Notes 2. Data management

Responsible data management is part of good research. For the collection/generation of data and the analysis of this data, timely measures need to be taken to ensure the storage and later reuse of the data. This means that prior to the start of the research project researchers must ascertain **a)** which data could be relevant and **b)** how these data could be stored so that they are accessible for reuse. After a proposal has been awarded funding, the researcher will draw up a detailed data management plan in which the researcher explains how all relevant data research data will be made findable, accessible, interoperable and reusable (FAIR).

NWO and ZonMw want to contribute to the development of good data management by asking researchers to make all relevant research data findable, accessible, interoperable and reusable (FAIR). Therefore, in the data management section, researchers will be asked before their research starts, to think about how the data collected should be ordered and categorised such that it can be made freely available. Researchers will often need to take measures to this effect during the production and analysis of the data.

NWO understands ‘data’ to include collected, unprocessed data as well as analysed, generated data. This includes all conceivable forms of digital and non-digital data (such as samples, completed questionnaires, sound recordings, etc.).

*NWO only requires the storage of data that are relevant for reuse*. NWO assumes that within disciplines there are widely held opinions about which data are relevant for storage and reuse. Research Data Netherlands offers a [checklist](https://researchdata.nl/) for the selection of data that can be eligible for archiving.

Research results should be stored in such a way that they can be retrieved and reused in the long term, also by researchers in disciplines and organisations other than those in which the research took place. The operating principle is that all stored data are, in principle, freely accessible and that access is only limited if required for reasons such as privacy, public security, ethical limitations, property rights and commercial interests.

The costs of data management are eligible for funding and should be included in the project budget. Important factors that determine the costs are:

- the type of data;
- the capacity needed for storage and backup;
- the amount of manual work needed to allocate metadata and the compilation of other documentation such as codebooks and the queries used in the statistical package;
- the extent to which the data needs to be protected;
- the hiring in of external data management expertise or other expertise.

With the data management section NWO mainly wants to raise awareness about the importance of responsible data management. The section is therefore not included in a committee's decision about whether or not a proposal should be awarded funding. NWO does, however, submit this section to the committee and referees for advice. After a proposal has been awarded funding, the researcher should elaborate the section into a data management plan.

2.1. Will data be collected or generated that are suitable for reuse?

[x]  **Yes** *(Please answer questions 2.2 to 2.4)*

[ ]  **No** *(Please explain below why the research will not result in reusable data or in data that cannot be stored or data that for other reasons are not relevant for reuse.)*

***Motivation:***

**Please explain the types of data/code you expect to generate and how these are relevant to your (or the wider) research community.**

2.2. Where will the data be stored during the research?

During the course of the research project, all data will be stored on local servers (project drive, SURFdrive) maintained and automatically backed up by TU Delft ICT. Lab documentation will be collected using paper/electronic lab notebooks.

Software and/or analyses scripts will be stored on GitLab (or GitHub), hosted and maintained by TU Delft ICT.

2.3. After the project has been completed, how will the data be stored for the long-term and how will the data be made available for use by third parties? For whom will the data be accessible?

All datasets will be published at [4TU.ResearchData](https://data.4tu.nl/info/), which is a trusted and certified research data repository (Data Seal of Approval certification). All datasets will be accompanied by metadata (compliant with DataCite, schema.org and Dublin Core) to ensure that all datasets are findable and accessible online. All datasets will be publicly available to anyone for re-use under an open licence (CC-BY). Every dataset will be also assigned a Digital Object Identifier (DOI), to make them citable and persistently available. 4TU.ResearchData preserves the data-sets for at least 15 years, maintaining their integrity and authenticity, in accordance with 4TU.Centre for Research Data preservation policy.

2.4. Which facilities (ICT1, (secure) archive, refrigerators or legal expertise) do you expect will be needed for the storage of data during and after the research? Are these facilities available?1

The project will not require use of any additional facilities on top of those discussed above. All the facilities discussed above are readily available to use. 4TU.ResearchData is free for TU Delft researchers up to 1 TB/per year/per researcher and we do not expect to exceed this.

 ICT facilities for data storage are considered to be facilities such as data storage capacity, bandwidth for data transport and calculating power for data processing.