PRACTICE PAPER

Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

Constanze Curdt[1,2]

[1] Institute of Geography, University of Cologne, DE

[2] Regional Computing Centre (RRZK), University of Cologne, DE

c.curdt@uni-koeln.de

Science conducted in cross-institutional, interdisciplinary, long-term research projects requires active sharing of data, documents and further information. Thus, within the Collaborative Research Centre/Transregio 32 'Patterns in Soil-Vegetation-Atmosphere Systems', funded by the German Research Foundation, research data management (RDM) services have been available since early 2007. These services were established to support all researchers during their entire individual research studies. They cover provision of general guidance, support and training for RDM. To fulfil the scientists' needs and requests with regard to storage, backup, documentation, search and sharing of data with other project members, a project-specific RDM system was designed and implemented. This system was developed and continuously modified in collaboration with the scientists to facilitate their system acceptance. Besides the mentioned services, the system supports further common services such as controlled access to data, rights management, data publication with DOI and data statistics (on repository and single data level). All RDM services provided for the scientists are thus bundled and available to the users in one system: a 'one-stop-shop'. After more than ten years of RDM service provision for the CRC/TR32, the repository statistics clearly visualize the use of the diverse RDM system services. Furthermore, it has been shown that an RDM adapted to the needs of interdisciplinary researchers can be fruitful and indispensable when scientists conduct their research study e.g. with a time lag. RDM services established at an early stage can contribute to a successful long-term research project.

Keywords: Research Data Management; Services; Data Repository; Geosciences

1 Introduction

In the last decade, the importance of research data management (RDM) has increased in many research areas. New methods and technologies have facilitated the rapid creation of (research) data and digital information. For example, in the geo- and atmospheric sciences, global climate data (incl. satellite imagery, model results and *in situ* measurements) have dramatically increased in volume and complexity (Overpeck et al. 2011). Thus, proper data storage and backup, as well as exchange, re-use and open access to data, have become even more important. Consequently, in recent years many RDM services have been devised and implemented by various institutions (Pryor et al. 2014). According to Jones (2014), these services can range from providing training and guidance to developing policies and strategies. Likewise, services cover the establishment of data repositories and infrastructures.

Science conducted in collaborative, cross-institutional, interdisciplinary, long-term research projects requires the active sharing of research data, documents and further information in a well-managed, controlled and structured manner. Usually, collaborating scientists depend not only on their own data but also on that of other colleagues to conduct their research. Consequently, infrastructures are required that can serve cross-disciplinary demands (Klar & Enke 2013). Mückschel et al. (2008) and Lotz et al. (2012) likewise emphasize the importance of appropriate RDM infrastructure within multi-disciplinary projects that

Art. 5, page 2 of 9

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

focus on environmental field studies and modelling. Appropriate infrastructures can improve collaboration between scientists and link research results, and are crucial for developing synergies. Thus they should support scientists during the research data life cycle (e.g. data collection, storage, sharing, documentation). Additionally, data metrics should be considered as they are becoming increasingly important within RDM (Whyte et al. 2014). These could include numerically tracked measures such as website visits or download and usage statistics of repository data. Infrastructures should thus be set-up according to the requests and needs of the scientists and should be in line with current standards and principles (e.g. metadata standards).

Within large research collaborations, the German Research Foundation (DFG) provides funding on request to establish infrastructures and RDM services according to needs of the scientists in 'Information Infrastructure' (INF) service projects (Redöhl 2016). This contribution presents the RDM practices for the interdisciplinary, DFG-funded, research project 'Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation', as exemplary RDM services developed within an INF service project. After a short introduction to the project background and data, the established RDM services and tools that have been developed to cover the project/scientists needs since 2007 is presented. Finally, the findings are discussed and a conclusion provided.

2 Project Background and Data

The Collaborative Research Centre/Transregio 32 (CRC/TR32, www.tr32.de) 'Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation' is an interdisciplinary, long-term research project. The project has been funded by the German Research Foundation (DFG) since 2007 and is currently in its final funding phase. Several research groups of the field of geosciences of the German universities of Aachen, Bonn and Cologne, as well as the Research Centre Jülich are involved. The main research aim is to yield improved numerical models to predict water, energy and CO_2 transfer by calculating patterns at various temporal and spatial scales. The study is conducted within the catchment of the river Rur located in western Germany, parts of The Netherlands and Belgium. Scientists from numerous disciplines (e.g. soil and plant sciences, geography, geophysics, meteorology and mathematics) are participating. They observe and model energy and water fluxes geographically from the single soil pores to the entire catchment scale and from the ground water to the atmospheric boundary layer (Simmer et al. 2015).

The CRC/TR32 scientists have collected a variety of research data in different spatial and temporal scales. These heterogeneous data have been obtained from various field campaigns (e.g. by meteorological or hydrological monitoring, airborne campaigns), laboratory studies (e.g. plant biomass measurements) or by following data modelling approaches. The data covers various spatial scales such as the entire Rur catchment, selected sub-regions or specific field plot area. Likewise, data have been collated in different temporal resolutions, e.g. every minute, daily, fortnightly, monthly, annually or at irregular interval. Moreover, all scientists have produced further documents during their study including publications, conference contributions (e.g. posters, talks) and PhD reports or dissertations.

3 Data Management Practices and Services

To support all CRC/TR32 scientist, mainly postdocs and PhD students, during the entire research life cycle, various research data management (RDM) services (**Figure 1**) have been established since 2007 in a funded INF service project.

One main issue and goal while developing the RDM services was the set-up of a RDM system, the CRC/TR32 project database (TR32DB, www.tr32db.de) (Curdt & Hoffmeister 2015). This self-designed system has been online since early 2008. It has continuously been refined and improved. The TR32DB was designed in close cooperation with the Regional Computing Centre (RRZK) of the University of Cologne. The scientists' needs and requests (e.g. various file sizes from kilobyte to gigabyte per file, heterogeneous data and formats) were considered during the set-up. Additionally, DFG requirements were taken into account. These included, for instance, the re-use of an available infrastructure and the cooperation with a computing centre or library. Likewise, a user-friendly system was established. The TR32DB supports common features of a RDM system and promotes most of the project RDM services, as described in the following.

3.1 Guidance, Training and Support

Guidance, training and support with regard to RDM are an important part of the established RDM services. Consequently, they have been provided for all project scientists during the project funding. They include practical training workshops, which run at irregular intervals relative to the demands of the project scientists. In general, these workshops cover an introduction to basic RDM principles, e.g. data organization, file formats and data documentation, and also provide a practical introduction to using the TR32DB system.
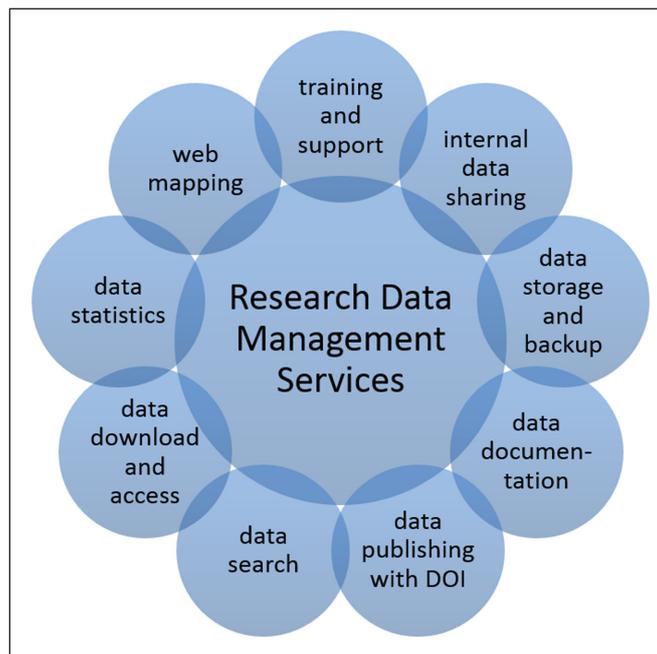
Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in
Soil-Vegetation-Atmosphere-Systems' by Data Management Services

Art. 5, page 3 of 9

Figure 1: Research data management services provided for CRC/TR32 members, modified from Curdt & Hoffmeister (2015).

The scientists are taught to upload various data types (e.g. research data, publications) to the TR32DB, enter metadata via a wizard and request data via the various data search options. Usually, attendance at these practical training sessions has been high since they provide a good opportunity to discuss and evaluate further developments of the system.

Additionally, user support and guidance is provided for the project members through various tutorials (e.g. short overview and expanded tutorials) and a frequently-asked-questions section on the website (http://www.tr32db.de/site/faq.php). Finally, user support and guidance is also offered by e-mail, phone or on a personal basis (e.g. at the various project meetings). Appropriate support is essential to ensure that the scientists use the services and infrastructure correctly and to their own benefit.

3.2 Internal Data Sharing

For internal exchange of research data and results with colleagues of other sub-projects, an internal, secure data exchange server platform was established within the TR32DB system. This special service is a peer-to-peer file sharing service that is secure and easy to use. This platform allows the scientists to exchange data up to 8GB per file. For example, preliminary research data, which are often in a raw format and are sometimes still in an unorganized form but nevertheless useful for other scientists (e.g. field measurement results that are needed for model input), can be shared. To share the data, scientist have to upload their raw or preliminary data to an internal separate data storage area within the TR32DB system that is provided for every single sub-project. A metadata documentation for these files is not requested. Data in the share storage are downloadable for all TR32DB users. Interested scientists can now directly download data from the file server or use the web-interface. In an internal, separate share section of the website, all files are presented sorted by sub-project. This file sharing service was requested by the scientists and is used frequently.

3.3 Data Storage and Backup

A secure, file-based data storage and backup (e.g. files up to 8GB) for all data created by the project participants is provided within the TR32DB system. Scientists can upload different types of data such as research data, geodata, publications and reports. These data are usually the finalized, processed and prepared data. Scientists are advised to upload their data in the standardized format of their scientific community. The data storage itself is organized in a hierarchical folder structure (e.g. clusters, sub-projects, different data types) based on the research project structure. All data are backed up every night in the hardware infrastructure of the RRZK using IBM's Tivoli Storage Manager backup system. This back-up ensures the long-term access and availability of the data. As a long-term preservation solution, the current plan is to use Ex Libris' Rosetta digital preservation system as a dark archive; it is provided by the North Rhine-Westphalian Library Service Centre (hbz) as a service.

Art. 5, page 4 of 9

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

3.4 Data Documentation

For the accurate documentation of all project data, the project specific, interoperable, multi-level TR32DB Metadata Schema (Curdt 2016) was designed and implemented within the TR32DB system. This schema is based on several existing metadata schemes (e.g. Dublin Core, Data Cite, ISO 19115). It enables all supported data types (e.g. data, geodata, publication, reports) to be accurately documented, which ensures that the data are findable, accessible and reusable. Physically, the metadata are stored in a MySQL-Database as well as in XML-format. Within the TR32DB system, the users provide the metadata for a specific dataset through a user-friendly metadata input wizard that is implemented in the systems website. This wizard was designed according to the metadata schema and supports a template feature to re-use existing TR32DB metadata. Additionally, the wizard promotes many drop-down lists for controlled vocabulary and info items with help texts including examples. After a successful metadata submission, the metadata can be modified by its creator as often as needed. A metadata export (i.e. to TR32DB-Schema, DataCite, ISO 19115) from the website was recently implemented. It can be accessed by interested users and allows the metadata to be sent to other services.

3.5 Data Publishing

With a successful submission of the metadata, a dataset is instantly visible and published on the TR32DB website (www.tr32db.de). The metadata of a dataset (compare **Figures 2** and **3**) are openly accessible for every website visitor. The data are arranged according to the structure of the metadata schema and are grouped in various categories covering information with regard to identification (e.g. dataset title, description, related data), responsible party (e.g. author/creator and contributor details) or topics (e.g. keywords). Furthermore, a recommendation for a citation of the dataset is provided, as well as metadata details about the metadata itself (e.g. metadata creator, metadata date and latest updates). By default, each dataset in the TR32DB contains an internal data identification number (e.g. ID 1471) and a corresponding permanent URL (e.g. http://tr32db.uni-koeln.de/data.php?dataID=1471) for access to the metadata of the dataset. A metadata export (e.g. to TR32DB-Schema, DataCite, ISO 19115) from the website that can be accessed by interested users and harvesting to other services were recently implemented.

Additionally, the data provider can apply a Digital Object Identifier (DOI) to his or her dataset to make it, for example, citable in a publication. This permanent identifier enables the access to corresponding TR32DB metadata and the underlying dataset. The TR32DB-DOI (e.g. 10.5880/TR32DB.20) provides access to the

Figure 2: Landing page of a TR32DB-DOI (exemplary presented: DOI 10.5880/TR32DB.20) including the link to the TR32DB metadata viewing webpage (A).

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

Art. 5, page 5 of 9

landing page (**Figure 2**). This website presents all relevant metadata of the dataset (e.g. data creator, title, description, data access details) and provides a citation recommendation for the registered dataset as well as a link to further metadata in the TR32DB. The DOI publication is conducted through the German Research Centre for Geosciences (GFZ) Potsdam. The GFZ offers a DOI registration service (Ulbricht et al. 2016) for GFZ repositories as well as for selected external geoscientific data repositories (e.g. such as TR32DB, TERENO, CRC 1211 Database). In the collaboration between GFZ and TR32DB, responsibilities were arranged such that TR32DB is in charge of the sustainable data storage, preparing and provision of quality controlled metadata according to DataCite Metadata Schema as well as of delivering a project-specific landing page (**Figure 2**). The GFZ offers, amongst other services, the technical infrastructure, software and interfaces for minting data DOIs as well as for metadata harvesting to other services and portals.

3.6 Data Search
Data search via metadata is an important service provided for the project participants. Consequently, several search options have been established within the TR32DB system to find data. The web-interface enables quick access to all project data through different search queries. These options include the data search according to predefined queries such as browsing for data by data types, topics, regions and sites, funding phases or project sections. Moreover, an advanced data search is provided that enables the combination of various search items such as a combination of predefined lists (e.g. data type, keywords, creator) and a free text search. Additionally, a spatial data search via a map is provided (**Figure 3**). The website visitor can pan and zoom to an area of interest and select a dataset represented by a red marker pin on the map. By moving over the pin or selecting it, the geographic coverage of the dataset is presented. Additionally, some details (e.g. data identifier, title) of the dataset are visualized. Further metadata is shown by selecting the dataset title.

3.7 Data Download and Access
Constraints with regard to download and access are important and were also demanded by the project members. Consequently, appropriate services were established within the TR32DB system. These services include, for instance, data download permissions that can be set by the data provider for each individual dataset. This service includes the option to make datasets only downloadable by sub-project members or by all TR32DB users (with a login), as well as free download for everyone. Moreover, different licenses can be applied to the datasets such as Creative Commons or Open Data Licenses. The default setting is the specific TR32DB Data Policy Agreement. The policy was developed in the project and approved by the project board.
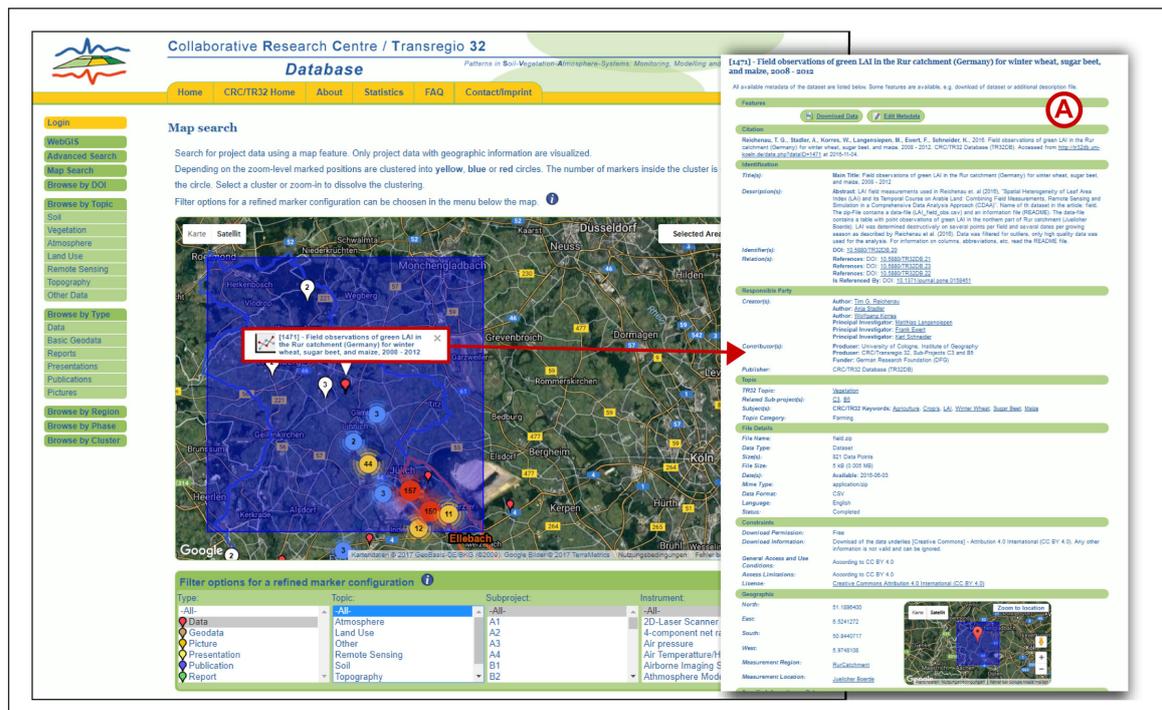

Figure 3: Spatial data search for data in the TR32DB system using the map search. Exemplary the link to the metadata viewing webpage (A) of the selected dataset ID 1471 is presented.

Art. 5, page 6 of 9

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

3.8 Data Statistics

In recent years, data metrics and statistics about datasets and repositories have become increasingly important (e.g. data views, downloads, mentions, bookmarks). Therefore, several data statistics services were developed for the TR32DB system. They have been available online via the website's statistics menu (http://www.tr32db.de/site/Statistics.php) since February 2016. Most of the metrics are openly accessible. In general, these metrics are numerically tracked measures for single datasets as well as for the entire content (all datasets) of the TR32DB data repository.

These statistics mainly support visualizing the TR32DB data content. For instance, charts are presented that display all available TR32DB data with respect to a specific value. Some dynamic stacked bar charts display the distribution of all TR32DB datasets to the assigned sub-project, subdivided according to the data types. Also, various statistical values are presented in relation to time such as the increase of datasets in the TR32DB over the years for the respective data types or at the project cluster level. Some pie charts display the entire repository data according to specific themes. The distribution of all data is, for example, presented according to all data types, topics, cluster/subprojects or locations. In addition, the most frequently downloaded datasets are listed. Further data metrics are available on the single metadata viewing webpage, such as downloads or metadata views of a specific dataset. The statistic functions allows the frequent use of the TR32DB system to be clearly shown. Data upload, sharing and download, as well as viewing of metadata, are visualized in the various charts.

3.9 Web Mapping

Several web mapping services are provided for project participants within the TR32DB system. One of these services is the map search for data (**Figure 3**) in which the geographic coverage for a dataset is provided via a web mapping application based on the metadata provision. The geographic coverage of the dataset is then presented on a map within the metadata overview website (compare **Figures 2** and **3**). Furthermore, an internal, integrated WebGIS was established within the TR32DB. It visualizes purchased geodata and climate data as well as providing data from project scientists (e.g. coverage of land use for the study area).

4 Discussion

The previous sections presented the RDM services and infrastructure components established within the INF project of the Collaborative Research Centre/Transregio 32 (CRC/TR32) project. Implementing such RDM support and services within interdisciplinary, long-term research projects such as DFG-funded CRCs is crucial. DFG therefore provides funding for these integrated INF services projects that can be used to support RDM and/or implement RDM infrastructures for CRCs. Besides the adequate funding of staff (e.g. postdoc researchers, technical staff), funding is also available for storage, hardware and software or licenses can be requested (Redöhl 2016).

The RDM service implementation in CRC projects can be diverse. Some CRCs focus on establishing technical infrastructures (e.g. Wang et al. 2016, Weber & Piesche 2016, Willmes et al. 2014), while other CRCs apply the concept of an 'embedded data manager' (Cremer et al. 2015). Nevertheless, only about 10% of all funded CRCs have successfully applied for and include an INF service project providing RDM support (Redöhl 2016). Although the added value of RDM in CRCs is tangible, Klar & Enke (2013) emphasize that RDM in CRCs is particularly important in achieving the overall research project goal. In many research projects, scientists are dependent upon the research results and data of other scientists to answer their research questions. Mückschel et al. (2008) highlight that this specifically applies e.g. for research projects in the earth and environmental sciences, where synergies have to be created between field data collectors and data modellers. For instance, within the CRC/TR32 the modelling community is dependent on results and data collected during field measurement campaigns since these are needed as input parameters for various models and simulation platforms. In this case, the open availability of all metadata, independent of data access rights of the data owner, clearly enhances the discoverability and re-use of data, in particular over the long project duration. Consequently, establishing RDM services and infrastructure such as data storage, backup, documentation, search and access is essential. RDM systems facilitate project data exchange within the project funding but also enable data re-use by future project participants (e.g. new PhD students or postdocs) and for future studies.

One part or the established RDM services for the CRC/TR32 scientists is the implemented project-specific RDM system. The TR32DB was established to support scientists during the entire research project life cycle (e.g. data storage, documentation, search, download). The CRC/TR32 scientists were involved in developing the RDM system and services design at an early stage, as recommended by Mückschel et al. (2008), thus facilitating

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

Art. 5, page 7 of 9

the use of the TR32DB. It is one major lesson learned that scientists will only use infrastructures if they are designed in compliance with their needs. Consequently, RDM systems should fulfil project challenges as well as be designed with regard to the project background. Additionally, the infrastructures should be easy, intuitive and simple if their acceptance and use is to be facilitated. It is essential to avoid overburdening the scientists. In addition, the system is implemented with a focus on data security (structure of the file system), reliability (backup in IBMs Tivoli Storage Manager (TSM), additional storage of metadata in XML-files) and interoperability, which enhance a scientist's trust in the system. Finally, continuous guidance, training and support (e.g. with descriptive tutorials, practical workshops) promote the acceptance and use of the system. Also patience as well as an understanding of the scientists and the reasons for their specific situation are necessary. These reasons could include, for instance, not sharing data that has not been published in a paper or are preliminary results. For this case, supporting internal, secure data-sharing platforms are important.

The successful use and acceptance of the TR32DB and the associated RDM services is clearly visible in the amount of data stored in the system. In August 2018, around 850GB of data were stored and around 1,800 datasets were accessible from the TR32DB metadata. These data mainly comprise final, processed research results. Only little raw data (e.g. from selected measurement instruments) is available. Additionally, TR32DB statistics (http://www.tr32db.uni-koeln.de/site/Statistics.php) show the frequent usage of the system, e.g. registered TR32DB users upload, share and download data. Likewise, internal statistics indicate that TR32DB visitors (without login) also access metadata and download openly available data. Data re-use is also sometimes presented within the metadata of a dataset. For example, the datasets ID 1497 (http://www.tr32db.uni-koeln.de/data.php?dataID=1497) and ID 1470 (http://www.tr32db.uni-koeln.de/data.php?dataID=1470) were created by means of other TR32DB data. The latter dataset is published with a TR32DB-DOI (DOI: 10.5880/TR32DB.23) and is supplemental material of a journal publication. To foster supplemental data with DOI in a journal publication, an early support and assistance with data organisation and documentation is necessary.

5 Conclusion and Outlook

Research that is conducted in collaborative, cross-institutional, large-scale, long-term, interdisciplinary research projects requires an active sharing of research ideas and their output (e.g. research data, publications, reports and other documents). This sharing should at best be conducted in a controlled and structured manner. Thus, it is important to establish appropriate, supporting research data management (RDM) services and infrastructure for the scientists according to their needs from the beginning of the project.

This paper presented an overview of established RDM services that were set up for the long-term, interdisciplinary DFG-funded research project Collaborative Research Centre/Transregio 32 (CRC/TR32), which focused on research in patterns in soil, vegetation and atmosphere systems. Over a period of 10 years, RDM services and a RDM system/repository were established, closely aligned to the requirements and needs of the scientist and DFG-recommendations. These services support the scientists during their research study with respect to secure data storage and backup and to data publishing, search and documentation. Statistics (on repository and data level) provide information e.g. on the provision and re-use of project data. Finally, workshops and practical training on the RDM system, as well as continuous support for all RDM services established since the project start, result in the use of the services and infrastructure. Is has been proven with the chosen concepts that establishing RDM services from the beginning of the project can contribute to the success of the entire research goal.

Before the end of the project funding in December 2018, a final user evaluation of the established RDM services is planned. Since the RDM service and system infrastructure are currently and will continue to be used in future research projects, the aim is to receive user suggestions for service improvements. Moreover, comprehensive support and training is essential for the final provision and publication of the remaining project data that has to be transferred to the TR32DB. With respect to the sustainability and the future data access, the TR32DB system infrastructure will be transferred to the 'TR32DB-Box' that will be hosted and maintained (e.g. update of single system components) in the hardware infrastructure of the Regional Computing Centre. Moreover, the current plan is to use the Ex Libris' Rosetta digital preservation system as a dark archive, a service that is provided by the North Rhine-Westphalian Library Service Centre (hbz).

Art. 5, page 8 of 9

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in Soil-Vegetation-Atmosphere-Systems' by Data Management Services

'Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation' funded by the German Research Foundation (DFG).

Competing Interests
The author has no competing interests to declare.

References

Cremer, F, Engelhardt, C and **Neuroth, H.** 2015. Embedded Data Manager – Integriertes Forschungsdatenmanagement: Praxis, Perspektiven und Potentiale. *Bibliothek Forschung und Praxis*, 39(1).

Curdt, C. 2016. Metadata Management in an Interdisciplinary, Project-Specific Data Repository: A Case Study from Earth Sciences. In: Garoufallou, E, Subirats Coll, I, Stellato, A and Greenberg, J (eds.), *Metadata and Semantics Research: 10th International Conference, MTSR 2016, Göttingen, Germany*, 357–368. Cham, Switzerland: Springer International Publishing.

Curdt, C and **Hoffmeister, D.** 2015. Research data management services for a multidisciplinary, collaborative research project: Design and implementation of the TR32DB project database. *Program*, 49(4): 494–512. DOI: https://doi.org/10.1108/PROG-02-2015-0016

Jones, S. 2014. The range and components of RDM infrastrucutre and services. In: Pryor, G, Jones, S and White, A (eds.), *Delivering Research Data Management Services: Fundamentals of Good Practice*, 89–114. London, UK: Facet. DOI: https://doi.org/10.29085/9781783300242.006

Klar, J and **Enke, H.** 2013. Forschungsdaten in der Gruppendomäne – Zwischen individuellen Anforderungen und übergreifenden Infrastrukturen. *Zeitschrift für Bibliothekswesen und Bibliographie (ZfBB)*, 60(6): 316–324. DOI: https://doi.org/10.3196/186429501360649

Lotz, T, Nieschulze, J, Bendix, J, Dobbermann, M and **König-Ries, B.** 2012. Diverse or uniform? — Intercomparison of two major German project databases for interdisciplinary collaborative functional biodiversity research. *Ecological Informatics*, 8: 10–19. DOI: https://doi.org/10.1016/j.ecoinf.2011.11.004

Mückschel, C, Weist, C and **Köhler, W.** 2008. Central data management in environmental research projects – selected problems and solutions. In: Müller, RAE, et al. (eds.), *28. Gil Jahrestagung*, 101–104. Kiel, Germany: Lecture Notes in Informatics.

Overpeck, JT, Meehl, GA, Bony, S and **Easterling, DR.** 2011. Climate Data Challenges in the 21st Century. *Science*, 331(6018): 700–702. DOI: https://doi.org/10.1126/science.1197869

Pryor, G, Jones, S and **Whyte, A.** 2014. *Delivering Research Data Management Services: Fundamentals of good practice*. London, UK: Facet Publishing. DOI: https://doi.org/10.29085/9781783300242

Redöhl, B. 2016. The DFG Perspective: Research Data Management with a Focus on Collaborative Research Centres (SFB). In: Curdt, C and Willmes, C (eds.), *Proceedings of the 2nd Data Management Workshop*, 96: 77–81. 28–29.11.2014. University of Cologne, Germany. Geographisches Institut der Universität zu Köln, Cologne, Germany. DOI: https://doi.org/10.5880/TR32DB.KGA96.12

Simmer, C, et al. 2015. Monitoring and Modeling the Terrestrial System from Pores to Catchments: The Transregional Collaborative Research Center on Patterns in the Soil–Vegetation–Atmosphere System. *Bulletin of the American Meteorological Society*, 96(10): 1765–1787. DOI: https://doi.org/10.1175/BAMS-D-13-00134.1

Ulbricht, D, Elger, K, Bertelmann, R and **Klump, J.** 2016. PanMetaDocs, eSciDoc, and DOIDB—An Infrastructure for the Curation and Publication of File-Based Datasets for GFZ Data Services. *ISPRS International Journal of Geo-Information*, 5(3): 25. DOI: https://doi.org/10.3390/ijgi5030025

Wang, W, Göpfert, T and **Stark, R.** 2016. Data Management in Collaborative Interdisciplinary Research Projects—Conclusions from the Digitalization of Research in Sustainable Manufacturing. *ISPRS International Journal of Geo-Information*, 5(4). DOI: https://doi.org/10.3390/ijgi5040041

Weber, A and **Piesche, C.** 2016. Requirements on Long-Term Accessibility and Preservation of Research Results with Particular Regard to Their Provenance. *ISPRS International Journal of Geo-Information*, 5(4). DOI: https://doi.org/10.3390/ijgi5040049

Whyte, A, Molloy, L, Beagrie, N and **Houghton, J.** 2014. What to Measure? Toward Metrics for Research Data Management. In: Ray, J (ed.), *Research Data Management: Practical Strategies for Information Professionals*, 275–300. West Lafayette, IN, USA: Purdue University Press.

Willmes, C, Kürner, D and **Bareth, G.** 2014. Building Research Data Management Infrastructure using Open Source Software. *Transactions in GIS*, 18(4): 496–509. DOI: https://doi.org/10.1111/tgis.12060

Curdt: Supporting the Interdisciplinary, Long-Term Research Project 'Patterns in
Soil-Vegetation-Atmosphere-Systems' by Data Management Services

Art. 5, page 9 of 9