Data Management Plan (DMP)

ICT PhD - 12 June 2024

Ruth Íñigo
ruth.inigo@upc.edu
The Research Plan will include a **Data Management Plan**

- Requirement for PhD students who joined any programme from 2023/2024 and onwards
- It has to be delivered during the first academic year from the first registration
a Data Management Plan (DMP) is...

A methodological document that describes the life cycle of the data collected, generated and processed during a research project, a doctoral thesis, etc.

DMPs are required by research funding agencies
A DMP contains information about...

- **What data** will be obtained to develop the research (text, numbers, image, audio, video, code, etc.)

- **How data will be**…
  - collected
  - organized
  - processed and analyzed
  - stored during the project
  - preserved when the project is over
  - shared and accessible
  - ready to be reused
Data Management Plan (DMP)

For PhD candidates and researchers a DMP can be useful as…

- a scheme to organize the data produced
- a way to document not only the data to be stored but a part of the research process
- a guide on how to reuse data in subsequent projects
- a way to prevent data loss
You can use **eiNa DMP**: a tool to create your Data Management Plan, developed by CSUC

- **Templates for PhD students are available in**
  - English; Catalan; Spanish

- Optional review at the end of the process
## Data Management Plan (DMP): Contents

### 0. General information

<table>
<thead>
<tr>
<th>General description</th>
<th>Of the expected data to manage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding Project</td>
<td>In case the thesis is related to a research project</td>
</tr>
<tr>
<td></td>
<td>Agency and grant number</td>
</tr>
<tr>
<td>Sensitive /personal data</td>
<td>The DMP of a project must be submitted to the <a href="http://schiller.com/projects/soundmanager/">UPC Ethics Committee</a> for review and approval, when it involves certain ethical issues (for example):</td>
</tr>
<tr>
<td></td>
<td>○ Fundamental systems or techniques in artificial intelligence (AI)</td>
</tr>
<tr>
<td></td>
<td>○ Special categories of personal data (ethnic or racial origin, genetic data, biometric data, data related to health, political)</td>
</tr>
<tr>
<td></td>
<td>○ Groups of vulnerable people (minors, people with disabilities...) etc.</td>
</tr>
</tbody>
</table>
## Data Management Plan (DMP): Contents

### 1. Data Collection

| Origin and description of data | ● Own data?  
● From a third party? (specify licenses) |
|-------------------------------|----------------------------------|
| Data type and formats         | ● Experimental; simulation, code; etc.  
● Spreadsheets, text, images, audio, video, etc.  
● **Preferably open formats: csv; txt; ods; etc. to facilitate preservation and re-use** (see [info on FAIR data](#)) |
| Data volume expected (at the end of the research) | ● Approximately  
● Datasets higher than 100 GB are considered large size (special authorization in CORA.RDR) |
## Data Management Plan (DMP): Contents

### 2. Data storage and security

| Restrictions that may affect data storage and security | • Contractual (agreements with third parts)  
| | • Ethical (personal data)  
| | • Commercial (patents)  
| | • Are special security measures needed? |

| Major data security risks | • Accidental deletion of data  
| | • Loss / theft  
| | • Consequences? |

| Storage of data during the research project | • Physical storage?; university infrastructure?, commercial cloud? |
## Data Management Plan (DMP): Contents

### 3. Data documentation

<table>
<thead>
<tr>
<th>Name and structure of the files and the folders</th>
<th>● File and folders naming</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version control</td>
<td>● Versions management</td>
</tr>
<tr>
<td></td>
<td>● What versions to keep?</td>
</tr>
<tr>
<td>Metadata standards</td>
<td>● To describe the datasets, if deposited</td>
</tr>
<tr>
<td></td>
<td>○ Geospatial; Astronomical; Biological; etc.</td>
</tr>
<tr>
<td></td>
<td>○ Dublin Core (by default)</td>
</tr>
</tbody>
</table>
## Data Management Plan (DMP): Contents

### 4. Access, share and reuse the data

| Restrictions on data sharing | ● General Data Protection Regulations (GDPR)  
● Ethics  
● Intellectual property or copyright  
● Commercial  
● Security |
| --- | --- |
| The potential users of data | ● You and/or other researchers in your RG / Department  
● Other researchers worldwide?  
● Industry? |
| Licenses that you will apply to the data to enable reuse | ● CC: CC-BY, CC-0  
● Code licenses: GNU, MIT, etc. |
### Data Management Plan (DMP): Contents

5. Deposit and conservation of the data

| Selection for long term conservation. What criteria to select files? | ● Relevance  
● Data linked to publications  
● Reusability |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>For how long the data would have to be preserved?</td>
<td>● International standards recommend a minimum of 10 years</td>
</tr>
</tbody>
</table>
| In which repository? | ● Consider volume of data, format; subject, etc.  
● Type: Institutional; Multidisciplinary; Specialized; Commercial |
Publishing research data

Publishing datasets in open access is not a requirement for doctoral students

You can do it optionally
Where to deposit data?

- Dataverse platform
- Free (gratis)
- Multidisciplinary
- Immediate DOI
- Datasets up to 100GB
- Requires a registration
- Metadata are reviewed by library staff

https://dataverse.csuc.cat/
Where to deposit data?: Other

- Find more information about how to publish in other repositories
- A comparative of repositories
Data Management Plan (DMP) and Open Science at UPC

2025 Roadmap for UPC Open Science

Governing Council Decision CG/2023/07/08, of 5 July 2023,
Library resources for PhD students

https://bibliotecnia.upc.edu/doctorands
Thanks!

All the information about DMP in Bibliotecnica:


Contact us: biblioteca.ferrate@upc.edu
ruth.inigo@upc.edu