# Data management plans

Data steward network meeting

November 8, 2018

Marleen Noomen - Library


WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

# DMP requirements at WUR

*https://www.wur.nl/en/Value-Creation_Cooperation/WDCC/Data-Management-WDCC/Planning/Institutional-requirements.htm*

WAGENINGEN UNIVERSITY & RESEARCH

Home ▾ WDCC ▾ Data Management ▾

Not only PhD students need a DMP, also chair groups need one!

## Institutional requirements

Wageningen Graduate Schools (WGS) requires that PhD candidates create a Data Management Plan for their research. Each graduate school has slightly different requirements on how to submit the DMP with other documents.

Once PhD candidates have created their DMP, it becomes an appendix to their research proposal and may be reviewed by Wageningen Graduate Schools.

Support and information

Wageningen Data Competence Center

Contact form

- Back to Data Management
- Planning your research
- Doing your research
- Finishing your research

# Data management plan

Not only WUR requires a data management plan, also funders.

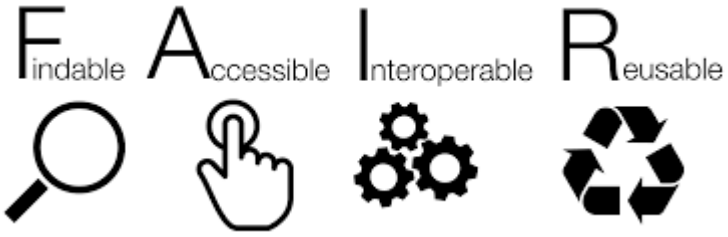
- NWO
- ZonMw
- Horizon2020
- ERC
- ...





Although they cover essentially the same subjects, they may vary in detail.

# FAIR data




What do all funders require?

→ Data is as **FAIR** as possible

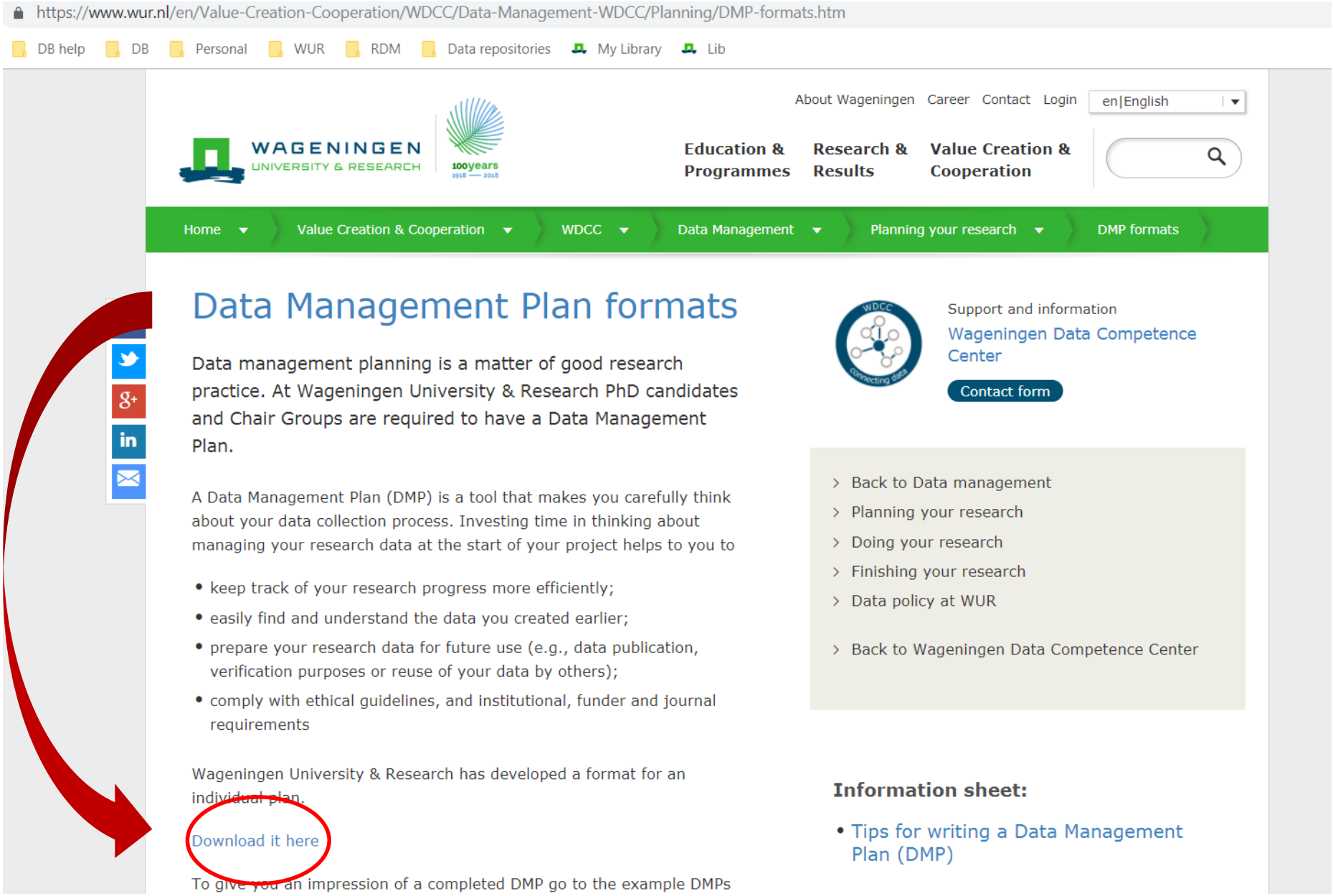
F = findable (e.g., archived, has a unique identifier)

A = accessible (e.g., Open Access or via authentication)

I = interoperable (e.g., can easily be combined with other datasets by computers or humans)

R = reusable (e.g., ready to be used for future research)

# WUR template



# Section 1: context

## 1. Describe the organizational context (info)

| | |
|---|---|
| **Name** | |
| **Date** | |
| **Chair group** | |
| **Graduate school** | |
| **Supervisor/ (co-)promotors** | |
| **Start date of project** | |
| **File name of this DMP** | |

# Section 2: description

**2. Give a short description of your research project (info)**

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

# Section 3: data management roles

### 3. Define data management roles (info)

Change/add roles over time

| **Roles** | |
|---|---|
| Who is **collecting** the data? | |
| Who is **analysing** the data? | List this as specific as possible |
| **Other**<br>(Is there a person in the research group with a specific responsibility for data management? Do other persons contribute, for example by writing code?) | It is important to keep track of who worked on the data at what time; for example by keeping an Electronic Lab Notebook |
| What is the role of your **supervisor?** | |

# Section 4: type and size of data, software

**4. Give an overview of expected types of research data (info), software choices (info), and data size & growth (info)**

| Data stage | Specification of type of research data | Software choice | Data size/ growth |
|---|---|---|---|
| Raw data | | | |
| Processed data | | | |
| Models/code | | | |
| Other? | | | |

# Section 4: example

These are closed formats. Preferred software/output is open format

| Data stage | Specification of type of research data | Software choice | Data size/ growth |
|---|---|---|---|
| Raw data | Video files (.mp4), interviews<br>.txt files | Media recorder, MS Word | Few GBs<br>20-40 GB |
| Processed data | Numerical files, statistical analyses<br>.csv files | Excel, SPSS<br>R | Few GBs<br>10-15 GB |
| Models/code | - | | |
| Other? | - | | |

Give an estimate of the size

# Section 5: short-term storage

## 5. Short-term storage solutions* (info)

Describe where the data will be stored physically and how the back-up is organised.

| Data stage | Storage location | Backup procedures (storage medium and location/ how often?) |
|---|---|---|
| Raw data | | |
| Processed data | | |
| Models/code | | |
| Other? | | |

# Section 5: example

Vulnerable to theft, loss, breaking

**Not the same location!**

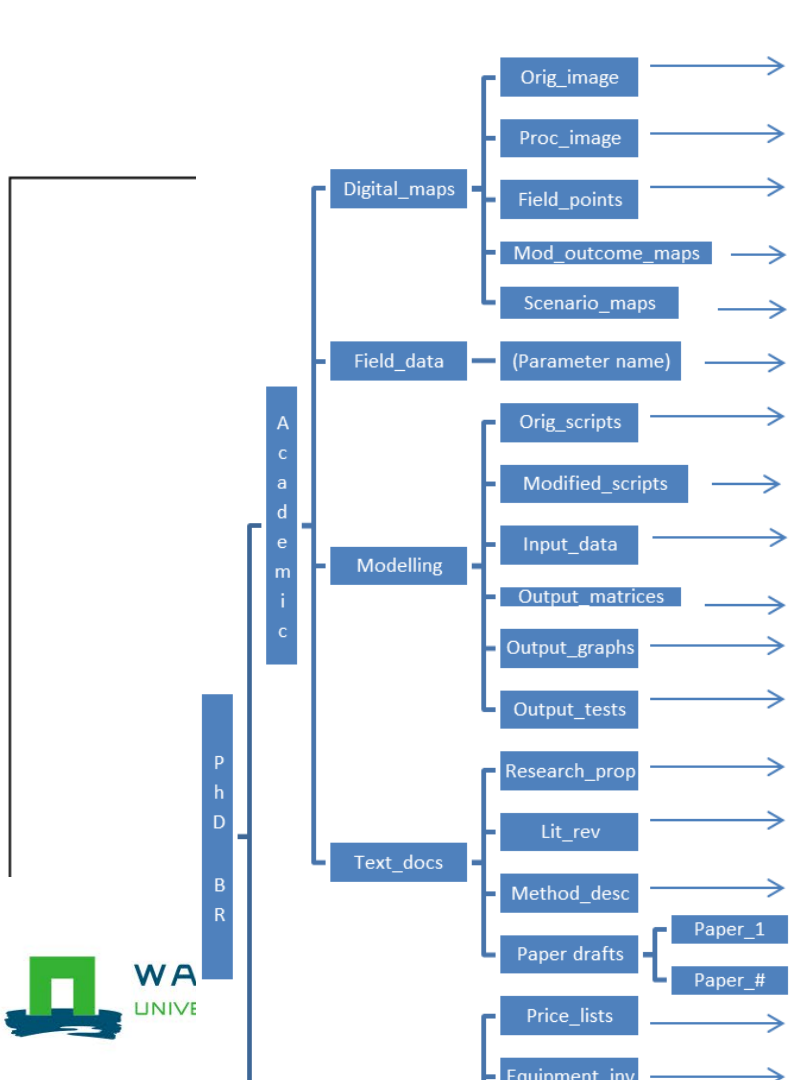
| Data stage | Storage location | Backup procedures (storage medium and location/ how often?) |
|---|---|---|
| Raw data | Laptop | Laptop & ~~external hard drive~~ / asap |
| Processed data | Network drive W-drive | Weekly |
| Models/code | | |
| Other? | | |

e.g. raw on lab computer, processed on W-drive

W-drive is backed up

# Section 6: structure of data

## 6. Structuring your data and information (info)




| Name | Example |
|---|---|
| Month_year_type | 01_1989_landsat |
| Month_year_type_process | 01_1989_landsat_orthorect |
| Field_points_date of actualization | Field_points_11_09_13 |
| Month_year_outcome type | 01_1989_wateryield |
| Year_parameter | 2050_wateryield |
| Parameter_date of actualization | Rainfall_11_09_2013 |
| Model_script_version | Shetran_script_v2.5 |
| Model_script_modification_trial# | Shetran_script_cloudwater_t1 |
| Parameter_col#_row#_trial# | Rainfall_4_250_t1 |
| Est. parameter_col#_row#_trial# | ET_4_250_t1 |
| Graph_x_y_trial# | Scatterplot_rainf_ET_t1 |
| Test_parameter_trial# | Senstest_rainfall_t1 |
| RP_trial# & _revised by | RP_t1_revLG |
| Topic_subtopic | Modelling_wateryields |
| Topic_method | Soil inf_infiltrometer |
| Paper code_date & revised by | TMCFvsPas_11_09_2014_LG |
| Company_date | Eijkelkamp_11_09_2013 |
| Equipment_list | Equipment_list |

# Section 7: documentation and metadata

## 7. Documentation and metadata* (info)

Describe how you are going to document your data collection process, what the resulting data files comprise and how they will be processed further. 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?)*

# Section 7: example

Think of:
- README.txt (NOT .docx!);
- data dictionaries to explain variables;
- adding metadata;
- (electronic) notebooks for research notes as well as roles

There will be a 'Readme directory file.d[illegible]' file in which all the specifications of the project/course (etc.) will be stored (se[illegible]ta structure above).

This file will contain the location of all relevant files, their specifications and their relations.

# Section 8: sharing, ownership, privacy

**8. Sharing, ownership and privacy*** (info)

| Sharing, ownership and privacy | (With) who(m), what and how? |
|---|---|
| Data **sharing**<br>- Do you expect that others may be interested in re-using your data? Do you have plans to share your data with these parties?<br>- How are you going to make sure your datafiles will be accessible once you leave the department? Who will take care of your data? | |
| Data **ownership**<br>- Any funders requirements to share your data, or to impose an embargo?<br>- Are there agreements on how the data will be used and shared within your group or with other parties involved in this research? (outside your group or outside Wageningen University & Research) | |
| **Privacy**<br>- Are there privacy or security issues, and if there are, how are you dealing with them? | |

= sharing during research and afterwards within WUR; not through a repository (9)

avoid conflicts with e.g. third parties

encryption, access, anonymisation...

# Section 9: long term storage

**9. Long-term storage*** (info)

Which part of your research data has value for long-term storage? Do you intend to preserve these data for the long term?

| Yes or no ? | Argumentation |
|---|---|
| | - 10 years retention obliged<br>- Most data (/parts) are of value to others |

Which data archive do you intend to use?

I intend to archive … data in …

- Browse **re3data.org**
- Repositories in your discipline?
- WUR archiving support: DANS / 4TU

# Section 9: example

ALL data, or part of the data, and which part then?

**9. Long-term storage*** (info)

Which part of your research data has value for long-term storage? Do you intend to preserve these data for the long term?

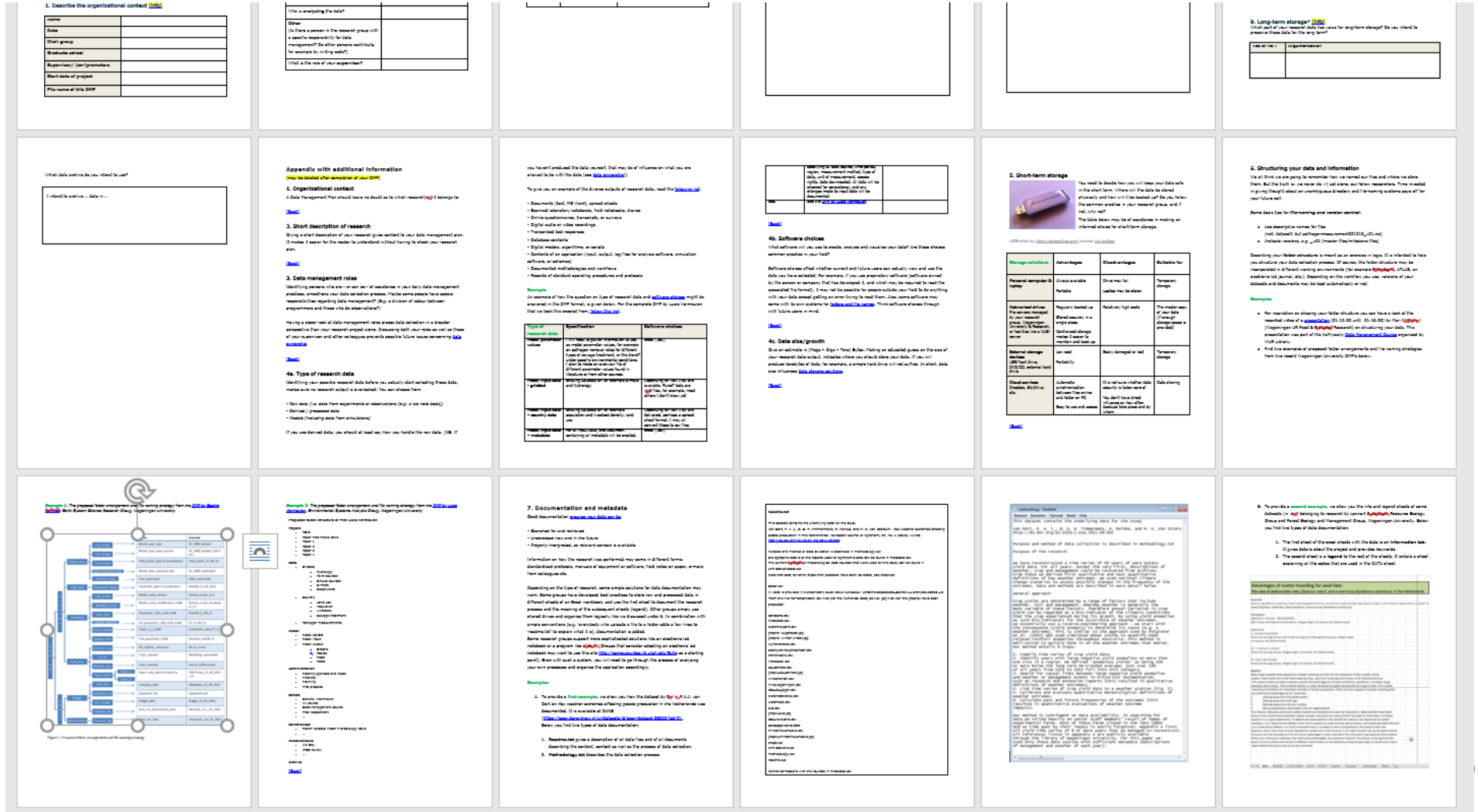
| Yes or no ? | Argumentation |
|---|---|
| Yes | Basic requirement of 10 years storage of all data and documentation. |

Not a data archive!

Which data archive do you intend to use?

I intend to archive all data on external hard drives.

# Lots of info in the DMP itself



# Data Management Support at WUR

**To get answers/support:**

- data@wur.nl (Data Management Support and WDCC)
- servicedesk.it@wur.nl (storage and security)

**To stay up to date:**

Intranet group Data@WUR

**For information:**

The Data Management website