

 **Sectorial Investigación. Bibliotecas**

**Plantilla Data Management Plan – FAIR DATA - Horizonte 2020**

1. **Data summary**

Provide A short introduction text explaining the purpose of the data collection/generation., a summary of the data addressing the following issues:

* Purpose of the data collection/generation and its relation to the objectives of the project
* [Explain the relation to the objectives of the project](https://dmp.csuc.cat/projects/my-plan-horizon-2020-fair-dmp--137/plans/269/edit%22%20%5Cl%20%22collapse-guidance-643-659-67)
* Types and formats of data that the project generate or collect (e.g.: quantitative, qualitative; generated from surveys, clinical measurements, interviews, medical records, electronic health records, administrative records, genotype data, images, tissue samples, experimental records / reports / designs / clinical validation).
* Are you reusing existing data?
* Origin of the data
* Expected size of the data
* To whom might it be useful -data utility-

Examples:

1. It is not envisaged that there will be any privacy issues with respect to the data as there is no personal data.
2. Mendeley database (.ris); MS Excel spreadsheet (.xlsx and comma-delimited .csv); MS Word for text documents (.docx); Microsoft Word 2007 for text-based documents. MP3 or WAV for audio files. Windows Media Video for video files. Quantitative data analysis will be stored in SAV file format (used by SPSS) from which data can be extracted using the open-source spss read Perl script. These file formats have been chosen because they are accepted standards and in widespread use. Files will be converted to open file formats where possible for long-term storage.
3. The research objectives require qualitative data that are not available from other sources. Some data exist that can be used to situate the findings of the proposed research and which will supplement data collection as part of the proposed research. Some quantitative data are available, but they are insufficiently detailed. In their current form, they would not permit as full a comparison across the cases as is desirable.
4. Data are collected through online questionnaires, structured observations and prototyping.
5. Existing sources of data: group previous research, academic collaborators, publicly available database/archive on x repository/archive/web.

**2. FAIR Data**

**2.1 Making data findable, including provisions of metadata**

* **Discoverability** of data/metadata standards: Reference to existing suitable standards of the discipline. If these do not exist, an outline on how and what metadata will be created. Use of persistent and unique identifiers.
* **Naming conventions** and clear **version numbers**
* Search **keywords** provided for reuse possibilities

Examples:

1. Generic metadata standard: Dublin core <http://dublincore.org> (used by the Repositori Digital de la UPF).
2. Once the data set is created and stored in the Repository, a persistent and unique identifier (handle) will be generated automatically.
3. No definite standards have been identified yet, the project X will investigate the existence of relevant available standards and metadata schemes.

**2.2 Making data openly accessible**

* **Which data** will be made openly available? Explain if certain datasets cannot be shared
* Specify **how** will the data be made accessible
* Specify **methods** or software tools to access the data: documentation, source code?
* **Where** the data and associated metadata, documentation and code will be deposited?
* **How** the access will be provided in case there are **restrictions**?

Examples:

1. The data generated under this project and information supporting information preservation and reuse will be deposited in RepositorY ….
2. Data which is able to be made publicly available will be offered to the XX Data archive at the end of the project.
3. Research papers written and published during the funding period will be made available with a subset of the (anonymised) data necessary to verify the research findings.
4. Sensitive data will not be publicly available, according to data protection law.
5. Data are not in English, which will limit the accessibility of the data.

|  |
| --- |
| → When processing personal data remember that you must abide by the specific European and Spanish legislation  |

**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

* Use of **standard vocabulary** for all the datasets to allow interoperability or mapping to the more commonly used ontologies

Examples:

1. The documents are based on XML according to a DTD. The vocabulary is represented in SKOS. The RDF data is based on an OWL ontology.
2. The data, metadata and documentation are compliant to disciplinary standards for easy interoperability and reuse.
3. The deposited data will be structured in compliance with community agreed, domain-specific data standards (when available) to ensure interoperability and reuse beyond the original purpose for which they were created. Information on tools and instruments need to use the data or to reproduce and validate results produced from them will be provided via the repository.

**2.4 Increase data reuse (through clarifying licenses)**

* How will the data be **licensed** for reuse?
* When will the data be **made available** for reuse?
* **Length of time** for which the data will remain reusable
* Describe data **quality assurance processes**

Examples:

1. The research data from this project will be issued under a Creative Commons Attribution (CC-by) license, meaning others can use, adapt and build upon our work non-commercially, as long as they credit us and license their new creations under the identical terms.
2. The dataset is made available under the Creative Commons CC0 public domain dedication (https://creativecommons.org/publicdomain/zero/1.0/)

**3. Allocation of resources**

Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered.

* Estimate the **costs** **for making your data FAIR** and how you intend to cover these costs. Will additional resources be needed to prepare data for deposit or meet charges from data repositories?
* Identify **responsibilities** for data management in your project
* Describe **costs** and potential value for **long-term preservation**

Examples:

1. The estimated costs for making data FAIR are XXX€, and they will be covered by the project itself, mainly by the budgeted personnel costs.
2. Currently there are no costs to provide FAIR data.
3. XXXX will be responsible for data management plan updates. XXXX will be responsible for data backup and storage and the project leader will be responsible for data archiving and publications in the repository.

**4. Data security**

* Address **data recovery** as well as **secure storage** and transfer of **sensitive data**

Examples:

1. Sensitive data will be separated as early as possible to create an anonymized dataset. Access to sensitive data is granted only for project members.
2. Data transfer is secured via https protocol.
3. Static data is backed up Two times a day (00:00, 180:00), locally, and retained for one week. Granularity: 7 (7x1)
4. Backups are processed automatically based on snapshot technology on a time-scheduled basis.

**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 with **personal data**?

Examples:

1. Informed consent for data sharing and long-term preservation is included during data collection. Sensitive data will be separated as soon as possible and kept secure.
2. The data management in this project do not affect to ethical issues.

|  |
| --- |
| → When processing personal data remember that you must abide by the specific European and Spanish legislation  |

**6. Other issues**

* Refer to **other** national/funder/sectorial/departmental **procedures for data management** that you are using (if any).

Example:

There are no further national, sectorial or departmental procedures for data management will be used.

**7. Further support in developing your DMP**

* Explain who has **supported** you in the creation of the DMP, if any.

Example:

* This DMP has been created on the basis of the g9 Library Group Template

|  |
| --- |
| Adaptado de [DMP word template (english)](http://guiesbibtic.upf.edu/ld.php?content_id=22376040) de la Biblioteca UPF, http://guiesbibtic.upf.edu/data/es/pgdatos, subjecta a una llicència de [Reconeixement-NoComercial-CompartirIgual 4.0 Internacional de Creative Commons](http://creativecommons.org/licenses/by-nc-sa/4.0/) |