UNIVERSITÉ
DE GENÈVE

SERVICE RECHERCHE

European Commission (EC) - Horizon 2020 (H2020) - Open Research Data Pilot (ORD)
Data Management Plan and Data Open Access

V05.10.2017

All Horizon 2020 projects starting from January 2017 are by default part of the Open Research Data Pilot (ORDP). The ORDP aims to make the research data generated by selected H2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access.

If your project is part of the pilot, you must take the following actions:

➔ **create a Data Management Plan (DMP)**

Data Management Plans (DMPs) are key element of good data management. A DMP is a formal document that describes the data management life cycle for the data to be collected, processed and/or generated by a Horizon 2020 research project.

More information on DMP:
- DMP template (EC) in Annex;
- How to create a DMP on the OpenAire website: https://www.openaire.eu/opendatapilot-dmp OpenAire recommends to refer the Principal Researcher to the Digital Curation Centre **DMP online tool** : https://dmponline.dcc.ac.uk/ , which offers DMP templates that match the demands and suggestions of the EC guidelines;
- EC guidelines on Data Management in Horizon 2020 :http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Contact at UNIGE: The Library offers Research Data Management personal support to help you with your DMP - researchdata-info@unige.ch

➔ **select a data repository**

If your Horizon 2020 project is part of the pilot, and your data meets certain conditions, you must deposit your data in a research data repository where they will be findable and accessible for others. That will preserve your data, metadata and possibly tools in the long term. It is advisable to contact the repository of your choice when writing the first version of your DMP.

More information: OpenAire website https://www.openaire.eu/opendatapilot-repository

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

UNIGE website and contact on data storage : http://www.unige.ch/researchdata/en/store-your-data/all/

Please note also:

You do not need to deposit all the data you generate during the project either

You must only disclose/deposit the data which underpins published research findings and/or has longer-term value. You are not expected to share sensitive data or breach any IPR agreements with industrial partners (please see the next point – opting out).

Please refer to the Digital Curation Centre website on *How to Select What Data to Keep* : http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep

Projects can opt out at any stage (either before or after signing the grant)

Opting out, which frees from the obligations associated with the conditions mentioned above, is possible if:

- • participation is incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited
- • participation is incompatible with the need for confidentiality in connection with security issues
- • participation is incompatible with rules on protecting personal data
- • participation would mean that the project's main aim might not be achieved
- • the project will not generate / collect any research data or
- • there are other legitimate reasons (you can enter these in a free-text box at the proposal stage).

Projects dealing with personal data need clearing by the Data Protection Officer

For approval by the institutional Data Protection Officer (DPO), please see our relevant information on ethics in H2020.

Annex

EC - H2020 FAIR Data Management Plan (DMP) template

This Horizon 2020 FAIR DMP template has been designed to be applicable 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 clearly spell this out.

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

It is not required 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. Rather, 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.

Template

1. Data Summary

- What is the purpose of the data collection/generation and its relation to the objectives of the project?
- 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 useful ('data utility')?

2. FAIR data

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

- Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of 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 clear version numbers?
- What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

2.2. Making data openly accessible

- Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.
- 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 made 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?
- Is it possible to include the relevant software (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, organisations, 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 be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?
- In case it is unavoidable that you 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 licences)

- How will the data be licensed to permit the widest 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/or used in the project useable by third parties, in particular after the end of the project? If the re-use 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:

3. Allocation of resources

- What are the costs for making data FAIR in your project?
- How will these be covered? Note that costs related to open access to research data are eligible as part of 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 as well as 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 have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and 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 make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?