Data management plan (DMP)

[projectname]

[project acronym]

|  |  |  |
| --- | --- | --- |
| Version | Effective date | Description of document/changes |
| 1.0 | [datever1] | First version of DMP – created for start of project (deliverable x.y) |
| 2.0 | XY/XY/XYXY | Second version of DMP – prepared for midterm review (deliverable x.z) |
|  |  |  |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| Level of distribution |  | This DMP is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/legalcode) (CC BY 4.0).  DOI: [xxx] |

**Project details**

|  |  |
| --- | --- |
| Project Coordinator Principal Investigator | [coordinatorname], [coordinatormail], [coordinatorid],  [coordinatoraffiliation], [coordinatorror] |
| Contact person (responsible for data management and DMP) | [contactname],[contactmail],[contactid],[contactaffiliation], [contactror] |
| DMP contributors | [contributors] |
| Start date | [startdate] |
| End date | [enddate] |
| Funding programme,  grant number | [grantid] |
| Internal project number  TU Wien | […] |

**List of acronyms**

|  |  |
| --- | --- |
| DMP | data management plan |
| RDM | research data management |
| … |  |
| … |  |
| … |  |
| … |  |

Content

**Inhaltsverzeichnis**

[Introduction 4](#_Toc66691660)

[Science Europe practical guide, FAIR data 4](#_Toc66691661)

[Relevant Policies and Guidelines 4](#_Toc66691662)

[1. Data description 4](#_Toc66691663)

[1a List of data sets that will be reused or produced 4](#_Toc66691664)

[1b Data generation and reuse 5](#_Toc66691665)

[2. Documentation and data quality 5](#_Toc66691666)

[2a Data organisation, metadata and documentation 5](#_Toc66691667)

[2b Data quality control 5](#_Toc66691668)

[3. Storage and backup during research process 6](#_Toc66691669)

[3a Storage and backup facilities 6](#_Toc66691670)

[3b Data security and protection of sensitive data 6](#_Toc66691671)

[4. Legal and ethical requirements 7](#_Toc66691672)

[4a Personal data 7](#_Toc66691673)

[4b Intellectual property rights and ownership 7](#_Toc66691674)

[4c Ethical issues 7](#_Toc66691675)

[5. Data sharing and long-term preservation 7](#_Toc66691676)

[5a Data publication, restrictions and embargo periods 7](#_Toc66691677)

[5b Long-term preservation and reusability 8](#_Toc66691678)

[6. RDM responsibilities and resources 8](#_Toc66691679)

[6a RDM-roles and responsibilities 8](#_Toc66691680)

[6b Resources 8](#_Toc66691681)

# Introduction

### Science Europe practical guide, FAIR data

A DMP is a structured document that keeps record of what research data is created and what happens to that data during and after a project. It helps with planning the research process and defining responsibilities in a research project involving several researchers or institutions.

For writing this DMP, we followed [the recommendations of Science Europe](https://www.scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/) as they reflect the guidelines agreed upon by the major funders in Europe.

To make our data FAIR, they generally will be treated according to the following criteria:

* We will make our data findable, by uploading it to a data repository that provides a persistent identifier, and adding relevant metadata.
* We will make our data accessible by providing open access to data, wherever possible. In cases, where open access is not possible, we will provide meaningful metadata plus contact information for access requests.
* We will make our data interoperable by providing and describing data in a way that is common within our domain by using the same file formats, schemas and vocabularies. We will provide good documentation for all our datasets.
* We will make our data reusable by adding metadata and comprehensive Readme files to all published data sets. The descriptions include details on the methodology used, analytical and procedural information. In case of publication, licenses for code and data will always be assigned and clearly marked.

### Relevant Policies and Guidelines

* TU Wien Policy for Research Data Management: <https://www.tuwien.at/index.php?eID=dms&s=4&path=Directives%20and%20Regulations%20of%20the%20Rectorate/Policy%20for%20Research%20Data%20Management.pdf>
* Directives and Regulations of the TU Wien Rectorate: <https://www.tuwien.at/en/tu-wien/organisation/central-divisions/data-protection-and-document-management/directives-regulations/>
* TU Wien Data Protection: <https://www.tuwien.at/en/tu-wien/organisation/central-divisions/data-protection-and-document-management/data-protection-at-tu-wien>
* European Commission’s document on Ethics and Data Protection: <https://ec.europa.eu/info/sites/info/files/5._h2020_ethics_and_data_protection.pdf>
* [Other (e.g. from a project partner)]

# Data description

### 1a List of data sets that will be reused or produced

Produced data sets

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataset ID | name | type | format | estimated volume | contains sensitive data |
| P1 | [dataset1name] | [dataset1type] | [dataset1format] | [dataset1vol] | yes/no |
| x | x | x | x | x | x |

Reused data sets

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| dataset ID | name | type | format | estimated volume | contains sensitive data |
| R1 | … | … | … | … | yes/no |
| R2 |  |  |  |  |  |
| R3 |  |  |  |  |  |

### 1b Data generation and reuse

**Data generation**

Dataset P1 will be collected in the laboratory by performing XYZ. …

Dataset P2 will be generated by different microscopes. …

Datasets P3 and P4 are results of simulations done using the VSC. …

**Data reuse**

Dataset R1 are statistics data of Statistik Austria on…, that will be used for a secondary analysis. The data are evaluated with the help of the statistics program R.

…

# Documentation and data quality

### 2a Data organisation, metadata and documentation

The filenames will follow the projects naming convention as defined in document [ABC] and include a timestamp of creation. Version control is automated.

As there are no domain specific metadata standards applicable, we will provide a README file with an explanation of all values and terms used next to each file with data. Additionally, we will provide common metadata such as title, description or keywords when publishing data in open access repositories. In such a case, we will follow the default template provided by the repository, such as Data Cite Metadata or Dublin Core.

A far as possible, we will use controlled vocabularies for our data to allow inter-disciplinary interoperability and machine-actionability.

### 2b Data quality control

Data quality checks will be done, e.g. checks of consistency of labels, logical errors in the data, data curation, and version control.

# Storage and backup during research process

### 3a Storage and backup facilities

For the duration of the project, storage and backup of data will be ensured by the project manager in cooperation with the responsible representative of TU.it.

**[Selection to choose from in Damap:]**

[All data sets/dataset P1, P2 and R2] will be stored on [TUproCloud](https://www.it.tuwien.ac.at/en/services/cooperation-and-communication/collaboration/tuprocloud-sync-and-share-for-projects), an owncloud-service operated by TU.it. Owncloud-data are kept independently and redundantly available at two different locations. This provides a daily disaster recovery backup and guarantees availability of 99,999%. (i.e. 5 min. service interruption per year). Only authorized staff members and project partners will have access.

[All data sets/dataset P1, P2 and R2] will be stored on [TUownCloud](https://www.it.tuwien.ac.at/en/services/cooperation-and-communication/data-storage/tuowncloud-sync-and-share), a cloud-service operated by TU.it. TUownCloud-data are kept independently and redundantly available at two different locations. This provides a daily disaster recovery backup and guarantees availability of 99,999%. (i.e. 5 min. service interruption per year). Only authorized staff members and project partners will have access.

[All data sets/dataset P1, P2 and R2] will be stored on [TUfiles](https://www.it.tuwien.ac.at/en/services/cooperation-and-communication/data-storage/tufiles-file-service-for-employees), central and readily available network drive with backup operated by TU.it. TUfiles is suitable for storing data with moderate access requirements, but high availability demands that allows full control of allocating authorisations. TUfiles-data are kept independently and redundantly available at two different locations. This provides a daily disaster recovery backup and guarantees availability of 99,999%. (i.e. 5 min. service interruption per year). Only authorized staff members will have access.

[All data sets/dataset P1, P2 and R2] will be stored on a server provided by TU.it ([Server Housing](https://www.it.tuwien.ac.at/en/services/network-infrastructure-and-server-services/server-services/server-housing)). The server is housed in one of four dedicated TU.it server rooms with limited access, and operated by our Institute.

For virtualisation, we will be using [TUhost (Virtual Server Housing)](https://www.it.tuwien.ac.at/en/services/network-infrastructure-and-server-services/server-services/tuhost-virtual-server-hosting), the central and highly available TU.it virtualisation platform, hosted on VMware ESXi. Hardware. Storage and backup software will be provided by TU.it, and our institute will be responsible for the server operation. Ask for information on backup routines and access control in Damap?

For distributed version management of files (especially of source code), we will be using [TUgitLab](https://www.it.tuwien.ac.at/en/services/cooperation-and-communication/data-storage/tugitlab), an application for managing repositories based on Git provided and managed by TU.it. Our institute’s administrators will manage GitLab groups, assign project permissions, and assign external project partners as additional GitLab users. This service is highly available and scalable on the Kubernetes platform.

[other]: free text in Damap

### 3b Data security and protection of sensitive data

All incidents will be handled individually by an incident response team that is maintaining the affected service.

Access to the data during research:

|  |  |  |  |
| --- | --- | --- | --- |
| dataset ID | selected project members | all other project members | the public |
| P1 | writing | reading only | no access |
| P2 | writing | writing | reading only |
| P3 | writing | reading only | no access |
| P4 | writing | no access | no access |
| R1 | reading only | reading only | reading only |

“Project members will not store sensitive data on computers in the lab or external hard drives connected to those computers. They will not carry data with them (e.g. on laptops, USB sticks, or other external media). All data centers where project data is stored have sufficient certifications. All project web services are addressed via secure http (https://...). Project members have been instructed about both generic and specific risks to the project.”

We pay strict attention to compliance with the relevant institutional and national data protection policies listed in the introduction of this document.

# Legal and ethical requirements

### 4a Personal data

At this stage, it is not foreseen to process any personal data in the project. If this changes, advice will be sought from the data protection specialist at TU Wien (Verena Dolovai), and the DMP will be updated.

### 4b Intellectual property rights and ownership

TU Wien will be the owner of the data generated and have the rights to control access. Further details will be covered in the consortium agreement.

The digital research data obtained will be published Open Access under a Creative Commons CC-BY license, provided that there are no data protection concerns. Further data will be made available with restrictive access.

### 4c Ethical issues

Ethical issues in the project have been identified and discussed with the Research Ethics Coordinator at TU Wien (<https://www.tuwien.at/en/research/rti-support/research-ethics/>). They relate to… and will be resolved by… Relevant ethical guidelines in this project are…

# Data sharing and long-term preservation

### 5a Data publication, restrictions and embargo periods

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ID | data and metadata will be shared | restrictions / embargo reasons | estimated publication date | location for publication (repository) | PID | license |
| 1 | [yes/ no/ partly/ metadata only] | According to the cooperation contract: exclusive use of the research data until one year after the end of the project | Underlying data: date of paper publication, all others: after project | TU data | DOI, provided by TU data | [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/legalcode) |
| 2 |  |  |  |  | DOI: … |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |

Blabla on table and selection process…

### 5b Long-term preservation and reusability

|  |  |  |  |
| --- | --- | --- | --- |
| ID | location for long-term storage (min. 10 years) | discipline specific metadata standard used  (if available) | foreseeable research uses and/or users |
| 1 | TU data |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |

[If applicable: Description of methods / software needed to access and use the data: …]

[If applicable: Description of protocol to access restricted data: …]

# RDM responsibilities and resources

### 6a RDM-roles and responsibilities

The [PI] / [data officer XY] will direct the data management process overall, with the research assistants responsible for ensuring metadata production, day-to-day cross-checks, back-up and other quality control activities are maintained. The lead country researchers will be responsible for routine supervision of the dataset development.

### 6b Resources

0.25 FTE are provided in the project for the appropriate preparation of the research data for preservation and publication. Preserving and publishing the data in the TU Wien data repository is free of charge.

Or table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Cost type | Description | Unit | Value |
| TUproCloud | File based storage |  | € | 250 |
| IPR | Legal advice | See details below | € | 1 300 |
| Data preparation, data management | Personnel | 0.25 FTE IVa for 3 years | € | 50 000 |
| … | Software license |  | € | … |
| Software development | Personnel | 0.5 FTE IVa for 2 years | € | 65 000 |
| **Estimated total costs** |  |  | **€** | **120 000** |

[Additional description (if required): …]