ESDT (zwaartekracht, 2019)

A Data Management Plan created using DMPonline

- Creator: Sander de Boer
- Affiliation: University of Twente
- Funder: Netherlands Organisation for Scientific Research (NWO)
- Template: Data Management Plan NWO (January 2020)
- ORCID iD: 0000-0002-4789-4588

Project abstract

Opkomende sociaal disruptieve technologieën (SDT's), zoals robots, kunstmatige intelligentie, synthetische biologie en klimaattechnologie, leveren de nodige sociale en ethische uitdagingen op. In dit filosofische onderzoeksprogramma worden nieuwe manieren ontwikkeld om daarmee om te gaan. De samenleving, cultuur en het dagelijks leven gaan veranderen door SDT's, en enkele van onze meest elementaire concepten en waarden komen erdoor op losse schroeven te staan. Wij gaan de technologie-ethiek innoveren om de ontwikkeling en invoering van deze technologieën kritisch te evalueren en richting te geven.

General Information

Name applicant and project number

Project leader: Philip Brey
Project number: 024.004.031

Name of data management support staff consulted during the preparation of this plan and date of consultation: Peter Noort, p.d.noort@utwente.nl (consulted while writing the original project proposal and again in April 2020 for the update of the DMP).

1. What data will be collected or produced, and what existing data will be re-used?

1.1 Will you re-use existing data for this research?

If yes: explain which existing data you will re-use and under which terms of use.

Yes

We will re-use existing data where possible, both data generated within the project itself and pre-existing research data generated outside the project. Since most of the subprojects have not yet been defined, it is impossible in this stage to provide a list of datasets that we will re-use.

1.2 If new data will be produced: describe the data you expect your research will generate and the format and volumes to be collected or produced.

The research programme will be carried out over a period of 10 years. We expect to to initiate about 30 subprojects carried out by a PhD and about 18 more subprojects carried out by a postdoc, in addition to the research projects carried out by the PI's. These subprojects have yet to be defined, which makes it impossible

to provide a list of datasets that we will create at this stage. But what we can already say is this:

Volume:
We expect the size of the data collected in the entire project to be relatively small. It will likely include (expert) interview data, field observations (e.g., in living labs), survey data, and data from targeted experiments involving human participants.

Format:
Data will be stored in accordance with the preferred formats as defined by the 4TU.
See https://researchdata.4tu.nl/fileadmin/user_upload/Documenten/preffered_file_formats.pdf

1.3. How much data storage will your project require in total?

10 – 100 GB

This is a first estimate. We estimate the low volume, because we expect most of the subprojects that will generate research data to involve either (expert) interview data, field observations (e.g., in living labs), survey data, or data from targeted experiments involving human volunteers. None of these usually take up substantial storage space.

Depending on the subprojects that may be initiated in the coming years, this number may be adjusted upwards.

2. What metadata and documentation will accompany the data?

2.1 Indicate what documentation will accompany the data.

For each subproject that will be initiated, we will draft a project-specific extension to this general DMP. That extension will list the datasets that will be created and/or re-used, and will include a specification of the accompanying metadata and documentation.

2.2 Indicate which metadata will be provided to help others identify and discover the data.

See under 2.1.

3. How will data and metadata be stored and backed up during the research?

3.1 Describe where the data and metadata will be stored and backed up during the project.

Institution networked research storage

During research, all digital data and metadata will be stored on the departmental servers of the host institutions which are backed-up automatically and daily, in line with the existing data management policies of the host institution.
For ongoing research projects, data may be shared with participating researchers through Dataverse (DANS). Data exchange between partners may also take place through another secure EU-based service such as Surfdrive. US based services such as Dropbox will not be allowed.

3.2 How will data security and protection of sensitive data be taken care of during the research?

Default security measures of the institution networked research storage.

All data storage practices will follow the host institutions' data policies as this will ensure the most secure storage possible.

4. How will you handle issues regarding the processing of personal information and intellectual property rights and ownership?

4.1 Will you process and/or store personal data during your project?

If yes, how will compliance with legislation and (institutional) regulation on personal data be ensured?

Yes

Some of the subprojects will likely collect and process personal information. In those subprojects where personal information is collected, data collection, storage, and sharing will proceed according to the conditions specified in the informed consent forms supplied in advance to participants. A subproject proposal, including these informed consents forms, will be submitted to the Ethics Committee of the PI's institution for approval, prior to conducting the research project. The programme manager of the Gravitation Consortium will ensure that this procedure is followed.
When constraints put forward by the ethics committee that reviewed the proposal for the subproject conflict with the intentions put forward in this datamanagement protocol or with the supplement to it written for the subproject in question, the constraints put forward by the ethics committee will take precedence. In those cases, we will update the DMP to reflect these additional constraints.

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

In this stage, we assume that managing rights to the data will follow from normal university practice (where the rights are shared between the PI and his/her host institution). Should the consortium decide to deviate in part from this practice, any such deviations will be recorded in the consortium agreement that is being drafted and this DMP will then be updated accordingly.
See also section 5.5.

5. How and when will data be shared and preserved for the long term?

5.1 How will data be selected for long-term preservation?

All data resulting from the project will be preserved for at least 10 years.

5.2 Are there any (legal, IP, privacy related, security related) reasons to restrict access to the data once made publicly available, to limit which data will be made publicly available, or to not make part of the data publicly available?

If yes, please explain.

Yes

Impossible to provide further details in this stage as it depends on the subprojects that have yet to be formulated.
We will in all cases strive to minimize any legal restrictions to sharing data in the case of collaborative projects (e.g. NDAs demanded by an external partner).
When we collect personal data, our first priority will be to safeguard the rights of participants under the GDPR, which also involves following any guidelines set by the Ethics committee that reviewed the subproposal. But given that, we will take all reasonable actions to minimize the restrictions, including

trying to obtain consent for re-use in specific cases and (statistical) data obfuscation.

5.3 What data will be made available for re-use?

Other (please specify)

Relevant data may include (expert) interview data, field observations (e.g., in living labs), survey data, and data from targeted experiments involving human volunteers.
Unless there are valid reasons which make research data unsuitable for sharing (in particular GDPR constraints, or legal constraints in the case of collaborations with industry partners), all data, underlying code and any other materials needed to replicate research findings will be appropriately documented, stored and shared in a research data repository for at least 10 years from the date of publication, in accordance with the FAIR (Findable, Accessible, Interoperable and Reusable) principles.
We recognize that there can be an exceptional case where publicly sharing all underlying research data immediately upon publication of the research findings is unreasonably onerous for the researcher of the subproject in question. In that case, the researcher may submit a written reasoned request to the Management board of the gravitation consortium to deviate from the above. The decision of the Management board is final.
When the data itself cannot be made publicly available, we will ensure that at least the relevant metadata is published in a suitable repository. Where appropriate, we will adhere to international standards of sharing meta-data (e.g., CERIF or Common European Research Information Format), which will improve the reusability of the project's archived data.

5.4 When will the data be available for re-use, and for how long will the data be available?

Data available upon completion of the project.

5.5 In which repository will the data be archived and made available for re-use, and under which license?

After completion of a subproject, data will be registered and archived, preferably in one the following two repositories:

- EASY (DANS)
- 4TU.Centre for Research Data

These archives assign a persistent identifier (DOI) to the datasets.
For all datasets where there are neither GDPR constraints nor NDA's, the data will in principle be made freely available under either the creative commons license CC0, which is the default option of 4TU.Researchdata, or the CC-BY 4.0 license.
For all datasets where GDPR constraints exists, data will only be made available to the extent that the GDPR and the informed consents forms signed by participants permits.

5.6 Describe your strategy for publishing the analysis software that will be generated in this project.

All code, scripts and algorithms that are required to interpret en re-use the data will be made available in addition to the data and metadata itself.

6. Data management costs

6.1 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)?

We expect that the size of the data generated in this project will be relatively small and therefore no additional budget is required for data storage/archiving.