WUR_DMProtocol_TemplateAndGuidance_v12-01

Authors: Irene Verhagen
https://orcid.org/0000-0001-5588-1333

Danny de Koning – van Nieuwamerongen
https://orcid.org/0000-0002-8264-4541

Shauna Ní Fhlaithearta
https://orcid.org/0000-0002-0101-8937

Date: 20230622

Organization: Wageningen university & Research
https://ror.org/04qw24q55

WUR data management protocol template

The WUR data management protocol template can be used as a basis for the data management protocol of the chair group (WU) / business unit (WR), hereafter called 'research group'. When using this template, fill in the required information at [...].

- This template is based on the WUR data policy requirements.
- A data management protocol is a 'living document' and, as such, can be edited and updated at any time. An annual review of the data management protocol is considered good practise.
- You are free to add (a) header(s) to this template to better align with the requirements of your research group, but the original headers must be retained.
- Answers provided are suggested texts / examples and can be tailored to the data management practices within the research group as long as compliance to the WUR data policy is maintained.
- To get to additional information in the appendix for each section, hold keyboard key CTRL + left-click [info] or right-click [info] and select 'open hyperlink'.
- Do you need support creating this protocol? Please, contact data@wur.nl, your privacy officer, your information security officer, your legal department, or visit https://www.wur.eu/rdm for more information.

Data management protocol - [name research group]

Title of document	Data management protocol [name research group]
Document first drafted	[date in YYYYMMDD]
Date of update + version	[date in YYYYMMDD + vxx (e.g. 20230309_v01)]
Previous versions	[list dates and version of previous versions of this protocol]
Document approved by	[name(s)]
Data management protocol created and written by	[names]
Contact	[data steward contact name + email] [research group secretariat contact name + email] [information security officer contact name + email] [privacy officer contact name + email

1. Purpose [info]

The purpose of this data management protocol is to set general data management practices within [name research group]. More specifically, these data management practices aim to prevent data loss, increase safe handling of personal or otherwise sensitive data, and increase reusability of data by complying with the Dutch Association of Universities (UVN) Code of Conduct, GDPR, WUR information security and data policy. The WUR data policy is guided by, amongst others, the FAIR principles (Findable, Accessible, Interoperable, Reusable) and the motto to share data 'as open as possible, as closed as necessary'. The WUR data policy requires:

- Chair groups to write and update a data management protocol (group-level). This is not required for business units, but highly recommended.
- PhD candidates to have a data management plan (project-level). Nevertheless, it is highly recommended to create a data management plan for each new research project regardless of function title or organisation.
- Secure, safe and shareable storage of the research data during research.
- Preserving of research data underlying publications (articles/reports/theses) for at least 10 years after research.
- Registration of preserved data in Pure.

The data management practices outlined in this data management protocol apply to all research data produced within research projects performed at [name research group]. This includes research data derived within BSc / MSc thesis projects, PhD thesis projects, postdoc projects, staff projects, as well as other research projects. Research data encompasses all information and / or resources required to reproduce the results and conclusions of research. Data can include audio, video, transcripts, analysis and processing scripts/code, tabular data, protocols, non-digital data etc.

Please note there is support available to help you with questions on the data management protocol at data@wur.nl or visit https://www.wur.eu/rdm for more information. For handling personal or otherwise sensitive data and legal issues, your privacy officer, information security officer, and legal department can be consulted.

2. Links to information used

Policies

- WUR data policy
- Information security policy
- Data classification
- Data sharing guidelines
- Personal data

Software / apps and storage media

- Storage finder
- Repository finder
- ApprovedApps tool
- VPN and W: drive access

Research data management support

- Research data management website
- data@wur.nl

Templates

- Research data management plans and protocols
- Research data documentation

Privacy, security, legal support

- SmartPIA (blue button)
- Data ownership
- Privacy and privacy officers
- Information Security and information security officers
- Legal departments (contacts under 'Who can draft me a contract?')
- Personal data
- GDPR

Data preservation

- Data preservation
- Data publication
- Data registration
- Data documentation
- Yoda metadata editor
- Schema.org
- Fairsharing.org

Data stewardship

- Data steward tasks

3. Definitions and abbreviations

RDM	Research data management
FAIR	Findable, Accessible, Interoperable, Reusable
DMP	Data management plan
PO	Privacy officer
ISO	Information security officer
GDPR	General Data Protection Regulation
Personal data	Privacy sensitive data; containing information of humans
Research group	Chair group (WU) / business unit (WR)
WU	Wageningen University
WR	Wageningen Research
WUR	Wageningen University and Research
Head of research group	Chair holder / business unit manager

4. Roles and responsibilities [info]

Suggested text / table (can be adopted / tailored): [name research group] identifies the following roles and responsibilities within the context of research data management:

Who	Roles and responsibilities
Head of research group	• The final responsibility for compliance with the WUR data policy lies with [name head of research group]. • Responsible that a DM protocol will be created, which adheres to WUR policies (including the WUR information security and data policy) and GDPR, but can delegate the execution (e.g. to the data steward). • Ensures a data steward is present within the group and transfers tasks when the current data steward leaves the group / passed on the role. • Is responsible for ensuring a contact point for the appropriate handling of data access requests for data published restricted access or archived by the research group.
Data Steward See data steward tasks	• Initiates and (co)-creates the data management protocol and provides advice on this matter to the management of the research group. • Updates the data management protocol when necessary in consultation with the (head of) the research group. • Functions as a primary contact point to members of the research group for questions about data management and refers them to (data management) support when necessary. • Informs new members of the research group about the data management protocol. • Communicates the WUR data policy to members of the group. • Advises group members on their DMP. • [Optional: add other group specific tasks/responsibilities].
Researcher Any researcher (e.g. PhD candidate, postdoc, associate professor, etc.)	• Responsible for data management during the entire life cycle of the research project. • Takes the WUR data policy into account when writing new projects. • Informs, where applicable, PhD candidates, or MSc / BSc thesis students about the data management protocol. • Supervises data management by a PhD candidate, MSc / BSc thesis student (where applicable). • Determines, in the case of being the data controller (see here), the data classification of the data to be collected and consults with the PO or ISO of the department where required (e.g. when personal or otherwise sensitive data is collected).

	• Uses storage solutions and software used within the research group (see 'Safe and shareable storage during research' below) that befits the data classification. • Consults the data steward or ISO of the department about the appropriateness of a storage solution or software that is used that is not in the ApprovedApps tool of WUR. • Ensures careful handling of any personal or otherwise sensitive data during and after the research project. Where applicable, the researcher ensures that the project is registered in SmartPIA by the project leader. • Makes appropriate arrangements for safe handling of (sensitive) data by the students. See information on VPN and W:drive. • Ensures that research data during research is safely stored and sufficiently backed-up according to the WUR information security policy, where multiple project members should have (restricted) access to the data. • Preserves data underlying (a) publication(s) or (a) report(s) after the research project for at least 10 years in an appropriate storage medium or data repository conform the WUR information security policy. • Registers data underlying (a) publication(s) or (a) report(s) in Pure via an email to data@wur.nl. • [Optional: add other group specific tasks / responsibilities].
PhD candidate Additional responsibilities	• Writes a DMP within 6 months of the start of the PhD. • Makes sure that their DMP is checked by the data steward and / or supervisor. It is recommended to have the DMPreviewed : request a review via data@wur.nl or request feedback via DMPonline (dmp.wur.nl). • Adds the DMP to the PhD project proposal and uploads the DMP in Hora Finita. • Consults with the supervisor when others are requesting access to the data during and after research. • [Optional: add other group specific tasks / responsibilities].
Supervisor Additional responsibilities	• Supervises the writing of a DMP by a PhD candidate or MSc / BSc thesis student (where applicable). • Responsible for data management during the research project of PhD candidates and / or students. • Co-authorises others for read / write access to the data (folders and / or files) and ensures they are up to date. • [Optional: add other group specific tasks / responsibilities].
BSc/MSc student	• Registers their MSc thesis in Osiris and acknowledges the thesis agreement in Osiris.

	• Uses storage solutions and software used within the research group (see 'Safe and shareable storage during research' below) that match the classification of the data. For example, solutions provided by WUR and / or in the ApprovedApps tool (WUR login required). • Requests approval from their supervisor to share data or store data outside of WUR managed platforms • Ensures careful handling of any personal or otherwise sensitive data during and after the research project. • Hands in their data to their supervisor when the project is finished. This only applies when [name research group] is leading in supervising the thesis. • [Optional: add other group specific tasks / responsibilities].
Other	• [Optional: add other group specific tasks / responsibilities].

5. Data (rights)holder, sharing and accessibility [info]

Suggested text (can be adopted / tailored): WUR is (rights)holder to all data created during research created by its employees and non-employee PhD candidates, excluding BSc and MSC students (see this page), unless otherwise stipulated in a consortium agreement or contract between WUR and other parties (the legal department will be contacted for such agreements). Sharing of and accessibility to the data created by [name research group] will be arranged on a case-to-case basis in consultation with the legal department (where required), and depends on sensitivity and contractual obligations (see this page). When data is created by a party other than WUR, the policies of that party should be consulted (such as (rights)holder policies). When data is being reused, the licence restrictions to that data are adhered to.

6. Safe and shareable storage during research [info]

Suggested text / table (can be adopted / tailored): The WUR data policy requires secure and shared storage of research data (including data documentation and metadata). The classification of the data determines how secure the storage has to be. For data classified as serious, contact the ISO whether additional security measures are needed. For data classified as disruptive, it is mandatory to contact the ISO to determine additional security measures. Several secure and shared storage solutions are provided by WUR. Using these storage solutions ensures compliancy with the WUR data policy for the data collected at [name research group]. As such, data is stored in locations that are backed up automatically, have secure access management, and integrated integrity checks.

Do **NOT** store research data on hard disks, USBs, personal laptops (exceptions excluded), and external third party cloud services such as Google Drive and Dropbox (exceptions excluded). Transient use of external hardware (USB, external hard-disks, voice/movie/photo recorders) is allowed for transport of data from the data collection site to the secure storage location, when appropriate security measures are applied that befits the data classification of the research data. Using WUR storage solutions prevents data loss and ensures that data can be accessed by authorised parties in case of emergency or departure from WUR. When using a platform to (temporarily) store data for processing, analysing, or sharing, make sure that the platform providers securely handle the data according to WUR information and security policies and adhere to the General Data Protection Regulation (GDPR, AVG in Dutch) where applicable. Contact the ISO to help in determining whether the chosen platform is appropriate.

What data should be stored where?	
Description of type of data	**Description of storage location(s)**
Storage location(s) used within [research group]	[e.g. W:\xxx\xxx…; Git@WUR project repository address; Yoda@WUR research-xxx-xxx and vault-xxx-xxx]
During research Raw data Processed data Analysed data Results / output Documentation Metadata	• Before using software to temporarily store, collect, process and analyse data, the data classification is determined and the ApprovedApps tool consulted. • At all times the raw research data and an up to date version of the research data is securely stored on above mentioned storage location(s). OneDrive, and M-Drive are only used transiently and for working copies.*

	• External hardware (USB-drive, external hard-drives, audio recorders, video recorders, etc.) are not allowed to permanently store and backup data. • Personal or otherwise sensitive data on hardware or platforms required for collection, processing, or analysis purposes, is removed as soon as possible from those platforms and hardware (after the data is safely transferred to the above mentioned storage solution(s)).
During research Data sharing without publication	• Before sharing data, the data classification is taken into account to determine whether the data can be shared and on which platform, and whether data sharing or processing agreements are needed. • For small data files that do not contain personal or otherwise sensitive data: [e.g. WUR Teams, WUR OneDrive, institutional mail, Yoda@WUR, SURFfilesender] • For Small and large data files: [e.g. WUR Teams, WUR OneDrive, Yoda@WUR, SURFfilesender].
Source code and scripts	• If applicable, source code and scripts are managed through Git@WUR. • When not applicable, source code and scripts will be managed just like other research data. • When using Git@WUR and publishing research data files in a data repository, the master branch is downloaded and added to the research data files in the data repository. •
[add other requirements where applicable]	

* Upon termination of a WUR account, all files on the M:-drive and/or Drive for Business are automatically deleted, including permanent deletion of data stored there. Therefore, research data in a personal folder/personal cloud storage should also be stored in a department or project folder accessible to the research group.

7. Data preservation and registration after research [info]

Suggested text / table (can be adopted / tailored): In compliance with the WUR data policy, all data underlying publications (articles / reports / theses) are preserved (via archiving or publishing) and registered to enable verification of the research and reusability of the data. Data underlying publications include:

- Raw data (where legally allowed, e.g. according to the GDPR).
- Processing and analysis scripts.
- Processed / cleaned data where required, e.g. when scripts are not sufficient or available.
- Analysed data where required, e.g. when scripts when scripts are not sufficient or available.
- Source code and software developed during the research project.
- Data documentation and metadata, which includes minimally the WUR recommendations.
- [add when applicable]

For continued access to the data for long-term periods (years to decades), data will (also) be preserved in preferred formats where possible (for examples see: https://dans.knaw.nl/en/file-formats/). When data cannot be stored using preferred formats, the research data documentation (see 'Data documentation and metadata' below) will elaborately describe contents of the data and the used software (including version and provider of software).

Both data published in a data repository and archived on the W-drive / Yoda@WUR (see table) are registered in Pure. Registration is performed through an email to data@wur.nl which includes:

- The persistent identifier of or pathway to the data.
- References or persistent identifiers to the accompanying publication(s) (journal or report).
- Creator / author name(s).
- Metadata (minimally the WUR recommendations).

Data that do not underly publications are not required to be registered, but it is recommended. If applicable, research models can be registered in the WUR model gallery.

What data should be preserved where?	
Data archive location used within [name research group]	e.g. [W:\xxx\xxx…; Git@WUR project repository address; Yoda@WUR research-xxx-xxx and vault-xxx-xxx] [contact name and email address] is the contact person within the research group for access and management to the mentioned archive location(s) (e.g. the data steward, the secretary of the group, the supervisor). An agreement with this contact person is present that personal or otherwise sensitive data is only accessed / viewed upon permission from a researcher or head of the research group.
Data underlying publication / report. Raw data Processed data Analysed data Results / output Documentation Metadata	• Research data files underlying a journal publication or report are published openly in data repositories when the data classification, contractual obligation, licences, and policies allow open access publication (see the WUR repository finder). • When research data cannot be made openly available, at least the metadata to the data is published, while the underlying data needs to be archived within WUR in the above mentioned archive location. • When research data cannot be made openly available, but access can be requested (i.e. restricted access publication), a data sharing agreement is set up. Additionally, the primary contact point for data requests is [names and email addresses of contact] which will be in charge of evaluating requests and providing access to the data for 10 years after publication of the data. [Note that this preferably is not one person as that person may leave WUR, it needs to be a contact point that is more likely to persist for at least 10 years and is able to evaluate approve requests with regards to the nature of the data.] • If source code or scripts are too large or have too many dependencies for it to be (sensibly) downloaded and archived, the code remains in Git@WUR with the adherence that the git project repository name is not changed or relocated to ensure that any links provided remain usable.** • Data is permanently and irretrievably removed from systems when legal, contractual obligation, licence, or policy requirements demand it. Each research project clearly states in the DMP the types of data that are to be preserved and the types of data that are not allowed to be preserved.
Data not underlying publication / report	Although not mandatory, the head of the research group encourages openly publishing data not underlying a publication or report where the data classification, contractual obligation,

	licences, and policies allow open access publication. When not allowed, we encourage archiving the data internally at WUR in the above mentioned archive location with documentation and metadata. If source code and scripts are not underlying a research data file publication, the master branch from Git@WUR is downloaded and added to the research data to be archived.**
[add other where applicable]	

** Git systems do not provide a persistent identifier like a DOI. Persistent access to an individual git project repository is only possible if the project repository is not renamed or relocated. Hence, it is good practice to at least download the master branch of your project repository and add it to your other research data files where possible.

8. Data documentation and metadata [info]

Suggested text / table (can be adopted / tailored):

Documentation

Data should be accompanied by documentation in the form of a README.txt file. Documentation should cover [e.g. methodology, explanation variables, analytical and procedural information etc.]. The minimum information supplied in the documentation will be equivalent to the WUR readme and codebook template (DOI: 10.5281/zenodo.7701727).

Metadata

Metadata should at least have the following components [e.g. title, author(s) and affiliation(s), keywords, licence, etc.]. When a discipline specific metadata standard is available and known or a discipline specific repository is used that incorporate their own metadata standard, those metadata standards will be applied [please list the metadata standards (can be found a.o. at schema.org or fairsharing.org)]. When publishing data in a general repository, such as DANS-EASY, 4TU.ResearchData, Zenodo or Yoda@WUR, the metadata standard of the repository is applied to the data. When there is (i) no discipline specific metadata standard, (ii) the data repository only allows limited metadata, or (iii) the data is archived internally at WUR, the minimum required metadata to be applied will be the Yoda metadata terms based on the Datacite metadata standard. These metadata can be created from within Yoda@WUR or through the Yoda metadata editor (when the Yoda@WUR platform is not being used by the group).

9. Data management plans for PhD candidates/research projects

[info]

Suggested text / table (can be adopted / tailored): When possible and allowed by the funder, the WUR DMP template is adhered to. This is possible for projects funded by NWO, ZonMW, and the European Commission. When the funder does not allow the WUR template, the template of the funder is adhered to. DMPs are created using DMPonline, which provides the WUR and funder templates with guidance. A manual to create an account and DMP is available via DOI: 10.5281/zenodo.7073740. When desired, and usually mandatory when a funder is involved, a review of the DMP can be requested in DMPonline. RDM support encourages a review of the DMP regardless of whether it is mandatory (by the funder). In general, the DMP should be updated and maintained throughout the research project when changes in data management practices are made.

10. Feel free to add (a) header(s) (if applicable)

Appendix: additional information

(may be deleted after completion)

Answers provided above are examples and can be tailored to the data management practices within the research group as long as compliance to the WUR information security and data policy is maintained.

Info: 1. Purpose

A group may choose to have the data management protocol also apply to (MSc/BSc) students.

Info: 4. Roles and responsibilities

Identifying persons who are (or can be) of assistance in data management practices helps to clarify the data collection process within the research group. Identifying these roles are also important when employees / PhD candidates / students have to hand over the data to WUR in the event of leaving WUR. In the table several (potential) roles and responsibilities have been provided. For the advice document on roles for data stewards, please see here. For information about data classifications, see the additional information at '4. Safe and shareable storage during research'. Feel free to remove, adjust or add roles.

Info: 5. Data (rights)holder, sharing and accessibility

Explain, in general, what the procedures / guidelines are considering the data rights(holder), data sharing and accessibility within the research group when:

1. Data is collected solely by WUR (employees).
2. Data is collected from / with / on behalf of third parties (other than WUR).

WUR is the rightsholder of the research data and entitled to data(bases) created by any WUR staff member within the scope of their employment. This means that WUR researchers are not the owner nor right holders of the data(base).

WUR as rightsholder should be discussed in the case the funder of the project, i.e. a third party (either public or commercial), imposes ownership / rightsholder conditions. This discussion takes place before the project starts. E.g. partner funding can condition that the data will become under (shared) ownership of the partners / partners are (shared) rightsholder. In this case a consortium agreement or a data sharing agreement (DSA) should be agreed upon before starting the project, where required, in consultation with the legal department (intranet WUR, login required, under 'Who can help me draft a contract?').

The party that is the rightsholder of the data decides what others are allowed to do with it; how the data is shared and accessibility to the data. If you collect data from / with / on behalf of an external party, you may have other agreements (e.g. consortium agreement, data sharing agreement) with that party, and these generally overrule WUR's IP Policy & Value Creation. The research data sharing guidelines can help with this.

Info: 6. Safe and shareable storage during research

Describe the general procedures / guidelines for data storage during research within the research group. If there are general pathways on the W-drive or Yoda@WUR, these can be added as well.

Ensure that storage practices within the research group comply with the WUR's data policy and the Information Security policy (e.g. data classification). Safe and shareable storage of data following these policies means storing research data (including code etc.) and data documentation during research on (a) storage solution(s) provided by WUR (and thus managed by WUR). Note that WUR's storage solutions are backed up automatically and data recovery is in place. Check the ApprovedApps list (intranet WUR, login required) if your data needs a storage solution or software with additional security measures. Any questions regarding the above mentioned can be addressed to your information security officer (intranet WUR, login required).

Note that all WUR employees and users need to properly and safely handle and store research data, but appropriate platforms may not be available by default for all users. For example, MSc and BSc students do not have access to all WUR managed platforms or hardware for handling and storing of (personal or otherwise sensitive) data. Specific arrangements must be made for these groups of WUR users.

Info: 7. Data preservation and registration after research

Describe the general practices for preserving and registering data within the research group.

WUR's data policy stipulates that research data underlying publications (articles, reports, theses) must be preserved for at least ten years, if possible in a data repository. Repositories are online archiving services that preserve data safely and make them findable. Data can be archived publicly, and with restricted access. However, there could be legitimate reasons not to deposit data in a repository, but to archive the data on the W-drive / Yoda@WUR (or archive the data with a third party owning / being the rightsholder of the data): there is personal or otherwise sensitive data, company interest involved etc.), WUR is not the rightsholder, Intellectual Property Rights, the data volume is too big etc. The research data sharing guidelines can help in setting the general practices within the research group for sharing research data.

The data underlying publications that are preserved should, also in compliance with the WUR data policy, be registered in Pure. Registration can be done by sending an email to data@wur.nl with the persistent identifier (e.g. DOI, accession number) or link/pathway to the data and associated metadata. To link the data to the accompanying publications, the publications involved and their links / DOIs need to be specified as well. Library staff will then ensure your research output is properly linked and becomes visible in Research@WUR.

Info: 8. Data documentation and metadata

Describe the general procedures / guidelines how data should be documented within the research group. Outline the metadata that will accompany the data within the research group and whether there are metadata standards to be used. It is possible to make a distinction between BSc / MSc projects and other projects.

It is essential to systematically document data during research, when archiving data internally (e.g. at WUR) and when depositing data in an online repository to make data understandable, discoverable, citable and reusable. The most common required form of documenting research data is by adding a readme file, which is further supplemented with metadata. Independently of where data is preserved (e.g. at WUR, in a data repository) it should be accompanied by:

- a readme file
The readme file contains information about e.g. the steps that have been undertaken in processing and analysing data. In short: all information necessary to understand the data, reproduce research and verify results. WUR Library advises to use the WUR readme file template as the minimum required documentation to add to the data. Feel free to add more documentation where appropriate and required.

- metadata
Metadata is machine-readable information about the data, according to fixed terms, which makes the data findable and searchable. WUR Library advises to use the Yoda metadata terms as the minimum required metadata to add to the data. You can fill in these terms in Yoda, when applicable as a storage solution, or use the Yoda metadata editor and download the metadata as a .json file.

- a codebook
WUR Library advises to use the WUR codebook template to explain variables, abbreviations, etc, because this template is in .csv format, which makes it easy to import into software such as R, Python, etc.

The templates for the readme file and codebook can be found here DOI: 10.5281/zenodo.7701727. Via this link a filled in example of a readme file, codebook and metadata file can be found as well.

Info: 9. Data management plans for PhD candidates/research projects

Describe the guidelines for how DMPs are implemented in the research group.

The answer provided is an example, but can be tailored.

The WUR data policy stipulates that every PhD candidate should have a DMP. The policy applies to WU, but not to WR. However, we highly recommend DMPs for any research projects (regardless of function or organisation). Additionally, we encourage BSc / MSc students to write a DMP to create awareness regarding handling data throughout the research cycle.

A DMP is a 'living' document in which data management practices on the project level are described. Funders more and more require a DMP for which the template of the funder can be used. When no funder is involved or the funder has no template available, the WUR provides a template in DMPonline or via DOI: 10.5281/zenodo.7233370