 

**WUR Data Management Plan template**

* Use this template to fill in a Data Management Plan for your research project.
* For PhD candidates: The Data Management Plan can be attached to your project proposal as an appendix.
* This template is based on Data Management Plan core requirements as defined by Science Europe: ‘[Practical Guide to the International Alignment of Research Data](https://www.scienceeurope.org/media/4brkxxe5/se_rdm_practical_guide_extended_final.pdf)’
* This template has been approved by NWO and ZonMw and can be used to fulfill the requirement of submitting a Data Management Plan to these research funders. Please, check the Appendix for specific guidelines for these funders.
* You are free to add topics to this template to better align with your project, however, the original topics must be retained.
* Right-click [info] and select ‘open hyperlink’ to additional information in the appendix.
* For more tips for filling in your Data Management Plan, take a look at [this document](http://edepot.wur.nl/458771).
* Questions? Contact data@wur.nl or visit the [Data Management website](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC.htm) for more information. For privacy questions contact your privacy officer or visit [the Privacy & Personal Data intranet page](https://intranet.wur.nl/umbraco/en/about-wur/policy-regulations/privacy-personal-data/).

1. Describe the research project [[info](#DescribeTheResearchProject_1)]

1a Describe the organizational context of your research project.

|  |  |
| --- | --- |
| Name researcher |   |
| DMP version (or date last modified) |   |
| Chair group / Business Unit |   |
| Graduate school  |   |
| Supervisor / (co-)promotors |   |
| Start date of project |   |
| Project number |   |
| Funding body |   |

1b. Give a short description of your research project

|  |  |
| --- | --- |
| Title |   |
| Summary |         |

1c. List the individuals responsible for the following data management tasks

|  |  |
| --- | --- |
| Data collection |   |
| Data quality |   |
| Storage and backup |   |
| Data archiving / publishing |   |
| Data stewardship / support |   |
| Any other role [......] |   |

1d. Name of data management support staff consulted during the preparation of this plan and date of consultation.

|  |
| --- |
|   |

## 2. Describe the data to be collected, software used, file formats and data size [[info](#DescribeDataSoftwareFileFormats_2)]

2a. Will you reuse existing data for this project?

☐ Yes. Please, specify below which data (e.g. DOI or url) and the terms of use (e.g. licence).

☐ No. Please, describe below any constraints to re-using existing data.

|  |
| --- |
|    |

2b. Will new data be produced?

☐ Yes. Please, describe the data you expect to generate, in terms of type software and format in the table below.

☐ No

|  |  |  |
| --- | --- | --- |
| Data type | Software  | (Open) file format |
| (e.g. numerical) | (e.g. Excel) | (e.g. .csv) |
|   |   |   |
|   |   |   |
|   |   |   |

2c. Estimate how much data storage you require in total

☐ 0-10 GB

☐ 10-100 GB

☐ 100-1000 GB

☐ >1000 GB

## 3. Storage of data and data documentation/metadata during research [[info](#DataStorageDuringResearch_3)]

3a. Where will the data and accompanying documentation/metadata be stored and backed up during the research project?

☐ W:drive (WUR network drive)

☐ OneDrive for Business (WUR cloud storage)

☐ SharePoint / Teams (WUR collaborative platform)

☐ Other (please specify location and back-up frequency below)

|  |
| --- |
|     |

## 4. Structuring your data and information [[info](#Structuring_4)]4a. Give a (visual) representation of the folder structure you intend to use.

|  |
| --- |
|  |

4b. Describe the file naming conventions you intend to use.

|  |
| --- |
|    |

4c. Describe the file versioning system you intend to use.

|  |
| --- |
|    |

## 5. Data documentation and data quality [[info](#Documentation_5)]

5a. Describe below what data documentation and metadata will accompany your data.

|  |
| --- |
|  |

5b. Describe what data quality controls will be used?

|  |
| --- |
|   |

6. Working with confidential data (personal data, ethics), data ownership, sharing and access [[info](#SensitiveData_6)]

6a. Are there reasons (privacy, ethics, contractual agreement, commercial interests, public security, IP rights) to restrict access to the data or limit which data will be made publicly available?

☐ Yes. Please, describe the reasons below.

☐ No

|  |
| --- |
|     |

6b. Will you process and/or store personal data during your research project?

☐ No, continue to question 6d.

☐ Yes. Please, specify below which measures you will take to ensure data protection and safeguard the privacy of the participants in your project.

|  |
| --- |
|  |

6c. Is this project registered in SmartPIA?

☐ No

☐ Yes

6d. Are there other ethical issues that need to be taken into account?

☐ No

☐ Yes

6e. Who has ownership and controls access over the data?

|  |
| --- |
|     |

6f. Will there be any intellectual property (IP) rights associated with the data?

☐ No

☐ Yes. Please, indicate below hoe IP rights will be managed.

|  |
| --- |
|      |

## 7. Data archiving and publishing [[info](#DataArchivingPublishing_7)]

7a. Do you have selection criteria, which allow you to determine which part of the data should be preserved once the project has ended?

☐ No

☐ Yes. Please, elaborate below.

|  |
| --- |
|     |

## 7b. What data will be archived internally (e.g. WUR network drive) for a minimum of 10 years?

☐ All (raw) data produced during the project will be archived internally.
☐ Other, please specify below.

|  |
| --- |
|     |

7c. What data will be published and made available for re-use via a data repository?
☐ All data produced during the project will be published in a data repository.
[ ]  Only the metadata is published in a data repository.
☐ Other, please specify below.

|  |
| --- |
|     |

7d. When will the data be available for re-use, and for how long will the data be available?
☐ Data available as soon as the article is published
☐ Data available upon completion of the project
☐ Data available after completion of the project (with embargo)
☐ Other, please specify below.

|  |
| --- |
|    |

7e. Which data repository do you intend to use to make the data findable and accessible?

|  |
| --- |
|    |

7f. Which metadata standard will be used to describe the data during archiving / depositing in a data repository?

|  |
| --- |
|  |

7g. Which licence/terms of use will be applied to the data?

|  |
| --- |
|   |

7h.If analysis software is generated in this project, describe your publishing strategy, below.

|  |
| --- |
|  |

## 8. Data management costs

8a. What resources (in time and/or money) will be dedicated to data management and ensuring that data is reusable?

|  |
| --- |
|    |

8b. If there are additional costs related to preparing the data for reuse, how will these costs be covered?

|  |
| --- |
|    |

**Appendix: additional information** *(may be deleted after completion)*

1. Describe the research project.

1a. Organisational context of your research project
A Data Management Plan should always include contextual information about the researcher and the project.

1b. Description of your research project
Giving a short description of your research helps the reader to understand your work and put the description of your data management into context.

1c. Data management responsibilities
Identifying persons who play a role in your daily data management practices helps clarify the data collection process. Identifying these roles are also important in the event you leave WUR, or on completion of your PhD. For data to be accessible for at least 10 years, responsibility for the archived data should lie with more than one person. Feel free to add roles.

1d. Data management support staff
You can consult with Data Management Support staff to assist with your Data Management Plan.

Is this project funded by NWO? If yes, NWO requires that researchers consult with Data Management Support at an early stage. Plans that have not been consulted with Data Management Support staff will not be considered by NWO. Please do so by contacting data@wur.nl before submitting.

2. Describe the data to be collected, software used, file formats and data size

2a. Re-use existing data
Think of the following potential reasons why the re-use of existing data was considered, but not implemented: is the data you want to re-use not publicly available or otherwise restricted in access? Would the costs be too high (e.g. acquiring specific software)? Is the data too big in size (e.g. TBs)?

2b. Producing new data
Data type refers to whether the data can be classified as for example textual, numeric, audio, film etc. Try to use software that let you also save files in an [open format](https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set_language=en), so that others can open the files (now and in the future) even if they don’t have the software.

3. Data storage during research
Ensure that your storage practices follow [WUR’s policy on data storage](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Data-policy/Storage.htm). Safe storage following the policy would be the W-drive or SharePoint / Team, possibly in combination with IT-approved cloud storage. Note that if you store data with a cloud service or on the university network, they are backed up automatically.

4. Structuring your data and information

4a. Folder structure
Write or visualise how you plan to organise your folders and files: which folders and subfolders will contain which files, and how you name these. Find an example and tips [here](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Organising-files-and-folders.htm).

4b. File naming conventions
Applying a consistent and descriptive file naming convention (i.e. a systematic file naming method) helps to

* identify the content of a (data)file without opening it
* easily and quickly locate, retrieve and filter (data)files, even if they have changed folders
* easily sort and browse through your (data)files
* identify missing (data)files

Find tips [here](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Organising-files-and-folders.htm).

4c. File versioning
Make sure that you have a system in place to keep track of file versions. A simple and effective system is to incorporate version numbers (e.g. v01, v02 etc.) in your file names. Additionally, depending on the type of files you work with, you could use systems that keep track of versions of individual files (e.g. Sharepoint) or multiple files (e.g. Git).

5. Data documentation and metadata
It is essential to systematically [document](https://wageningenur4.sharepoint.com/sites/WURResearchDataManagement/Gedeelde%20documenten/General/DMPonline-WUR/It%20is%20essential%20to%20systematically%20document%20your%20data%20during%20your%20research%20and%20when%20depositing%20your%20data%20set%20into%20an%20online%20repository.%20This%20allows%20you%2C%20as%20well%20as%20current%20and%20future%20users%2C%20to%20find%2C%20understand%20and%20reuse%20your%20data.) your data during your research and when depositing your data set into an online repository. Documentation is information added to data to ensure that this data is understandable to yourself and to others, both during and after your research. Metadata is also documentation, but in a structured form.

5a. Data documentation:

* the study level: e.g. explain the study purpose and methodology used
* the data level: e.g. define the meanings of variables and codes in your files

You can document your data using, for example:

* README files: these often list the files in a data set and what these contain
* exports from a (lab) notebook

Is this project funded bed ZonMw? If yes, ZonMw requires that researchers archive a ‘Replication package’, which should contain:

* the raw data (if you have reused existing data, you may have processed data rather than ‘raw’ data; in that case, the existing data must already be permanently archived, or must now be archived);
* the data on which publications are based;
* documentation on the research methodology used (such as code books, data manuals, metadata compilation, machine settings, use of SOPs, version management etc.), project proposal, approval from ethics committees such as METC, all stakeholders (researchers, laboratory assistants, test subjects etc.) ) in short, everything needed to ‘retrace the steps’.
* When storing qualitative data, describe the procedure used to transcribe the data (including conventions and symbols).

5b. Data quality controls
Describe how data quality will be controlled and documented. This can include practises like calibration, repeated measurements, standardised data capture, data entry validation, peer review of data, or the use of controlled vocabularies.

6. Working with sensitive data (personal data, ethics), data ownership, sharing and access
Working with sensitive data requires a considered approach to data sharing, storage, analysis and use, to mitigate any data security or privacy risks (see 6b). For questions, please consult the privacy officer ([Privacy & personal data - Intranet WUR](https://intranet.wur.nl/umbraco/en/about-wur/policy-regulations/privacy-personal-data/)).

6a. Restrictions to data access
Indicate whether there are any restrictions on [sharing the data](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Data-policy/Data-Sharing-guidelines.htm). If it is necessary to restrict access to certain parts of the data or to apply a data sharing agreement, explain how and why (e.g. licences, embargo). Explain what actions will be taken to overcome or to minimise restrictions, e.g. drawing up a file describing the terms of use.

6b. Processing and/or storing personal data
Personal data is any information that relates to an identified or identifiable living individual. Personal data must be properly handled and protected. To process personal data, a legal basis and purpose are required. The purpose is the reason for which personal data will be processed.

The following principles always apply:
 (a) be accurate and transparent; inform the person concerned as fully as possible about the processing of their personal data
 (b) process personal data only for the purpose for which it has been collected; further processing is generally not allowed
 (c) limit processing only to the personal data that is necessary to achieve the purpose
 (d) ensure that data are accurate and up to date
 (e) store personal data only as long as required for the purpose
 (f) ensure that personal data are secure.

Working in accordance with the privacy law, or GDPR, requires careful processing and storage of personal data. Risks can be mitigated by using data minimisation, informed consent forms, managed access to the data, secure data storage (IT managed) and transfer (encryption), pseudonymisation, anonymisation etc. Do the [Privacy Check](https://sharepoint.wur.nl/sites/Privacy_and_Personal_Data/_layouts/15/WopiFrame.aspx?sourcedoc=%7bDA11E946-0788-4CC3-9626-7FBB0EC8ED7A%7d&file=Do%20the%20privacy%20check.pdf&action=default&_gl=1*1y4v698*_ga*MTkzMDY2OTM3LjE2MzI4MTI5ODU.*_ga_M3YT587VEZ*MTYzMjgyMTI5My4zLjEuMTYzMjgyMTU0OS4w) to see if your activities are already privacy proof.

See [Privacy & personal data - Intranet WUR](https://intranet.wur.nl/umbraco/en/about-wur/policy-regulations/privacy-personal-data/) for more information about personal data.

6c. Registration in SmartPIA
Check if this project has been registered in WUR’s central register for data processing activities ([SmartPIA](https://smartpia.wur.nl)). All processing of personal data at WUR must be recorded in SmartPIA by completing a questionnaire assessing the privacy risks of the project. Based on the type of personal data used and the purpose of that use, the project managers will hear if any safety measures need to be taken.

6d. Other ethical issues
Here, consider whether you need ethical clearance from e.g. an animal experimental committee or a medical or social ethics review committee.

6e. Data ownership and access
The party that has ownership over the data decides what others are allowed to do with it. [WUR’s policy](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Data-policy/Data-ownership-Policy.htm) stipulates that WUR has ownership over data collected by its staff. However, if you collect data from/with an external party, you may have other contracts or agreements (e.g. consortium agreement, data sharing agreement) which overrule WUR’s ownership policy. Explain which parties have (joint)ownership and thus control data sharing and access.

7. Data archiving and publishing
WUR requires that all research data underlying publications is archived for at least 10 years. Research data should be made publicly available for re-use, unless there are valid reasons not to do so. In that case, research data should at least be archived within WUR.

Is this project funded by NWO? If yes, NWO expects researchers to preserve the data resulting from their projects for at least ten years, unless legal provisions or discipline-specific guidelines dictate otherwise. As much as possible, research data should be made publicly available for re-use. As a minimum, NWO requires that the data underpinning research papers should be made available to other researchers at the time of the article’s publication, unless there are valid reasons not to do so. The guiding principle here is 'as open as possible, as closed as necessary'. Due consideration is given to aspects such as privacy, public security, ethical limitations, intellectual property rights and commercial interests. In relation to research data, NWO recognises that software (algorithms, scripts and code developed by researchers in the course of their work) may be necessary to access and interpret data. In such cases, the Data Management Plan will be expected to address how information about such items will be made available.

7a. Data selection
At least all the research data necessary to verify and validate research results should be preserved. However, you do not need to preserve all data. In some situations it is easier or less costly to generate exactly the same data again than to store that same data. In some circumstances, it may also be in the interests of privacy to generate data again rather than storing it. To help you determine which research data to preserve, the Digital Curation Centre provides ‘[Five steps to decide what data to keep](https://www.dcc.ac.uk/sites/default/files/documents/publications/Five%20Steps%20to%20decide%20what%20data%20to%20keep.pdf)’.

7b. What data is internally archived
It is not always possible to publish data in a repository due to legitimate reasons (see 7). However, that does not mean that this data should not be preserved. Additionally, data that has not been used in a publication can still have (future) value for reuse.

7c. What data will be published
The findability of your data set for other potential users will be enhanced by publishing the data set in a data repository. If you cannot publish your data set in a data repository (see 7 for legitimate reasons), you can still publish the metadata. The metadata about the data set may include title, creator(s), affiliation(s), description, research goal, contact information and conditions for getting access to the data, etc (see also 5.).

It this project funded by ZonMw? If yes, ZonMw expects researchers to publish the metadata in the case the data themselves cannot be published due to e.g. privacy, public security, ethical limitations, intellectual property rights and commercial interests.

7d. When and for how long will the data be available for re-use
The data policies of WUR and NWO require, as a minimum, that data underlying publications should be made available to other researchers at the time of publication (unless there are valid reasons not to do so) and for a minimum of 10 years. Explain when the data will be made available. Indicate the expected timely release. Explain the reason and duration of any embargo periods. Explain whether exclusive use of the data will be claimed and if so, why and for how long. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents.

7e. Choosing a data repository
Repositories are services that preserve data safely and make them findable. You can archive publicly or with restricted/closed access. WUR Library can help you with [depositing](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Finishing/Why-and-where-to-publish-your-data-set.htm) data in the repositories DANS-EASY, 4TU.ResearchData and Zenodo. You can also look for other repositories, for example on [www.re3data.org](http://www.re3data.org/). Use a repository that is certified (check this [here](https://www.coretrustseal.org/why-certification/certified-repositories/)), that assigns persistent identifiers (e.g. DOIs) to data sets and/or the metadata, that offers the [licence](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Finishing/Data-Licenses.htm) you want to put on your data (this may differ between repositories). Additionally, it is recommended to use a repository that is (often) used in your discipline, as such a repository is familiar with the data type(s) and volume within your discipline.

7f. Metadata
Metadata is highly structured, machine-readable information to describe datasets (or data within datasets). A metadata standard is a set of fixed fields. Common fields to describe a dataset are author, year, title, spatial coverage, etc. Many repositories use an existing metadata standard, such as [Dublin Core](https://en.wikipedia.org/wiki/Dublin_Core) or DataCite. This way, any dataset you deposit to the repository will use that as a standard. However, you can always add relevant metadata and there are discipline-specific metadata standards available, which we recommend you to use.

7g. Licences
Indicate under which license the data may be re-used. Check the commonly used [Creative Commons licences](https://creativecommons.org/about/cclicenses/). WUR’s data sharing guidelines state that WUR researchers should at least be attributed for their scientific effort (i.e. CC BY). CC BY means that other researchers should credit you as a creator of the data set. Note: funders could differ in the requirements for which licence to put on a data set and repositories differ in which licences they offer.

7h. Publishing the analysis software generated
Indicate whether potential users need specific tools or software (e.g. specific scripts, codes or algorithms developed during the project) to access, interpret and (re-)use the data. Indicate how these items will be made available. Consider the sustainability of software needed for accessing and interpreting the data. Check the Five Recommendations for FAIR Software (<https://fair-software.nl/>).