Guide for data management plans (DMPs)

Version: 1.1

Date: 17.3.2021

Contents

1) Introduction **2**

2) Why do I benefit from a DMP? **2**

3) Funding agencies and their DMPs **3**

4) Structure of a DMP **4**

a) General overview 4
b) DMP sections 4

Contact **7**

1) Introduction

*A data management plan (**DMP**) is a document that describes the intended management of research data. This includes activities during the research process as well as after completion of the project. The plan contains all information necessary to describe and document the collection, processing, storage, archiving and publication of research data. The scope of a DMP can vary from a few paragraphs to several pages.*[1]

2) Why do I benefit from a DMP?

A data management plan consumes resources in its creation and offers many advantages at the same time.

A data management plan:

- *creates a firm basis for a uniform handling of data in the research process*
- *facilitates the understanding of your own data*
- *facilitates the coordination between project partners*
- *helps to identify potential problems at an early stage and to outline solutions for them*
- *defines responsibilities*
- *regulates access rights*
- *helps to avoid data duplication, data loss and security loopholes*
- *can be (sometimes obligatory) part of a grant application*

A useful initial consideration and argument for a data management plan is to think backwards, i.e. where and how should the data be archived or published? These considerations make it necessary to set the course in the data management workflow early on, e.g. with regard to formats, standards, metadata, licenses, etc.[1]

[1] Biernacka et al., “Train-the-Trainer Concept on Research Data Management”, FDMentor, 2020, Accessed on: 4.3.2021, https://doi.org/10.5281/zenodo.4071471

3) Funding agencies and their DMPs

Data management plans are already required by a number of research funders. For funding under the European Commission's H2020 programme, a DMP is required at the start of funding.[2] For the Austrian Science Fund (FWF), the DMP must be submitted together with the funding agreement and is a prerequisite for the start of the project. A template with assistance is provided.[3]

For the Austrian Research Agency (FFG), there is currently no uniform regulation across all programmes. Whether a DMP is required is regulated at the tender level. For example, in the current call for proposals for the thematic programme "ICT of the Future", a DMP must be submitted.[4] Table 1 shows a summary of the essential criteria.

Table 1: Comparison of the funders' requirements regarding the data management plan

Funders	Plan required?	Levy with appliation?	Content	Updates?
EC Horizon 2020	*Data management plan*	*No, first plan within the first 6 project months*	*Contents of the Horizon 2020 Template*	*Update if significant changes occur and at the end of the project.*
FWF	*Data management plan*	*No, must be submitted together with the grant agreement for an approved project.*	*Content of the FWF template*	*Can be changed throughout the project; final version must be submitted with final report.*
FFG	*Not in all programme lines*	*Yes, together with the project description*	*According to the tendering guide (DMPonline is recommended)*	*The following changes must be reported within the framework of the reporting process*

[2] European Commission, H2020 Programme. „Guidelines on FAIR Data Management in Horizon 2020." Version 3.0. Accessed on: 4.3.2021, https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
[3] FWF – Forschungsdatenmanagement. Accessed on: 4.3.2021, https://www.fwf.ac.at/de/forschungsfoerderung/open-access-policy/forschungsdatenmanagement/
[4] IKT der Zukunft: Ausschreibungen 2020. Accessed on: 4.3.2021, https://www.ffg.at/ausschreibung/ikt-der-zukunft-ausschreibungen-2020

4) Structure of a DMP

a) General overview

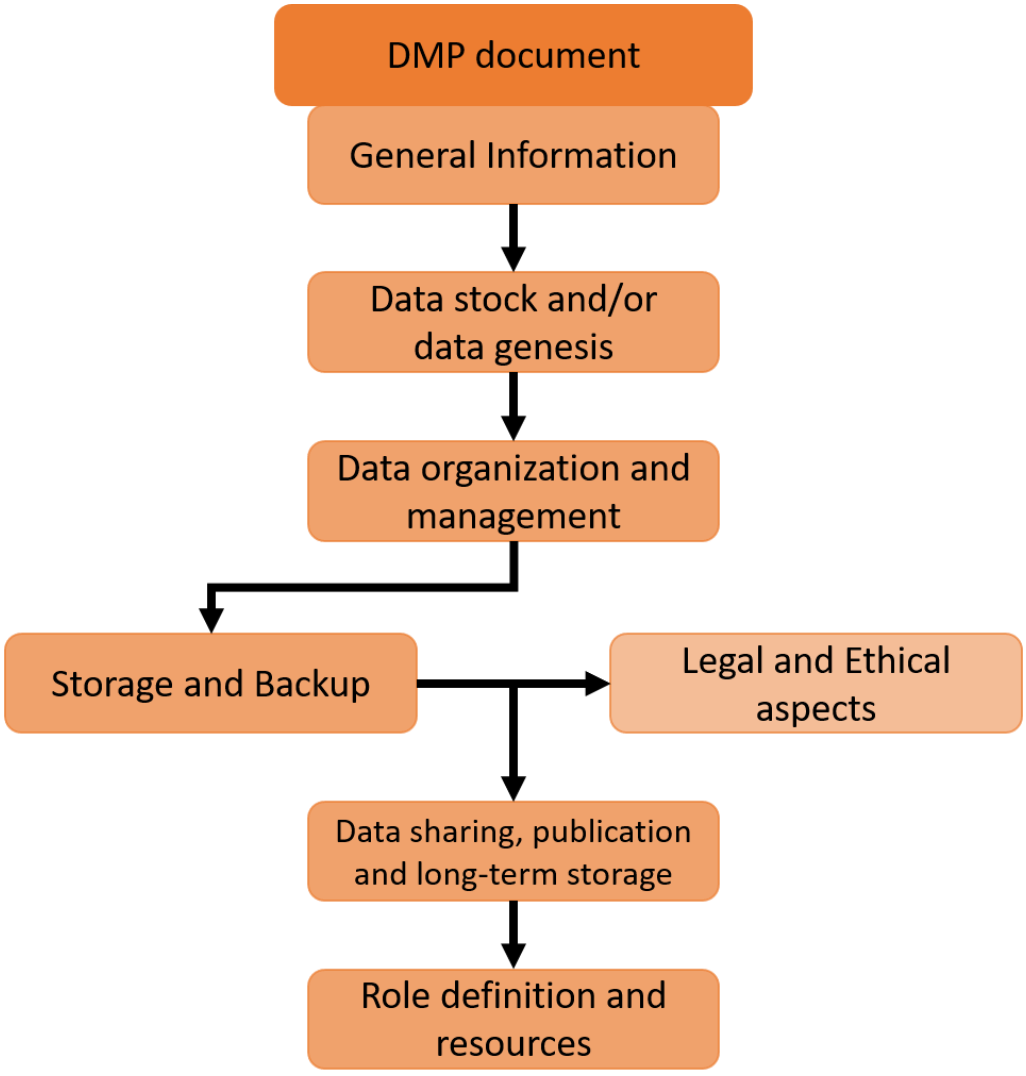

Figure 1: Schematic overview of a DMP[5]

b) DMP sections

a. *General Information*

Contains some general information as project name, project number, name of applicant, funding programme, version/date of DMP and license.

[5] Inspired by: Elemente eines DMP. Accessed on: 5.3.2021,https://www.forschungsdaten.info/themen/informieren-und-planen/datenmanagementplan/

b. Data stock and/or data genesis

Description of the origin and quality of existing or newly acquired data and their integration into the project.

- ***How will existing data be re-used and/or new data be collected or produced?***
  State any constraints on re-use of existing data and explain if the re-use of any existing data sources has been considered but discarded. Describe methodologies or software for new data and how data provenance will be documented.[6]
- ***What data (e.g. the kind, formats and volumes), will be collected or produced?***
  Give details on the kind (e.g. numeric, textual, image, audio, video, and/or mixed media) of data and the format (e.g. pdf, xls, doc, txt, or rdf). Justify the use of certain formats (e.g. preference for open formats, standards accepted by data repositories, widespread usage within the research community, or on the software or equipment that will be used). Give any details on the storage space required.[6]

c. Data organization and management

Information on data management, quality control measures, documentation for easy reuse of data and metadata.

- ***What metadata and documentation (e.g. the methodology of data collection and way of organising data) will accompany the data?***
  Indicate which metadata (e.g. DDI, TEI, EML, MARC, CMDI) will be provided to help others identify and discover the data. Think about what information and how it is captured (e.g. a 'readme' text file, file headers, code books, or lab notebooks), is needed to enable re-use.[6]
- ***What data quality control measures will be used?***
  Describe processes (e.g. calibration, peer review of data, ...) that ensure how the consistency and quality of data collection will be controlled and documented.[6]

d. Storage and Backup

Describes where the data is stored and backed up, and how the security of sensitive data is ensured over the life of the project.

- ***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. Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution.[6]
- ***How will data security and protection of sensitive data be taken care of during the research?***

[6] Practical Guide to the international Alignment of Research Data Management (Extended Edition), January 2021; Science Europe

Explain who will have access to the data during the research and how access to data is controlled (consider data protection). Mention the institutional data protection and how data will be recovered in the event of an incident.[7]

e. *Legal and Ethical aspects*

Covers topics such as property rights and copyrights, ensuring data protection and legal and ethical barriers in terms of accessibility.

- ***If personal data are processed, how will compliance with legislation on personal data and on security be ensured?***
  Ensure that when dealing with personal data, data protection laws are complied and you also might need to conclude a data processing contract. Explain whether there is a managed access procedure in place for authorised users of personal data.[7]
- ***How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?***
  Explain who will be the owner of the data (rights to control access) and whether intellectual property rights are affected. Indicate whether there are any restrictions on the re-use of third-party data.[7]
- ***What ethical issues and codes of conduct are there, and how will they be taken into account?***
  Demonstrate awareness while consider whether ethical issues can affect how data are stored and transferred, who can see or use them, and how long they are kept. Follow the national and international codes of conducts and institutional ethical guidelines, and check if ethical review (for example by an ethics committee) is required for data collection in the research project.[7]

f. *Data sharing, publication and long-term storage*

Details of how interoperability is sought and how data exchange is to take place. Information about DOIs and long-term storage.

- ***How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?***
  Explain when the data will be made available (expected release) and how the data will be discoverable and shared. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents. Indicate who will be able to use the data and if it is necessary to restrict access to certain communities or to apply a data sharing agreement, explain how and why. Explain the plan for data preservation and give information on how long the data will be retained.[7]

[7] Practical Guide to the international Alignment of Research Data Management (Extended Edition), January 2021; Science Europe

- ***How will data for preservation be selected, and where data will be preserved long-term (e.g. a data repository or archive)?***
  Indicate how it is decided which data are to be to be retained and which must be destroyed (contractual, legal or regulatory purposes). Explain the foreseeable research purposes for the data and which data need to be preserved long-term. Specify where the data should be deposited so that I can be curated effectively beyond the lifetime of the grant.[8]
- ***What methods or software tools are needed to access and use data?***
  Indicate whether potential users specific tools for accessing and (re-)using the data. Indicate the mechanism (e.g. repository, directly processed requests) through which the data will be shared.[8]
- ***How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?***
  Persistent identifiers (PIDs) should be applied so that data can be reliably and efficiently located and referred to. PIDs also help to track citations and re-use. Typically, long-term repository will provide a persistent identifier for pursuing.[8]

g. Role definition and resources

Includes information on those responsible for data management and details of the expected costs of complying with the DMP.

- ***Who (for example role, position, and institution) will be responsible for data management (e.g. the data steward)?***
  State the roles and responsibilities for data management activities and name responsible persons (specially for collaborative projects). Regular update the DMP, if possible.[8]
- ***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)?***
  Explain how the necessary (e.g. storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges) and additional (for preparation of data for deposit or to meet any charges from data repositories) resources to prepare the data for sharing/preservation have been costed in. Carefully consider and justify any resources needed to deliver the data.[8]

Contact

If you have any questions about this document or need additional help, just write to us (RDM Team):

rdmteam@tugraz.at

[8] Practical Guide to the international Alignment of Research Data Management (Extended Edition), January 2021; Science Europe