phone

Data Contact Person

☒ Same as Principal Investigator[Save](#)

Data Summary

- Will you re-use any existing data and what will you re-use it for?
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

NOT «A HUGE
AMOUNT OF
DATA»

COSTS

Data volume

- Note what volume of data you will create in MB/GB/TB. Indicate the proportions of raw data, processed data, and other secondary outputs (e.g., reports).
- Consider the implications of data volumes in terms of storage, access and preservation. Do you need to include additional costs?
- Consider whether the scale of the data will pose challenges when sharing or transferring data between sites; if so, how will you address these challenges?

USE A TABLE FOR
DIFFERENT DATA
FORMATS

Data format

- Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif, .tiff).
- Explain why you have chosen certain formats. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centres or widespread usage within a given community.
- Using standardised, interchangeable or open formats ensures the long-term usability of data; these are recommended for sharing and archiving.
- See UK Data Service guidance on [recommended formats](#) or DataONE Best Practices for [file formats](#).

YOU DIRECTLY FILL THE
FORM IN (USE BULLET
POINTS AND TABLES)

YOU CAN SAVE AND GO BACK
LATER IN ANY STEP

USE STANDARD
FORMATS (LINK TO
THE RDA
DIRECTORY)

PREVIEW OF ALL THE QUESTIONS

FAIR data

- 2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?
- 2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- 2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to help others identify and discover the data?
- 2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that

- What metadata will be provided to help others identify and discover the data?
- Researchers are strongly encouraged to use community metadata standards where these are in place. The Research Data

Alliance offers a [Directory of Metadata Standards](#). Data

repositories may also provide guidance about appropriate metadata standards.

- Consider what other documentation is needed to enable reuse. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data and software used to collect and/or process the data.
- Consider how you will capture this information and where it will be recorded, e.g., in a database with links to each item, in a 'readme' text file, in file

FAIR data (0 / 25)

2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?

Guidance

DCC

Metadata & documentation

SPECIFIC GUIDES

Save

2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Guidance

DCC

Metadata & documentation

Allocation of resources (0 / 4)

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?

B I [list icon] [list icon] [link icon] [table icon]

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

B I [list icon] [list icon] [link icon] [table icon]

General > Article 6.2.C.3 Other goods, works and services

Costs related to research output management (HE) – Costs for research output management (e.g. *management of research data*) are eligible if the eligibility conditions are fulfilled, including open access to peer-reviewed publications (but see the additional eligibility condition referenced immediately below), research data and other outputs.

COSTS FOR DATA
MANAGEMENT ARE ELIGIBLE
(GRANT 6.2.C.3)

WHO IS RESPONSIBLE?

Who will be responsible for data management in your project?

B I [list icon] [list icon] [link icon] [table icon]

Roles & responsibilities

- Outline the roles and responsibilities for all activities, e.g., data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Individuals should be named where possible.
- For collaborative projects you should explain the coordination of data management responsibilities across partners.
- See UK Data Service guidance on [data management roles and responsibilities](#) or DataONE Best Practices: [Define roles and assign responsibilities for data management](#).

Perugia Prova

Project Details

Contributors

Plan overview

Write Plan

Share

Download

Format

pdf

csv

html

pdf

text

docx

json

☐ project details coversheet

☒ question text and section heading

☒ unanswered questions

PDF formatting

Font

Face

Arial, Helvetica, Sans-Serif

Size (pt)

10

Margin (mm)

Top

25

Download Plan

DIFFERENT FORMATS
FOR YOUR FINAL
DOWNLOAD

1. Data summary

Provide a summary of the data addressing the following issues:

- State the purpose of the data collection/generation
- Explain the relation to the objectives of the project
- Specify the types and formats of data generated/collected
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- State the expected size of the data (if known)
- Outline the data utility: to whom will it be useful

Purpose: _____

Relation: _____

Data will be created as the research is original.



Reuse: videos watched on YouTube, chats coming from social platforms

Origin: interviews and observations; content analysis




Type of data	Format (ongoing)	Format (preservation)	Expected Size	Consent collected	Reuse Rights
A Textual	.doc	.txt	10 M		
A1 [Interview template]					
A2 [Consent form]					
A3 [Interview transcriptions]				x	
A4 [Code book]	.txt				
A5 [Report]					
A6 [Video transcriptions]					CC BY
B Tabular data	.xls	.csv	50 M		
B1 [Participants description]					
B2 [Thematic analysis]					
B3 [Content analysis]					
C Audio			10 G		
C1 [Recording]	.mp3	still to be decided		x	
D Video			100 G		
D1 [Recorded interviews]	.mp4			x	
D2 [Downloaded]					CC BY


FAIR wizard

FAIR WIZARD FROM 2024
ONLY INSTITUTIONAL
ACCESS (ON A FEE BASIS)

 cuap-unito 

DATA MANAGEMENT PLANNER

-  Dashboard
-  Projects
-  Book a call

 Elena Giglia
Researcher


Create Project

Name

Knowledge Model



Common DSW Knowledge Model 2.6.5

DSW Knowledge Model originating from mindmap made by Rob Hooft 

Question Tags

You can either use all questions from the knowledge model or filter them by question tags.

☐ Use all questions

☒ Filter by question tags

☐ ELSI

☐ Horizon 2020 DMP

☐ Horizon Europe DMP

☐ Science Europe DMP

☐ maDMP

Cancel

Create

FAIR wizard

Prova CUAP Horizon Europe

Questionnaire Metrics Preview

View Import replies

I. Administr...

1

- Contributors
- Research Project(s)
- To execute the D...
- Do you require h...
- Describe nationa...

II. Re-using ...

1

III. Creating ...

5

IV. Processin...

1

V. Interpreti...

✓

VI. Preservin...

2

VII. Giving ac

3

I.3 To execute the DMP, is additional specialist expertise required?

Horizon Europe DMP

☒ Desirable: Before Submitting the DMP

- ☐ a. No
- ☐ b. Yes, trained support staff is available
- ☐ c. Yes, we will be training existing staff
- ☐ d. Yes, we will be hiring new people with additional expertise

I.1 Contrib...

Horizon Europe DMP

Each person contributing to creating or executing the data management plan should probably should have a Contact Person, and a Data Curator.

☒ Desirable: Before Submitting the DMP

+ Add

I.2 Research Project(s)

Horizon Europe DMP

Add each of the research project(s) that you are (or will be) working on and for which the DMP. Give each project a small identifying name for yourself.

☒ Desirable: Before Submitting the Proposal

FAIR wizard

BAREND MONS'
BOOK CHAPTER

Is there pre-existing data?

What's up?

For many decades if not centuries, virtually every experiment started with the collection or creation of 'observations' and created earlier, in all kinds of surveys and increasingly of course from sources such as social media maybe already somewhat new data specifically generated to answer a hypothetical question is still so commonplace that careful thinking about data creation will need to continue, but increasingly we have to ask the question whether such new data are absolutely available in reusable format, there may well be existing data collections 'Other People's' Data and associated Services question or least may be crucial for the interpretation of your own data.

Do

- Search for data sets (OPEDAS) that may be re-usable and can help you to reduce the number of new data sets you may create.
- Include annotated collections of data and curated databases in your search.
- Check the accessibility and license situation attached to the relevant data sets you found.
- Check their interoperability. They may be relevant but not interoperable with your analysis pipelines. In that case you may need to transform the data so they are not reusable for your purpose.
- Ensure that using OPEDAS will not restrict in any way the use of your results later on, including copyright and freedom of information.
- Check how to cite and acknowledge OPEDAS.
- Consider to actively involve OPEDAS owners in your research in order to make optimal use of their data.
- Speak to colleagues who did similar experiments before to find out about potential OPEDAS you may consider to use.

Don't

- Assume no OPEDAS exist without thorough checking using all your possibilities.
- Start an experiment without properly checking with colleagues about the best approach and OPEDAS out there.
- budget for data generation in your study without justifying to the funder why the generation of the data is necessary.
- Move into actual experimentation without consulting a data expert.

Chapters

I. Administ... 1

II. Re-using ... 3

Is there any pre-existing data?

WIZARD
(DIFFERENT
PATHS)

V. Interpreti... ✓

VI. Preservin... 2

VII. Giving ac... 3

✓ II.1 Is there any pre-existing data?

Horizon Europe DMP

Are there any data sets available in the world that you can re-use?

☒ Desirable: Before Submitting the Proposal

☒ Data Stewardship for Open Science: atq

☒ External links: [Google dataset search](#), [Datacite Search](#), [RDMkit on Reusing Data](#), [RDMkit on Existing Data](#)

☐ a. No

☒ b. Yes

Clear answer

Answered less than a minute ago by Elena Giglia.

LINK TO
USEFUL TOOLS

⌚ II.1.b.1

What existing openly available standard reference data did you consider re-using?

FAIR wizard

Data management

Data life cycle



Your role



Your domain



Your tasks



Compliance monitoring

Costs of data management

Data analysis

Data brokering

Data management

Your tasks

Data organisation



What is the best way to name a file?

Description

Brief and descriptive file names are important in keeping your data files organised. A name is the principal identifier for a file and a good name gives information what the file contains and helps in sorting them, but only if you have been consistent with the naming convention.

Considerations

- Best practice is to develop a file naming convention with elements that are important to your project already when the project starts.
- When working in collaboration with others, it is important to follow the same file naming convention.

Current Phase

Before Submitting the P...



III.2

How will you do file naming and file organization?

Horizon Europe DMP

Putting some thoughts into file naming can save a lot of trouble later.

☒ Desirable: *Before Submitting the DMP*

☒ External links: [RDMkit on data organisation](#)

☐ a. Explore



Chapters

I. Administrative inf... 1

II. Re-using data 3

III. Creating and colle... 5

III.3.a.1 Data format/type

Horizon Europe DMP

☒ Desirable: Before Submitting the Proposal

STANDARD DATA TYPE

a systems biology model.

ENCODE peak information Format type model and format

The ENCODE peak information Format is used to provide called regions of signal enrichment based on pooled

Biological Imaging methods Ontology (FBbi) type terminology

A structured controlled vocabulary of sample preparation

Newick tree Format type model and format

The Newick Standard for representing trees in computer-
in 1857 by the famous English mathematician Arthur Cayley

BioModels Ontology type terminology artefact

The BioModels Ontology is an OWL Representation of the
BioPortal, and no official homepage can be found, therefore

III.3.a.3 Does this data format enable sharing and long term archiving?

Horizon Europe DMP

Complicated (binary) file formats tend to change over time, and software may not stay compatible
patents or being hampered by restrictive licensing.

Ideally a format should be simple, text only, completely described, not restricted by copyrights, and

☒ Desirable: Before Submitting the Proposal

☐ a. No

☒ b. Yes Interoperability

Answered less than a minute ago by Elena Giglia.

- LONG TERM?
- DATA VOLUME

- FAIR SCORE

III.3.a.4 What volume of data of this type will you be working with?

Horizon Europe DMP

☒ Desirable: Before Submitting the Proposal

☐ a. So small that it is not a problem

☐ b. I can specify the total amount

☐ c. I can specify the number of files/subjects and the size of each

FAIR wizard

VII. Giving access to data

This chapter deals with the information needed by people who will re-use your data, and

THE SYSTEM ASKS YOU THE
QUESTIONS A DATA STEWARD
WOULD ASK...

✓ VII.1 Will you be working with the philosophy 'as open as possible' for your data?

Horizon Europe DMP

The FAIR principles do not contain any direction toward intellectual property protection.

The true goal of funding agencies is to create the maximum

☒ Desirable: *Before Submitting the Proposal*

☒ Data Stewardship for Open Science: [jvm](#)

☐ a. No Openness

☒ b. Yes Openness

↺ Clear answer

Answered less than a minute ago by Elena Giglia.

✓ VII.2 Can all of your data become completely open over time?

Horizon Europe DMP

Some data may be subject to a temporary embargo, or need to stay closed for specific reasons.

☒ Desirable: *Before Submitting the Proposal*

☒ a. No Openness

☐ b. Yes Openness

↺ Clear answer

Answered less than a minute ago by Elena Giglia.

✎ VII.2.a.1 Are there legal reasons why (some of your) data can not be completely open?


Horizon Europe DMP






☒ Desirable: *Before Submitting the Proposal*

☐ a. No

☐ b. Yes ⋮

FAIR wizard

Prova CUAP Horizon Europe 

 Questionnaire  Metrics  Preview  Documents  Settings

Summary Report

Answered (current phase): 6/24 
Answered: 6/42 

Metrics

Measure

Interoperability

1.00

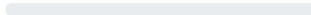
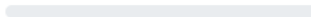


Openness

0.50



Administrative information

Answered (current phase): 0/1 
Answered: 0/5 

Re-using data

Answered (current phase): 1/4 
Answered: 1/5 

Creating and collecting data

Answered (current phase): 3/10 

IN THE END
YOU GET A
FAIR SCORE

FAIR wizard

DS Wizard

Knowledge Models

LABORATORIO PROVA

Questionnaire

Metrics

Preview

Documents

Settings

Section A: Data Collection

1. What data will you collect or create?

Re-used datasets

We will use the following reference datasets:

- **databae COVID del Ministero salute**
(<https://www.salute.gov.it/portale/nuovocoronavirus/dettaglioPubblicazioniNuovaLingua-italiano&id=3147>)

We will use version "bollettino 9/12/2021 (ver 2.2)" of this dataset. If a new version becomes available during the project, we will stay with the old version.

Data formats and types

We will be using the following data formats and types:

- **tabellari**

It is a standardized format. This is not a suitable format for long-term archiving; however, we plan to convert it to a suitable format before the end of the project. We expect to have 30 GB of data in this format.

- **testuali**

2. How will the data be collected or created?

There will be no instrument dataset in this project.

Storage and file conventions

We will use a filesystem with files and folders with the following folder

Data Management Plan

LABORATORIO PROVA

Contact person: There are no contact people specified yet

Based on: Common DSW Knowledge Model, 2.3.0 (dswwork:2.3.0)

Project phase: Before Submitting the Proposal

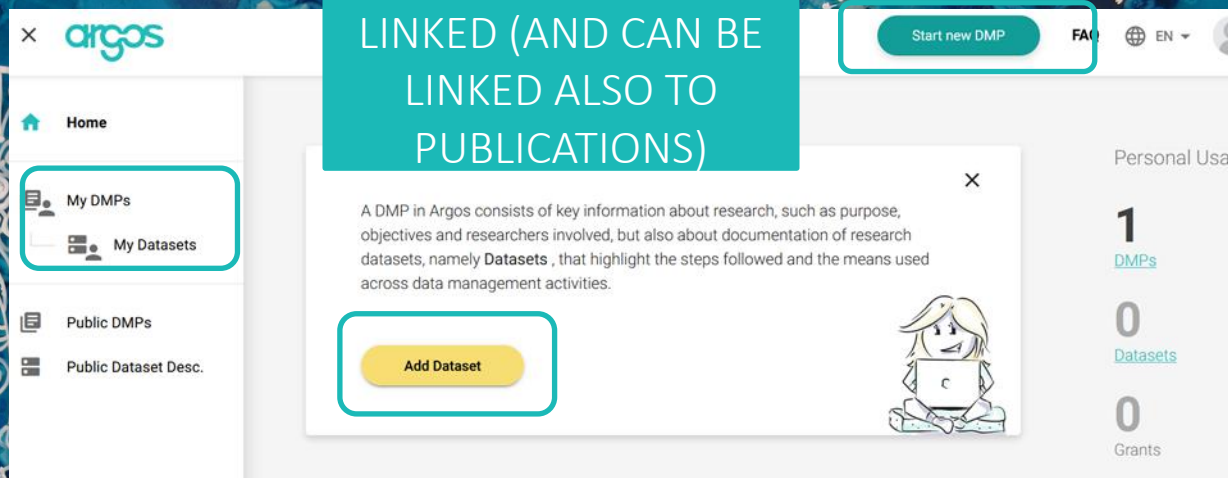
Created by: Elena Giglio (elena.giglio@unibo.it)

Generated on: 21 Dec 2023

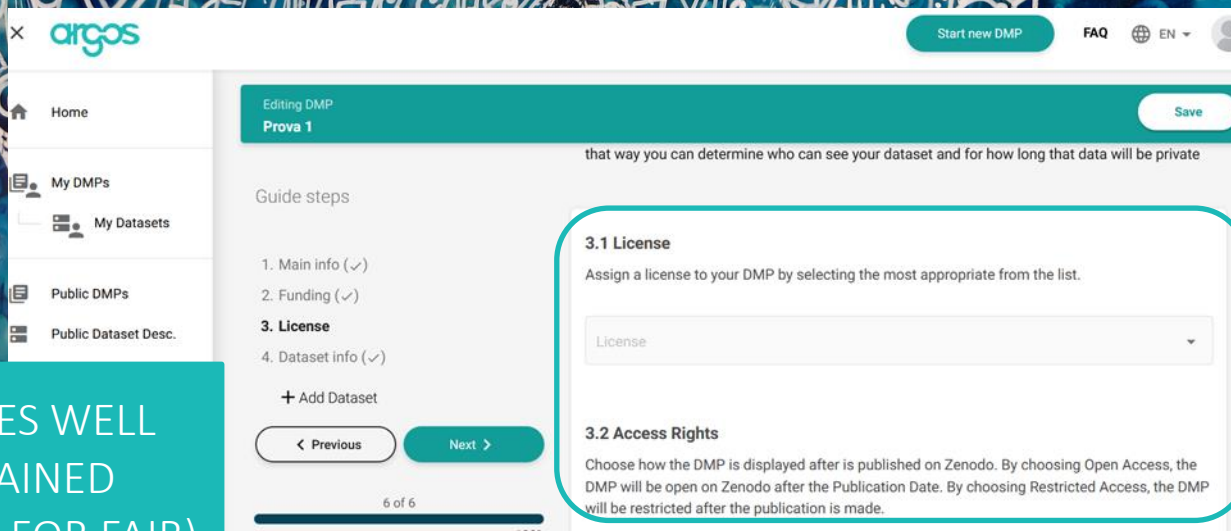
IN THE END YOU CAN DOWNLOAD A COMPLETE DOCUMENT... WITHOUT WRITING A SINGLE WORD. THE SYSTEM EXTRACTS THE RELEVANT INFORMATION FROM YOUR WIZARD'S CHOICES

ARGOS

DMP AND DATASET
LINKED (AND CAN BE
LINKED ALSO TO
PUBLICATIONS)



LICENSES WELL
EXPLAINED
(CRUCIAL FOR FAIR)



ARGOS

UNIBO CUSTOMIZATION

<https://www.openaire.eu/argos-community-calls>

CRUCIAL FOR FAIR
PRINCIPLES

The screenshot displays the 'Editing DMP' interface for 'Prova 1'. The top bar is teal with the title 'Editing DMP' and a 'Save' button. Below the title, a teal box contains the text: 'that way you can determine who can see your dataset and for how long that data will be private'. The main content area is divided into two columns. The left column, titled 'Guide steps', lists four steps: '1. Main info (✓)', '2. Funding (✓)', '3. License', and '4. Dataset info (✓)'. Below the list is a '+ Add Dataset' button and navigation buttons for '< Previous' and 'Next >'. A progress indicator at the bottom shows '6 of 6'. The right column, titled '3.1 License', contains the instruction 'Assign a license to your DMP by selecting the most appropriate from the list.' and a dropdown menu labeled 'License'. Below this, the section '3.2 Access Rights' provides instructions: 'Choose how the DMP is displayed after is published on Zenodo. By choosing Open Access, the DMP will be open on Zenodo after the Publication Date. By choosing Restricted Access, the DMP will be restricted after the publication is made.'

ARGOS

WIZARD WITH
DIFFERENT PATHS

Start new DMP

FAQ

EN

Home

My DMPs

My Datasets

Public DMPs

Public Dataset Desc.

About

Terms Of Service

Adding dataset
PROVA (unsaved changes)

To DMP: Prova 1

Discard

Save

Finalize

Select from the list of options

Guide steps

0. Main info (✓)

< Previous

Next >

3 of 19

16%

1.1.2 Is it physical or digital?

☐ Physical ☐ Digital ☐ Other

1.1.3 Are you generating or re-using it?

Apart from data produced within your research, you might want to repurpose data that have been produced and shared by others, in different research context.

If you are planning to use other researchers' data in your research there are a couple of things that you need to take into consideration, such as copyright of datasets, permissions provided through licenses, ethical aspects for reuse, etc. If you are unsure of the license of a dataset that you are re-using, you may use the [License Clearance Tool](#).

...WITH MANUAL
INTEGRATIONS IN
FREE TEXT

Start new DMP

FAQ

EN

Home

My DMPs

My Datasets

Public DMPs

Public Dataset Desc.

About

Terms Of Service

Glossary

User Guide

Editing Dataset
PROVA (unsaved changes)

To DMP: Prova 1

Discard

Save

Finalize

Back to DMP

Guide steps

< Previous

Next >

6 of 21

29%

1.1.3 Are you generating or re-using it?

Apart from data produced within your research, you might want to repurpose data that have been produced and shared by others, in different research context.

If you are planning to use other researchers' data in your research there are a couple of things that you need to take into consideration, such as copyright of datasets, permissions provided through licenses, ethical aspects for reuse, etc. If you are unsure of the license of a dataset that you are re-using, you may use the [License Clearance Tool](#).

Learn more: <https://www.openaire.eu/can-i-reuse-someone-else-data>

☐ New ☒ Re-used

SOMETIMES A BIT
GENERIC/STIFF

Research Data

Software

Workflows

Protocols

Models

1.1.1 What kind of research output are you describing?

Research data is information (particularly facts or numbers) collected to be examined and considered, and to serve as a basis for reasoning, discussion or calculation. Open access to research data - the right to access and reuse digital research data under the terms and conditions set out in the Grant Agreement.

Other research outputs refer to outputs that are produced or reused during the research and data management lifecycle. They might be objects, instruments and materials in digital form, such as software, workflows, protocols, models, etc. Equally, in their analog form they might include new materials, antibodies, reagents, samples, etc.

ARGOS

ATTENTION! NO WARNING ON THE «FAIRNESS» LEVEL. I CAN SIMPLY PUT «NO» WITHOUT A WARNING TELLING ME THAT THE DATA WILL BE NOT FAIR (AND BASICALLY USELESS!)

Editing Dataset
PROVA (unsaved changes)

To DMP: Prova 1

Discard Save

3.1.1.2 Will you provide metadata for the described dataset / output?

Metadata is data about data and is an essential set of information describing scientific outputs, in the form of either physical or digital objects, in a machine-readable format. According to the expected use, metadata can be given different attributes. Most common type which enables discovery and identification are descriptive metadata. Descriptive metadata contain information about key aspects needed to search for and successfully find a given scientific output, e.g. by its title, author/creator, abstract, keywords. Moreover, metadata may be used for describing a service or a scientific instrument.

☐ Yes ☒ No Required

Provide more details for your answer

argos

Start new DMP FAQ EN

Home My DMPs My Datasets Public DMPs Public Dataset Desc.

About Terms Of Service Glossary User Guide Contact Support

Editing Dataset
PROVA

To DMP: Prova 1

Save Finalize

Guide steps

6 of 21
29%

2.1.1 Does the described output support any scientific publication?

Create the links between scientific publications and research data.

☒ Yes ☐ No

The Multipillar System for Health Care Financing: Thirteen €

List of values provided by external source(s)

Discrete Mathematics & Theoretical Computer Science

List of values provided by external source(s)

ZENODO

List of values provided by external source(s)

YOU CAN LINK DATASETS AND PUBLICATIONS (USEFUL FOR OPENAIRE RESEARCH GRAPH)

ARGOS

WARNING: IF YOU
«FINALIZE», THE DMP
GOES STRAIGHT TO
ZENODO!

Back

Description

PROVA

Owner . Edited : 11 January 2023



Part of

Prova 1



Grant

**Preparing open access in the european research area
through scholarly communication**



FINALIZE



EXPORT

Description authors



Elena Giglia (you)
Owner



Invite



Home



My Plans



My
Descriptions



Public Plans



Public Descriptions

About

Terms Of Service

ARGOS

ULTIMATE - Data Management Plan Version 0

Funder
European Commission|EC

Grant
ULTIMATE: indUstry
water-utiLiTy symbiosis
for a sMarter wATer
society/ No 869318

Researchers
Aitor Corchero (orcid:0000-0002-8463-4128), Joep van
den Broeke (orcid:0000-0002-5707-740X)

IN THE FINAL OUTPUT YOU
CAN SEE AUTOMATICALLY
EXTRACTED TEXT AND
MANUALLY INPUT TEXT (-)

Datasets

Title: Case Studies Dataset

Template: Horizon 2020

The Case Studies in ULTIMATE are collecting experimental data from laboratory experiments and pilot scale water treatment installations. This dataset contain the public data collected as part of the Case Studies.

Dataset Description

1.1 Data Summary

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

- To keep on record
- To develop a product

1.1.2 What are the types of the described generated/collected data?

- sample or specimen data
- observational (e.g.
- sensor data
- data from surveys)
- experimental (e.g.
- gene sequencing data)
- simulation (e.g.
- climate modeling data)

1.1.3 What are the formats of the described generated/collected data?

- Text files
- Numerical
- Models

1.1.3 What are the formats of the described generated/collected data?

- Text files
- Numerical

Batch datasets in CSV and JSON formats

1.1.4 What is the origin of the described data?

Primary data

1.1.5 What is the expected size of the described data?

MB (megabyte)

The size of data is around 100 to 500 MB including electricity, gas and meteorological datasets

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

- Researchers
- Research communities

INTERESTING
SECTION ON REUSE
POTENTIAL

your
best

ENJOY!!!!