

# Université de Montpellier: University of Montpellier - DMP template (english)

## 1. PROJECT INFORMATION

1.1 Project members and contact details

1.2 Project dates, duration and timetable (estimated or defined)

1.3. Theme of the call for projects

*Exemple de réponse:*

Life Sciences domain (HAL nomenclature).

*Recommandations:*

- Use keywords.
- Specify the nomenclature used.

## 2. DESCRIPTION OF THE PROJECT DATA

2.1 Presentation of the generated and collected data

Which data (types, formats and volumes for example) will be collected or produced during the project?

- Nature, type, scale
- Methods of data production and processing: collection of new data or reuse of data (data conversion/transformation; data sharing/exchange; data purchase)
- Estimated volume of data
- Period of data collection
- Other data and associated publications
- Quality control

*Recommandations:*

- Indicate the software and methodologies used for the collection of the new data.
- Describe the type of data produced (database, table, text, image, audio, video...).
- There are open format standards for data ([https://doranum.fr/wp-content/uploads/FS2\\_liste\\_indicative\\_formats\\_V1.pdf](https://doranum.fr/wp-content/uploads/FS2_liste_indicative_formats_V1.pdf)).

*Exemple de réponse:*

Types of data: experimental data, observational data, survey data, textual data, genomic data, samples, images, audiovisual data, models, etc.

Detail the format of the data: the way in which the data are coded for storage, usually reflected by the file name extension (e.g. pdf, xls, doc, txt, or rdf).

Data processing tools: Excel, Stata, SPSS, Matlab v8 with Signal Processing toolbox, etc.

Indicate the projected volume.

2.2 Reuse of existing data

Are the data used in the project based on the reuse of existing data? If so, cite their origin.

*Recommandations:*

- State any restrictions on the reuse of pre-existing data. If applicable, briefly state the reasons why the use of existing data sources was considered but rejected.
- If material protected by specific rights is used during the project, specify any constraints on access and/or reuse.
- Will you combine your new data with these old datasets? If necessary, specify the terms of reuse and the means used to comply with these terms.

*Exemple de réponse:*

FAOSTAT statistical data on sorghum production, free to use, Land Cover Map 2000 data, Centre for Ecology and Hydrology, Forest Research has licence agreement, etc.

2.3 Which data quality control measures will be implemented?

*Recommandations:*

- Explain how the quality and conformity of data collection will be controlled and documented.
- Use the data documented in the laboratory notebook for quality purposes rather than for certification (time of survey, calibration, repeat samples or measurements, standardised data capture, data entry validation, description of experimental processes...).

*Exemple de réponse:*

Copy out the data referenced in the laboratory notebook

## 3. DESCRIPTION OF THE PROJECT'S DATA AND METADATA

3.1 Which metadata (e.g. ontologies, taxonomies) and documentation (e.g. data collection methodology and organisation) will accompany the

data?

*Exemple de réponse:*

The metadata are specified by a scientist and is organised according to the fields recommended by the MIAPPE standard provided in OWL: General metadata, Timing and location, Biosource, Environment, Treatments, Experimental design, Sample collection, processing, management, Observed variables.

*Recommandations:*

- Indicate which metadata standards and formats will be used (e.g. DDI, TEI, EML, MARC, CMDI).

3.2 How were the metadata produced (GPS, type of instrument, manual entry etc.)? By whom?

*Recommandations:*

- Indicate which metadata will be provided to assist in searching and identifying data.
- Indicate how the data will be organised during the project (naming conventions, version control and folder structures...).

3.3 Are the data easy to find, freely accessible, interoperable and reusable (FAIR principles)?

*Recommandations:*

- Think about the documentation that would be needed to enable the reuse of the data by other researchers or for commercial purposes.
- This may include information on the methodology used to collect the data, the analysis procedures and methods used, the definition of variables, units of measurement, etc.

## 4. RESPONSIBILITIES, ETHICAL REQUIREMENTS AND INTELLECTUAL PROPERTY RIGHTS

4.1 Which institution will own the intellectual property rights for the data created during the project?

What is the relevant legislation?

*Recommandations:*

- Explain the collaborative project: include all the institutions involved, in which way(s), these are involved, etc.
- Have the conditions of data ownership been studied? In any case, data ownership is stipulated in the consortium agreement, the terms of which must be negotiated before the project and the DMP: give the link to the consortium agreement if possible.

*Exemple de réponse:*

See article 5 "Intellectual Property" in the consortium agreement of the XXX project (LIRMM).

4.2 How will compliance with the provisions of the legislation on personal data and data security be ensured if personal data are processed?

*Recommandations:*

- When handling personal data, ensure that data protection regulations are applied (e.g. GDPR <https://www.cnil.fr/fr/comprendre-le-rpgd>).
- Consider anonymisation of personal data for preservation and/or sharing (correctly anonymised data are no longer considered as personal data).
- Explain whether a specific access procedure has been set up for users authorised to access personal data.
- The DPO of your research unit can help you with the writing of this part and answer your questions.

*Exemple de réponse:*

Anonymised personal data (collected during surveys); use of CNIL (French National Commission on Informatics and Liberty)'s PIA software to describe the sensitivity of the data, etc.

4.3 How will ethical issues be taken into account and codes of conduct respected?

*Recommandations:*

- Determine whether ethical issues will affect the way data will be stored and transferred, who will be able to view or use them, and which retention periods will apply. Demonstrate that these aspects are well considered and planned.
- Adopt national and international codes of conduct and the institutional code of ethics and check whether a review of practices (e.g. by an ethics committee) is required in relation to data collection for the research project.

*Exemple de réponse:*

Data collected during animal experiments validated by the xxxx ethics committee in its yyyy decision, research carried out in partnership with countries outside Europe: the recommendations of the H2020 manual have been respected; respect of the Nagoya protocol (ABS), etc.

## 5. STORAGE, SHARING AND REUSE OF DATA DURING THE PROJECT

5.1 How will data and metadata be stored and backed up throughout the research process?

*Recommandations:*

- Describe where the data will be stored and backed up during the research process and how often the backup will be done. It is recommended to store the data in at least two separate locations.
- Focus on data access and how it will be secured (frequency, automatism...).
- Discuss the resources (budget and time allocation) dedicated to data storage

*Exemple de réponse:*

- For the physical hosting of the data: the data will be stored on the Meso@LR mesocentre data infrastructure with replication. The environment is hosted at Cines in the ZRR zone. The system is natively protected against ransomware.
- On the geographical location of the data: in France, in the EU, outside the EU, don't know (cloud without geographical storage details), etc.

5.2 How will data sharing be ensured between partners throughout the research process?

*Recommandations:*

- Explain who will have access to the data during the research process and how data access is controlled, especially during collaborative research.
- Make priorities of data sharing between the different partners in the research project and organisation in order to make collaborative work more fluid.

*Exemple de réponse:*

- On data sharing: the data collected will be shared with the research group [...] of the X University... in charge of the theme Y.
- On data access rights during research, access control policy; agreement between partners established; formalised process; one-off access authorisation request, etc.
- On data access rights by partner researchers: free access; declaration of identification; password; certificate; private key, etc.
- To monitor the production and analysis of data during the project: laboratory notebooks, research protocols, logbooks, time stamping of data and results, personal authentication on tools, etc.

## 6. DISSEMINATION, SHARING AND REUSE OF DATA AFTER THE PROJECT

6.1 How will the data and metadata be disseminated after the research project?

*Recommandations:*

- The principle that applies is: "As open as possible, as closed as necessary".
- Explain if there are any restrictions on data sharing (embargo, publication deadlines...).
- Will you allow reuse, redistribution or the creation of new tools, services, datasets or (derived) products? Will commercial use be allowed?

6.2 Where will the data be stored for dissemination?

Depositing the data in a research data repository allows it to be stored after the project for access and reuse. There are thousands of repositories of different types: disciplinary, multidisciplinary, publisher-specific, institutional, research project-specific, etc.

*Recommandations:*

- Is a repository recommended to you by one of the stakeholders of your project (your funder, publisher, institution, etc.)?
- If no repository is recommended to you, you need to find one that is adapted to your needs. You can use repository directories (re3data <https://www.re3data.org/>) depending on your search criteria (disciplinary, European repository, etc.)
- The National Plan for Open Science (2018, <https://scienceouverte.couperin.org/category/politiques/pnso/>) invites researchers to deposit data in certified data repositories: you can consult the CoreTrustSeal site to find such repositories (<https://www.coretrustseal.org/>).

*Exemple de réponse:*

In genomics, the scientific community uses Unitprot or GenBank as a reference database.

## 7. ARCHIVING AND LONG-TERM PRESERVATION

7.1 Which data should be preserved in the medium and long term?

*Recommandations:*

- Define the data preservation and sharing plan and provide information on how long the data will be archived permanently (3-5 years, 10 years, etc.).
- Explain when the data will be made available.
- Indicate whether sharing of the data will be deferred or limited, for example for publication reasons, to protect intellectual property or patenting.
- Discuss the resources (budget and time allocation) dedicated to archiving the data.

7.2 How will the data to be preserved be selected and where will it be preserved in the long term (e.g. in a data repository or an archive)?

*Recommandations:*

- Indicate which data should not be disclosed or should be destroyed for contractual, legal, or regulatory reasons.
- Will the data need to be anonymised or converted to more stable file formats?
- Indicate where the data will be deposited. In a discipline-specific or a broader repository? (<https://coop-ist.cirad.fr/gerer-des-donnees/deposer-des-donnees-dans-un-entrepot/5-comment-choisir-un-entrepot-de-donnees>).

7.3 Which methods or software tools will be required to have access to the data and use them?

*Recommandations:*

- Indicate whether potential users will need specific tools or software for data access and (re)use.

- Indicate whether the data will be shared via a data portal (e.g. institutional portal), whether access requests will be handled on-line, or whether another mechanism will be used.
- Indicate whether a unique and persistent identifier (such as a DOI) will be provided.

*Exemple de réponse:*

The data stored within the Meso@LR infrastructure will be accessible via the NextCloud tool of the xxx laboratory to any user present in the LDAP directory of the structure. All the people associated with the project will have an account. At the end of the project, the data will be uploaded to the open data repository via the University of Montpellier's Dataverse portal and a DOI will be assigned. Whenever possible, data papers will be published. The software associated with the data processing will be contributed to the Software Heritage project.