LIBER Case Study:
Research Data Management at Radboud University

AUTHOR: Mijke Jetten, University Library, Radboud University, m.jetten@ubn.ru.nl

KEYWORDS: generic, institutional, policy, support, software services, infrastructure, data management plans, training

1 What was the starting point?

In November 2013, the Executive Board at Radboud University declared that all research data underlying a publication, together with the information necessary for potential reuse of data (metadata), have to be stored at the time of publication for a minimum period of ten years. The University's policy includes storage of the data of approved Bachelor's and Master's theses. At present, each research institute is in the process of supplementing the University's policy regarding the storage and management of discipline-specific research data.

The policy is grounded in the University Workgroup for Academic Integrity, which final report states that there is a relationship between the careful handling of research data and honest academic practice. Radboud University reflects this principle in their formulated policy with the ambition that all published research is reproducible.

Research data generated at Radboud University are stored, managed and made accessible in accordance with legal, academic, and ethical requirements as well as the requirements of financial providers. The University's policy reflects national and European developments in terms of data management and storage.

In two ways, the University is attentive to the required conditions for proper data management:

√Pilot 1. Support. The University Library is appointed to provide knowledge, advice, and support to researchers for the purpose of data management, including support in writing a Data Management Plan. Presently, in a two-year (2014-15) pilot, the University Library is giving shape to the services related to this new Research Data Management task (by establishing a Research Data Expertise Centre).

√Pilot 2. Infrastructure. Furthermore, the Research Data Management project includes conducting more in-depth research into a data management and storage infrastructure that is ideal for Radboud University, while focussing on big data as well as long tail data. A two-year (2014-15) pilot infrastructure is currently under way at the Donders Institute for Brain, Cognition and Behaviour (establishing a Donders infrastructure, in the near future possibly developed into a more broader Radboud University infrastructure).

Both pilots are financed internally by Radboud University itself. Besides the University Library (establishing a Expertise Centre Research Data) and the Donders Institute for Brain, Cognition and Behaviour (establishing a Donders infrastructure), the participants are the Executive Board (CvB), Market Research, Strategy and Development (MSO), Concern Information Management (CIM), and the ICT Service Centre (ISC).

June 2014

2 What kind of research data is targeted?

The policy is aimed at researchers of all Radboud University's research institutes. In supplementing the University's policy, each research institute has to provide additional policy and / or protocols about, amongst others, the kind of data that is gathered, managed and stored.

These supplements involve specifying both the kind of data gathered within the particular research field (e.g. interview data, video data, survey data, quantitative data, MRI data, observational data, experimental data, etc.) and the specific demands in terms of storing the data (e.g. storing raw data, reused data, processed data, data analyses, metadata). The argumentation is that research institutes themselves know best the specifics and demands of the research data at their institute. Regarding establishing metadata, research institutes are advised to link as much as possible to the standards that apply to their particular fields.

The research institutes also determine how long data are stored (minimum of ten years), in which repository the data are stored, accessibility and permission to reuse the data, the legal and ethical aspects involved in storing data, and support and training for researchers that are needed / available. The Data Management Plan for each research project is in accordance with the policy and protocols set out by the research institute. In providing knowledge, advice and support, the University Library cooperates with existing national and discipline-specific data repositories to make sure that the data are stored in the best-possible way (e.g. using preferred formats).

The research institutes will deliver these supplements in the coming months. As Radboud University is a broad university, focused on alpha, gamma, and beta sciences, the supplements being prepared by the research institutes are expected to be diverse and interdisciplinary, and involve all kinds of research data.

3 What is the organisational framework?

Roles and responsibilities

Final responsibility for and decision-making powers in the University's Research Data Management project is to be found in the RDM Steering Committee, which is chaired by a member of the University's Executive Board and includes Principal Investigators from various disciplines as well as the University Library's Director and the ICT Director, and an adviser from the department of Concern Information Management.

√Pilot 1. Support. The support pilot (establishing a Research Data Expertise Centre) is managed by an Information Specialist (data librarian) from the University Library. Both pilots have their own project group. The University Library's project manager is also part of the Donders Infrastructure project group.

√Pilot 2. Infrastructure. A Donders Institute researcher and an ICT developer manage the infrastructure pilot (establishing a Donders infrastructure). The endeavour to develop the infrastructure into a Radboud University infrastructure is managed by an Information Architect.

The primary responsibility for data storage and proper data management lies with the researcher / research project leader. The Director of the research institute is ultimately responsible for data storage by contributors to research that falls under their responsibility. The primary responsibility for data storage relating to approved Bachelor's and Master's theses lies with the Supervisor. The Director of Education is ultimately responsible for data storage for approved Bachelor's and Master's theses.

√Pilot 1. Support. The services in the Research Data Expertise Centre (information) are already outlined as they have been developed on the basis of detailed interviews with researchers about Research Data Management practices and needs within the various research institutes.

These services include sharing knowledge and giving advice and support with regard to, first, the storage of research data before and at the moment of publication (the Research Data Expertise Centre fulfills a front-office role with regard to the existing national data repositories Dutch Dataverse Network, DANS and 3TU.Datacentre; a front-office role might be established as well for discipline-specific repositories); second, support in writing a Data Management Plan (a University format is being tested); third, advice concerning the provision of rich metadata of research data; fourth, realising enhanced publications; fifth, providing legal and ethical advice related to the storage and management of research data; sixth, instructing researchers, particularly students and PhD students, in good data management and writing a Data Management Plan, both general and tailored, for instance within the context of graduate and doctoral schools. These services are still in development and the support that has been offered since the establishment of the University's policy represents a pilot. Information for researchers is provided on a website (http://www.ru.nl/library/services/expertisecentre/).

During the pilot phase, research institutes can call upon the University Library for advice and support in developing their own policy and protocols.

√Pilot 2. Infrastructure. At present, the specific components of the Donders Infrastructure are undecided as the requirements have to be discussed with researchers in the research institute.

Policies

Both the Donders Infrastructure pilot and the University Library's support pilot are grounded within the University's policy. The policy and protocols of the various research institutes are going to be derived from this University policy. The main reason for the University's Research Data Management policy is the relation between careful handling of research data and honest academic practice (scientific integrity and reproduction of research), as well as legal, academic, and ethical requirements and the requirements of financial providers (e.g. safe and sustainable preservation of data, sharing and reusing data, and enhanced publications).

The full University's policy can be found at http://www.radboudnet.nl/scientificintegrity/data-storage/university-policy/ (Intranet, only accessible by employees). The policy is summarised at http://www.ru.nl/library/services/expertisecentre/radboud-university/ (general access). Below is the aforementioned summarised form:

1. Data are stored at the time of publication of the research (including dissertations) at the latest, together with at least all the information necessary for potential reuse of data (metadata). A plan for managing data should be drafted prior to the commencement of data collection.

2. The data of approved Bachelor's and Master's theses are stored.

3. The retention period for research data is a minimum of ten years.

4. The primary responsibility for data storage and proper data management lies with the researcher/project leader. The Director of the Research Institute is ultimately responsible for data storage by contributors to research that falls under their responsibility. The primary responsibility for data storage for approved Bachelor's and Master's theses lies with the Supervisor. The Director of Education is ultimately responsible for data storage for approved Bachelor's and Master's theses.

5. The University is mindful of the required conditions for proper data management:
a) To provide knowledge, advice, and guidance for the purpose of data management and provide a Data Management Plan.
b) To provide an adequate infrastructure for data storage and management, during and after

research, insofar as the researchers have not organised adequate alternative facilities for data storage themselves.

6. A list of stored datasets will be included in the self-evaluation of the Standard Evaluation Protocol.

7. The University policy regarding the storage and management of research data will be supplemented by each research institute; the policy of the research institute should include agreements on at least the following:

- The responsibilities within the institute for data management and storage;
- Which data are included;
- How metadata are established, linking as much as possible to the standard and best practices that apply to the field;
- Where data are stored in the long term as well as the short term;
- How the data are protected in the event of technical problems (or a reference to the guarantees made by the supplier of the infrastructure);
- A possible longer minimum retention period (longer than ten years);
- Which maximum retention periods will apply (a minimum of ten years);
- Accessibility and reuse of datasets (defining for each dataset or research project how, under which conditions, when, and by whom the data can be reused);
- Privacy of sensitive data (processing, storage, access, security);
- Support and training for researchers.

The Data Management Plan for each research project is in accordance with the agreements made by the research institute.

4 What kind of support services are provided to researchers?

√Pilot 1. Support. Support and training provided to researchers by the University Library (Research Data Expertise Centre) are described in the previous paragraph.

The current practices and needs were collected by non-randomised interviews with researchers from the various research institutes, as described above.

Part of the University's Library pilot (establishing a Research Data Expertise Centre) involves training library staff. Currently, the focus is on awareness within the Library of this new Research Data Management task as a shared task for all units within the Library (User, Collection, and Information Services), and thus the necessity for all employees to gain (some basic) knowledge of the Research Data Management services (to be developed). In a later phase, part of the staff will be trained to be able to integrate Research Data Management related tasks in their regular activities.

5 What kind of infrastructure is provided?

Primarily, the storage of research data, both during a research project and at the moment of publication, will be performed using existing infrastructures, such as local network disks, national archives such as Dutch Dataverse Network, DANS and 3TU.Datacentre and discipline-specific archives. This is aimed at sharing, preserving and reusing data, as well as the enrichment of research data and publication. For archiving, discovery and registration of research data, we first and foremost rely on functionalities provided by these national archives. Access limitations to stored data are determined by individual researchers, based upon the specific policy of his or her research institute with regard to reusability.

Interoperability of these national and discipline-specific archives with other repositories (such as our Radboud Repository for publications) and other databases for research data (including

the ones used to import metadata and register scientific output such as Metis) is in fact a concern. Therefore, an aspect of the Research Data Management project at Radboud University is strengthening this interoperability, for instance by exploring the possibilities of a minimum metadata standard and a centralised CMS based on the CERIF model that communicates with other repositories. At present, options are only being explored, but one realises that this is the right time to take steps in the direction of integration of and operability between the various infrastructural systems, both within and outside the University.

√Pilot 2. Infrastructure. As previously mentioned, the functionalities of the Donders Infrastructure are currently undecided, as the requirements are being established, together with researchers.

6 What have you learned so far? What's next?

From the pilot we have learned that:

- Research data services are best developed by holding in-depth conversations with researchers.
- Research data management is tailor-made and consequently time-consuming for both researchers and supporting staff.
- Different types of research data require different advice and support. This reaffirms the claim that research data management is tailor-made.
- Research data management policies are best made within research institutes as they are aware of the different practices and needs with regard to research data management.
- Big data and long tail data require different infrastructural solutions as existing archives in the Netherlands do not offer solutions for big data (yet) due to their size, only for long tail data.

What are the challenges ahead?

- Migrating from pilot to embedded services, both regarding infrastructure (embedding the Donders infrastructure in the Research Institute) and support (embedding the Expertise Centre Research Data in the University Library of the future)
- Migrating from a Donders Infrastructure to a more broader Radboud University infrastructure.
- Creating awareness among researchers, which will be stimulated by the research institute policies and protocols in progress as well as by funder requirements with regard to Research Data Management.

Further information

Website: University Library's Research Data Expertise Centre (information), http://www.ru.nl/library/services/expertisecentre/

Contact: Natalia Grygierczyk (n.grygierczyk@ubn.ru.nl) & Mijke Jetten (m.jetten@ubn.ru.nl), University Library, Radboud University