TEMPLATE HORIZON 2020 DATA MANAGEMENT PLAN (DMP)

- **Instructions and footnotes in blue must not appear in the text.**
- **For options *[*in square brackets*]*: the option that applies must be chosen.**
- **Enter the appropriate data for fields in [grey in square brackets] (even if they are part of an option as specified in the previous item).**

Introduction

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

Guidelines on FAIR Data Management in Horizon 2020 are available in the Online Manual.

FAIR data management

Generally, your research data should be 'FAIR,' which is findable, accessible, interoperable, and reusable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or solution.

This template is not intended as a strict technical implementation of the FAIR principles; FAIR somewhat inspires it as a general concept.

More information about FAIR:

FAIR data principles (FORCE11 discussion forum)

FAIR principles (article in Nature)

Structure of the template

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

It is optional to provide detailed answers to all the questions in the first version of the DMP that needs to be submitted by month 6 of the project. Instead, the DMP is intended to be a living document in which information can be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. Therefore, DMPs should have a precise version number and an updated timetable. As a minimum, the DMP should be updated in the context of the

periodic evaluation/assessment of the project. If no other regular reviews are envisaged within the grant agreement, an update must be made in time for the final assessment at the latest.

The following main sections to be covered by the DMP are outlined. At the end of the document, Table 1 summarizes these elements in bullet form.

This template itself may be updated as the policy evolves.

Project[1] Number: [insert project reference number]

Project Acronym: [insert acronym]

Project title: [insert project title]

DATA MANAGEMENT PLAN

[1] The term 'project' used in this template equates to an 'action' in certain other Horizon 2020 documentation

1. Data Summary

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

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

Will you re-use any existing data, and how?

What is the origin of the data?

What is the expected size of the data?

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

2. FAIR data

2. 1. Making data findable, including provisions for metadata

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

What naming conventions do you follow?

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

Do you provide exact version numbers?

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

2.2. Making data openly accessible

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

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

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

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

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

Can the relevant software be included (e.g., in open-source code)?

Where will the data and associated metadata, documentation, and code be deposited? Preference should be given to certified repositories which support open access where possible.

Have you explored appropriate arrangements with the identified repository?

If there are restrictions on use, how will access be provided?

Is there a need for a data access committee?

Are there well-described conditions for access (i.e., a machine-readable license)?

How will the identity of the person accessing the data be ascertained?

2.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organizations, countries, etc. (i.e., adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

What data and metadata vocabularies, standards, or methodologies will you follow to make your data interoperable?

Will you use standard vocabularies for all data types in your data set to allow interdisciplinary interoperability?

If you are unavoidable to use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

2.4. Increase data re-use (through clarifying licenses)

How will the data be licensed to permit the most extensive re-use possible?

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Are the data produced and used in the project useable by third parties, in particular after the end of the project? If the reuse of some data is restricted, explain why.

How long is it intended that the data remains re-usable?

Are data quality assurance processes described?

Further to the FAIR principles, DMPs should also address the following:

3. Allocation of resources

What are the costs for making data FAIR in your project?

How will these be covered? Open access to research data costs is eligible for the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

Who will be responsible for data management in your project?

Are the resources for long-term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

4. Data security

What provisions are in place for data security (including data recovery, secure storage, and transfer of sensitive data)?

Is the data safely stored in certified repositories for long-term preservation and curation?

5. Ethical aspects

Are there any ethical or legal issues that can impact data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and the ethics chapter in the Description of the Action (DoA).

Is informed consent for data sharing and long-term preservation included in questionnaires dealing with personal data?

6. Other issues

Do you use other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

7. Further support in developing your DMP

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.

The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.

Valuable listings of repositories include:

Registry of Research Data Repositories

Some repositories like Zenodo, an OpenAIRE and CERN collaboration), allow researchers to deposit publications and data while providing tools to link them.

Other valuable tools include DMP online and platforms for making individual scientific observations available such as ScienceMatters.

SUMMARY TABLE 1

FAIR Data Management at a Glance: issues to Cover in your Horizon 2020 DMP

This table summarizes the Data Management Plan (DMP) issues to be addressed, as outlined above.

DMP component	Issues to be addressed
1. Data summary	• 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
2. FAIR Data 2.1. Making data findable, including provisions for metadata	• Outline the discoverability of data (metadata provision) • Outline the identifiability of data and refer to the standard identification mechanism. Do you use persistent and unique identifiers such as Digital Object Identifiers? • Outline naming conventions used. • Outline the approach toward search keywords. • Outline the process for explicit versioning • 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

2.2 Making data openly accessible	• Specify which data will be made openly available. If some data is kept closed, provide a rationale for doing so • Specify how the data will be made available. • Specify what methods or software tools are needed to access the data. Is documentation about the software required to access the data included? Can the relevant software be included (e.g., in open-source code)? • Specify where the data and associated metadata, documentation, and code are deposited • Specify how access will be provided in case there are any restrictions
2.3. Making data interoperable	• Assess the interoperability of your data. Specify what data and metadata vocabularies, standards, or methodologies you will follow to facilitate interoperability. • Specify whether you will be using standard vocabulary for all data types in your data set to allow inter-disciplinary interoperability. If not, will you provide mapping to more commonly used ontologies?
2.4. Increase data re-use (through clarifying licenses)	• Specify how the data will be licensed to permit the most comprehensive reuse possible • Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed • Specify whether the data produced and used in the project is useable by third parties, in particular after the end of the project. If the re-use of some data is restricted, explain why • Describe data quality assurance processes • Specify the length of time for which the data will remain re-usable
3. Allocation of resources	• Estimate the costs for making your data FAIR. Describe how you intend to cover these costs • Identify responsibilities for data management in your project • Describe the costs and potential value of long-term preservation
4. Data security	• Address data recovery as well as secure storage and transfer of sensitive data
5. Ethical aspects	• To be covered in the context of the ethics review, ethics section of DoA, and ethics deliverables. Include references and related technical aspects if not covered by the former
6. Other	• Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

HISTORY OF CHANGES		
Version	**Publication date**	**Change**
1.0	13.10.2016	▪ Initial version
1.1	1.07.2023	Grammatical Changes (Richard Dennis)