Guide to writing a Data Management Plan

This document is designed to help **researchers** and **project managers** to write **Data Management Plans** (DMPs). It accompanies the **DMP template document** developed by the CeRIS library. The template document consists of a **set of questions** that researchers should think about at the beginning of a research project. For each question, we provide **examples of answers** and **advice** given by the relevant services of the Institut Pasteur. The structure of the DMP is based on the template provided by the European Commission, with additional material from several services of the Institut Pasteur: the library and archives of the CeRIS, the Information Systems Department, the Legal Affairs Department, the Ethics Unit, the Patents and Inventions Department, the Quality Department, and the Center for Translational Science.

What are we talking about?

Research data are factual records (figures, texts, images, and sounds), used as primary sources for scientific research and commonly accepted in the scientific community as necessary to validate research findings (OECD definition, 2007).

A scientific **dataset** is a collection of data forming a coherent whole. The use of datasets requires metadata explaining how the data were produced and should be used.

A **Data Management Plan** (DMP) is a document written at the start of a research project which defines how the generated data will be managed, during and after the project: from data collection until data sharing and archiving. This is a draft document that should be regularly updated during the research project.
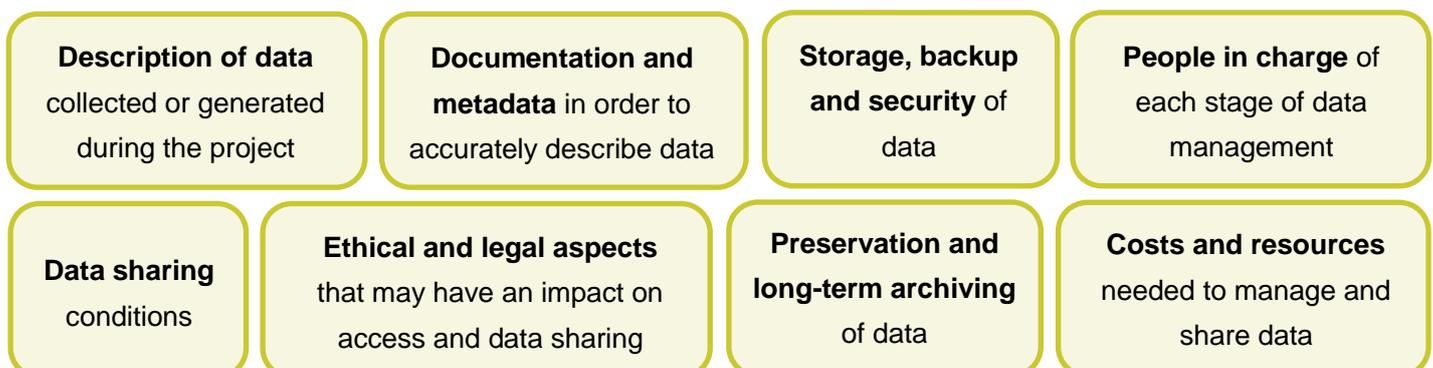
Who should write the DMP?

The **research team coordinating the project** is responsible for writing the DMP. However, various services at the Institut Pasteur are available to help researchers at each stage of the process. **Feel free to contact them** if you need help:
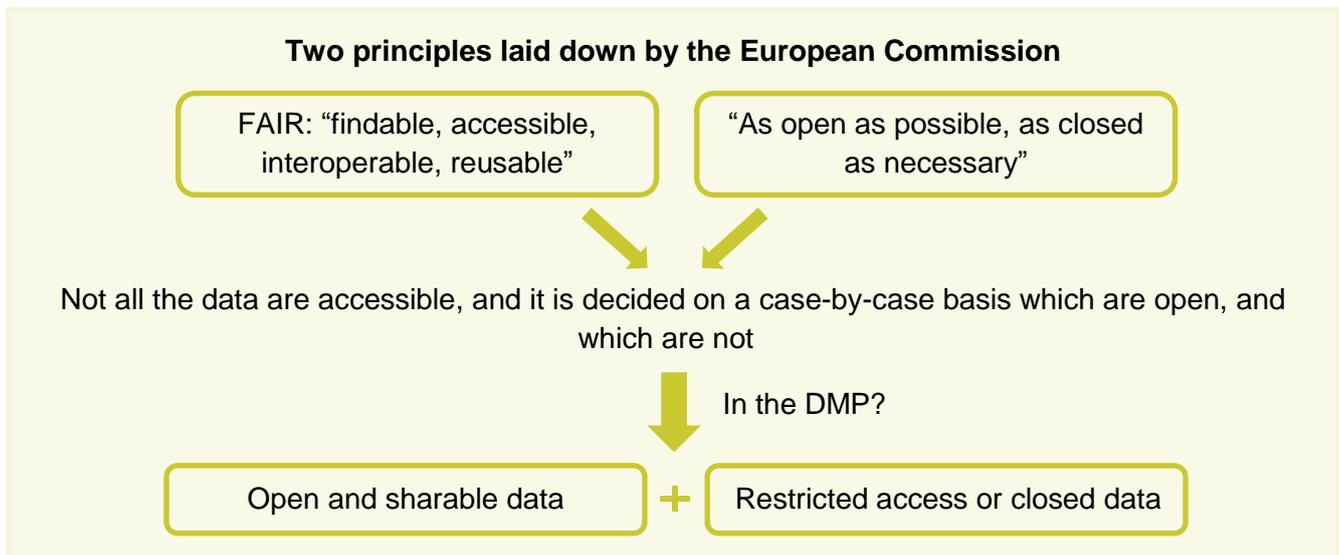
The CeRIS Library	bibliotheque@pasteur.fr
The CeRIS Archives	archives@pasteur.fr
The Legal Affairs Department	• anne-laure.morin@pasteur.fr for general questions • rgpd@pasteur.fr for questions concerning personal data
The Quality Department	qualite@pasteur.fr
The Ethics Unit	samira.ouchhi@pasteur.fr
The Information Systems Department	• informatique@pasteur.fr for general questions • rssi@pasteur.fr for questions concerning data security
The Patents and Inventions Department	sbi@pasteur.fr
The Grants Office	go-all@pasteur.fr
The Center for Translational Science	crt-opendesk@pasteur.fr

What information do you need to include in the DMP?

The writing of the DMP aims to formalize and compile **information that is often already available,** to optimize the preparation of all stages of the data lifecycle. The DMP addresses the following topics:

Description of data collected or generated during the project	**Documentation and metadata** in order to accurately describe data	**Storage, backup and security** of data	**People in charge** of each stage of data management
Data sharing conditions	**Ethical and legal aspects** that may have an impact on access and data sharing	**Preservation and long-term archiving** of data	**Costs and resources** needed to manage and share data

Is the DMP restricted to open data?

Two principles laid down by the European Commission

FAIR: "findable, accessible, interoperable, reusable"

"As open as possible, as closed as necessary"

Not all the data are accessible, and it is decided on a case-by-case basis which are open, and which are not

In the DMP?

Open and sharable data **+** Restricted access or closed data

A DMP can therefore be drawn up for both **data sharing** and **data with partially restricted or closed access**. In the latter case, the DMP should mention the reasons for not sharing data.

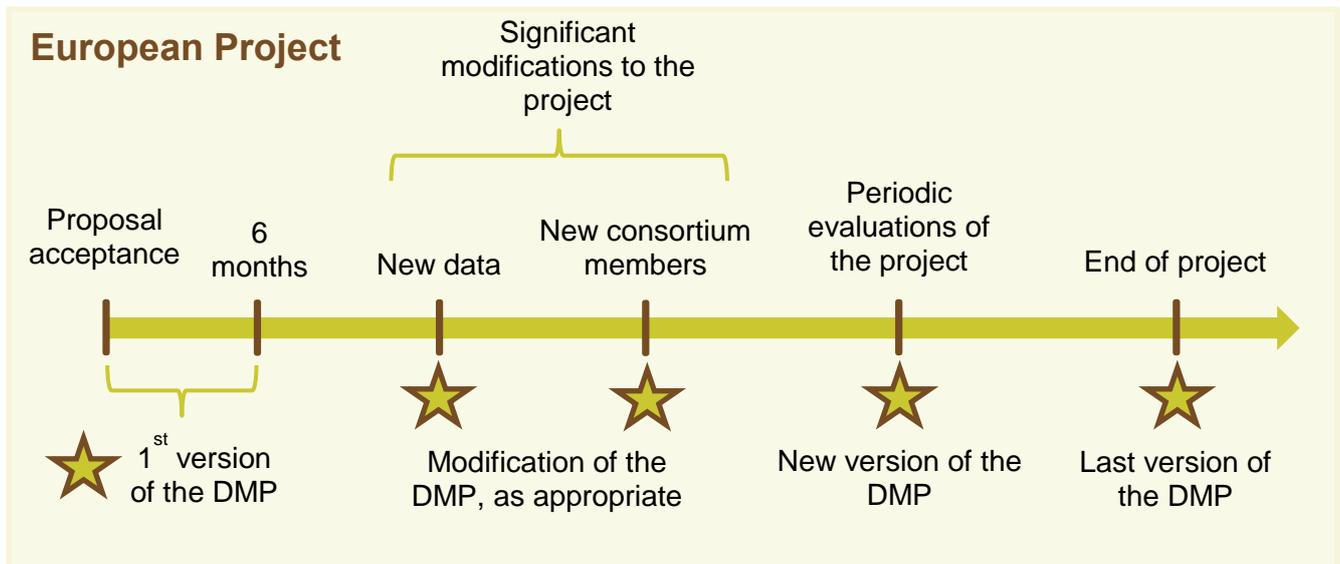## When should the DMP be written?

For European projects:
The first version of the DMP must be written within six months of the proposal being accepted. It is not essential to provide detailed answers to all the questions in this version. Regular updates are then required:

- Every time that the project is significantly modified (e.g. new data, decision to file for a patent, new members joining the consortium)
- Before each of the scheduled evaluations of the project
- Before the final project review

The DMP must include a visible version number and may be accompanied by an update calendar.

For other projects:
Other funding organizations do not provide recommendations. We therefore suggest following the recommendations of the European Commission.

European Project

Significant modifications to the project

| Proposal acceptance | 6 months | New data | New consortium members | Periodic evaluations of the project | End of project |

1st version of the DMP — Modification of the DMP, as appropriate — New version of the DMP — Last version of the DMP

Why write a Data Management Plan?

⭐ REFLECTION: ask yourself the right questions

⭐ ANTICIPATION: prepare all stages of the data lifecycle

⭐ MANAGEMENT: optimize the management of your data

⭐ OBLIGATION: the DMP is a deliverable required by the European Commission

Why improve data management?

⭐ REPRODUCIBILITY: ensure the integrity and reproducibility of the research

⭐ QUALITY: improve the precision, completeness, authenticity and reliability of the data

⭐ SAVE TIME: avoid duplicating data, thus saving time and resources

⭐ PRESERVATION: improve data security and minimize the risk of data loss

⭐ VALORISATION: increase the visibility and impact of researchers' work

⭐ INNOVATION: foster reuse and innovation through sharing

⭐ COLLABORATION: contribute to the setting up of scientific collaborations

Where should the DMP be stored/disseminated?

During the project, the various versions of the DMP can be posted on the project website. At the end of the project, the final version of the DMP can be deposited in an open data repository, such as Zenodo.

Other questions?

Contact the CeRIS library for more information:
- Fanny SEBIRE: fanny.sebire@pasteur.fr
- Anne LASSAILLY-BONDAZ: anne.lassailly-bondaz@pasteur.fr

Development

Fanny SEBIRE (CeRIS library, Institut Pasteur). This document is based on documents from other universities, research institutes [1], [2], [3], and the European Commission [4].

References

[1] Reymonet N., Moysan M., Cartier A. et Délémontez R. *Réaliser un plan de gestion de données « FAIR » : modèle*. (V2, 2018). Licence Creative Commons CC-BY-NC-SA. https://hal.archives-ouvertes.fr/sic_01690547

[2] CEA. *Guide « Plan de Gestion de Données Scientifiques »* (V5, septembre 2017). Licence Creative Commons CC-BY-NC-SA.

[3] IFSTTAR. *PGD - Plan de Gestion de Données Scientifiques, DMP –Data Management Plan* (janvier 2018).

[4] Commission européenne, direction générale pour la recherche et l'innovation. *Guidelines on FAIR Data Management in Horizon 2020 (V3.0, 26 juillet 2016).* http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf