
Plan Overview

A Data Management Plan created using DMPonline

Title: TPM - DMP Manual

Creator: Nicolas Dintzner

Affiliation: Delft University of Technology

Template: TU Delft Data Management Plan template (2021)

Project abstract:

How to use this manual: open it, and open your own DMP at the same time, fill your DMP by copy/pasting and adjusting the content from this DMP into your own.

Most important: Unless the cooperation is a publicly advertised fact (you could *tweet* it), do not use DMPonline for activities related to the military, national security, or ministry of justice, or similar public bodies.

Feel free to use any abstract you already have, and copy it here.

Ideally, include information regarding the type of activities you will have (e.g. : interviews, surveys, workshop, lab/classroom intervention, writing software, using web tools, ...)

If possible: be aware of the contract you are working before starting (e.g.: internship agreement for MSc student, internal TUD MSc project, participation in consortium - Horizon projects, specific grants) - this will help

Optional: research domain - if you find something nice.

Optional: give rough estimate of the time for the activity described above in the "Project start" and "end" date -

ID: 145971

Last modified: 05-07-2024

Grant number / URL: Funder information (above included) is optional.

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

TPM - DMP Manual

0. Administrative questions

1. Name of data management support staff consulted during the preparation of this plan.

The review by the data steward is optional. If you don't request feedback - write it here, and leave the field below empty. To request feedback, use the "Request feedback" tab above. If I don't reply within a few working days, do send me a reminder directly to me (you should find my email somewhere online).

2. Date of consultation with support staff.

1901-01-01

1. Data description and collection or re-use of existing data

3. Provide a general description of the type of data you will be working with, including any re-used data:

This table is the most important part of the plan.

Ideally, by reading this, I can understand the flow of data during your project. The table is not perfect for this, so bear with it.

Quick notes:

- Type of data: consider all the data that is acquired, the steps they will go through, and what will be the output. You can be broad. Segregate data that may be confidential, private or sensitive into different types: Check the example below.
- File format(s) is particularly important if you are using very specific tools, or closed-formats. Open formats (.txt, csv,) are preferred but not mandatory.
- Source of data is important to check re-use rights on the material (license, terms of use,...) - for the time being, the source is important, you can figure out the licensing later - name of the source at least.
- Purpose of processing: the purpose of this step of your study - what are you trying to find out (mostly relevant for personal data) - if unsure, check the example below.
- Storage location - for your workspace: standard options are :
 - TUD OneDrive (with your TUD netID) convenient for most work
 - SurfDrive (TUD employee only, folders can be shared with students and external parties) - like dropbox, simple.
 - TUD MS Teams (mostly used for recording, but you can consider storing things in a MS Team "Team")
 - Project Storage : acquired through TopDesk, will appear as a shortcut in your "U:\:" drive - for anything you consider somewhat sensitive (and everything you consider very sensitive).
- Who will have access: in general, make sure at least 2 people have access to important material - anything else can be "on a need" basis. Make clear what you intend to make public during the project (e.g. "storage location: github", "access: everyone").

Simple interview example:

You are welcome to leave notes around the table to help understand things a bit, if you think it helps.

| Type of data | File format(s) | How will data be collected (for re-used data: source and terms of use)? | Purpose of processing | Storage location | Who will have access to the data |
|------------------------------|--------------------------|---|---|------------------------------|---|
| Interview recording | .AVI, .MP4, .MP3, .MKV | From participant, during interview | Understanding the barriers and drivers to the adoption of X and Y, in the domain of ABC | TUD OneDrive TUD MS Teams | Bob Bob (thesis supervisor). Alice Alice (MSc student) Eve Eve (Phd student in the group) |
| Interview transcript | .DOCX, .TXT, .RTF, .PDF, | Generated during interview | - | - | - |
| Interview anonymous summary* | same | derived from interview transcripts and videos | - | TUD OneDrive | - |

* Note for the reader: VERY *de-identified* summaries (as in: almost bullet-point list, only a participant number, a generic job description, and the domain of activities) - and as long as we have the consent forms, this is not *strictly* speaking (GDPR) anonymous. Only use that if you can rely on it for analysis - otherwise skip - all the data above is personal data, we can't publish it, keep it carefully - more below on this.

Below more examples.

| Type of data | File format(s) | How will data be collected (for re-used data: source and terms of use)? | Purpose of processing | Storage location | Who will have access to the data |
|------------------------------------|----------------|---|--|---|---|
| Survey questions | word | created | Investigate perception of people on the topic of public transport. | TUD OneDrive | Bob bob (PhD student) Alice Alice (Supervisor) |
| Survey answers | .csv | survey | - | TUD OneDrive Qualtrics (survey platform) | Bob bob (PhD student) Alice Alice (Supervisor) |
| Analysis script | .py | created | Statistical analysis of the result | TUD GitLab | Bob bob (PhD student) Alice Alice (Supervisor) Alice's team |
| Public transportation usage report | .pdf, .docx | Ministry of transportation of The Netherlands public website (No copyright) | Base line for transportation use in that area | TUD OneDrive | Bob bob (PhD student) Alice Alice (Supervisor) |

You are welcome to leave notes around the table. A good note here would be (as it indicate more "data processing"):

"We'll do 2 round of survey: 1) we propose a few options to people 2) after adjusting our options, we do a check to see what changed"

4. How much data storage will you require during the project lifetime?

- > 5 TB

If you need more than 250GB, you will have to use the "TUD Project Storage" for your data

If you need more than 5TB, you need assistance from our faculty information coordinator (? Ralph...)

II. Documentation and data quality

5. What documentation will accompany data?

- README file or other documentation explaining how data is organised
- Other - explain below
- Documentation in an Electronic Lab Notebook
- Data will be deposited in a data repository at the end of the project (see section V) and data discoverability and re-usability will be ensured by adhering to the repository's metadata standards
- I will adhere to disciplinary metadata standards - please explain which standards in the box below
- Data dictionary explaining the variables used
- Methodology of data collection

That's what will accompany the "supplementary material" of the publication. So, the answer depends on what you intend to make public at the end of the study.

Usually:

"1" for MSc thesis appendix, the methodology is in the thesis too... with the data.

"2" if any data is made available, regardless of how. (extra work: a small file to document the data, additional work is minimal).

"3" if you use tabular data, or models with structured types (extra work here: production of the "dictionary" of the data - description of each column of the spreadsheet, the variable in the model). Such dictionaries would be nice to have with the data, even if it's in

the publication as well. Up to you!

"4" use this if you intend to follow a specific format, or vocabulary for your data (that requires more than space than I have here - unless you are following a specific metadata standard, don't click)

"5" if you intend to publicly archive data at the end of the project.

6 relates to eLab Journals, which we do not have at TPM.

7 if you don't publish anything. Click that, and leave a note: "no public data"

III. Storage and backup during research process

6. Where will the data (and code, if applicable) be stored and backed-up during the project lifetime?

- Another storage system - please explain below, including provided security measures
- OneDrive
- SURFdrive
- Git(lab)/subversion repository at TU Delft
- Project Storage at TU Delft

Click the storage solutions you intend to use above, and quick recap' of the other storage you may need. Consider any external online tools you may need as well

Qualtrics, storage of partner company (internship provider or other), Internal group storage, MiroBoard, MentiMeter, ...

IV. Legal and ethical requirements, codes of conduct

7. Does your research involve human subjects or 3rd party datasets collected from human participants?

- Yes

8A. Will you work with personal data? (information about an identified or identifiable natural person)

If you are not sure which option to select, first ask your [Faculty Data Steward](#) for advice. You can also check with the [privacy website](#) . If you would like to contact the privacy team: privacy-tud@tudelft.nl, please bring your DMP.

- Yes

*Here understand the question as: **will you handle personal data for the project, at any time** (email addresses, for participant management and recording? Signature on consent forms? pictures of people from the web ?....)*

Example answers (if it's obvious, you can be short) - more questions to come:

- interview
- survey
- pictures/videos of people collected from [source]
- Input from partner: household-level hourly energy consumption, with location
- census data from [somewhere]

8B. Will you work with any other types of confidential or classified data or code as listed below? (tick all that apply)

If you are not sure which option to select, ask your [Faculty Data Steward](#) for advice.

- No, I will not work with any confidential or classified data/code
- Yes, I work with other types of confidential or classified data (or code) - please explain below

- Yes, politically-sensitive data (e.g. research commissioned by public authorities, research in social issues)
- Yes, data which could lead to reputation/brand damage (e.g. animal research, climate change, personal data)
- Yes, data related to competitive advantage (e.g. patent, IP)
- Yes, confidential data received from commercial, or other external partners
- Yes, data falling under export control regulations
- Yes, national security data (e.g. nuclear research)

Usually at TPM: "8 - No", or "5 - reputation and brand damage", and perhaps "6 - politically-sensitive issues".

If the risks you see are inherent to your topic, I recommend you chose for a "TUD Project Storage" for your data. If you are very concerned by the security of your data, contact our Faculty Information Security Coordinator (Ammar).

If the risks are purely accidental (which is most often the case), use any TUD storage - just be careful.

9. How will ownership of the data and intellectual property rights to the data be managed?

For projects involving commercially-sensitive research or research involving third parties, seek advice of your [Faculty Contract Manager](#) when answering this question. If this is not the case, you can use the example below.

Indicate here what contract frame the activity described here - that's what defines the IP rights for the activity. A couple of standard options are provided below:

- "internal TUD MSc project" (for students doing their MSc here)
 - Warning: the exact status of the IP of MSc student is not clear. if you, student, are contributing to a larger research project, we would like it if you agreed to transfer your data to use at the end! Or make it public. But it's your choice.
- "MSc graduation agreement" (for MSc students doing an internship - the agreement covers general IP clauses - it's a standard template)
- "Consortium agreement" (indicate funder if you did not put it on the front page)
- "Collaboration agreement"
- "Internal TUD project" - when it's just us!

That will help a lot in case personal data is transferred between partners. If we do not have a legal agreement with the partner with which we want to share data, we will have to "create one", so that the agreement for personal data can be valid (GDPR agreements are addendum to existing contracts - at TUD).

10. Which personal data will you process? Tick all that apply

- Special categories of personal data (specify which): race, ethnicity, criminal offence data, political beliefs, union membership, religion, sex life, health data, biometric or genetic data
- Email addresses and/or other addresses for digital communication
- Copies of passports or other identity documents
- Citizen Service Number (BSN)
- Financial information, such as bank account numbers
- Access or identification details, such as personnel number, student number
- IP addresses
- Telephone numbers
- Other types of personal data - please explain below
- Data collected in Informed Consent form (names and email addresses)
- Signed consent forms
- Photographs, video materials, performance appraisals or student results
- Gender, date of birth and/or age
- Names and addresses

Tic what you need to do the full experiment - names, email addresses are common for interviews, gender and age also but only if relevant for the study.

In general: only collect what you need for the study - data minimization.

Special categories of personal data: it's a GDPR concept - anything that can be used to discriminate people, and medical data. That requires special attention (TUD Project Storage)

Usual answers here include:

[] tick of relevant boxes above and

- Job description, domain of activity
- age (*please indicate*)
 - age (in ranges)
 - age (year of birth)
 - age (date of birth)
- Area of living (region / city / zip code with 4 digit / ...)
- GPS position through out the day (tracking ? monitoring ?)

Some metadata standard can be useful. See the example below.

Sex/Gender : we will rely on meta-data standard proposed by ZonMW for sex and gender

(<https://bioportal.bioontology.org/ontologies/ZONMW-CONTENT?p=classes&conceptid=http://purl.org/zonmw/covid19/10061>)

11. Please list the categories of data subjects

Indicate here who you will interact directly or indirectly. Indicate the country of where they are located if possible. Common answers are similar to:

Experts in the field of energy production, including members of start-ups, private company, and local public representative (municipality) in the Netherlands

Survey respondents recruited through personal and professional network (LinkedIn)

Members of NGO active in the field of food safety located anywhere in the world.

Employees of KLM operating in San Martin, working on emergency systems

Participants in the event/workshops/lecture ...

Bike users of Delft - interviews, field work

Note for "The citizen of city X", or "General inhabitant for neigh". Please provide info on recruitment of the participants here.

Example:

Citizen of the neighborhood X, contacted through the local pet owner group.

Citizen of Amsterdam, survey distributed by city hall to specific bike users

12. Will you be sharing personal data with individuals/organisations outside of the EEA (European Economic Area)?

- Yes

13. To which countries will you be transferring personal data:

- Other (please indicate in the box below)

This is usually a "NO" (in the DMP). Please indicate in the list in which country we will be getting or sending information.

If personal data is transferred to or received from a partner located in a country that is not part of the EEU, and not on the "adequacy list" of the European Commission, more legal work is required.

14. Please contact the privacy team (privacy-tud@tudelft.nl) for advice on data transfer. Please also bring your draft DMP when contacting the privacy team.

Please record below their advice, the data transfer mechanism used and agreed security measures:

Update this field if the privacy team has provided advice on this.

15. What is the legal ground for personal data processing?

- Informed consent
- Other - please explain and contact the privacy team (privacy-tud@tudelft.nl). If you have already contacted the privacy team and received their advice, please record their advice below

"Other" - we can't get direct consent from people, we will need to justify "why we have the data" - GDPR.

And that depends on the project - the privacy team is in charge of determining what "legal ground" we can best use for your experiment.

Usual tense situations:

- *personal data collected from social media*
- *medical/patient data: we can't get consent from the patient themselves.*

contact privacy team once your data processing is clearly defined - minimize as much as possible the personal data you collect and what you store/publish.

16. Please describe the informed consent procedure you will follow:

All study participants will be asked for their written consent for taking part in the study and for data processing before the start of the [interview | workshop | survey]

Verbal consent is reasonable as well - make sure you ask the permission for everything (including the right to quote! and the publication of the data/result!) - written consent is better.

Digital and paper versions are acceptable.

17. Where will you store the signed consent forms?

- Same storage solutions as explained in question 6

Make sure you know where the proof of consent are, we need them for as long as we have the personal data (names, emails, recordings, transcripts, signature)

18. Does the processing of the personal data result in a high risk to the data subjects?

If the processing of the personal data results in a high risk to the data subjects, it is required to perform [Data Protection Impact Assessment \(DPIA\)](#). In order to determine if there is a high risk for the data subjects, please check if any of the options below that are applicable to the processing of the personal data during your research (check all that apply).

If two or more of the options listed below apply, you will have to [complete the DPIA](#). Please get in touch with the privacy team: privacy-tud@tudelft.nl to receive support with DPIA.

If only one of the options listed below applies, your project might need a DPIA. Please get in touch with the privacy team: privacy-tud@tudelft.nl to get advice as to whether DPIA is necessary.

If you have any additional comments, please add them in the box below.

- None of the above applies

Those risks are related to people -

"Evaluation and scoring" - of people, or their performance at specific tasks.

Usual answers are:

None of the above applies

Sensitive personal data (medical data)

Data concerning vulnerable data subjects (children)

Data concerning vulnerable data subjects (patient of a specific service at a hospital)

19. Did the privacy team advise you to perform a DPIA?

- No

Update this field with the advice of the privacy team.

22. What will happen with personal research data after the end of the research project?

- Personal research data will be destroyed after the end of the research project
- Anonymised or aggregated data will be shared with others

Just tick the right boxes here. Usually:

"Anonymised or aggregated data will be shared with others": *result section in publication, quotes,*
"Personal research data will be destroyed": *we get to chose "when", but participants must be informed - to be dealt with in the associated "informed consent form"*

23. How long will (pseudonymised) personal data be stored for?

- Other - please state the duration and explain the rationale below

Usual answers (the project can be a MSc thesis, or a 5 year consortium):

- the personal data will be destroyed at the latest 1 month after the completion of this project
- the personal data will be preserved for up to 2 years after the completion of this project.

Note: you get to pick the duration for which the data is preserved, but you, then, also need to manage it for that duration.

24. What is the purpose of sharing personal data?

- Other - please explain below

Usual answer:

We do not intend to share personal data.

25. Will your study participants be asked for their consent for data sharing?

- Yes, in consent form - please explain below what you will do with data from participants who did not consent to data sharing

Note that you intend to publish at least the result of the study, so they are consenting to "some" data sharing (to be dealt with in the consent form).

Usual answers:

- People who do not consent will not be interviewed
- People who do not consent will not take part in the survey
- Participants will be asked if the summary of the interview can be published. If they refused, the summary will be considered as personal data and will not be shared.

V. Data sharing and long-term preservation

27. Apart from personal data mentioned in question 22, will any other data be publicly shared?

- All other non-personal data (and code) underlying published articles / reports / theses

List here the type of material that can be made publicly available at the end of the study

Usual answers:

- Interview questions/protocol, informed consent form template, (include anonymous summaries if possible/useful)
- survey questions, anonymized survey answers (no free text, limited demographic info), analysis script, data tables underlying graphs.
- software package, with input and output data of the study
- software package, with test/toy examples (validation/replication)

29. How will you share research data (and code), including the one mentioned in question 22?

- All anonymised or aggregated data, and/or all other non-personal data will be uploaded to 4TU.ResearchData with public access

Zenodo and 4TU Research Data are good repositories.

More importantly, consider here how you will preserve the personal / confidential data at the end of the project - where is it going to be stored for 2 years, 5 ...

Usual answer:

For MSc thesis:

"My data will be shared in a different way"

the data will be in the appendices of my MSc thesis (if it fit in a PDF...)

Otherwise:

.Option 1 : "4TU.ResearchData "

30. How much of your data will be shared in a research data repository?

- < 100 GB

31. When will the data (or code) be shared?

- As soon as corresponding results (papers, theses, reports) are published

You can always delay the publication - embargo, confidentiality ...

32. Under what licence will be the data/code released?

- Other - Please explain
- AGPL-3.0
- LGPL-3.0
- GPL 3.0+
- GPL-2.0
- EUPL-1.2
- Apache
- BSD
- MIT License
- CC BY-NC-ND
- CC BY-NC-SA
- CC BY-NC
- CC BY-ND
- CC BY-SA
- CC BY
- CC0

Usually:

CC-BY (for research data)

MIT or BSD (for software, scripts)

1 license per dataset ! You can mix software and data content in a single dataset, but pick the license that matches "the most important of the two". If you can't decide, make 2 different dataset with different license when you upload your data.

In very short for data :

CC0: public domain equivalent

CC-BY: public domain, but invite proper attribution of the work (citation)

CC-NC: public domain, no commercial usage

CC-SA: public domain, remix/modifications must be public

CC-ND: public domain, no remix/modification allowed

For MSc thesis, if the data in the appendices of the thesis:

- Tick "other" and indicate "Dutch copyright laws."

VI. Data management responsibilities and resources

33. Is TU Delft the lead institution for this project?

- Yes, leading the collaboration - please provide details of the type of collaboration and the involved parties below

*For all PhD, MSc project, done in cooperation with other institutions, TUD is "leading the collaboration" - for your activity anyway!
For DMPs for specific tasks in project WP, we are most of the time leading those experiment as well.*

Usual answers are:

- *for MSc student:* internship at NS
- This project is done in cooperation with TNO
- This project is part of the Scale-Up project, funded by the ERC, with 24 partners.

34. If you leave TU Delft (or are unavailable), who is going to be responsible for the data resulting from this project?

The name of your back up in case of accident. Usual answers are:

- Master thesis supervisor: Bob Bob (email)
- PhD supervisor: Alice Alice (email)

35. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

In case you need access to specific resources you can use this question here - possible answers are:

- None
- Delft Blue Super computer access
- Hosting platform for project website/tool
- Support from the "Digital Competence Center of Delft"
- Software/system support (system administrator, developer,...)