



**Individual and project data management plan (DMP) format**

**For CPT**

**Version 301220**

**Format for a data management plan**

* *Data management is required for PhD projects as well as other research projects. DMPs are therefore required for all research projects.*
* *This format consists of 9 sections.*
* *This format is intended to give you a helping hand in writing your DMP. You are free to add content elements when your particular research project requires it. The DMP should, however, include answers to the nine sections below.*
* *Hover over (info) for more information about each section. It will give you a hyperlink (CTRL-click) to additional information in the Appendix.*
* *For more information about the requirement of filling in a data management plan for PhDs, visit* [*this URL*](http://www.wageningenur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/Data-Management-Planning-1.htm)*. The filled-out DMP will be an appendix to your research proposal and will be subject to review by Wageningen Graduate Schools.*
* *Much additional information can be found at:* [*https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC.htm*](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC.htm)
* *Any remaining questions? Contact* *datamanagement.support@wur.nl* *or visit* [*www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub.htm*](http://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub.htm). In case of remaining issues or concerns: *approach the data management steward for your Chair Group. For KTI: Annemarie van Paassen. For COM: Margit van Wessel*
* *Elaborate the Data Management Plan before you start the field work and send it to the Data Management steward of the respective chair group, as information. Research project coordinators and PhD supervisors are responsible for the safe and systematic storage of data, but the stewards will monitor with staff once a year and facilitate where possible.*

**1.** **Describe the organizational context** [(info)](#organizationalcontext" \o "information)

|  |  |
| --- | --- |
| **Name** |  |
| **Date of this version** |  |
| **Chair group** |  |
| **Graduate school (if applicable)** |  |
| **Supervisor/ (co-)promotors (if applicable)** |  |
| **Other researchers involved - mention also Chair Group or name of institute (if applicable)** |  |
| **Commissioning organization (if applicable)** |  |
| **Start date of project** |  |
| **File name of this DMP**  |  |

**2. Give a short description of your research project** [(info)](#shortdescriptionofresearch" \o "information)

|  |  |
| --- | --- |
| **Title** |  |
| **Abstract** |  |

 **3.** **Define data management roles**

[(info)](#Datamanagementroles" \o "Information)

Data management is primarily the responsibility of the researcher(s), with supervisors in a coaching role. Data management stewards have a role facilitating and overseeing data management at Chair Group level. Please specify roles below.

|  |  |
| --- | --- |
| **Roles** |  |
| Who is **collecting** the data? |  |
| Who is **processing & analysing the data**, taking care of day to **day data storage and back-up** of data in systemic and safe way? |  |
| Who takes care of **data documentation** to make the data sets understandable? |  |
| How is data ownership and storage responsibilities regulated between WUR and non-WUR **research partners**? For info look at WUR data ownership decision aid <https://edepot.wur.nl/477968> |  |
| Who takes care the data are safely stored **at Chair Group/CPT level (W:drive)** during the research project? Note: the structure for data management on the W-drive is still being developed and will be made available early 2021. This document will be updated then and CPT colleagues will be informed.  |  |
| Who takes care the data are registered in[**PURE**](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Data-policy/Registering.htm) **and saved in an external repository** after completion of a research project? (if applicable – see also section 8 + CPT DMP) |  |
| Name data management steward |  |

**4. Give an overview of expected** **types of research data** [(info)](#Typeofresearchdata" \o "Information), **software choices** [(info)](#Softwarechoices).

If you are working with datasets that are extraordinarily large for SSG, also indicate data size & growth [(info)](#Datasizeandgrowth)

Note: Though privately kept diaries and observation notes contribute to your analysis and publication, it is often hard to make it easily accessible and relevant to outsiders. You may also see these as deeply personal. Hence you can well consider keeping these out of W.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data stage** | **Specification of type of research data** | **Software choice**  | **Data size/growth**  |
| Use of existing data |  |  |  |
| Formats used to generate data (interview formats etc.) |  |  |  |
| Raw data, qualitative |  |  |  |
| Raw data, quantitative |  |  |  |
| First processed data (e.g. Atlas-ti etc.) |  |  |  |
| Analysed data for publication |  |  |  |
| Other? |  |  |  |

**5.** **Short-term storage solutions** [(info)](#Shortermstorage)Describe where the data will be stored and how the back-up is organised.

|  |  |  |
| --- | --- | --- |
| **Data stage** | **Storage location** | **Backup procedures**(storage medium and location/ how often?)  |
| Use of existing data |  |  |
| Formats used to generate data (interview formats etc.) |  |  |
| Raw data, qualitative |  |  |
| Raw data, quantitative |  |  |
| First processed data (e.g. Atlas.ti etc.) |  |  |
| Analysed data for publication |  |  |
| Other? |  |  |

**6.** **Structuring your data and information** [(info)](#Directoryandfilenaming" \o "Information and examples)

|  |
| --- |
| Include with your dataset a document entitled ‘Description of data’ which includes the following elements: Title of the dataset:Creators:Contributors:Related publications:Description:Keywords:Spatial coverage:Temporal coverage:This dataset contains the following files:......... |

Below, provide a visual representation of the system for directory- and file names you intend to use. (See [examples](#ExamplesDirectory) in appendix for inspiration)

|  |
| --- |
|  |

Does your workflow provide for version control? If not, describe how you intend to keep versions apart.

|  |
| --- |
|  |

 **7.** **Documentation and metadata** [(info)](#DocumentationandMetadata" \o "Information)

Describe how you are going to document your data collection process, what the resulting data files comprise and how they will be processed further. During and after your project, this documentation is to be included with your dataset. Think about documenting the:

*1. content (what does your dataset contain?)*

*2. context (who, what, why, where and how will the data be collected and analysed?)*

*3. process (are there specific processes and does it make sense to organise your notes according to these processes?)*

|  |
| --- |
|  |

**8.** **Personal data protection, sharing and ownership** [(info)](#SharingandOwnership" \o "Information)

Note that you are asked to archive the data related to your publications for 10 years on the W-drive. The respective folder with your data, processed data and meta data will be protected and only accessible to the respective research team. Private and sensitive, but relevant data for your publication are safe as the data management steward and chairholder will not access it without an urgent need and your agreement.

|  |  |
| --- | --- |
| **Informed consent & personal data protection** | **Description of considerations, actions and implications**  |
| Have you attained informed consent of the research respondents? How?  |  |
| Have you consulted and applied ethical rules for personal data protection (e.g. EU [General Data Protection Regulation - GDPR](https://www.openaire.eu/blogs/gdpr-and-the-research-process-what-you-need-to-know?Itemid=2068&%3Bhighlight=WyJzZW5zaXRpdmUiLCJkYXRhIiwiZGF0YSdzIiwiJ2RhdGEiLCJkYXRhJyIsInNlbnNpdGl2ZSBkYXRhIl0%3D))? |  |
| Have you anonymised and encrypted research data? (if applicable) |  |
| **Sensitivity, Sharing and Ownership** | **(With) who(m), what, why, when and how?**  |
| Data **sharing** - Given data confidentiality consideration: what level of confidentiality apples to your data? Based on what considerations and/or agreements? See <https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Data-confidentiality.htm>- Do you expect that certain specified parties (like project partners) may be interested in re-using your data?Do you have plans to share your data with these parties? - Do you plan to share your data via external repositories like [DANS-EASY](https://easy.dans.knaw.nl/ui/home)? and [4TU.Centre for Research Data](https://researchdata.4tu.nl/?file%3A788=)? Making your data accessible in this way is standard good practice (unless inhibited by confidentiality requirements) and is required by VSNU. See <https://www.go-fair.org/fair-principles/> for the guidelines adhered to by VSNU. This form of sharing may also be a requirement from journals you publish in. If yes, which data archive do you intend to use for sharing data? -Do you plan to share data with project partners during the project through a Cloud service? Teams is the recommended option. Indicate your plans on this front. |  |
| Data **ownership** - Who owns the data? - Are there funders’ requirements to share your data, or to impose an embargo? - Are there agreements on how the data can be used and shared further within your group or with other parties involved in this research? (outside your group or outside Wageningen University & Research) |  |

**9. Long-term storage** [(info)](#Longtermstorage)
For many if not all projects, long-term storage will be required.

|  |  |
| --- | --- |
| Do your research data need to be stored long-term? Which part of your data? Based on what argumentation?Who takes care the data are safely stored **at Chair Group/CPT level (W:drive)** after the research project? |  |
| Who will have access to the data stored on the W:drive after the completion of the research project? Make sure this includes the group’s data management steward. |  |

Data need to be stored long-term on the W-drive. In addition, data (including confidential data) can also be stored in external repositories like [DANS-EASY](https://easy.dans.knaw.nl/ui/home)? and [4TU.Centre for Research Data](https://researchdata.4tu.nl/?file%3A788=).

|  |
| --- |
|  I intend to archive … data in … |

**Appendix with additional information**(may be deleted after completion of your DMP)

**1. Organizational context**
A Data Management Plan should leave no doubt as to which research(er) it belongs to.

[[Back]](#organizationalcontextback)

**2. Short description of research**Giving a short description of your research gives context to your data management plan. It makes it easier for the reader to understand without having to check your research plan.

[[Back]](#shortdescriptionofresearchback)

**3. Data management roles**Identifying persons who are - *or can be* - of assistance in your daily data management practices, smoothens your data collection process. Maybe some people have special responsibilities regarding data management? (E.g. a division of labour between programmers and those who do observations?)

Having a closer look at data management roles places data collection in a broader perspective than your research project alone. Discussing both your roles as well as those of your supervisor and other colleagues prevents possible future issues concerning [data ownership](#SharingandOwnership).

[[Back]](#Datamanagementrolesback)

**4a. Type of research data**Identifying your possible research data before you actually start collecting those data, makes sure no research output is overlooked. You can choose from:

*- Raw data (i.e. data from experiments or observations (e.g. a lab note book))*

*- Derived / processed data*

*- Models (including data from simulations)*

If you use derived data, you should at least say how you handle the raw data. (NB: If you haven’t produced the data yourself, that may be of influence on what you are allowed to do with the data (see [data ownership](#SharingandOwnership))).

To give you an example of the diverse outputs of research data, read the [following list](http://www2.le.ac.uk/services/research-data/documents/an-introduction-to-managing-research-data):

• Documents (text, MS Word), spread sheets

• Scanned laboratory notebooks, field notebooks, diaries

• Online questionnaires, transcripts, or surveys

• Digital audio or video recordings

• Transcribed test responses

• Database contents

• Digital models, algorithms, or scripts

• Contents of an application (input, output, log files for analysis software, simulation software, or schemas)

• Documented methodologies and workflows

• Records of standard operating procedures and protocols

[[Back]](#Typeofresearchdataback)

**4b. Software choices**What software will you use to create, analyse and visualize your data? Are these choices common practice in your field?

Software choices affect whether current and future users can actually view and use the data you have collected. For example, if you use proprietary software (software owned by the person or company that has developed it, and which may be required to read the associated file format), it may not be possible for people outside your field to do anything with your data except getting an error trying to read them. Also, some software may come with its own systems for [folders and file names](#Directoryandfilenaming). Think software choices through with future users in mind.

[[Back]](#Softwarechoicesback)

**4c. Data size/growth**Give an estimate in (Mega – Giga – Tera) Bytes. Making an educated guess on the size of your research data output, indicates where you should store your data. If you will produce terabytes of data, for example, a simple hard drive will not suffice. In short, data size influences [data storage solutions](#Shortermstorage).

[[Back]](#Datasizeandgrowthback)

**5. Short-term storage**

You need to decide how you will keep your data safe in the short term. Where will the data be stored physically and how will it be backed up? Do you follow the common practice in [WUR](https://www.wur.nl/en/Value-Creation-Cooperation/WDCC/Data-Management-WDCC/Doing/Storage-solutions.htm#:~:text=Therefore%20WUR%20policy%20on%20data,or%20on%20a%20Team%20Site.&text=Every%20WUR%20account%20holder%20can,to%20the%20IT%20Service%20Desk.) and especially your research group, and if not, why not?
The table below may be of assistance in making an informed choice for short-term storage. You can also consult the the intranet page with [Frequent Asked Questions on file management](https://intranet.wur.nl/umbraco/en/practical-information/it-services/file-storage), the Data Storage Decision Aid from <https://edepot.wur.nl/446448> and for confidential data look at <https://edepot.wur.nl/446447>.

*USB-stick by* [*http://dpconline.org/*](http://dpconline.org/) *shared* [*via twitter*](https://twitter.com/digitalfay/status/411444578122600450/photo/1)

|  |  |  |  |
| --- | --- | --- | --- |
| **Storage solutions** | **Advantages** | **Disadvantages** | **Suitable for** |
| **Personal computer & laptop****(WUR computers are encrypted for safe storage)** | Always availablePortable | Drive may failLaptop may be stolen | Temporary storage |
| **Networked drives**M: and W: drive servers managed by your research group, Wageningen University & Research, or facilities like a NAS-server | Regularly backed upStored securely in a single place. Centralized storage makes it easier to maintain and back up.  | Relatively high costs | The master copy of your data(if enough storage space is provided)  |
| **Cloud services**For WUR [OneDrive for Business](https://sharepoint.wur.nl/sites/it/dataopslag_storage/EN-OneDriveForBusiness.pdf), MS Teams and SURFfilesender are preferred above Dropbox, SkyDrive, etc.  | Automatic synchronization between files online and folder on PCEasy to use and access | It is not sure whether data security is taken care ofYou don’t have direct influence on how often backups take place and by whom | Storage when internet does not allow access to M: or W; drives, and data sharing  |
| **External storage devices**USB flash drive, DVD/CD, external hard drive | Low costPortability  | Easily damaged or lost  | Temporary storage |

[[Back]](#Shortermstorageback" \o "Back)

**6. Structuring your data and information**We all think we are going to remember how we named our files and where we store them. But the truth is: we never do :-) Let alone, our fellow researchers. Time invested in giving thought about an unambiguous directory and file-naming systems pays off for your future self.

A ***file name*** is a principal identifier for a file. Good file names provide useful cues to the content, status and version of a file, uniquely identify a file and help in classifying and sorting files. File names that reflect the file content also facilitate searching and discovering files. In collaborative research, it is essential to keep track of changes and edits to files via the file name. File names should be independent of the location of the file on a computer.

**Best practice is to:**

* create meaningful but brief names
(not: dataset1 but pathogenmeasurement021213\_v01.xls)
* Indicate versions, e.g. \_v01 (master files/milestone files)
* avoid using spaces, dots and special characters (& or ? or !)
* use hyphens (-) or underscores (\_) to separate elements in a file name
* avoid very long file names
* reserve the 3-letter file extension for application-specific codes of file format (e.g. .doc, .xls, .mov, .tif)

Whilst computers add basic information and properties to a file, such as file type, date and time of creation and modification, this is not reliable data management. It is better to record such essential information in the file name.

**Examples of useful file names:**

* FG1\_CONS\_2010-02-12.rtf : interview transcript of the first focus group with consumers, that took place on 12 February 2010
* Int024\_AP\_2008-06-05.doc : interview with participant 024, interviewed by Anne Parsons on 5 June 2008
* BDHSurveyProcedures\_00\_04.pdf : version 4 of the survey procedures for the British Dental Health Survey

**Examples of poor file naming:**

* SrvMthdDraft.doc, SrvMthdFinal.doc, SrvMthdLastOne.doc and SrvMthdFridaynight.doc
* Focus group consumers 12 Feb?.doc
* Health&Safety Procedures1

Making a ***folder structure*** is meant as an exercise in logic. Think carefully how best to structure files in folders, in order to make it easy to locate and organise files and versions while working. When working in collaboration with others, the need for an orderly structure is even higher. Consider the best hierarchy for files, deciding whether a deep or shallow hierarchy is preferable.

Underneath you find an example for a PhD research project but could easily do for another joint research project. First you make a data management plan and store it. Then you systematically describe and store the data gathered in the field work in the ‘data and documentation’ folder. These data serve for various publications, hence in the folder Research methodology and analysis you make a folder per publication in which you store the applied research methodology and analytical framework for this respective publication, the processed data files as well as the final publication.

During the research the main researcher stores the data in M and in OneDriveforBusiness (or other agreed safe application) more accessible in the field and possible to share with research partners. There are 3 moments in which you are asked to put your information in the CPT archive : W.

1. When you accomplished the data management plan
2. When you have finalised (a reasonable part of) the data collection
3. When you have finished the analysis for a respective publication

Figure 1. Proposed folder arrangement and file naming strategy.



[[Back]](#Directoryandfilenamingback)

# **7. Documentation or so-called metadata**

# Good documentation [ensures your data can be](http://www2.le.ac.uk/services/research-data/organise-data/metadata):

*• Searched for and retrieved*

*• Understood now and in the future*

*• Properly interpreted, as relevant context is available.*

Metadata documents other data, it’s ‘data about data’. It concerns a ‘readme.txt’ or Word file, providing information about an item and its relevance so that it can easily be found and understood when needed.

1. It provides the reference to the article published plus DOI, and structured information explaining the purpose, origin, time references, geographic location, creator, access conditions and terms of use of a data collection.
2. It enables correct interpretation; hence it describes the type of data gathered and the applied research methodology and analysis:

*a. The content (what does your dataset contain?)*

*b. The context/research methodology (who, what, why, where and how are the data collected and analysed?)*

*c. How are the data stored: what data sets can be found where (structure of the sub-folders in which you stored the different type of data at W)*

[[Back]](#DocumentationandMetadataback)

**8. Personal data protection, sensitive data, sharing and ownership**
Sensitive data is data that must be protected against unwanted disclosure. Access to sensitive data should be safeguarded. Protection of sensitive data may be required for legal or ethical reasons, for issues pertaining to personal privacy, or for proprietary considerations.

Examples of sensitive data are:

* **Personal data:** identifiers such as names or identification numbers, physical, physiological, genetic, mental, economic, cultural or social characteristics, it also includes location data from GPS or mobile phones
* **Confidential data**: trade secrets, investigations, data protected by intellectual property rights Security: passwords, financial information, national safety, military information...
* **Combination of different datasets** that can be combined into sensitive or personal data
* **Personal and sensitive metadata** 

When handling and dealing with sensitive data, keep in mind that special attention should be given to collecting, processing, handling and storing data throughout the research process. In particular research data that contains personal data with which a living person can, directly or indirectly, be identified has to be handled with care. This concerns both textual data and image and sound data. Examples of direct data are someone’s name and address, but it could also be a photo or an interview. An indirect fact is, for example, someone’s employer. For personal data fully informed consent should be given for collecting, processing and storing data.

The EU has strong regulations regarding personal (e.g. [General Data Protection Regulation - GDPR](https://www.openaire.eu/blogs/gdpr-and-the-research-process-what-you-need-to-know?Itemid=2068&%3Bhighlight=WyJzZW5zaXRpdmUiLCJkYXRhIiwiZGF0YSdzIiwiJ2RhdGEiLCJkYXRhJyIsInNlbnNpdGl2ZSBkYXRhIl0%3D)) and sensitive non-personal data. Based on GDPR, the Netherlands Authority of Data Protection formulated the General Regulations Data Protection ([Algemene Verordening Gegevensbescherming (AVG))](https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-nieuwe-europese-privacywetgeving/mag-u-persoonsgegevens-verwerken). Background information may be found on the website of the country coordination point for research data management (Landelijk Coördinatiepunt Research Data management, LCRDM) https://www.lcrdm.nl/en, especially on the page of [advice and tips](https://www.lcrdm.nl/en/rdm-advice-tips) which contains the [rich picture underneath](https://surfdrive.surf.nl/files/index.php/apps/richdocuments/public?fileId=3402401273&shareToken=BPRxchnZ44NZAgW)





More info on the Dutch regulations (in Dutch) can be found at:

* De [Central Committee Human related research](https://www.ccmo.nl/onderzoekers/wet-en-regelgeving-voor-medisch-wetenschappelijk-onderzoek)
* The Ethical, Legal, Social Implications ([ELSI) servicedesk](https://www.elsi.health-ri.nl/) for biomedical and Health research
* In the second half of 2020 the Dutch Association of Universities VNSU is expected to launch an update of the code of conduct personal data for scientific research.

[International](https://autoriteitpersoonsgegevens.nl/nl/over-privacy/wetten/internationale-privacywetgeving): The protection of personal data differs from country to county. Within the EU the rules are quite similar, as they are based on the EU General Data Protection Regulation GDPR Outside the EU the rules differ. To exchange personal data  [from the EU](https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/internationaal-gegevensverkeer/doorgifte-naar-derde-landen) , hence also from the Netherlands, to a country outside te EU is only allowed when this country provides adequate data protection.

### How to prepare sensitive data for storage and, if desirable, sharing?

When storing sensitive data, the first concern is to find a good security strategy for your type of data. To share and research on sensitive data an easy solution is to give open access to the metadata, but restrict access to the corresponding [sensitive data sets](https://edepot.wur.nl/446447). In this context, one should also consider that not only sensitive data, but also personal and sensitive metadata might exist. Additional solutions are the anonymization of data and metadata, [access restrictions](https://edepot.wur.nl/446448), licenses for usage, and permissions granted from copyright holders.

**Anonymization**

Anonymization irreversibly destroys any way of identifying the data subject. Personal data that has been rendered anonymous in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible.

OpenAIRE provides researchers with a tool to anonymise data: [Amnesia](https://amnesia.openaire.eu/). The guide for which you can find [here](https://www.openaire.eu/amnesia-guide).

**Pseudonymization**

Pseudonymization substitutes the identity of the data subject in such a way that additional information is required to re-identify the data subject. The pseudonym allows tracking back of data to its origins, which distinguishes pseudonymization from anonymization, where all person-related data that could allow backtracking has been purged. Pseudonymised data are still legally considered as sensitive data because the data can be linked back to a person, but it's considered as a secure approach since personal identifiers are stored somewhere else.

**Encryption**

Encryption is a very generic term and there are many ways to encrypt data. All WUR computers are encrypted, for safe data storage. The key to a good encryption strategy is using strong encryption and proper key management. Encrypt sensitive data before it is shared. Encryption will make your data totally unintelligible to those who may try to access it which might reduce re-usability.

In case none of these options are available for your dataset, data should not be made open and be archived under a closed license in a [Trustworthy Repository](https://www.openaire.eu/find-trustworthy-data-repository). You can however publish a description (i.e. public metadata) of your data without making the data itself openly accessible, which enables you to place conditions around access to the data.

 [[Back]](#SharingandOwnershipBack)

**9. Long-term storage**The code of conduct for scientific practice requires that you ***retain your data for ten years after you have published your article*** and make it available upon request for verification purposes. You may be able to fulfil such requests while you are in your present job, but to make data available for a longer period for re-use and verification, you should store it in a data archive with proper documentation and in a sustainable data format.

Research data should always be retained for the short term. Additional long-term storage is at least recommended if:

*1. data underlies publications
2. Long-term storage and sharing is required by funders
3. there is a legal requirement
4. it is likely that others may want to re-use the data*

In the flow chart below, you can check whether your data are a possible candidate for long-term archiving and re-use. If your answer(s) to the question(s) is/are YES, you should at least consider storing these data for the long term.



Of course, practical considerations may make you decide not to pursue long-term storage even though your data might be of future value. Think about the following:
*- Can the data be understood? Is the data documentation clear enough? If not, long term archiving is of no use.*

*- Are there any legal requirements NOT to share the data? (privacy, etc.)*

*- Does the expense of archiving justify the long-term value?*

If you choose long-term storage, is there a common practice/institution in your field? Or do you intend to use the services provided by Wageningen UR? Wageningen UR Library has agreements with the repository from The Royal Netherlands Academy of Arts and Sciences, called [DANS EASY](https://easy.dans.knaw.nl/ui/home), as well as with the data repository for the technical sciences, [4TU.ResearchData](http://researchdata.4tu.nl/en/home/).
You can contact datamanagement.support@wur.nl for more information.

# **Examples**

# - For more inspiration, [see how Frits van Evert](http://www.wageningenur.nl/en/show/Dataarchivingdoubleinterview.htm), researcher at Agrosystems Research, supports archiving research data in a sustainable way.- The dataset by dr. ir. P.A.J. v Oort on *Key weather extremes affecting potato production in the Netherlands* is a good example of what sustainable data archiving in a data repository looks like [https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset:55620](https://easy.dans.knaw.nl/ui/datasets/id/easy-dataset%3A55620).

**Long term storage of sensitive data**

Sensitive Data management in social science is practiced by DANS, the Data Archiving and Networked Services in the Netherlands that promotes sustained access to research data form the areas of humanities and social sciences research. Data protection is needed to hide the identity of respondents of interviews or test subjects, in particular in research addressing wars, armed conflicts, medical or psychiatric treatments, handicaps (specially handicapped children). Such sensitive data can only be accessed by authorized users of which the identity has been checked (authentication) and who have signed special agreement for the conditions of use.

Legally, the ownership of research data isn’t very clear. So, it is important to have a sound understanding of what you are allowed to do with the data and how you will leave your data behind when the time comes to pursue your career at another organisation. Therefore, ownership isn’t so much about ‘property’ (to whom do the data belong). It’s about *custodianship*: What is going to happen to the data when your project is finished? Who is the person responsible for taking care of your data and ensuring it can be accessed when you are gone? Can you still publish about the data and use them for further research when you have left the university or research centre? With whom are you going to make these arrangements and how is the rest of the world going to know?

These questions are generally best discussed with you supervisor and funders.

[[Back]](#LongtermstorageBack)