ESFRI-FED
PROGRAMME

IN SUPPORT OF THE VALORISATION OF BELGIAN FEDERAL

COMPONENTS IN ESFRI DISTRIBUTED AND VIRTUAL INFRASTRUCTURES
PHASE 1 – [2021-2025]

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

DATA MANAGEMENT PLAN : [ACRONYM]
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

## **INTRODUCTION**

##### **WHAT IS UNDERSTOOD AS RESEARCH DATA?**

Research data are the evidence that underpin the answer to research questions and can be used to validate findings. Data can be quantitative information or qualitative statements collected by researchers in the course of their work by experimentation, observation, modelling, interview or other methods, or information derived from existing evidence.

For the purpose of BELSPO's data management policy, research data also includes digital information extracted from physical objects such as scientific and archaeological collections, physical artworks or biobanks.

Software is not included in the definition. However, BELSPO recognises that software (algorithms, scripts and codes developed by researchers in the course of their work) may be necessary to access and interpret data. In such cases, the data management plan needs to address how information about such items will be made available.

##### **WHY IS A DATA MANAGEMENT PLAN NECESSARY?**

Data Management Plans document and sustain your research project by explaining how it deals with copyright / open access requirements and ethical issues, and describe the plan for acquisition, long-term data preservation and sharing modes. They contribute to increasing the impact and visibility of your research data and ensure that the way you are handling data complies with the Open Data principle applied by BELSPO.

##### **WHAT IS EXPECTED FROM THE DATA MANAGEMENT PLAN?**

The Data Management Plan (DMP) should describe how you as a researcher **deal with the collected data before, during and after the project.** It is a key element of a good data management.

As part of making research data findable, accessible, interoperable and re-usable (**FAIR**), the DMP shall include information on:

* how the data will be collected,
* the type, size and format of the generated data,
* when, where and in what format the data will be made accessible,
* how the data will be curated and preserved for ulterior use (including after the end of the project).

It will clearly specify which categories of users are likely to benefit from access to the data.

The DPM must also contain information regarding the legal and ethical aspects of data.

In this respect, researchers shall use to the maximum existing platforms having the highest standard of preservation, curation, deposit and reuse.

## **8.1. Collected, reused and/or generated data**

* ***Data content:*** *Refers to the type of data regarding what it contains. E.g. numeric (databases, spread sheets), textual (documents), image, audio, video, mixed media…*
* ***Data format:*** *Refers to the technical format of data; to the way in which the data is encoded for storage, often reflected by the filename extension. For example: pdf, xls, doc, txt, rdf…
Whenever possible, give preference to open and standard formats as they facilitate sharing and long-term re-use of data.*
* ***Data volume:*** *You may roughly estimate this using the following categories: From 0 – 10GB; From 10 – 100 GB; From 100 – 1000 GB; More than 1000 GB.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A. My proposal will: | [ ]  | Collect data | [ ]  | Reuse existing data | [ ]  | Generate new data |

*Please describe:*

* *Which data you will collect/reuse/generate*
* *How data will be collected / from which source it will be reused / how will it be generated*
* *Its content, technical format and estimated volume.*
* *Any existing constraints regarding its use.*

|  |
| --- |
|  |

## **8.2. Legal issues**

***Legal issues*** *include* ***personal data*** *and* ***intellectual property*** *issues. Regarding personal data, you must ensure when dealing with personal data that* ***Data Protection Laws*** *(i.e. GDPR) are complied with.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. My proposal will use / process / store **personal** data | [ ]  | Yes | [ ]  | No |

*If your answer is ‘yes’: shortly describe the kind of personal data.*

*Add the process and reference to your file in your host institution’s privacy register.*

|  |
| --- |
|  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| B. The work undertaken in the project will possibly result in **research data** with potential for **technology transfer** and **valorisation** | [ ]  | Yes | [ ]  | No |

*If your answer is ‘yes’, your proposal must consider possible intellectual property issues.*

*Explain who will be the owner of the data (who will have the rights to control access).
Indicate whether there will be intellectual property rights/restrictions for the data you created, and if applicable, describe how these will be managed.*

|  |
| --- |
|  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| C. Will agreements with 3rd parties **restrict the dissemination or exploitation** of the data the project will (re)use? | [ ]  | Yes | [ ]  | No |

*If your answer is ‘yes’: explain which data are affected by this agreement. State the restrictions that are in place.*

|  |
| --- |
|  |

## **8.3. Documenting the data of the project**

*Data must comply, as much as possible with* ***FAIR principles****; it must be findable, accessible, interoperable and reusable. For this purpose, data must be accompanied by descriptive information in the form of* ***metadata****. Metadata is the information that describes, explains, locates, and /or makes the use of an information source**easier to retrieve. Where metadata are in place, researchers are advised to use and* ***mention metadata standards****.*

|  |
| --- |
| A. What **documentation** will be provided **to enable understanding and reuse** of the data collected / generated in this project? |

|  |
| --- |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| B. **Metadata standards** will be used: | [ ]  | For all data | [ ]  | For some data | [ ]  | For none of the data |

* *If your answer is ‘for all data’ or ‘for some data’, please describe in detail which standards will be used.*
* *If your answer is ‘none of the data’, please state in detail which metadata will be created to make the data easy/easier to find and reuse.*

|  |
| --- |
|  |

## **8.4. Data storage and back-up during the project**

* *Give preference to the* ***use of robust, managed storage with automatic backup****, such as provided by IT support services of your* ***home institution****. Most research institutions have networked drives, which offer ample storage space and data security for most purposes.*
* *Consider* ***data protection, particularly if your data is sensitive*** *– for example, containing personal data, politically sensitive information or information relating to religion and health. If this is your case, enquire with your institution's research support staff whether your intended storage solution meets your institution's data security policy.*
* *Costs related to data* ***storage and backup during the project can be covered by the project budget*** *providing these are fully justified and relate to the project.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. The datawill be **stored** in... | [ ]  | Institution Networked Research Storage | [ ]  | Other |

*If your answer includes 'other':*

* *Specify which storage solutions you will use during the project, in addition to / instead of the institutional networked research storage.*
* *Explain the reasons for using these solutions. E.g. because you need more space than offered by your institution; to facilitate data sharing with collaborators; or because your data requires additional security.*

|  |
| --- |
|  |

|  |
| --- |
| B. How will the data be **backed up**? |

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| C. How will **data security** and protection of sensitive data be taken care of during the project? | [ ]  | Not applicable (there are no sensitive data) |
| [ ]  | Default security of the institution networked research storage |
| [ ]  | Additional security measures |

*If your answer is other than ‘Not applicable’: Describe the main risks and how these will be managed.*

|  |
| --- |
|  |

|  |
| --- |
| D. What are the expected costs for data storage and back-up during the project? How will they be covered? |

|  |
| --- |
|  |

## **8.5. Data preservation in the long term - after the project has ended**

* *BELSPO expects the data generated during the project to be preserved (archived) in the long term, in as far as legal and contractual agreements allow. As rule of thumb,* ***long-term storage is considered to be at least 10 years****, unless legal provisions or discipline-specific guidelines dictate otherwise.*
* *Costs related to data* ***preservation in the long term can be covered by the project budget*** *providing these are fully justified and relate to the project.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. All the data will be **preserved** in the long term (at least 10 years) | [ ]  | Yes | [ ]  | No |

*If your answer is ‘NO’: clearly describe what data will be preserved long-term and what data will be destroyed for contractual, legal or regulatory purposes, or for physical preservation issues. Indicate how you will decide which data to keep.*

|  |
| --- |
|  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| B. The datawill be **archived** within... | [ ]  | Institution Networked Research Storage | [ ]  | Other |

*If your answer includes 'other':*

* *Specify which storage solutions you will use in the long term, in addition to / instead of the institutional networked research storage.*
* *Explain the reasons for using these solutions. E.g. because you need more space than offered by your institution; to facilitate data sharing with collaborators; or because your data requires additional security.*

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| C. How will **data security** and protection of sensitive data be taken care of in the long term? | [ ]  | Not applicable (there are no sensitive data) |
| [ ]  | Default security of the institution networked research storage |
| [ ]  | Additional security measures |

*If your answer is other than ‘Not applicable’: Describe the main risks and how these will be managed. Inquire with your institution's research support staff whether your intended storage solution meets your institution's data security policy if your research involves sensitive data.*

|  |
| --- |
|  |

|  |
| --- |
| D. What are the expected costs for data **preservation in the long term**? How will these costs be covered? |

|  |
| --- |
|  |

## **8.6. Data sharing and reuse**

* *Data must comply, as much as possible with* ***FAIR principles****; it must be findable, accessible, interoperable and reusable.*
* *Data sets will be* ***linked to the scientific publication*** *they underpin, and which have either been deposited in,* ***or linked to Orfeo****, BELSPO's central Open Access Repository for publications.*
* *The* ***data available for sharing and reuse may differ from the preserved data****, since there may be legal, IP, privacy or security related reasons preventing or restricting the access to data, or lacking of space for large data volumes to deposit them in a repository in its entirety. This could be the case for part or the entirety of data; in the short, mid or long term.*
* *For* ***data requiring protection, BELSPO therefore observes the “as open as possible, as closed as necessary”*** *principle. A staged approach will provide access for communities of certified users, adapting the degree of certification of users to the sensitivity of the data.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A. Are there any **factors restricting or preventing the sharing** or reuse of the data (e.g. agreements with 3rd parties)? | [ ]  | Yes | [ ]  | No |

*If your answer is ‘YES’: explain which data are affected by this agreement. State the restrictions that are in place.*

|  |
| --- |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| B. **Which** data will be made **available to the public**? | [ ]  | All | [ ]  | Some part | [ ]  | None |

*If your answer is ‘some part’ or ‘none’:*

* *Indicate the restrictions on the sharing of the data (why can’t it be shared)*
* *Explain what data sharing agreement will be implemented*
* *Explain what actions will be taken to overcome or to minimize restrictions.*

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| C. **Where/how** will data be made **available to the public**? | [ ]  | Open Access repository |
| [ ]  | In a restricted access repository |
| [ ]  | Upon request by mail |
|  | [ ]  | Other (specify) |

*If your answer is other than ‘Open Access repository’: Indicate where and how access will be provided.*

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| D. **When** will data be made **available to the public**? | [ ]  | As soon as corresponding communications are published |
| [ ]  | After the project is finished |
| [ ]  | After the completion of the project (with embargo) |

*If your answer is other than ‘as soon as corresponding communication(s) are published’: Indicate the reasons for the restrictions on the time release of data (embargo periods). For example, to publish, protect intellectual properties, or seek patents.*

|  |
| --- |
|  |

|  |
| --- |
| E. **Who** will be able to **access the data** and under which conditions? |

|  |
| --- |
|  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| F. **Which** data will be made **available for re-use**? | [ ]  | All | [ ]  | Some part | [ ]  | None |

*If your answer is ‘some part’ or ‘none’: Indicate the restrictions on the re-use of the data. Explain what actions could be taken to overcome or to minimize restrictions.*

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| G. Under what **license** will be data shared **for re-use**? | [ ]  | Creative Commons CCO |
| [ ]  | Creative Commons CC-BY |
| [ ]  | Other (specify) |

*If your answer is ’other’ : Indicate which license will the data have for reuse, and why.*

|  |
| --- |
|  |

|  |
| --- |
| H. What are the **expected costs** for data sharing? How will these costs be covered? |

|  |
| --- |
|  |

## **8.7. Responsibilities**

*By default, the* ***Principal Investigator (PI)*** *bears the overall responsibility for updating & implementing this DMP.*

|  |
| --- |
| A. Who will be responsible for the **data documentation & metadata**? |

|  |
| --- |
|  |

|  |
| --- |
| B. Who will be responsible for **data storage & back up** during the project? |

|  |
| --- |
|  |

|  |
| --- |
| C. Who will be responsible for ensuring **data preservation and sharing**? |

|  |
| --- |
|  |

|  |
| --- |
| D. Who bears the end responsibility for **updating & implementing this DMP**? |

|  |
| --- |
|  |