Research Data Management Regulations
Leiden University

Focal points and explanations of the various articles appear in italic.

A. Definitions and basis

1. **These regulations are intended as a framework for a University research data management policy and require further elaboration at disciplinary level. Each Faculty Board must decide how the regulations will be further elaborated upon within the faculty and which role the institutes will play.**

2. **Research data is understood to mean the following: all data that is gathered and generated during academic research.**
   This also includes the data that is acquired by processing and analysing (raw) research data. We follow NWO here, which states, 'we understand "data" to mean both gathered, unprocessed data and analysed, generated data.' This does not mean that all data must also be preserved. The decision concerning which data to preserve is recorded at the disciplinary level (see article 16).

3. **These regulations apply to all digital and non-digital research data.**
   We follow NWO policy here. Non-digital data can mean, for instance, preserved bio-material, audio recordings, paper questionnaires and paper lab journals. NB: this stipulation does not mean that all non-digital data must be digitised per se, but it does mean that the same requirements apply to non-digital data concerning findability, reusability, long-term preservation, etc.

Before the research

4. **A data management plan (DMP) must be drawn up before data collection for a research project begins. The DMP elaborates upon the data management policy of the faculty/institute for the specific research project in question.**

During the research

5. **During the research, research data must be securely preserved. This means that the integrity, availability and – if required – confidentiality of the data must be guaranteed.**
   Please note: once the research has been completed, the research data must be preserved for the long term together with the metadata, software and other documentation required for reuse (Article 7). This must also be borne in mind during the research, for instance because it will generally be impossible to provide satisfactory metadata for all the data once the research has been completed.

After the research

6. **Research data must be managed in such a way that, at the latest,**
   - at the point in time of the last publication arising from the research, or
   - at the point in time of the formal completion of the research project
   it is preserved so that it is at least findable, accessible, comprehensible and reusable in the long term.
   Two instances are mentioned here because not all research culminates in a publication. However, it is then also important that the research data is preserved and is findable. Externally funded research projects will generally formally conclude upon their financial-administrative completion. An explanation of the criteria that are used:
   - *Findable*
     The data must be findable for other researchers and involved parties. Its findability improves if it is deposited in a (discipline-specific) data archive or repository. The information about the data (metadata) is then registered in a standardised manner and is findable, also for search engines. Future findability is guaranteed by assigning a persistent identifier to the data. A commonly used identifier for data is DOI (digital object identifier). A DOI is assigned by a data archive or data publisher and is a unique number for a digital object, in this case a dataset. The DOI remains the same even if the location of the dataset (URL) changes. A DOI or other persistent identifier is used in citations or references to the dataset.
   - *Accessible:*
     Data accessibility does not necessarily mean that the data must immediately be made fully open. The data will have to be made accessible in some instances, for example for verification by a grant provider or journal, but will remain inaccessible to the wider public. With sensitive data, full publication will never be an option. Research data can be retained under embargo; then only those who have deposited the data have access to it. The duration of the embargo is determined in consultation with the data archive. A further option is only to grant access if a request for access is submitted to the researcher. The researcher then knows who is consulting the data and can reach agreements about its use and reuse. NB: if the decision is made to limit access to the data, it is essential that the metadata (description of the data) is findable.
   - *Comprehensible*
     Metadata and any supplementary documentation must describe the data in such a way that other researchers will also be able to understand and use it.
   - *Reusable*
     This criterion is further elaborated upon in Article 7.
   - *Long-term retention*
     This criterion is further elaborated upon in Article 8.

7. **Research data must be stored together with the metadata, other documentation and possibly the software and version of the software required for its potential reuse.**
   Research data must be stored in such a way that it is independent of the underlying equipment/hardware, such as microscopes, scanners or recording equipment. Long-term data formats that are supported by data archives should be used if possible. Retention of hardware can be considered in certain cases, for instance in the case of software that is only compatible with an obsolete computer operating system. Close attention must be paid to the cost/benefit ratio here.

8. **Once the research has been completed, the research data must be retained securely for the long term. This means that the integrity, availability and – if required – confidentiality of**

the data must be guaranteed. The data should be retained according to international guidelines. For digital data this can be achieved by preserving it in a *Trusted Digital Repository* (an online depot). Non-digital research data must be preserved according to the prevailing standards in the discipline.
A Trusted Digital Repository possesses equipment, hardware and software for the long term and reliable retention of digital information. This facility must be embedded in a reliable organisation that has a long-term mission and funding.
There are three prevailing international standards for Trusted Digital Repositories and the manner in which they retain digital data for the long term. In ascending order of reliability requirements these are: the Data Seal of Approval[1], the Nestor Seal[2] (also known as DIN 31644) and the RAC standards[3] (also known as ISO 16363). Data preservation at a repository with a Data Seal of Approval (the lowest category) currently suffices for NWO and Horizon 2020.
It is recommended to digitise non-digital data as far as possible for long-term retention.

9. **The minimum retention term for research data is ten years.**
Here we follow The Netherlands Code of Conduct for Scientific Practice of the VSNU (2014 version).

B. Responsibilities

10. **The Faculty Board is responsible for the further elaboration of these regulations and determines the level at which this takes place: at faculty and/or institute level.**
These regulations are a framework within which the faculty can determine its own policy. This will be sometimes only at faculty and sometimes only at institute level; sometimes a combination of both will be the most appropriate. With regard to the latter, this could mean general faculty regulations that have a specific appendix for each institute.

11. **The academic director of the institute is responsible for the correct implementation of the policy.**
The DMPs for the research projects in the institute describe how the policy is implemented in practice (see below).

12. **The DMP for a research project must record which data-management responsibilities are assigned to the various members of staff who are working on the project.**
Many research projects involve several researchers, for instance PhD candidates, postdocs and members of the permanent staff. The DMP records which responsibilities have been assigned to whom. Projects in which researchers from multiple institutes and/or external parties are involved require special attention. It is also important to record what will happen to the data and the responsibility for it if one of the researchers leaves.

13. **All DMPs must be stored centrally within the faculty or institute for at least 20 years.**
Please note: if a DMP is destroyed, the researcher or institute can no longer demonstrate how certain decisions concerning the destruction of data were made (with regard to academic integrity). The retention term is longer for a DMP than for the data, because the DMP can then be used in

[1] http://www.datasealofapproval.org
[2] http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html
[3] http://www.iso16363.org/

any instances that arise to demonstrate that the destruction of a certain dataset (for instance after the standard ten years) was in line with the DMP.

14. **A data protection officer is appointed at university level who is responsible for adherence to the policy with regard to the security of research data.**
If research proposals are being granted funding within the Horizon 2020 framework, the data protection officer may be asked to provide a statement to the effect that the 'technical data protection procedures' will be adhered to within the project. The data protection officer is thus responsible for the protection of the research data during the project.

15. **The Executive Board is responsible for providing the facilities and support that make reliable data management possible.**
This does not mean that all facilities and support must necessarily be offered within the University; if required and beneficial, agreements can be reached with national or international parties who offer facilities and/or support in the field of data management.

C. Elaboration

16. **In the elaboration of these University regulations, the following aspects at least must be addressed. Derogations from these regulations must be recorded and well-motivated. The elaboration of the regulations should correspond with the standards that currently prevail in the discipline.**

- How data-management responsibilities are assigned within the institute;
- Which data needs to be retained, given the type of research conducted in the institute;
- What the policy is concerning data arising from bachelor's, master's and research master's programmes (and theses in particular);
- How the requirement that the integrity, availability and confidentiality of the data must be guaranteed will be met during the research;
- How, once the research has been completed, the requirement will be met that at the time of the research's publication the data must be retained in such a way that it remains findable, accessible, comprehensible and reusable in the long term.
- A possible maximum retention term for data and DMPs.

17. **The following aspects at least must be covered in the DMP for a research project:**

- How data-management responsibilities are assigned within the project and what will happen if the researcher, or one of the researchers, leaves;
- The types of data that will be generated and collected (with format and scope);
- The collection method(s) or origin of the data (including hardware and software);
- The standards and metadata for the documentation of the data (discipline-focused and/or in line with standards set by the repository where the data is preserved);
- Where the data will be preserved during the research and how security and access will be arranged;
- The measures that will be taken for the long-time preservation of and access to the data;
- Who will have access to the data at which point;
- How sensitive or otherwise confidential data is dealt with.

Researchers can use the templates that grant providers such as NWO and the European Commission use to draw up a DMP. The University has also created such a template. In the DMP, *researchers can in part refer to or draw upon the data management policy of their faculty or institute.*

18. **All the above articles have an implementation term of three years from the time of establishment, with the exception of the appointment of the data protection officer.**
    The proposed policy has such an impact that it is not realistic to make it compulsory within the short term. Opting for an extended implementation period means that faculties and institutes each have the opportunity to apply their own phasing-in plan. This does not alter the fact that in particular researchers who wish to submit grant proposals will probably already need to follow the greater part of this policy in the short term.

19. **These regulations do not have retroactive effect.**
    In principle, data collected in research projects that have already been completed therefore does not fall under the regulations. There is a grey area here if a new project builds upon previously collected data; it will be necessary to examine on a case-by-case basis whether it is possible to preserve the old data for the long term.

20. **If there is due cause, these regulations will be revised by the Executive Board during the implementation phase above.**
    At least the Research Council, the ICT and Research Steering Group and the Information Managers Platform must be consulted for such a revision.

The Research Data Management Regulations were established by the Executive Board on 19 April 2016.