

# European Commission (Horizon): Horizon Europe DMP - Grant proposal

## 1. Data summary

### 1.1 Types of data/other research outputs

#### *Guidance:*

List which types of research data the project expects to generate, collect or use and the estimated volume for each data type.

Also indicate if data will be newly generated or reused. In case the project will reuse existing data, provide information on their provenance (i.e. the source/origin of the data).

In addition, give an overview of any other digital or physical research outputs (except for publications) if applicable.

## 2. FAIR principles

### 2.1 Findability of data/research outputs

#### *Guidance:*

Indicate for each data/output type in which **trusted repository** the project plans to deposit the data/outputs, and which kind of **persistent identifier** (PID) the repository will assign to them (e.g. DOIs, handles, ...).

Horizon Europe requires beneficiaries to deposit research data (including raw data, to the extent technically feasible) in a trusted repository as soon as possible (at the latest by the end of the project, or at the time of publication if the data underpin a scientific publication), or immediately in case of public emergencies.

Such repositories help make data findable by assigning them a persistent identifier (i.e. a globally unique and long-lasting reference to digital objects) and by making rich dataset descriptions ('metadata') available online in a searchable resource.

In Horizon Europe, the **following are considered trusted repositories:**

- Certified repositories (e.g. CoreTrustSeal, nestor Seal DIN31644, ISO16363), or disciplinary and domain repositories commonly used and endorsed by research communities, and recognised internationally
- General-purpose repositories or institutional repositories that present the essential characteristics of trusted repositories (incl. assigning persistent identifiers and providing sufficient metadata to enable discovery, reuse and citation)

Personal websites and databases, publisher websites, cloud storage services (Dropbox, Google drive etc.), and platforms such as Academia.edu and ResearchGate are NOT considered repositories.

For Horizon Europe calls with a condition relating to the European Open Science Cloud (EOSC), data must be deposited in an **EOSC-federated repository** (find services via the [EOSC Portal](#)).

### 2.2 Accessibility of data/research outputs

#### *Guidance:*

For each data/output type, indicate whether and when open access will be provided to data/outputs. If not, explain why (e.g. IPR/commercial exploitation considerations).

If applicable, elaborate on provisions for access to restricted data for verification purposes.

Note that Horizon Europe requires beneficiaries to make research data accessible via a data

repository as soon as possible following the principle ‘as open as possible, as closed as necessary’. In other words, open access is expected as the default, unless there are justified reasons for restricting or closing access to some or all of the research data. Consult the [AGA](#) (art. 17, p.160) on legitimate reasons for restricting access and examples of valid justification. Also, the ‘A’ in FAIR does not mean that data have to be fully openly accessible (i.e. free to access and reuse by anyone for any purpose). Data can (and under Horizon Europe should) be FAIR, even when access is restricted. Providing restricted access means that there are limits on who can access and reuse, and for what purpose.

### 2.3 Interoperability of data/research outputs

#### *Guidance:*

Describe which data and metadata standards, formats and/or vocabularies will be used to allow data/output exchange and reuse within and across disciplines, systems, etc. (e.g. standards common in the relevant research community or for the data/output type in question)

### 2.4 Reusability of data/research outputs

#### *Guidance:*

Indicate under which **license** data and/or other research outputs will be shared (e.g. Creative Commons, Open Data Commons). Licenses enhance reusability by making explicit how research data/other outputs may be reused (i.e. they clarify terms of use).

Horizon Europe requires that open access research data are licensed under a CC-BY, CC0, or equivalent license. Metadata should be open access under a CC 0 license or equivalent. This is also recommended in cases where data must be closed or restricted but there are no compelling reasons for metadata not to be findable and accessible.

Horizon Europe requires beneficiaries to provide information via the data repository about any research output or any other tools and instruments needed for data reuse and validation (e.g. software, algorithms, protocols, models, workflows, electronic notebooks etc.). This information should include:

- a detailed description of the research output/tool/instrument
- how to access it
- any dependencies on commercial products
- potential version/type
- potential parameters, etc.

In addition, beneficiaries are encouraged to provide open access to these research outputs, tools and instruments under a CC-BY or CC0 license or equivalent, unless legitimate interests or constraints apply. This might even be an obligation in certain work programmes/call conditions or in case of public emergencies.

## 3. Resources and responsibilities

### 3.1 Curation and storage/preservation costs

#### *Guidance:*

Discuss what costs associated with making research data/outputs FAIR (e.g. for curating, storing and/or preserving them) the project is expected to incur, and provide an estimate.

Note that costs related to data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions), and should be budgeted in the proposal.

### 3.2 Person/team responsible for data management and quality assurance

*Guidance:*

Describe who will be responsible for assuring data quality and for the various aspects of managing research data/outputs.

# European Commission (Horizon): Horizon Europe DMP - Full DMP

## Version information

Action number

*Guidance:*

Provide the project reference number.

Action acronym

*Guidance:*

Mention the project acronym

Action title

*Guidance:*

Provide the full project title.

DMP version number

*Guidance:*

Indicate the current version of the DMP. E.g. v1.0, 2.0, ...

Date

## 1. Data summary

1.1 Will you re-use any existing data and what will you re-use it for?

*Guidance:*

State the reasons if re-use of any existing data has been considered but discarded.

1.2 What types and formats of data and other research outputs will the project generate or re-use?

*Guidance:*

List which **types of research data** the project expects to generate, collect or use and the estimated volume for each data type.

Also indicate if data will be **newly generated or reused**. In case the project will reuse existing data, provide information on their provenance (i.e. the source/origin of the data).

In addition, give an overview of any **other digital or physical research outputs** (except for publications) if applicable.

1.3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?

*Guidance:*

Refer back to question 1.2 if applicable.

1.4 What is the expected size of the data that you intend to generate or re-use?

*Guidance:*

Refer back to question 1.2 if applicable.

1.5 What is the origin/provenance of the data, either generated or re-used?

*Guidance:*

Refer back to question 1.2 if applicable.

1.6 To whom might your data be useful ('data utility'), outside your project?

*Guidance:*

Speculate and describe to whom your research output may be useful after publishing.

## **2.1 FAIR data: Making data findable, including provisions for metadata**

2.1.1 Will data and other research outputs be identified by a persistent identifier?

- Yes: describe below
- No: describe below

*Guidance:*

Persistent Identifiers (PIDs) must be provided for the data, for all author(s) involved in the action and, if possible, for their organizations and the grant.

**Examples** of commonly used PIDs include:

- DOI or handle for data and other research outputs
- ORCID or ResearcherID for authors
- ROR ID for organizations
- grant DOI for grants
- but also accession numbers within specific disciplines, notably in the Life Sciences e.g. an accession number in a database such as the Protein Data Bank.

Persistent Identifiers for data or other research output may be provided by trusted repositories where the research output is deposited (see also question 2.2.3).

2.1.2 Will rich metadata be provided to allow discovery?

What metadata will be created?

What disciplinary or general standards will be followed?

In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

2.1.3 Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

- Yes: describe below
- No: describe below

2.1.4 Will metadata be offered in such a way that it can be harvested and indexed?

- Yes: describe below
- No: describe below

## **2.2 FAIR data: Making data accessible**

2.2.1 Will the data and other research outputs be deposited in a trusted repository?

- Yes: describe below
- No: explain why

*Guidance:*

Horizon Europe requires beneficiaries to deposit research data (including raw data, to the extent technically feasible) in a trusted repository as soon as possible (at the latest by the end of the project, or at the time of publication if the data underpin a scientific publication), or immediately in case of public emergencies.

Such repositories help make data findable by assigning them a persistent identifier and by making rich dataset descriptions ('metadata') available online in a searchable resource.

In Horizon Europe, **the following are considered trusted repositories:**

- Certified repositories (e.g. CoreTrustSeal, nestor Seal DIN31644, ISO16363), or disciplinary and domain repositories commonly used and endorsed by research communities, and recognised internationally
- General-purpose repositories or institutional repositories that present the essential characteristics of trusted repositories (incl. assigning persistent identifiers and providing sufficient metadata to enable discovery, reuse and citation), see the [Annotated model grant agreement](#) for more details.

Personal websites and databases, publisher websites, cloud storage services (Dropbox, Google drive etc.), and platforms such as Academia.edu and ResearchGate are NOT considered repositories.

For Horizon Europe calls with a condition relating to the European Open Science Cloud (EOSC), data must be deposited in an **EOSC-federated repository** (find services via the [EOSC Portal](#)).

2.2.2 Have you explored appropriate arrangements with the identified repository where your data and other research outputs will be deposited?

- Yes
- No

2.2.3 Does the repository ensure that the data and other research outputs are assigned an identifier? Will the repository resolve the identifier to a digital object?

2.2.4 Will all data and other research outputs be made openly available?

- Yes
- No, certain datasets cannot be shared openly for the following reasons:

*Guidance:*

If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the [Grant Agreement](#).

2.2.5 Is an embargo applied to give time to publish or seek protection of the intellectual property (e.g. patents)?

- Yes
- No

2.2.6 If an embargo is applied (see question 2.2.5), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

2.2.7 Will the data and other research outputs be accessible through a free and standardized access protocol?

- Yes: describe below
- No: describe below

2.2.8 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

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

2.2.10 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

- Yes
- No

2.2.11 Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why.

- Yes
- No: explain why

*Guidance:*

Metadata of deposited data must be open under a Creative Common Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded).

2.2.12 Will metadata contain information to enable the user to access the data?

- Yes
- No

2.2.13 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

2.2.14 Will documentation or reference about any software needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

## **2.3 FAIR data: Making data interoperable**

### 2.3.1

What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines?

Will you follow community-endorsed interoperability best practices? Which ones?

2.3.2 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?

Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

2.3.3 Will your data and other research outputs include qualified references to other data (e.g. other data from your project, or datasets from previous research)?

- Yes

- No

*Guidance:*

A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: <https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/>).

## 2.4 FAIR data: Increase data re-use

2.4.1 How will you provide documentation needed to validate data analysis and facilitate data re-use?

*Example Answer:*

E.g. The data package deposited in the trusted repository will include README files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.

2.4.2

Will your data and other research outputs be made freely available in the public domain to permit the widest re-use possible?

Will your data and other research outputs be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

*Guidance:*

Research data should be made [open access by default](#) and licensed under the latest version of **CC BY** (attribution required) or **CC 0** (public domain), or equivalent.

However, it is recognized that data should be 'as open as possible, as closed as necessary', and **exceptions** can be made when providing open access to data:

- Is against the beneficiary's legitimate interests, including regarding commercial exploitation;
- Is contrary to any other constraints, such as data protection rules, privacy, confidentiality, trade secrets, Union competitive interests, security rules, intellectual property rights or;
- Would be against other obligations under the Grant Agreement.

In such cases, data can be kept restricted, closed or under embargo, but beneficiaries **must explain in the DMP the legitimate exception(s) under which they choose to restrict access to (some of the) research data.**

If your project generates other forms of research output, consider to what extent and how they will be made openly available.

2.4.3 Will the data and other research output produced in the project be useable by third parties, in particular after the end of the project?

- Yes
- No

2.4.4 Will the provenance of the data and other research outputs be thoroughly documented using the appropriate standards?

- Yes
- No



2.4.5 Describe all relevant data quality assurance processes.

### 3. Other research outputs

3.1 Do you have any additional information, that was not addressed in the previous sections, which you wish to provide regarding other research outputs that are generated or re-used throughout the project?

*Guidance:*

Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

### 4. Allocation of resources

4.1 What will the costs be for making data and other research outputs FAIR in your project?

*Guidance:*

Consider direct and indirect costs related to storage, archiving, re-use, security, etc.

4.2 How will these be covered?

*Guidance:*

Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions).

4.3 Who will be responsible for data management in your project?

4.4 How will long term preservation be ensured?

*Guidance:*

Discuss the necessary resources to accomplish this: costs and potential value, who decides and how, what data will be kept and for how long.

### 5. Data security

5.1 What provisions are or will be in place for data security?

*Guidance:*

Include data recovery as well as secure storage/archiving and transfer of sensitive data.

5.2 Will the data be safely stored in trusted repositories for long term preservation and curation?

- Yes
- No

### 6. Ethics

6.1 Are there, or could there be, any ethics or legal issues that can have an impact on data sharing?

- Yes
- No

*Guidance:*

Ethics or legal issues 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).

6.2 Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

- Yes
- No
- Not applicable

## **7. Other issues**

7.1 Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

- Yes: describe below
- No

## **European Commission (Horizon): Horizon Europe DMP - GDPR record**

### **GDPR record**

Have you registered personal data processing activities for this project?

- Yes
- No
- Not applicable

## **European Commission (Horizon): Horizon Europe DMP - DPIA**

### **DPIA**

Have you performed a DPIA for the personal data processing activities for this project?

- Yes
- No
- Not applicable