

## DEVELOPMENT OF A METADATA MANAGEMENT SYSTEM FOR AN INTERDISCIPLINARY RESEARCH PROJECT

C. Curdt \*, D. Hoffmeister, G. Waldhoff, C. Jekel, G. Bareth

Institute of Geography, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany  
(c.curd, dirk.hoffmeister, guido.waldhoff, jekelc, g.bareth)@uni-koeln.de

Commission IV, WG IV/1

**KEY WORDS:** Metadata, Research, Management, Database, Spatial Infrastructures, Internet/Web, Soil, Vegetation, Atmosphere

### ABSTRACT:

In every interdisciplinary, long-term research project it is essential to manage and archive all heterogeneous research data, produced by the project participants during the project funding. This has to include sustainable storage, description with metadata, easy and secure provision, back up, and visualisation of all data. To ensure the accurate description of all project data with corresponding metadata, the design and implementation of a metadata management system is a significant duty. Thus, the sustainable use and search of all research results during and after the end of the project is particularly dependent on the implementation of a metadata management system. Therefore, this paper will describe the practical experiences gained during the development of a scientific research data management system (called the TR32DB) including the corresponding metadata management system for the multidisciplinary research project Transregional Collaborative Research Centre 32 (CRC/TR32) 'Patterns in Soil-Vegetation-Atmosphere Systems'. The entire system was developed according to the requirements of the funding agency, the user and project requirements, as well as according to recent standards and principles. The TR32DB is basically a combination of data storage, database, and web-interface. The metadata management system was designed, realized, and implemented to describe and access all project data via accurate metadata. Since the quantity and sort of descriptive metadata depends on the kind of data, a user-friendly multi-level approach was chosen to cover these requirements. Thus, the self-developed CRC/TR32 metadata framework is designed. It is a combination of general, CRC/TR32 specific, as well as data type specific properties.

### 1. INTRODUCTION

An important task in every interdisciplinary, long-term research project is the management and archiving of research data, including sustainable storage, description with metadata, easy and secure provision/exchange, back up, and visualisation of all collected or created project data. In particular, projects that focus on environmental field studies and regional modelling in a spatial context have to consider this issue (Mückschel and Nieschulze, 2004). The overall success of such projects depends, besides the scientific research, on a well organized management of all scientific research data. Jensen et al. (2011) notice that standardised metadata are a necessary requirement for the documentation and persistent safeguarding of research data. Consequently, metadata are supporting access and reuse of data-based research results. Moreover, funding organisations like the German Research Foundation (DFG) accentuate the description of scientific data with metadata. All data should be at least described using the Dublin Core metadata standard. In addition, all primary data should be described with metadata referring to the content of the dataset to reuse them in the context of other research questions (DFG, 2009b).

Therefore, a main priority within scientific research data management should be the establishment of a corresponding metadata management system to guarantee the accurate description of all project data. Jensen et al. (2011) emphasize the significance of a management system or software to create extensive metadata documentations. These applications should support input, management, and versioning of metadata.

Furthermore, Jensen et al. (2011) point out that most applications are adapted to the needs of a special discipline. Thus, the sustainable use and search of all research results during and after the end of the project particularly depend on the implementation of a metadata management system with focus on the scientific background of the project.

In the field of environmental sciences, several research projects have already solved their problems of metadata management. These are for example: Shumilov et al. (2008), Göttlicher et al. (2010), Heimann et al. (2010), Damm and Götze (2009), Kralisch and Zander (2010), and Michener (2006).

Unfortunately, many data managers gain the experience of 'empty archives' after setting up their systems (Nelson, 2009) and notice the unwillingness of scientists to share their data (Borgman, 2010; Piwowar, 2011) including metadata. There are several approaches to solve these problems. With focus on metadata, data managers should generate tools that automatically store data and their metadata in one system (Rajabifard et al., 2009). Data and corresponding metadata should be captured at the same time, when submitted to a repository (Downs and Chen, 2010).

In the context of metadata management, some scientific research projects focus their work on Digital Object Identifiers (DOI, <http://www.doi.org>). In general, the aim of a DOI is to cite and link to electronic resources, which is widely used for scientific articles. The usage of DOIs for scientific primary data is getting more and more popular. A DOI offers a persistent and

---

\* Corresponding author.

stable reference to scientific data and is therefore an easy way to link an article with the underlying scientific data (Brase and Farquhar, 2011).

In this contribution, we present the development of a metadata management system, which is an important part of the entire scientific research data management of the Transregional Collaborative Research Centre 32 (CRC/TR32) 'Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation' funded by the DFG. First of all, we will give an overview about the research project and the corresponding project data. Then, we will focus on the development of the metadata management system, which includes an outline on the entire research data management. Finally, the paper ends with a discussion and conclusion.

## 2. PROJECT BACKGROUND AND DATA

The CRC/TR32 ([www.tr32.de](http://www.tr32.de)) is an interdisciplinary, joint research project between the German Universities of Aachen, Bonn, Cologne, and the Research Centre Jülich. The project funding started in 2007 and is expected to expire at the end of 2018. The involved research partners cover several scientific disciplines like hydrology, soil and plant science, geography, meteorology, and geophysics. They work in different project sections on exchange processes between the soil, vegetation, and the adjacent atmospheric boundary layer (SVA). The overall research goal of the CRC/TR32 is to yield improved numerical SVA models to predict CO<sub>2</sub>-, water- and energy transfer by calculating the patterns at various spatial and temporal scales. The hypothesis of the CRC/TR32 covers the explicit consideration of structures and patterns, which lead to a common methodological framework. The project participants focus their study on the catchment area of the river Rur, situated in western Germany, parts of Belgium and the Netherlands (CRC/TR32, 2011).

The overall CRC/TR32 is subdivided into different project areas, the so-called clusters. These are four scientific clusters and a central, coordination cluster. The scientific clusters differ by the sub-system (Soil-Vegetation-Atmosphere) and by the spatial scale range (local, sub-basin, regional/basin), on which they concentrate. Each cluster is split up into project sections (SP). Currently, the DFG is funding 23SPs for research. Within each SP around 1-2 PostDocs and 1-2 PhDs plus Master/Bachelor students are working on the research goal of the CRC/TR32.

As a result, the numerous project members create a lot of different data in various spatial and temporal scales. In particular, a huge amount of data is collected during several field measurement campaigns, from various sensor networks, hydrological or meteorological monitoring or airborne measurement campaigns. Results from field observations or flight campaigns are for example, wind temperature, air CO<sub>2</sub> concentration, soil water content, soil CO<sub>2</sub> concentration, leaf characteristics, root growth, land use analyses, crop surface or volume models (Hoffmeister et al., 2009; Waldhoff and Bareth, 2009; Korres et al., 2009). In addition, some research partners produce results from laboratory measurement methods or model simulations. As a consequence of the various results from field measurements and airborne campaigns, corresponding analyses and model simulations, all project participants create a great number of publications, pictures, presentations, and reports.

Within the framework of the CRC/TR32, adequate geodata of different scales were purchased to support and cover the needs of the project participants concerning their modelling and estimation purposes. For example, geodata (e.g. soil, elevation, climate, topographic, land use, or remote sensing data) was ordered from different local and national institutions like the National Survey Agency of North-Rhine Westphalia, Germany's National Meteorological Service, and the National Agency of Geology.

Format and size of the project data varies due to the different research background of the project participants. Therefore, the project participants provide data files for example in MS Excel, ASCII, NetCDF, binary format, PDF JPEG, TIF, or geodata format. In addition, the size of a single file varies from only few kilobytes to several gigabytes.

## 3. DEVELOPMENT OF A METADATA MANAGEMENT SYSTEM

The developed metadata management system is a basic component of the scientific research data management or the so-called project database of the CRC/TR32 (TR32DB). Therefore, in this section at first it is essential to describe the TR32DB and then the focus will be on the metadata management system.

### 3.1 Demands and design of the TR32DB

The entire CRC/TR32 scientific research data management (Curdt et al., 2010), which includes the metadata management system, is physically located at and implemented in cooperation with the Regional Computing Centre of the University of Cologne (RRZK). The TR32DB is accessible online ([www.tr32db.de](http://www.tr32db.de)). It is developed following specific DFG guidelines, like the 'Proposals for Safeguarding Good Scientific Practice' (DFG, 1998) and a specific bulletin concerning 'Service-projects for information management and information infrastructure in CRC – INF' (DFG, 2009a).

Moreover, the TR32DB has to consider the demands and needs of the interdisciplinary background of the project. Therefore, data from different research fields (SVA) and various spatial scales (local/point to regional/basin) have to be handled. Besides collected or created research data, research results like publications, presentation, reports, or pictures need to be observed. Additionally, a multitude of purchased geodata has to be managed.

In detail, the TR32DB design (Figure 1) comprises a file system, a database, and a web-interface with integrated web mapping application. The next section will give a short overview about the three components.

The file system is the physical storage of the CRC/TR32 project data. It is implemented by the Andrew File System (AFS), which is a distributed networked file system. The AFS was chosen in cooperation and discussion with the RRZK. Reasons for the AFS are for example security and scalability, cross platform access, and location independence. Furthermore, it is simple to archive and backup data. As a result of the cooperation with the RRZK, all project data are available during and beyond the end of the project.

The AFS is connected to a database. This is put into practice by storing the path of the datasets in a MySQL database. The main task of the database is the physically storage and management of the entire metadata, which belong to the project data. In addition, the MySQL database has to handle administrative data like user details or user rights.

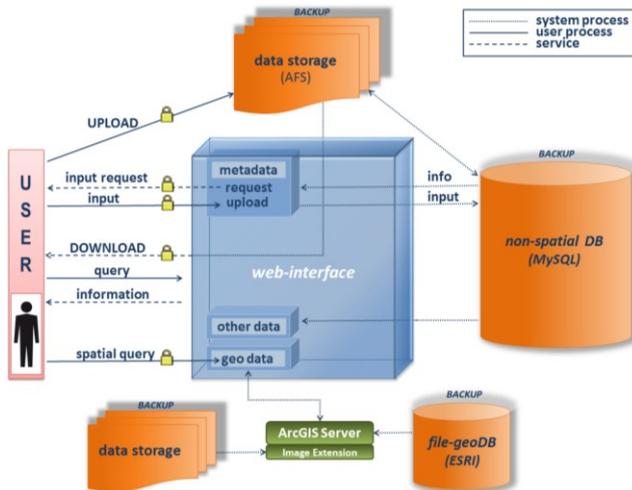


Figure 1. TR32DB design (modified after Curdt et al., 2011)

The self-developed TR32DB web-interface ([www.tr32db.de](http://www.tr32db.de)) is the connecting component between AFS and MySQL database. Management and visualisation of all project data and purchased data (e.g. climate data or geodata) via metadata are the main duty of the web-interface. In other words, the web-interface is the user-interface to the metadata management system.

The web-interface is implemented in a simple layout. The user can access general information on the top menu. The side menu enables just access to the project data, for example via a search tool or according to data topic, data type, research region/location or in order by the project structure. Besides the functions, which are connected to the metadata management system, authorized users can also use specific tools to access geodata, search for climate data, and share their project data.

By means of the implemented internal Web-GIS, authorized users can visually search and explore existing geodata and their attributes. For example, users can match various geodata together like land use and soil data of a specific region. Moreover, they can search for climate data stations in the Rur catchment and display corresponding attributes (e.g. station name, measuring period, parameters or measurement gaps).

A specific climate data tool, also implemented in the internal area of the web-interface, enables users to quick and easy access the purchased climate data (e.g. from German Weather Service or Meteomedia Group). Via a search form, users can query climate data attributes by selecting multiple stations in combination with a specific time extent. Finally, they can export the result. In addition, authorized users can generate a diagram of a climate data query by selecting multiple stations, climate data parameters, and a temporal extent.

Project members can temporary share their data via a special exchange area in the AFS. Data in the share area are available for all authorized users via the web-interface.

### 3.2 The TR32DB metadata management systems

The metadata management system was designed, realized, and implemented according to different requirements, to describe all project data with accurate metadata. First of all, the needs of the project background and participants have to be considered. It is important to arrange a simple and user-friendly metadata management system, which does not overburden the users. Furthermore, it has to cover all demands on the interdisciplinary data. All data types, collected/created by the project participants have to be included. Finally, recent metadata standards have to be noticed. Since quantity and sort of descriptive metadata depend on the kind of data, a user-friendly multi-level approach was chosen to cover all requirements.

Thus, the self-developed multi-level CRC/TR32 metadata framework is designed (Figure 2). It is a combination of general, CRC/TR32 specific, and data type specific properties.

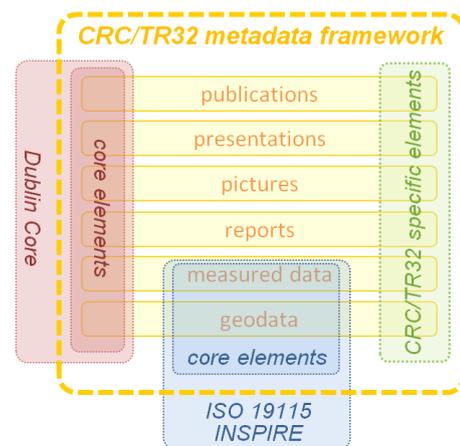


Figure 2. CRC/TR32 metadata framework

The qualified Dublin Core Metadata Element Set (DCMES, <http://www.dublincore.org/>) was chosen as a 'basic' level for the metadata model. The DCMES is a simple and widely accepted standard, which covers all general requirements and information to describe a dataset. This assures, that all types of data can be at least described with the 'basic information'. These are (mandatory elements are marked with an asterisk):

- Title: *Main title of the dataset (\*)*
- Creator: *Person who created the dataset (\*)*
- Subject: *TR32 specific topic of the dataset (\*)*
- Description: *Abstract describing the dataset (\*)*
- Publisher: *Organization responsible for making dataset available*
- Contributor: *Person and/or institution responsible for making contributions to dataset content*
- Date: *Date of publication/creation of dataset (\*)*
- Type: *Type of dataset. Recommended setting is dataset. Dataset collection is recommended for dataset series (\*)*
- Format: *Format of the dataset (\*)*
- Identifier (incl. Identifier-Type): *Identifier of the dataset, e.g. URN, DOI*
- Language: *Primary language of the dataset (\*)*
- Relation (incl. relation-type): *A reference to a related resource including type of relation*
- Coverage (temporal): *Time (period) covered by the content of the dataset, depending on the data type (\*)*
- Rights: *Access-permission of the dataset (\*)*

In addition, CRC/TR32 specific metadata properties with focus on SVA can be added to reach the specific needs and demands of the project. For example, the datasets can be described with specific keywords with focus on SVA or project special CRC/TR32 data topics (e.g. Soil, Vegetation, Atmosphere, Land Use, Remote Sensing or Topography).

Furthermore, to fulfil the requirements of the six data types (publications, presentations, pictures, reports, measured/modelled data, and geodata), specific data properties can be added. For example, a 'report' can be described with additional attributes like: the report date, the report type, the report city, the report institution, the report volume or the report pages. A 'picture' can be described with: the recorded date (begin/end), location, region, event, orientation (landscape, portrait), size, camera, and copyright information. A 'presentation' can be completed with information like: presenter, presentation date, presentation type (e.g. keynote, poster, talk, other), presentation event (e.g. conference, cross group meeting, project meeting, other), the event title, event location, and event period (begin/end).

The supplementary attributes of the data type 'publication' make an exception, because different publication types require various attributes, which has to be considered. Therefore, the additional properties of the data type 'publication' are initially: year, status (e.g. in print, in review, published, submitted, and unpublished), review (e.g. yes, no), and publication type (e.g. article, event paper, book section). Further attributes for a 'publication' are now dependent on the choice of the publication type. For example, an 'article' can be described with article type (e.g. electronic, journal, magazine, newspaper), publication source, publication source URL, publisher, volume, issue, pages, and page range.

The data types 'measured/modelled data' and 'geodata' also represent a particular case. To fulfil the requirements of spatial related datasets, demands and requirements of geographic metadata standards like ISO 19115 or the INSPIRE directive have to be considered. Therefore, both spatial related dataset can be described with attributes like: temporal extent (begin/end), lineage, reference system name and system identifier, or a geographic bounding box (north, east, south, west coordinates). Moreover, to 'measured/modelled data' properties like measurement/model region, measurement/model location, as well as the used instrument (e.g. equipment group/method, model, manufacturer, registered office, URL) and resolution distance, temporal frequency and measured/modelled parameter can be added. To enable, that the dataset can be understood without contacting the data creator, it is possible to add an extra description file in PDF format. Thus, a user can provide complete information about the content or creation of a dataset, which are not covered by the metadata elements. This are for example background information about the measurement or a detailed description of measurement parameters and units.

The so-called 'metadata on metadata' of a dataset are stored automatically. These are for example: data file storage path, metadata creator (e.g. name, institution, email, phone), metadata changing date, as well as funding phase and project section, where the dataset was created.

As already described in section 3.1, to realize the metadata management system, a MySQL database was implemented to store the entire metadata of the project. This database is

connected to the CRC/TR32 web-page to provide a user-friendly interface. Via the web-interface, authorized users are able to link their project data with adequate and specific metadata. Users, who are logged in, are requested to add metadata to a specific dataset, which they have just uploaded to the data storage (AFS). The input of the corresponding metadata is carried out by a simple structured metadata input form, which is adjusted according to the multi-level CRC/TR32 metadata framework. In the current online version (Figure 3), users at first choose the dataset, whom they want to add metadata to.

Figure 3. Metadata input form for data type 'data'

The input form will be automatically arranged according to the chosen data type. Now, the user has to enter and submit the metadata to the system. To make the input form more user-friendly, a template function was implemented. Here, the user can reuse metadata of an older dataset and modify the details for the new dataset. Hence, the users are saving time and avoiding mistakes. After submitting the metadata, the dataset details are immediately available online and searchable via the web-interface. In addition, authorized users can view and edit their uploaded metadata in their 'user home'.

Another function of the metadata management system is the application of CRC/TR32 DOIs. Users, who have already submitted their data to the TR32DB, can apply for a CRC/TR32 DOI to make the scientific dataset citable as a publication. The user has to follow and accept some conditions, for example he has to describe the dataset with specific metadata. Both, the CRC/TR32 metadata framework and the DOI system (Brase, 2004), are based on the Dublin Core metadata standard. Thus, the already connected metadata to a dataset can be reused and just have to be completed according to further DOI requirements. Consequently, the application of a CRC/TR32 DOI is designed very comfortable and user-friendly. As a result, a primary scientific dataset, stored in the TR32DB, can now be cited in a publication using the following citation:

Waldhoff, G. (2010): Land use classification of 2009 for the Rur catchment. doi: 10.1594/GFZ.TR32.1.

For example, Waldhoff (2010) is the underlying scientific dataset to describe the study in Bolten and Waldhoff (2010).

Every visitor of the web-interface is able to search for project data via metadata by the left menu according to the topic and type of data or project phases and clusters. In addition, a search tool can be used. Via the search form the user can combine various queries with each other like for example a full-text search, data type, funding phase, temporal extent, CRC/TR32

keywords, creator, and regions/sites. As result of all searches by the web-interface, a list of datasets will be displayed (Figure 4).



Figure 4. Request on type 'data' with metadata info window

By clicking on the info button, the metadata details of the data file will be shown in a pop-up window. The metadata are again arranged in a list according to the multi-level approach. First, the 'basic' metadata details are displayed, followed by the specific metadata of each data type. Every user is able to view the additional description file of the dataset, but only authorized users have permission to download the data file. By downloading the file, the user is encouraged to obey the CRC/TR32 data policy, for example concerning the use and adequate citation of the dataset.

#### 4. DISCUSSION AND CONCLUSION

The centralised management of scientific research data of an interdisciplinary research project involves several problems, due to various demands on the system. Mückschel et al. (2007) point out that all heterogeneous data should be stored in an explicit structure including the corresponding metadata. Büttner et al. (2010) complete the importance of metadata. Scientific data should be described with sufficient information to enable the reconstruction and verifiability of the dataset. As a consequence, the dataset can be used as a basis to answer further scientific questions.

We have implemented a running system (TR32DB) to manage all scientific research data of the CRC/TR32, including an accurate metadata management system. This was developed according to the funders requirements (DFG, 1998; DFG, 2009a; DFG, 2009b). In addition, all project data and metadata are captured at the same time and in one system according to Rajabifard et al. (2009) and Downs and Chen (2010).

The TR32DB was designed according to the needs and demands of the users. The system can handle huge amounts of heterogeneous data files from various disciplines in different data file sizes (kb to GB). All collected or created project data can be described with accurate metadata via the user-friendly metadata input form, which is integrated in the TR32DB web-interface. This enables users to enter metadata and edit their metadata wherever they have internet access, as well as access and download their data. For example, project participants who started their work in the second funding phase are able to access project data including accurate metadata, which was collected during the first funding phase via the TR32DB.

By implementing a multilevel metadata framework, we have considered the needs of the project participants. Every metadata provider can choose for his specific dataset, how much extra metadata he wants to enter in addition to the mandatory fields. Therefore, users are not overburden by the multiplicity of metadata input fields. Furthermore, the demands on the data types are considered, as well as recent geographic standards and principles (e.g. ISO19115 or INSPIRE directive). In particular for spatial related data, metadata are vital, because they will provide the user with information about its purpose, accuracy, quality, and actuality (Rajabifard et al., 2009).

To solve the problem of 'empty archives' described by Nelson (2009) and to create an additional value to the TR32DB, all users can apply for a persistent CRC/TR32 DOI for their data, e.g. Waldhoff (2010). This corresponds to other interdisciplinary projects like 'Scientific Drilling Database' ([www.scientificdrilling.org](http://www.scientificdrilling.org)) or 'PANGAEA' ([www.pangaea.de](http://www.pangaea.de)).

However, the TR32DB poses disadvantages concerning data connection. Although, all data files can be linked to other data files using a relation, which is a voluntary element from the DCMES, the TR32DB data are not connected like in a semantic web (e.g. Heimann et al., 2010; Willmes et al., 2012). The interdisciplinary background of the CRC/TR32, the produced heterogeneous data files, and the time-consuming development are the main reason against the semantic web approach. For future developments, the design of an ontology for selected frequent data (e.g. Eddy covariance stations) is possible, particularly with regard to Sensor Networks.

Currently, all project data are just stored in the TR32DB and are not represented in an open data repository, data library, or a data centre, yet, like PANGAEA (Data Publisher for Earth & Environmental Science) or the WDC (World Data Center for Climate). This is another task for the future work. At the moment, the metadata management system is in rearrangement. A more user-friendly wizard to enter metadata according to the CRC/TR32 metadata framework is in development, as well as an improved representation and search of all metadata on the web-interface. Currently, the TR32DB design including metadata framework is transmitted to another interdisciplinary research project with focus on resilience, collapse and reorganisation in social-ecological system of African Savannas. Due to the developed multi-level metadata approach, it was very easy to adapt the metadata management system to another discipline.

#### REFERENCES

- Bolten, A., Waldhoff, G., 2010. Error Estimation of ASTER GDEM for regional applications - Comparison to ASTER DEM and ALS elevation models. Proceedings of the Digital Earth Summit, Nessebar, Bulgaria.
- Borgman, C.L., 2010. Research Data: Who will share what, with whom, when, and why?, China-North America Library Conference, Beijing.
- Brase, J., 2004. Using Digital Library Techniques – Registration of Scientific Primary Data, *Lecture Notes in Computer Science*, pp. 488-497.

- Brase, J., Farquhar, A., 2011. Access to Research Data. *D-Lib Magazine* 17.
- Büttner, S., Hobohm, H.-C., Müller, L., 2011. Research Data Management. In: S. Büttner, H.-C. Hobohm and L. Müller (Eds.), *Handbuch Forschungsdatenmanagement*. Bock u. Herchen, Bad Honnef, pp. 13-24.
- CRC/TR32, 2011. CRC/TR32 Funding proposal, 2<sup>nd</sup> funding phase, Bonn, Germany. <http://www.tr32.de> (10 Jan. 2012).
- Curdt, C., Hoffmeister, D., Jekel, C., Brocks, S., Waldhoff, G., Bareth, G., 2011. TR32DB - Management and visualization of heterogeneous scientific data. Proceedings of the 19th International Conference on Geoinformatics, Shanghai, China. doi: 10.1109/GeoInformatics.2011.5981092.
- Curdt, C., Hoffmeister, D., Jekel, C., Udelhofen, K., Waldhoff, G., Bareth, G., 2010. Implementation of a centralized data management system for the CRC Transregio 32 'Patterns in soil-vegetation-atmosphere-systems', in: Curdt, C., Bareth, G. (Eds.), *Proceedings of the Data Management Workshop*. Kölner Geographische Arbeiten, Cologne, Germany, pp. 27-33.
- Damm, T., Götz, H.-J., 2009. Modern Geodata Management: Application of Interdisciplinary Interpretation and Visualization in Central America. *Intern. Journal of Geophysics* 2009, 13.
- DFG, 1998. Proposals for Safeguarding Good Scientific Practice. DFG, Bonn. [http://www.dfg.de/download/pdf/dfg\\_im\\_profil/reden\\_stellungnahmen/download/empfehlung\\_wiss\\_praxis\\_0198.pdf](http://www.dfg.de/download/pdf/dfg_im_profil/reden_stellungnahmen/download/empfehlung_wiss_praxis_0198.pdf) (15 Jan. 2012).
- DFG, 2009a. Merkblatt-Service-Projekte zu Informationsmanagement und Informationsinfrastruktur in Sonderforschungsbereichen INF. DFG, Bonn. [http://www.dfg.de/foerderung/programme/koordinierte\\_programme/sfb/module/modul\\_inf/index.html](http://www.dfg.de/foerderung/programme/koordinierte_programme/sfb/module/modul_inf/index.html) (15 Jan. 2012).
- DFG, 2009b. Recommendations for Secure Storage and Availability of Digital Primary Research Data. DFG, Bonn. [http://www.dfg.de/download/pdf/foerderung/programme/lis/ua\\_inf\\_empfehlungen\\_200901\\_en.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901_en.pdf) (15 Jan. 2012).
- Downs, R.R., Chen, R.S., 2010. Self-Assessment of a Long-Term Archive for Interdisciplinary Scientific Data as a Trustworthy Digital Repository. *Journal of Digital Information* 11.
- Göttlicher, D., Dobbermann, M., Nauss, T., Bendix, J., 2010. Central data services in multidisciplinary environmental research projects - the data-management of the DFG research unit 816, in: Curdt, C., Bareth, G. (Eds.), *Proceedings of the Data Management Workshop*. Kölner Geographische Arbeiten, Cologne, Germany, pp. 59-64.
- Heimann, D., Nieschulze, J., König-Ries, B., 2010. A flexible statistics web processing service - Added value for information systems for experiment data. *JIB*, 7(1): 140.
- Hoffmeister, D., Bolten, A., Curdt, C., Waldhoff, G., Bareth, G., 2009. High-resolution Crop Surface Models (CSM) and Crop Volume Models (CVM) on field level by terrestrial laser scanning. *Proceedings SPIE*, Vol. 7840, 78400E. doi:10.1117/12.872315.
- Jensen, U., Katsanidou, A., Zenk-Möltgen, W., 2011. Metadaten und Standards. In: S. Büttner, H.-C. Hobohm and L. Müller (Eds.), *Handbuch Forschungsdatenmanagement*. Bock u. Herchen, Bad Honnef, pp. 83-100.
- Korres, W., Koyama, C.N., Fiener, P., Schneider, K., 2009. Analysis of surface soil moisture patterns in agricultural landscapes using Empirical Orthogonal Functions. *Hydrol. Earth Syst. Sci.*, 14, 751-764, doi:10.5194/hess-14-751-2010.
- Kralisch, S., Zander, F., 2010. Environmental Data Management with the River Basin Information System (RBIS), in: Curdt, C., Bareth, G. (Eds.), *Proceedings of the Data Management Workshop*. Kölner Geographische Arbeiten, Cologne, Germany, pp. 83-91.
- Michener, W.K., 2006. Meta-information concepts for ecological data management. *Ecological Informatics* 1, 3-7.
- Mückschel, C. and Nieschulze, J., 2004. Editorial zum Schwerpunktthema dieser Ausgabe: Datenmanagement in interdisziplinären Umwelt-Forschungsprojekten. *Zeitschrift für Agrarinformatik*, 4, p. 68.
- Mückschel, C., Nieschulze, J., Weist, C., Sloboda, B., Köhler, W., 2007. Herausforderungen, Probleme und Lösungsansätze im Datenmanagement von Sonderforschungsbereichen. *eZAI*, 2.
- Nelson, B., 2009. Empty archives. *Nature* 461, 160-163.
- Piowar, H.A., 2011. Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data. *PLoS ONE* 6, e18657.
- Rajabifard, A., Kalantari, M., Binns, A., 2009. SDI and metadata entry and updating tools, in: van Loenen, B., Besemer, J.W.J., Zevenbergen, J.A. (Eds.), *SDI Convergence. Research, Emerging Trends, and Critical Assessment*, Netherlands Geodetic Commission, Delft, pp. 121-136.
- Shumilov, S., Rogmann, A., Laubach, J., 2008. GLOWA Volta GeoPortal: an interactive geodata repository and communication system, Digital Earth Summit on Geoinformatics: Tools for Global Change Research, Potsdam, Germany.
- Willmes, C., Brocks, S., Hoffmeister, D., Hütt, C., Kürner, D., Volland, K., Bareth, G., 2012. Facilitating integrated spatio-temporal visualization and analysis of heterogeneous archaeological and palaeoenvironmental research data. *Proceedings of the XXII Congress of the ISPRS*, Melbourne, Australia.
- Waldhoff, G., 2010. Land use classification of 2009 for the Rur catchment. doi: 10.1594/GFZ.TR32.1.
- Waldhoff, G., Bareth, G., 2009. GIS- and RS-based land use and land cover analysis - case study Rur-Watershed, Germany. *Proceedings SPIE*, Vol. 7146, 714626. doi:10.1117/12.813171.

#### Acknowledgements

We gratefully acknowledge financial support by the CRC/TR32 "Patterns in Soil-Vegetation-Atmosphere Systems: Monitoring, Modelling, and Data Assimilation" funded by the German Research Foundation (DFG).