

Data Management Plan template

Name of student/researcher(s)	
Name of group/project	
Description of your research	Briefly summarise the type of your research to help others understand the purposes for which the data are being collected or created.
Funding body(ies)	
Grant number	A grant number provides unique identification for the grant.
Partner organisations	
Project duration	Start: MM-DD-YYYY End: MM-DD-YYYY
Date written	MM-DD-YYYY
Date last update	
Version	A new version of the DMP should be created whenever important changes to the project occur due to inclusion of new data sets, changes in consortium policies or external factors.
Name of researcher(s) with roles/responsibilities for data management	Naming anyone with specific roles and responsibilities for data management is especially important for collaborative projects that involve many researchers, institutions, and/or groups.

1. Data Collection

Can you describe the data you will be creating/collecting?

Checklist:

- *How will data be collected?*
- *What type of data will be collected? (Measurements, observations, models, software....)*
- *In what file formats?*
- *Will it be reproducible? What would happen if it gets lost or becomes unusable later?*
- *What is the estimated size of the data, and what growth rate?*
- *How do you handle version control to maintain all changes that are made to the data?*
- *Which tools or software are needed to create/process/visualize the data?*
- *Will you also use pre-existing data? From where?*

2. Data Storage and Back-up

How do you ensure that during your research all research data are stored securely and backed up or copied regularly?

Checklist:

- *How will the raw data be stored and backed up during the research?*
- *How will the processed data be stored and backed up during the research?*
- *Which storage medium will you use for your storage and backup strategy? Network storage; personal storage media (CDs, DVDs, USBs, portable hard drives); cloud storage*
- *Are backups made with sufficient frequency so that you can restore in the event of data loss?*
- *Are the data backed up at different locations?*

3. Data Documentation

How will your data be documented to help future users to understand and reuse it?

Checklist:

- *What standards will be used for documentation and metadata? If there is not a standard already available for your data, outline how and what metadata will be created.*
- *How will your data be documented during your research and for long-term storage?*
- *What directory and file naming convention will be used to enable the titling of your folders, documents and records in a consistent and logical way?*
- *What project and/or data identifiers will be assigned? (e.g. DOI or Digital Object Identifier)*

4. Data Access

How will you manage access and security?

Checklist:

- *How will you manage copyright and Intellectual Property Rights issues? E.g. Who owns the data? How will the data be licensed for reuse?*
- *Are there any limitations on the access of your data?*
- *What are the access criteria for the data (open/restricted access, embargo period, etc.)?*
- *Who controls data access (e.g. PI Principal Investigator, student, lab, University, funder) ?*

5. Data Sharing and Reuse

How will you share the data?

Checklist:

- *If you allow others to reuse your data, how will the data be shared? In case the dataset cannot be shared, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).*
- *Any sharing requirements (e.g., funder data sharing policy) ?*
- *Audience for reuse? Who will use it now? Who will use it later?*
- *When will you publish it and where? Is your data underlying a scientific publication?*
- *Which tools/software are needed to view/visualize/analyse the data?*

6. Data Preservation and Archiving

Which data should be preserved, and/or shared?

Checklist:

- *Which criteria will you use to decide which data has to be archived for preservation and long-term access. Which data has to be destroyed?*
- *How long should it be preserved (e.g., 3-5 years, 10-20 years, permanently)?*
- *What file formats? Are they long-lived?*
- *Which data repository is appropriate for archiving your data (subject-based or 3TU.Datacentrum)?*
- *What costs (if any) will your selected repository charge?*

Guidance

Data Management Plan

This template has been adapted from the 3TU.Datacentrum. The template may be appropriate for principal investigators collaborating in a research project or for research students working on a PhD or Masters project. The Data Management Plan consists of six sections. Each section is accompanied with a Checklist, providing the most important questions to be answered in your Data Management Plan. Furthermore, more detailed information on each section will show up when you 'open' the info button.

1. Data Collection

Give a short description of the data, including amount (if known) and content. If possible give a rough estimate of the number of files.

What type of data you're generating impacts how you will manage the data and how long it needs to be preserved. There are four main types of research data:

- **Observational data:** *captured in real time, typically cannot be reproduced exactly*

- **Experimental data:** *from labs and equipment, can often be reproduced but may be expensive to do so*

- **Simulation data:** *from models, can typically be reproduced if the input data is known*

- **Derived or compiled data:** *after data mining or statistical analysis has been done, can be reproduced if analysis is documented*

Data types *could include text, numbers, images, 3D models, software, audio files, video files, reports, surveys, etc.*

Provide information on the existence (or not) of similar data and the possibilities for integration and reuse.

File formats

In planning a research project, it's important that you consider which file formats you will use to store your data. In some cases, this will be dictated by the software you're using or the conventions of your discipline, but in other cases you may have to make a choice between several options. These are likely to be some of the key factors in your decision-making:

- *What software and formats you or colleagues have used in past projects.*
- *Any discipline-specific norms (and any peer support that comes with them).*
- *What software is compatible with hardware you already have.*
- *Whether you have funding for new software for the job.*
- *How you plan to analyse, sort, or store your data.*

But you should also consider:

- *What formats will be easiest to share with colleagues for future projects.*
- *What formats are at risk of obsolescence, because of new versions or their dependence on particular software and/or hardware*
- *What formats it will be possible to open and read in the future.*

• *What formats will be easiest to annotate with metadata so that you and others can interpret them days, months, or years in the future. In some cases, you may be best off using one format for data collection and analysis and converting your data to a standard format for archiving once your project is complete. After conversions, data should be checked for errors or changes that may be caused by the export process.*

Version control

Because digital research data can so easily be copied, over-written or changed, researchers need to take steps to protect its authenticity. Research time is wasted and valuable data put at risk if researchers work with outdated versions of files.

Version control can prevent this. Control is particularly important if data is being used by multiple members of a research team, or if research files are shared across different locations.

A regime to synchronize different copies or versions of files will improve research efficiency and help guarantee the authenticity of the data. Good practice generally involves the keeping of a single master file, to which all changes are recorded. Version control mechanisms should be established and documented before any data is collected or generated.

2. Data Storage and Back-up

It is the responsibility of the researcher to ensure that their research data is regularly backed-up and stored securely for the life of the project. It is good practice to store only what you need to keep and keep at least three copies of crucial data. It is recommended that data is stored on the University's networked file servers and copies kept on remote storage and /or portable storage.

Generally there are four options for data storage:

- **Networked drives:** *University fileserver – As these are secure and backed-up regularly, they are ideal for master copies of your research data.*
- **Local drives:** *PCs and Laptops – Data can be lost because local drives can fail, or the computer may be lost or stolen. These are convenient for short-term storage and data processing but should not be relied upon for storing master copies, unless backed-up regularly.*
- **Remote or Cloud storage** – *commonly used services, such as Dropbox and Google Drive, will not be appropriate for sensitive data, and their service level agreements should be studied before using them to store your research data.*
- **External portable storage devices** – *External hard drives, USB drives, DVDs and CDs. These are very convenient, being cheap and portable, but not recommended for long-term storage as their longevity is uncertain and they can be easily damaged.*

Consult the ITService for advice about secure storage options.

You may choose to only back up certain data, or to back up files you use every day more regularly than others. The basic rule of thumb is: The more

important the data and the more often they change, the more regularly they need to be backed up.

If your files take up a large amount of space and backing up all of them (or backing them up sufficiently frequently) would be difficult or expensive, you may want to focus on backing up specific key information, programs, algorithms, or documentations that you would need in order to re-create the data in case of data loss.

3. Data Documentation

Describe the types of documentation that will accompany the data to help secondary users to understand and reuse it. This should at least include basic details that will help people to find the data, including who created or contributed to the data, its title, and date of creation and under what conditions it can be accessed.

Documentation may also include details on the methodology used, analytical and procedural information, definitions of variables, vocabularies, units of measurement, any assumptions made, and the format and file type of the data. Consider how you will capture this information and where it will be recorded. Wherever possible you should identify and use existing community standards.

File naming

Organising your files and folders effectively and efficiently can save you time and make collaboration easier by ensuring you are working on the correct version of the data. A good file name makes it easy to identify, locate and retrieve your data. There is no one recommended way to name your files and folders, but you should name your files consistently. If you work as part of a research group, you should decide on a file and folder naming system with your colleagues.

Identifiers

An identifier is a reference number or name for a data object and forms a key part of your documentation and metadata. To be useful over the long-term, identifiers need to be unique (globally unique if possible) and persistent (the identifier should not change over time).

The emerging identifier standard for publicly available datasets is the Digital Object Identifiers (DOIs). Although DOIs have been traditionally used for journal articles, they can now be assigned to datasets.

4. Data Access

During the research project you will want to keep your research data safe and secure. You will want to determine who has access to your data and what they are authorised to do with it. Data security is needed to prevent unauthorised access or disclosure and changes to or destruction of data. The principle investigators are responsible for ensuring data security. The level of security required depends upon the nature of the data – personal or sensitive data need higher levels of security.

It is possible that you will need remote access to your data, if you are working from more than one location, or not at the university. A number of

individuals may require access to the data, possibly with different privileges to read, write, update or delete. This may be accomplished by using the Maastricht Dataverse Network of a Virtual Research Environment. The use of cloud storage to share data depends upon the level of security needed. It is possible that your project may need to arrange for access to third party data that may have specific limitations in how they can be distributed (based on IP or the agreement by which your project obtained the data). When your research project has received data under confidentiality or other restrictions you will have to identify and explain these restrictions in your data management plan.

Ownership of research data must be clarified prior to, or at the beginning of a project. Future storage and reuse are directly affected by the intellectual property rights of research data. Ownership of the data and copyrighted datasets will depend on whether the project was created as part of sponsored research; the employment status of the creator; whether third-party data has been utilised during the conduct of research, and, in case of an 'encoded work' whether substantial university resources were used in the creation of the encoded work.

The conditions under which the data may be made available by the data repository to other researchers are determined by you as the principal investigator depositing the data.

5. Data Sharing and Reuse

Your research is valuable and important, and so is the data that it is based on. By publishing your data, you make it available to the scholarly community, who can study and build upon your work. Your work will become more visible and typically be cited more frequently. At the end of your research project, your funder may require you to share your research data, by publishing it with no access restrictions (open access). Some journal publishers also require the data supporting the research article to be published.

When disseminating your data, you need to think about who would be interested in your research findings, and the way how to reach this audience (by newsletter, community website, press release, attending seminars or conferences, etc.). You also may need to think about how you want others to reuse your data.

If data is generated using specifically-developed software, it may be necessary to provide a copy of the software, noting operating requirements, with the data.

6. Data Preservation and Archiving

You need to decide what data to preserve and what data to dispose of after the end of the project. You, as researcher, may have your own view on how long you want to retain the data. This will be influenced by the discipline, the type of data created and whether further work or publications will be based on it. Where you are unsure what data might need to be held you should seek advice from your supervisor.

Data selected for long-term preservation will normally be submitted to a funder established data centre, disciplinary data repository or an institutional data repository. Most funders regard costs for archiving the data or preparing it for archive as allowable as long as they are justified and incurred within the life of the project.

For any questions arising during writing the data management plan, you can contact the Data Management Helpdesk [<http://www.maastrichtuniversity.nl/web/Library/datamanagementhelpdesk.htm>] for support and advice.