

UT Research data policy

July 2015

Introduction

The careful handling of research data within the UT deserves more attention for the following reasons:

1. the importance of the quality, reliability, replicability and verification of scientific research
2. the possibility to improve the efficiency of research data availability thanks to digitalization and internationalization
3. the importance of characterization and visibility of the UT research, e.g. by linking data to the corresponding publications, and publishing citable data sets
4. the funders' requirements for managing research data
5. the fact that handling research data is part of research evaluations done by the Standard Evaluation Protocol (SEP)
6. the growing need for support in data management because of the growing amounts of data.

The way in which research data is being handled differs per academic discipline. A research data policy needs to take these differences into account. Therefore the responsibility for the research data policy primarily lies with the faculties, research groups and/or research institutes.

The goal of the research data policy is

1. providing clear criteria and minimum requirements for proper research data management which do justice to the differences between disciplines
2. stimulating the durable storage of research data and making it as widely accessible as possible both during and after the research
3. offering a framework for the continuous improvement of the infrastructure for and management of research data
4. stimulating the reuse of data, which corresponds to the goals of research funders.

The UT research data policy is about dealing with the data that is being collected for scientific research at the UT either directly (primary data) or indirectly (secondary data), and that UT researchers have authority over. The data policy also includes the description of experimental artefacts, measuring instruments and sample survey procedures used to collect and/or analyse the research data, as this information is often important when duplicating the research or reusing the data. This policy document summarizes the collected data including the corresponding descriptions and used set of instruments in the single term 'research data'.

Part of the research conducted at the UT is subject to previously fixed legal requirements, contractual conditions from third parties, etc. Regarding the research data policy, it is sufficient in those cases to refer to, for example:

- guidelines/ policies/ protocols concerning the management of research data by a programme, project or external party and a corresponding, adequate research data infrastructure
- the Netherlands Code of Conduct for Scientific Practice, which states that raw research data should be stored for at least 10 years
- rules set by funders, companies or other external owners of research data
- a Medical Ethical Examination Commission (METC in Dutch)
- legal requirements, e.g. concerning privacy

Responsibilities

The shared responsibility for managing research data lies with the researcher and the organization(s) affiliated with the research. Proper management of research data is not only a matter for the researcher(s) themselves, but also for the research group, the research institute, the faculty and the university. This is not just because of their reputation, but also because of the quality (verification) and efficiency (reuse) of the research. The starting point for the research data policy to be formulated is therefore that it should be in keeping with the existing responsibilities regarding research within the UT.

The primary responsibility regarding the policy and the management of research data is as follows.

Policy: faculties/ research institutes / research groups (mutual agreement)

- Every researcher is responsible for the way in which they deal with research data, in some cases together with a lead researcher (PI); the professor or chair of a research group has the final responsibility.
- The faculty or research institute supports the creation of a data policy for every research group, and monitors the use of this policy (for elements of the research data policy, please see below).
- The Executive Board sees to it that every faculty or research institute has a research data policy and that it is complied with. Faculties and institutes do not only include the minimum requirements in their policy, but can also add measurements aimed at characterization.

Awareness and knowledge: research institutes / research groups and study programmes

- The research institutes and groups are responsible for the sensible use of research data which at least conforms to the agreements and rules within the discipline and the institute's or group's formulated research data policy, and see to it that, if necessary, researchers will take additional relevant courses or training.
- The faculties (study programmes) are responsible for offering education in the field of managing research data in the academic discipline concerned. Usually this will be embedded in the teaching of research methodology.

L&A and ICTS will take care of:

- the offering of general courses or modules in the field of managing research data, both introductory and in addition to discipline specific courses and education (in consultation with study programs and TGS),
- the offering and technical management of an infrastructure for storing research data,
- the offering and, if needed, functional management of facilities for storing, documenting, making available and archiving of research data (archiving for example at 3TU.Datacentrum or DANS),
- the support of researchers, research groups, research institutes, faculties and the Executive Board in the field of research data management and policy.

Research data policy

The primary goal of the research data policy is offering clear criteria for the proper management of research data. Additionally, it shows which facilities concerning awareness, knowledge, infrastructure and support could be necessary to achieve proper research data management.

Criteria for proper research data management

The research group, faculty and/or research institute has a research data policy which at least includes

1. who is responsible for executing the research data policy of the group/ faculty/ research institute,
2. the fact that research data are stored in a durable way, and in accordance with legal regulations, any third party contractual requirements, etc.
3. where data will be stored both short-term and long-term (see #2 below), and how the data is described,
4. the criteria for the selection of the research data that will be durably stored,
5. the conditions under which research data are made available to others (e.g. anonymization).

Linked to the stages of a research, the following applies:

1. Prior to the research the way in which research data are handled both during and after the research has to be established. This can be done in a **Data Management Plan (DMP)** or a different type of document which is more common in the discipline and/ or the research partners concerned. Each PhD student sets up a DMP, which will be handed in as an appendix to their research proposal.
2. During the research the research data **will be saved in a central repository which is available to at least the members of the research group/ institute** and which is managed by this research group/ institute. Storage and access should be managed in accordance with legal regulations, any third party contractual requirements, etc.
3. Preferably during the research, but not later than 1 month after finishing the research, the research data are **archived in a trusted repository** (e.g. DANS or 3TU.Datacentrum). The research data are, taking legal regulations, any third party contractual conditions into account, preferably

publicly available. This covers at least the research data that form the basis of publications about the research, but can also comprise the full set of raw and/or edited research data.

4. After the research all durably stored research data and **the publications based on those data are linked**. This is at least the case for PhD dissertations.

Awareness and knowledge

The sensible use of research data and knowledge of research data management and policies is something that should be addressed explicitly in the methodology courses of the UT programmes. This includes the training of PhD students. This can be done through courses at TGS or other graduate schools.

Infrastructure

In dealing with research data a distinction can be made between the management of the data during the research and after. Often the available infrastructure is separated along this line; whereas during the research one usually speaks of storing the data, after the research one speaks of archiving the data.

When no (inter)national subject specific facilities or national services for the storage and sharing of research data during the research are available or needed, researchers can use UT facilities with backup and restore options (see data storage policy). ICTS offers a range of services for storing, backup and restoring your data, both within their standard service package, as well as custom solutions.

To store, describe (including metadata) and share research data, such as by IGS-datalab, researchers can use the Dataverse Network (DataverseNL) facility offered by the UT. This service allows you to store data and share it with specific people. DataverseNL can also be used to publish data, as a persistent identifier is assigned.

In order to be able to replicate research and for possible reuse of the data, it is important to archive this in a durable way after the research has ended. The designated archive for the technical and natural sciences is the 3TU.Datacentrum, for the social sciences this is DANS. It is also possible to use other discipline-specific data archives, as long as they have a Data Seal of Approval like the two aforementioned archives.

These aforementioned archives offer the option to link the research data to the publications based on them, so as to improve both the visibility and quality of the research.

Support

While conducting research, support on how to handle research data is crucial. L&A offers the following:

- Information on conditions set by research funders, legal requirements, etc. concerning the handling of research data.
- Support in setting up a Data Management Plan.
- Advice on how to store, describe and share research data. If necessary support in writing the metadata for research data is offered.
- Advice and support in the durable archiving of research data in DANS and 3TU.Datacentrum
- Advice and support in the publishing and making citable of research data, if necessary in connection with articles, dissertations, etc.
- General courses on research data management, aimed both at awareness and knowledge.

Costs and funding

Managing research data costs money, mainly for the storing of data during research and the archiving of data afterwards. The amount of costs depends, among other things, on the amount of data involved and the level of protection against abuse and data loss. Costs for data management can be reduced by being very selective in deciding which data should be saved in order to meet the requirements concerning the quality, reliability, replicability and verification of the research, as well as the desire to be able to reuse data. In order to save expenses it is also important to define the level of data protection depending on the requirements per type of data.

The costs for data management should be explicitly listed in the research project budget and data management plan.

In principle all expenses for data management, both during and after the research, should be funded by the research project budget. When costs for data management are made because of requirements set by the institution, such as archiving research data used for a publication, a researcher can apply for a grant from a centrally established incentive fund, helping with the characterization and visibility of UT research. Other (yet to be determined) conditions, such as other grants not being available, may apply.

Appendix: Approach

This policy document has been established after a series of interviews about data policy and data management with the following persons:

Prof. dr. P.M.G. Apers
Prof. dr.ir. A. van den Berg
Prof. dr.ir. E.C. van Berkum
Prof. dr.ing. D.H.A. Blank
Prof. dr.ir. A. de Boer
Prof. dr. S.A.H. Denters
Prof. dr. P.H. Hartel
Prof. dr.ir. H.J. Hermens
Prof. dr. G.J. Hospers
Prof. dr. F.M.G. de Jong
Prof. dr.ir. H.F.J.M. Koopman
Prof. dr. D. Lohse
Prof. dr. D. van der Meer
Prof. dr.ir. B. Nauta
Prof. dr. A. Need
Prof. dr.ir. A. Stein
Prof. dr.ir. A. Veldkamp
Prof. dr. M. Versluis
Dr.ir. P.W. de Vries
Prof. dr. R.J. Wieringa

These persons have been chosen based on the diversity and completeness of the scientific disciplines. Some persons have also been interviewed because of their position as scientific director or NWO board member.

In addition, data management policies from other universities and the data storage policy from UIM have been used.