LIBER Case Study:
Piloting Research Data Support at the University of Goettingen

AUTHOR: Dr Birgit Schmidt, Jens Ludwig, Goettingen State and University Library, bschmidt@sub.uni-goettingen.de, ludwig@sub.uni-goettingen.de

KEYWORDS: institutional, interdisciplinary

1 What was the starting point?

The University of Goettingen's Principles of Good Research Practice (Georg-August-University Göttingen 2012) stipulate that all primary data, which are the basis of a scientific publication, have to be stored on a stable and secured medium for a minimum of 10 years. In Germany such policies are typically closely aligned with the respective policy of the German Research Foundation (DFG). Since 2011, the university's research support service has been involved in the coordination of the universities' activities in research data management. It provides basic advice to researchers for data management planning and collaborates with the library and the scientific computing centre (GWDG). A policy for research data management is currently under development and will be complemented by a roadmap for building up institutional data management infrastructures and services.

Goettingen State and University Library (SUB) has a strong tradition of research and development in the field of information management and long-term preservation, including that of research data (e.g. Neuroth et al 2013). In particular, SUB is involved in the development of research infrastructures for the arts and humanities (DARIAH, TextGrid, CENDARI, etc.) and open access infrastructures for publications (OpenAIRE). Over the last few years, a demand for infrastructure and support measures at an institutional level has emerged within the context of large cross-institutional research projects as well as smaller research units on campus.

Since 2010, the German Research Foundation (DFG) has requested basic data management plans in project proposals and since 2007 large collaborative research projects can apply for embedded information infrastructure (INF) sub-projects. Since 2012, SUB's Research & Development Department has been contributing to the EFForTS project (Collaborative Research Centre 990: Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (Sumatra, Indonesia), https://www.uni-goettingen.de/en/310995.html) and provides practical advice and support for the development of a tailored data management and information infrastructure. Further collaborative projects with embedded data management support are currently under development.

Based on the growing demand for IT and data-based eResearch services, the University of Goettingen forges an eResearch Alliance (http://www.uni-goettingen.de/en/482513.html) which is aimed at providing institutional support to researchers, scientists and scholars. The eResearch Alliance is run mutually by the Goettingen State and University Library and the Computing and IT Competence Centre Goettingen (GWDG). The offerings being developed by the eResearch Alliance in the coming years include setting up digital infrastructures for the humanities and the natural and social sciences. Furthermore, counselling and training courses will be offered and services set up for research data management.

June 2014

**Ligue des Bibliothèques Européennes de Recherche
Association of European Research Libraries**

2 What kind of research data is targeted?

The infrastructure of EFForTS is dedicated to serve an interdisciplinary team involving researchers from forestry, agriculture, biology, geosciences, socio-economic sciences, anthropology and information science. The data are collected by groups of researchers during their research visits to Sumatra, Indonesia. In addition, data are collected by three universities in Indonesia which collaborate with Goettingen University in the EFForTS project. The collected data cover all technical and methodological kinds of data and include, for example, household surveys and case studies, plant, animal and soil samples, sound recordings of birds, geographical information, etc. The data are further analysed and processed to determine, e.g. the effects of land use on soil-cycling rates, etc.

Data creators must deliver research data and metadata in an appropriate form according to common standards and good academic practices. In particular, all data will be available according to the Recommendations for Secure Storage and Availability of Primary Research Data (DFG 2009). With regard to quality control, these recommendations stipulate that disciplinary criteria and processes shall apply.

3 What is the organisational framework?

Roles and responsibilities

The INF sub-project is led by the Department of Ecoinformatics, Biometrics and Forest Growth, in close collaboration with SUB's Research and Development Department. The technical infrastructure combines CRC infrastructure (e.g. a platform for project-internal data exchange and for documenting events on research plots), university infrastructure (e.g. long-term storage infrastructure, persistent identifier infrastructure) and external components (e.g. external data repositories). Support services include the provision of templates for research data and documentation as well as support with the creation of individual data management plans.

Policies

Project Level

To secure a common understanding of the sharing and collaborative use of the research data across sub-projects and across institutions, the EFForTS project has agreed on a Data Exchange Agreement. All partners of the EFForTS project agree that the research data collected and compiled are fundamental to the success of the project and other projects working towards sustainable development. In order to achieve a better understanding of the interdisciplinary research topics, the sharing and collaborative use of the data is considered crucial. Some datasets relevant to many sub-projects are created centrally. The agreement clarifies the terms of data delivery, access and use as well as restrictions and conditions. For example, researchers are obliged to provide metadata to the project's information system before the beginning of data collection, and the research data must be deposited as soon as possible, at the latest one year after the field sample has been taken or the laboratory analysis has been completed. All metadata will be publicly available and released under a Creative Commons CC0 licence. Each member of the project should have access to all research data via the data creator. At the latest, research data will be automatically available internally two years after collection of the research data, and publicly available five years after collection of the research data. Researchers are encouraged to make their data internally/publicly available earlier than these periods. Exceptions are of course possible, e.g. for datasets with sensitive information.

Institutional Level

Goettingen University's institutional data management policy (currently under development) defines the responsibilities of the researcher, typically the principal investigator. It requests the development

of individual data management plans and advises on the handling of intellectual property rights and other legal constraints. The policy also outlines infrastructure and support services that will be provided by the university. Goettingen University encourages its researchers to make their research data openly accessible.

4 What kind of support services are provided to researchers?

As described above, the INF sub-project is embedded in the Collaborative Research Project and works closely with representatives of every research team. Three data managers supported by technical and management staff provide practical advice, collect information about researchers' needs, and translate this into requirements and solutions provided by the information system.

Basic training has been provided to researchers covering the basics of research data management, such as backup and storage, metadata and documentation, formats and quality control, names, identifier and versions, as well as ethics, rights and licences. These training sessions have been instrumental in raising awareness.

Basic research data training courses for graduate students at Goettingen University are also in preparation.

5 What kind of infrastructure is provided?

The EFForTS project uses BExIS (http://fusion.cs.uni-jena.de/bexis) as a system to support data sharing and reuse. Publication and archiving workflows will include technical quality assurance, submission to preservation facilities, registration in international catalogues, etc. Access restrictions depend on the data exchange agreement as described above and are facilitated through the information system's technical setting.

The scientific computing centre of the university provides archival storage and generic computing resources.

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

A typical challenge for institutions providing embedded research data management services is the imbalance of supply and demand for these services. Institutions usually do not have the capacity to initiate the foundations of the infrastructure tasks before the project starts even though basic services are often required early on in the project. It takes time to develop tools such as standards-compliant data creation, whereas researchers usually cannot delay the creation of data. And not all requirements can be defined or identified before the actual project starts.

Another challenge is to motivate researchers into actually providing data. Policies are only of limited use for this even if researchers willingly accept them. This depends much more on the acceptance of data managers within the project context. One small, incentivising example in the EFForTS project involves conferring symbolic awards like "data sharer of the month".

Further information

Website: University of Goettingen, Data Management and Archiving, http://www.uni-goettingen.de/en/221506.htm; Goettingen State and University Library, Research and Development Department, http://www.rdd.sub.uni-goettingen.de/www/pages/rdd_public/index.php?switch_language=en

Contact: Jens Ludwig, Goettingen State and University Library, ludwig@sub.uni-goettingen.de

References

Neuroth, H.; Strathmann, S. Oßwald, A. & Ludwig, J. [Eds.] (2013). Digital Curation of Research Data – Experiences of a Baseline Study in Germany, Verlag Werner Hülsbusch, http://nbn-resolving.de/urn:nbn:de:0008-2013112701

German Research Foundation (DFG) (2009). Recommendations for Secure Storage and Availability of Digital Primary Research Data, Committee on Scientific Library Services and Information Systems, Bonn, http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf

Georg-August-University Göttingen (2012). Statute of the Georg-August-University Göttingen for ensuring good scientific practice, in force from 21 December 2012, https://www.uni-goettingen.de/en/223832.html