WUR_DMP_TemplateAndGuidance_v08-02

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: 20230510

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

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'.

❖ 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 guidance 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.

❖ 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'.

❖ For a filled in DMP example have a look at DOI: 10.5281/zenodo.7096699.

❖ You can also find this DMP template in DMPonline (https://dmp.wur.nl/). Have a look at DOI: 10.5281/zenodo.7073740 on how to get started in DMPonline.

❖ For a review of your DMP, contact data@wur.nl through email or through DMPonline.

❖ Questions? Contact data@wur.nl or visit the WUR research data management website at https://www.wur.eu/rdm for more information. For privacy questions, contact your privacy officer or visit the Privacy & personal data WUR intranet page (includes a list of privacy officers). For information security questions, contact your information security officer or see the Information security WUR intranet page.

A. Describe the research project ^[info]

1. Name researcher:

```



```

2. What is the name of your department?

☐ Agrotechnology and Food Sciences

☐ Animal Sciences

☐ Environmental Sciences

☐ Plant Sciences

☐ Social Sciences

☐ Wageningen Food Safety Research

☐ Wageningen Food & Biobased Research

☐ Wageningen Centre for Development Innovation

☐ Wageningen Economic Research

☐ Wageningen Plant Research

☐ Wageningen Environmental Research

☐ Wageningen Marine Research

☐ Wageningen Bioveterinary Research

☐ Wageningen Livestock Research

3. What is the name of your chair group or business unit?

Please copy-paste the English name and abbreviation for:

- chair groups from this page.
- business units from this page.

Example: Bioprocess Engineering (BPE) or Contract Research Organization (CRO).

4. Describe the organisational context of your research project.

DMP version (or date last modified)	
Supervisor / (co-)promotors	
Graduate School (WU only)	
Start date of project	
End date of project	
Project number	
Funding body	

5. Give a short description of your research project.

Title	
Summary	

6. 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 [......]	

7. I have requested a review of this data management plan from:

☐ WUR Library – Data Management Support (data@wur.nl)

☐ The data steward of my chair group / business unit

☐ No review requested

8. Name of data management support staff and / or data steward consulted during the preparation of this plan and date of consultation.

B. Describe the data to be collected, software used, file formats and data size [info]

9. Will you reuse existing data for this project?

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

☐ No. Please describe below any constraints to reusing existing data.

10. Will new data be produced?

☐ Yes.

☐ No.

11. Please describe the data you expect to generate and / or use in the table below. Include reused existing data as well (as these are files that you manage and store).

File contents	Data type	Software	(Open) file format	Estimated size of each file (range)	Estimated number of files (range)
(e.g. lab analysis, gene sequence, interviews, lesion scores, etc.)	(e.g. numerical)	(e.g. Excel)	(e.g. .csv)	(e.g. 20 – 50 Mb)	(e.g. 50 – 100)

12. Estimate how much data storage you require in total (e.g. by using the information in the table at question 10).

☐ 0-10 GB.

☐ 10-100 GB.

☐ 100-1000 GB.

☐ >1000 GB.

C. Storage of data and data documentation / metadata during research [info]

13. Where will the data, code and accompanying documentation / metadata be stored and backed up during the research project (see the WUR Data Storage Finder)? Include platforms you use to share data, collect data on, or send data to for processing or analysis.

☐ W:drive Enterprise File Storage (WUR network drive).

☐ W:drive Massive File Storage Disaster Recovery (WUR network drive).

☐ W:drive Massive File Storage without Disaster Recovery (WUR network drive).

☐ Yoda (data management platform; SURF hosted WUR instance).

☐ WUR OneDrive for Business - only when an up to date version of the research data is also safely stored on the W:drive or Yoda.

☐ WUR SharePoint / Teams - only when an up to date version of the research data is also safely stored on the W:drive or Yoda.

☐ Git@ WUR (GitLab locally hosted at WUR)

☐ Other, please specify below the storage medium / system and describe back-up frequency, access management, and geographic location (e.g. within or outside the EU).

D. Structuring your data and information [info]

14. Give a (visual) representation of the folder structure you intend to use.

15. Describe the file naming conventions you intend to use. Please give one or multiple example(s).

16. How will you distinguish between versions of files (multiple answers possible)?

☐   Dates within file names are updated when files are modified.

☐   A version number in the format 'v01' will be added to all file names which increases after file modification.

☐   The designation 'vRAW' is added to file names that contain raw unaltered data (before any processing and cleaning). Any alteration of RAW data is done on a copy of the RAW data and appended with a version number which increases with each file modification (e.g. v01, v02, v03, etc.).

☐   We will use Git versioning for code / scripts.

☐   Other, please specify below.

E. Data documentation and data quality ^[info]

17. Describe below what [data documentation](#) and metadata will accompany the data to help make the data findable, understandable, and reproducible.

☐ The WUR readme file template (see template at
https://doi.org/10.5281/zenodo.7701727).

☐ The Yoda metadata form (see Yoda metadata editor at
https://utrechtuniversity.github.io/yoda-portal/).

☐ The WUR codebook template (see template at
https://doi.org/10.5281/zenodo.7701727).

☐ Electronic lab notebooks (e.g. eLabJournal).

☐ Elaborate documentation and notes within scripts / code.

☐ Other, please specify below.

18. Describe what data and analysis quality controls will be used?

☐ We will perform preliminary (pilot) experiments to validate intended experimental methods.

☐ We will use standard and validated protocols where appropriate.

☐ We will use repeated measurements to validate results (e.g. duplicate or triplicate analysis, multiple observer agreement, measurements taken over time, etc.).

☐ Statistical model assumptions are adhered to and assessed (e.g. (residual) distribution analysis, outlier analysis, (accounting for) independence, homogeneity of variance , etc.).

☐ We will use a statistical power analysis before and after the experiment.

☐ We will consult statisticians.

☐ Supervisors or peers will review the data and results for any anomalies (e.g. unexpected inconsistencies, outliers, correct labeling of data and / or treatments, correct and consistent coding applied, etc.).

☐ We will use standardised coding and terms of data throughout all experiments so that data descriptions are equal throughout various datasets created.

☐ We will use discipline specific community standards for labelling and coding of data. Please specify the community standard used.

☐ Other, please specify below.

F. Working with sensitive data (personal data, ethics), data ownership, sharing and access [info]

19. Who is the (rights)holder of the data (commonly known as the owner of the data)?

☐ WUR is the (rights)holder of the data.

☐ WUR is not the (only) (rights)holder of the data and a WUR approved formal (consortium) agreement or contract between WUR and other parties is present.

☐ Other, please specify below.

20. What is the data classification for your project (for example as specified in SmartPIA) taking into account the (privacy) sensitivity of the data?

☐ Negligible.

☐ Some.

☐ Serious.

☐ Disruptive.

21. Is this project registered in SmartPIA?

☐ Yes.

☐ No. Please register in SmartPIA in the case (privacy) sensitive data is collected (when applicable: via your supervisor, the project manager, see guidance).

22. Please specify the (sensitive) data and privacy protection measures. Note that any measures undertaken should be consulted with the Information Security Officer (ISO) and Privacy Officer (PO).

☐ Data is classified as negligible and standard WUR security measures are undertaken.

☐ Only WUR provided storage, processing, and analysis platforms are used as consulted with the ISO and PO where applicable.

☐ Access management to the data is either managed or approved by the project leader / supervisor of the project and contains clear documentation of who has access.

☐ Informed consents are present when information from humans are involved.

☐ Personal or other sensitive data will be removed when not required for verification of research.

☐ Personal and other sensitive data will not be made openly available and will at most be shared under formal agreements for which the ISO and PO are consulted.

☐ Personal and / or other sensitive data will be separately stored where possible to increase difficulty of linking data for those with unauthorised access to data.

☐ We will consult with the ISO and PO for appropriate measures to undertake.

☐ Other, please specify below.

23. Are there other ethical issues that need to be taken into account which may include approval from [ethical committees](#)?

☐ No.

☐ I work with animals and will seek / have approval of the ethics committees involved (Animal Welfare Body (IvD), Animal Tests Committee (DEC), Central Animal Testing Committee (CCD)).

☐ I work with humans in a social sciences aspect or that doesn't fall under Medical Research Involving Human Subjects Act (WMO) and will seek / have approval of the ethics committees involved (WUR Research Ethics Committee).

☐ I work with humans in a biological / physiological / medical / intervention aspect falling under the Medical Research Involving Human Subjects Act (WMO) and will seek / have approval of the ethics committees involved (Medical Ethics Committee-East (METC-Oost)).

☐ I work with genetic resources covered by the Convention on Biological Diversity and need to comply with the Nagoya Protocol.

☐ The data and results of my research could negatively impact nature / wildlife conservation efforts. Please explain below.

☐ The data and results of my research could negatively impact humans / society. Please explain below.

☐ Other, please explain below.

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

☐ No.

☐ Yes. Please explain how they will be managed and who will be contacted to discuss such rights handling.

G. Data archiving and publishing ^[info]

25. Are there reasons to restrict access to the data or limit which data will be made publicly available?

☐ No.

☐ Privacy / GDPR.

☐ Ethics.

☐ Contractual agreement.

☐ Commercial interests.

☐ Public security.

☐ IP rights.

☐ Other, please specify below.

26. Describe what data from question 10 will be archived internally (e.g. WUR network drive / Yoda) and not published, for a minimum of 10 years? Include the exact name of the storage medium chosen (see WUR Data Storage Finder).

☐ Not applicable as data will be published.

☐ Due to sensitivity of data we will need to archive (part of the) data underlying publications or reports internally. Please specify below which data and the chosen storage medium.

☐ Other, please specify below.

27. What data will be published and made available for reuse via a data repository?

☐ Data underlying publications or reports. Please specify below which data listed in question 10.

☐ Only the metadata is published in a data repository as the data are too sensitive to openly share.

☐ Data not underlying an article or report will also be published. Please specify below which data listed in question 10.

☐ Other, please specify below.

28. When will the data be available for reuse, and for how long will the data be available?

☐ Data will be available for at least 10 years as soon as the article or report is published and not required for any other article publication.

☐ Data will be available for at least 10 years upon completion of the project.

☐ Published data becomes accessible for at least 10 years after expiring of the embargo. Please specify below the reason for an embargo.

☐ Publication of data not underlying an article or report will be considered at the end of the project.

☐ Other, please specify below.

29. Which data repository do you intend to use to make the data findable and accessible (see the [WUR Repository Finder](#))?

☐ DANS-EASY.

☐ 4TU.ResearchData.

☐ Zenodo.

☐ Other, please specify below.

```

```

30. Which metadata standard will be used to describe the data during internal archiving and / or depositing in a data repository?

☐ Yoda metadata (DataCite metadata standard).

☐ Metadata standard from DANS-EASY, 4TU.ResearchData and / or Zenodo (which often are the DublinCore or DataCite standard).

☐ A discipline-specific metadata standard. Please specify below.

☐ Other, please specify below.

```

```

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

☐ Open access (Creative Commons Attribution licence (CC BY); anyone can access and reuse with attribution).

☐ Restricted access (custom licence text or data sharing agreement is required, dictating restrictions of access and reuse). When a data sharing agreement is required, the Privacy Officer or Information Security Officer is consulted.

☐ Closed access (only metadata published, data is not allowed to be requested or reused).

☐ Other, please specify and/or provide link to the licence below.

32. If software is generated in this project, describe your publishing strategy below.

H. Data management costs [info]

33. What resources (in time and / or money) will be dedicated to data management, data archiving or publication, and ensuring that data is reusable? Indicate as well how these costs will be covered.

☐ All costs for 10 year data storage and access management to that data after journal publication or report are covered by the research group / project.

☐ Costs relating to publications up to 100 GB in DANS-EASY, 4TU.ResearchData, Zenodo are covered by WUR Library.

☐ Costs relating to publications over 100 GB are covered by the research group / project.

☐ The chosen data repository does not involve extra costs.

☐ The chosen storage solution (e.g. Yoda) does not involve extra costs.

☐ The PhD candidate and supervisor will spend at least 10% of their time on research data management to approach the FAIR principles as much as possible.

☐ Other, please specify below.

A guidance. Describe the research project.

1. Name researcher

Please add your full name.

2. Name department

Please choose your department. If you are working from multiple departments, you can choose multiple answers.

3. Name chair group or business unit

Please add the English name of your chair group or business unit exactly as specified on the [chair group](#) or [business unit](#) webpage. Please carefully check the spelling. Preferably copy and paste directly from aforementioned websites.

4. Organisational context of your research project

A data management plan should always include contextual information about the researcher and the project. Then, it is clear to whom / which project the filled in data management plan belong and to which data the described data management practices apply.

5. 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.

6. 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.

7. Requested review

You can request a review of the DMP from data@wur.nl and / or from your local data steward. Provide the contact details of the reviewer in the next question.

8.   Data management support staff

You can consult with Data Management Support staff and / or a data steward to review your data management plan. For data management support staff to review your data management plan, contact data@wur.nl or request feedback via the 'Request feedback' tab in DMPonline.

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 / reviewed by Data Management Support staff will not be considered by NWO. Please do so by contacting data@wur.nl before submitting or via the 'Request feedback' tab in DMPonline.

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

9.   Reuse existing data

When others publish data, this gives you the opportunity to search for and find data to reuse within your research. However, make sure that you are aware of the terms of use (e.g. the licence) of the data that you want to reuse. The terms of use provide you with what you are allowed to do with the data. Want to find existing data? Have a look here for places to start.

In the case of not reusing data: think of the following potential reasons why the reuse of existing data was considered, but not implemented. Is the data you want to reuse 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)?

10. New data

Indicate whether you will collect / generate new data in the research project.

11. Description newly generated data.

Data type refers to whether the data can be classified as for example textual, numeric, audio, film etc. Note that data includes any processing or analysis scripts, protocols, or any other output resulting from research (figures, statistical output, processed data etc.). Try to use software that let you also save files in an open format, so that others can open the files (now and in the future) even if they don't

have the software. Providing the estimated number of files and file size helps to indicate for example the storage requirements and costs.

12. Storage requirement

According to the data table in question 10, estimate the amount of storage space you expect to need.

C guidance. Data storage during research

13. Where will data be stored

Ensure that you, in compliance with the WUR data policy, store data safely. Safe storage solutions would be for example the W-drive, Yoda or SharePoint / Teams, 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. The WUR Data Storage Finder enables researchers to find a suitable storage solution to store and manage digital research data during research.

D guidance. Structuring your data and information

14. Folder structure

Designing a logical folder structure ensures that you and fellow researchers can easily locate data now and in the future. Provide how you plan to organise your (sub)folders. If you already have a folder structure in place, instead of having to go through the whole folder structure and write it down is via the following powershell workflow (note: the following steps only work on Windows / does not work on zip files):

- Go to your file explorer
- Open the parent folder
- In the address bar, type and press enter afterwards: powershell
- A blue screen pops up. Continue in that screen
- Type and press enter afterwards: Get-ChildItem | tree > foldertree.txt
- There is now a file your folder [foldertree.txt], from which you can copy the content in the answer field.

If you are still designing a folder structure, find tips and an example here.

15. 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.

16. 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). Note: for files in Git, dates and version numbers should not be added to the filenames. Git is version control software, keeping track of versions for you and without having to apply file versioning yourself. When applying versioning by changing file names, Git would not recognise a new version as such, but as a completely different file.

E guidance. Data documentation and metadata

It is essential to systematically document your data during your research and when preserving data, whether done by archiving at WUR or depositing data into a data repository. Documentation is information added to data to ensure that the data is understandable and reusable to yourself and to others, both during and after your research. Metadata is also documentation, but in a structured form, describing the data in a way which facilitates cataloguing and discovery of the data.

17. Data documentation and metadata

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.

You can find the templates for the readme file and codebook 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.

Is this project funded by 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).

18. Data quality controls

Describe the performed practices to ensure the quality of the data being collected / generated. Due to a broad range of disciplines, data varies in type, volume, format, etc. As such, there is no 'one size fits all' when it comes to data quality control practices. Overall, data can be checked for completeness, duplication, statistical anomalies, etc. This can include practises like calibration, repeated measurements, standardised data capture, data entry validation, peer review of data, or the use of controlled vocabularies.

F guidance. 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, public image, operational, financial or privacy risks (see data classifications). For privacy questions, please consult your privacy officer (WUR intranet page Privacy & personal data – login required). For information security questions, contact your information security officer or see the WUR intranet page information security (login required).

19. Data ownership and access

The party that is the (rights)holder / has ownership over the data decides what others are allowed to do with it. WUR's policy on IP & Value Creation 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 are (joint) (rights)holders / have (joint)ownership and thus control data sharing and access.

20. Data classification

Classifying the data according to the WUR data classification scheme is important to determine the sensitivity of your data. Personal data may be one of the types of sensitive data one works with. Personal data is any information that (indirectly) 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:
- be accurate and transparent; inform the person concerned as fully as possible about the processing of their personal data
- process personal data only for the purpose for which it has been collected; further processing is generally not allowed
- limit processing only to the personal data that is necessary to achieve the purpose
- ensure that data are accurate and up to date
- store personal data only as long as required for the purpose
- ensure that personal data are secure

Other types of sensitive data (other than personal data) may exist (e.g., (financial) data of companies, or government data). Classifying your data will help you determine the proper steps to take when handling and storing your data. Your Information Security Officer and Privacy Officer can help you with the data classification.

21. Registration in SmartPIA

Check (with your supervisor, project manager) if this project has been registered in WUR's central register for data processing activities. If not registered, register your project via the blue button on [this WUR intranet page](#) (login required). All processing of personal data at WUR must be recorded in SmartPIA by the project manager. Based on the type of personal data used and the purpose of that use, the project manager will be notified if any safety measures need to be taken.

22. Data and privacy protection measures

Working in accordance with the WUR policy, laws, privacy law, or GDPR, requires careful processing and storage of personal data. Risks should 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. See [WUR intranet page Privacy & personal data](#) (login required) for more information about personal data and who is the Privacy Officer of your Science Group.

23. 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.

24. IPR associated with the data

Indicate whether intellectual property rights are affected. If so, explain which and how they will be dealt with. Think about the licence (how others can use the data) and a potential embargo period, which means that others can see that the data exists, but is made available later (e.g. when you want to publish finding or a patent first).

G guidance. Data archiving and publishing

WUR requires that all research data underlying publications is preserved for at least 10 years. Research data should be made publicly available for reuse, 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 reuse. 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 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.

25. Reasons to restrict access or limit public availability.

It could be that (a part of) the data cannot be made publicly available or that access to the data should be restricted. For example, because (part of the) data contains personal data, there are commercial interests involved, or making the data publicly available could infringe on public security.

26. What data is internally archived

It is not always possible to publish data in a repository due to legitimate reasons. However, that does not mean that the data should not be preserved. Additionally,

unpublished data that has not been used in a publication can still have (future) value for reuse.

27. What data will be published

The findability of data for potential users will be enhanced by publishing the data in a data repository. A data repository is a data infrastructure accessible via an online platform. Data repositories compile, manage, preserve, share and provide access to data, metadata and documentation. If you cannot publish the data in a data repository, you can still publish the metadata. Because metadata is machine-readable, the existence of the data can be discovered, adhering to the 'F' (findability) of the FAIR principles. In addition, it becomes more common that research funders (e.g. Horizon Europe) require that, in the case data cannot be published, the metadata associated to the data is published.

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.

28. When and for how long will the data be available for reuse

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. Indicate when the data will be made available. Explain the reason 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.

29. Choosing a data repository

Repositories are data infrastructures that are accessible for via an online platform. Data repositories compile, manage, preserve, share and provide access to data, metadata and documentation. Which repository to choose depends on for example the required level of access to the data, whether software/script/code needs to be published, and if a discipline-specific repository is available. WUR Library supports the multi-disciplinary repositories DANS-EASY, 4TU.ResearchData and Zenodo, which means that data librarians can help publish data and documentation in these repositories. Additionally, deposition costs For DANS-EASY and 4TU.ResearchData

are covered by WUR Library. The [WUR Repository Finder](#) can help you make a decision about which repository to choose to deposit data in. Choose a repository that is certified, that assigns persistent identifiers (e.g. DOIs) to data and / or the metadata, that offers the [licence](#) 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.

30. Metadata

Metadata is highly structured, machine-readable information to describe datasets (or data within datasets). A metadata standard is a set of fixed fields to facilitate findability and searchability of data. Common fields to describe data are author, year, title, spatial coverage, etc. Many repositories use an existing metadata standard, such as [Dublin Core](#) or [DataCite](#). This way, any data you deposit to the repository will use that as a standard. However, when not publishing data (due to legitimate reasons), you should always add relevant metadata (and publish it). 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 storage solution, or use the [Yoda metadata editor](#) and download the metadata as a .json file. However, if there are discipline-specific metadata standards available, these are recommended.

31. Licences

Indicate under which license the data may be reused. Check the commonly used [Creative Commons licences](#). WUR Library advices the CC BY licence for data, because WUR adheres to the motto 'as open as possible, as closed a necessary' and researchers should at least be attributed for their scientific effort. CC BY means that other researchers should credit you as a creator of the data. Note: the requirement to attribute the creator(s) of the data when data carries the CC BY licence depends on a certain level of intellectual creativity displayed. Simple representation of factual data is not protected by copyright and the CC BY license does not apply to that part of the data. However, independently of whether attribution is required or not, attribution is standard good research practice and is part of the principles of scientific research integrity. Note: funders could differ in the requirements for which licence to put on a data and repositories differ in which licences they offer.

32. Generated software

Within research the terms software (also called program or app), code, or scripts may be used. For the purpose of this DMP we mean the following:

- **Software:** A standalone set of compiled scripts and code into machine language instructions that can run independently for users to interactively use said scripts and code. Compiled and documented in such a way that a variety of users can use the software in their own workflow. Software examples are MS Word, R and Rstudio, Python and PyCharm.
- **Code:** A set of scripts that users can run interactively for their own purposes and different types of research data or disciplines. Often compiled, but always documented in such a way that a variety of users can interactively reuse the code in their own workflow. Code examples: R libraries, Python libraries, Shiny Apps.
- **Scripts:** A set of lines to 'automate' processes often very specific for single use purposes or specific to a type of research data, cannot run independently, used within software to be interpreted by that software and control a certain aspect of that software, often created for one's own specific purpose and not interactive, terminates when done with its work. Script examples are research data processing / cleaning scripts, (statistical) data analysis scripts, figure and table creation scripts, scripts to call on an API to extract data – all for one's own specific set of research data.

H guidance. Data management costs

33. Resources required

Think about investment in both money and time when it comes to data storage, data preparation, data archiving or publishing etc. Take into account the work required during research for proper handling, storage and documentation, as well as the work after research to make your data as FAIR as possible.