in: Constanze Curdt and Christian Willmes (eds.) (2016):
Proceedings of the 2nd Data Management Workshop. Kölner Geographische Arbeiten, **96**. Köln/Cologne: 115-126

CRC806-DATABASE: INTEGRATING TYPO3 WITH GEONODE AND CKAN

C. Willmes, Y. Yener, A. Gilgenberg, G. Bareth
GIS & RS Working Group, Institute of Geography, University of Cologne
(c.willmes, yyener, agilgenb, g.bareth)@uni-koeln.de

Abstract

The Collaborative Research Centre 806 database (CRC806-Database, http://crc806db.uni-koeln. de) is online and operating since 2011. The architecture consists of a Typo3 based website frontend, a CKAN based metadata storage and an OGC compliant Spatial Data Infrastructure (SDI). It was decided to update the system with some major changes to the overall architecture, by preserving the current API functionality and the URLs of the datasets in the database. This paper describes the system architecture of the partly new implementation of the CRC806-Database. The SDI part of the system is migrated from the current MapServer, GeoServer, MapProxy and pyCSW based implementation to a GeoNode based system. Additionally the Typo3 based frontend of the web portal is changed to use mostly server side Extbase & Fluid based content handling and rendering, instead of the current AgularJS based frontend. Due to stability and consistency difficulties of client side rendering we decided to build a more robust system and move to server side rendering. The reasons for migrating to GeoNode for the SDI stack and away from JavaScript based client side to a server side rendering are discussed by taking into account pro and contra of both approaches, as well as a list of lessons learned from the ongoing development and operation of the CRC806-Database.

Keywords: Research Data Management, Spatial Data Infrastructure, Data publication, Data archive, Interdisciplinary research.

1. Introduction

The web based Research Data Management Infrastructure (http://crc806db.uni-koeln.de) of the Collaborative Research Centre 806 (CRC 806, www.sfb806.de), further referred to as the CRC806-Database, was undergoing some major improvements, since the release of the first version, as described in detail in (Willmes et al. 2014).

The CRC 806 is a large interdisciplinary, German Research Foundation (DFG) funded research project, at the Universities of Cologne, Aachen, and Bonn, concerning Culture-Environment Interaction and Human Mobility in the Late Quaternary (Richter et al. 2012). The CRC806-Database serves as the Research Data Management Infrastructure of the CRC 806, that implements the funders demands (Effertz 2010, DFG 2009, DFG 1998) to guarantee publication and archiving of research data produced by the project. An overview of the DFG demands is given in Curdt et al. (2012). Thus, the aim of the CRC806-Database is to implement a platform for 1.) data publication and dissemination, 2.) data archival, and 3.) data discovery, mainly for the researchers and members of the CRC 806, but also for the research communities adjacent to the knowledge and data domains of the CRC 806.

This short paper aims to describe the technical evolution and progress between version 1 and version 2 of the CRC806-Database system, and can be seen as a practical lessons learned report from a research data repository implementation and its ongoing operation and maintenance. In section 2, it is asked why we change or develop the working system further by identifying and describing problems which lead to the developments and new implementations described thereafter in section 3, where we explain what feature, interface or functionality was improved or added to the system. In section 4 we present the results of the before outlined improvements. This is followed by a discussion of the new implementations, as well as some conclusions of the redesign process and its outcome, in section 5.2.

2. Why Changing a Running System?

The CRC806-Database web application is online and running since 2010, the system (see Fig. 1) consisting of a Typo3 (Typo3 Contributors 2014) based web frontend, a CKAN (Open Knowledge Foundation 2014) based metadata catalogue, and a MapServer (Mapserver Contributors 2014), MapProxy (Tonnhofer et al. 2014) plus pyCSW (Kralidis et al. 2014) based SDI, as well as a GeoExt (GeoExt Contributors 2014) based WebGIS, was described in detail in previous publications (Willmes et al. 2012, Willmes et al. 2014). It was decided to develop the system one step further and apply some major technological changes to the system to handle some more or less urgent problems. This decision was triggered by the experiences with, and observations, as well as problems of the system during the last 3 years and from feedback of the scientists working with the system. Another obvious purpose for the here presented developments is the evolution of the technology stack applied and of course the general progress in the field of Research Data Management.

2.1 Observed Issues and Problems

In this section, we take a closer look at the main issues and problems we observed during the last four years of operating the CRC806-Database system. The issues are ranging from technical instability, accessibility, and lack of capabilities to policy issues and cumbersome maintenance or updating procedures.

Instability Issues of the Catalogue Interface: Several stability and thus reliability issues were observed using the AngularJS (Google Inc. 2014) based catalogue frontend. One problem was the rendering of cryptic appearing AngularJS markup (see Fig. 2), if the response from the catalogue had delays or other issues. This is a very ugly experience in particular for the not-tech-savvy end users who most probably won't understand the cause of these errors, that is primarily network issues causing the application issues, and in this case in particular client side rendering issues.

Lack of DOM Markup Capabilities: AngularJS renders the page content into several Document Object Model (DOM) objects in a client-side procedure after

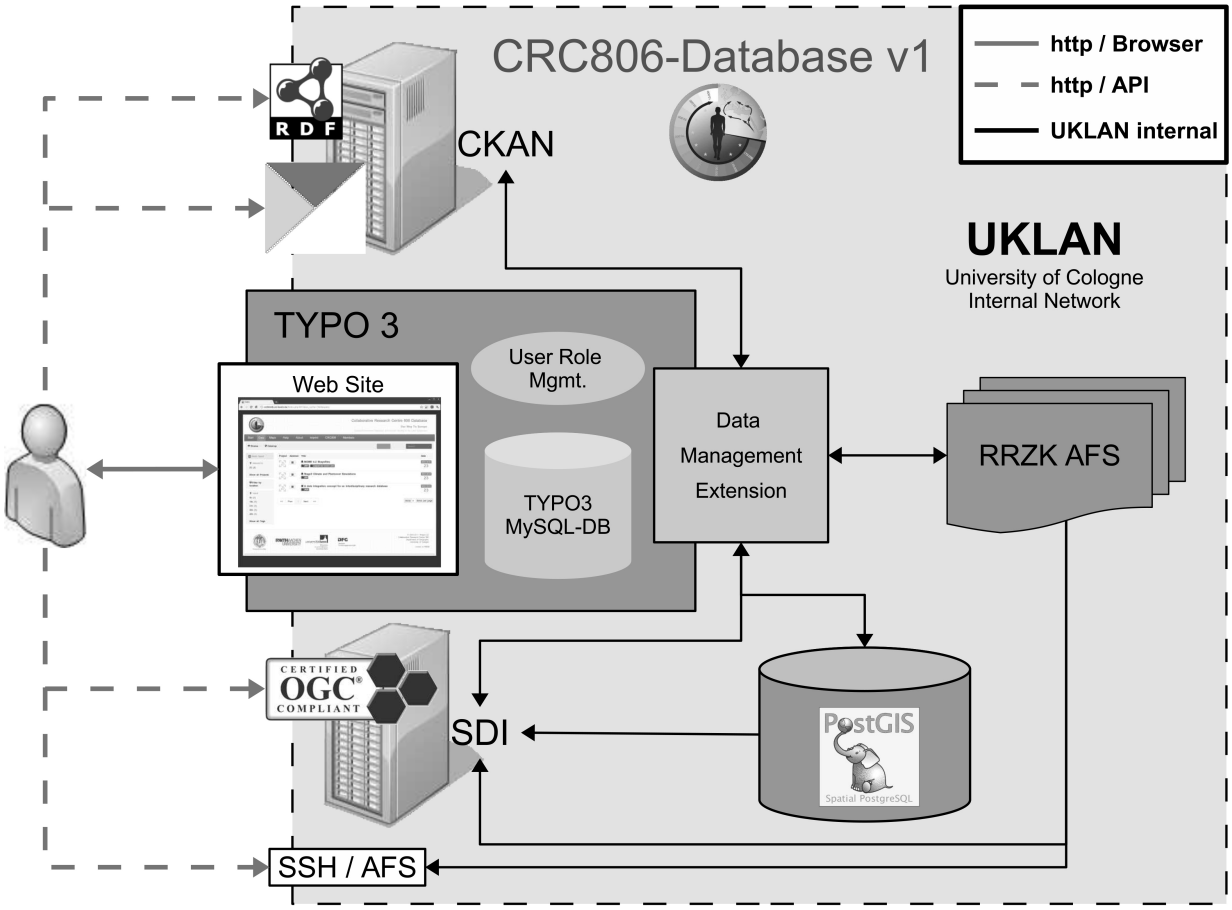

Fig. 1: Architecture of the first CRC806-Database version (WILLMES et al. 2014)

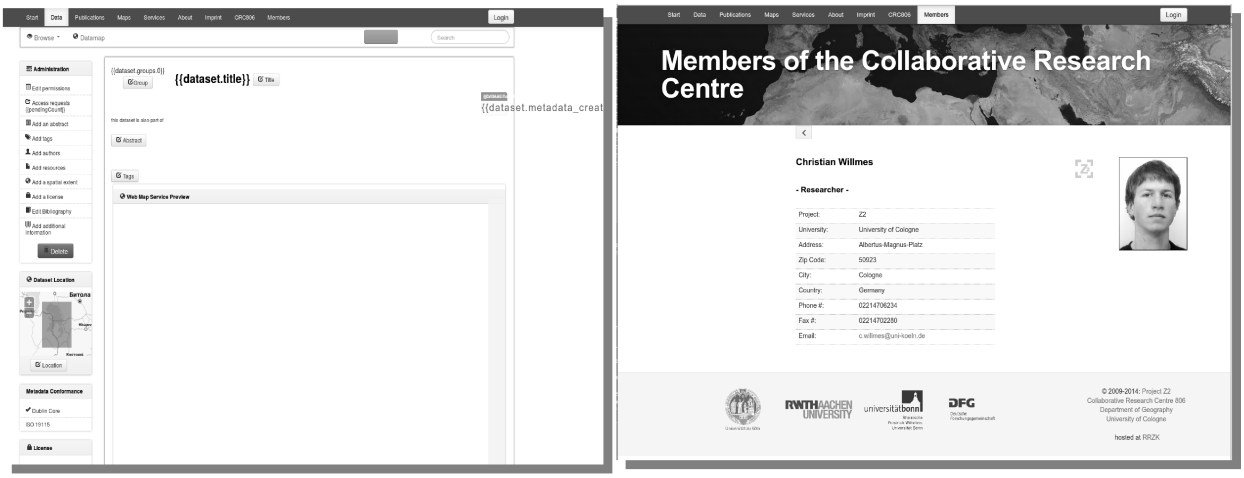

Fig. 2: Example screen shots of rendering instabilities. On the left, error and security implications of AngularJS frontend rendering issues. On the right, problems of the members directory, showing random user profiles

the HTML document itself, including a static not modifiable DOM, is served to the user. This issue is problematic in two aspects; first, the possibility to do Search Engine Optimization, which heavily builds on markup of content using XHTML tags, is drastically limited. Secondly, it is not possible to implement RDFa (Adida et al. 2008, Brickley et al. 2004, Carroll et al. 2004) markup in the dataset landing pages, to annotate the content and its metadata. The problem of limited Search Engine Optimization (SEO) capabilities was not that pressing as the problems of having almost no capability of modifying the DOM markup of the dataset web pages.

Accessibility Issues for Several Client Configurations: It was also observed, that for some configurations of browser and operating system, as well of course, if the user allows the execution of JS in his browser were problematic for accessing the catalogue as well as the WebGIS interface. Some AngularJS directives were differently interpreted by Chrome (Google Inc. 2014b), Safari (Apple Inc. 2014), and FireFox (Mozilla Found. 2014) web browsers, and apparently by Internet Explorer internetexplorer. Although, the incompatibilities of IE with CRC806-Database web application are reduced significantly, by the adaptation of mainly server side and less client side interaction handling and rendering.

Updating and Maintaining the WebGIS and SDI: The initial setup of the SDI was a custom configured MapServer and a 'from scratch' written GeoExt based WebGIS frontend, as described in (Willmes et al. 2012). This setup proved to be quite cumbersome and laborious to maintain. For example, if a new spatial dataset had to be exposed, the web service had to be implemented for MapServer, which involves advanced skills on several ends (MapServer configuration, GIS data handling, and thus advanced knowledge of GIS and WebGIS concepts, and a learning curve to work with the according software). To add a new web service to the WebGIS, three JavaScript files containing several hundred lines of code had to be maintained and adjusted accordingly. This setup caused a relatively high cost for adding new datasets to the SDI and WebGIS, because it is a highly specialized task, which resulted in almost not updating the SDI because of its high costs.

Buggy Members Directory: The previous, Typo3 frontend_users extension based members directory implementation had some serious bugs also. There were some caching problems, that resulted in showing random user profile detail pages on the main members directory overview page. This could only be resolved by manually clearing the Typo3 cache. Another shortcoming of the previous members directory was its limited capability of searching and filtering for members. It was more or less a hierarchical list ordered by the cluster and project structure of the CRC 806. But the main problem was the maintenance of the profile by the members. This was not possible in the first version. If some information for members needed to be changed, only Typo3 administrators of CRC806-Database team could do this task.

Issues of the News and Blog System: The Typo3 extension tx_news was used for handling and displaying the small blog entries and news posts on the frontpage of the CRC806-Database web site. The tx_news extensions had to be configured in kind of complex TypoScript notation. The interface for editing the news and blog entries lacked the possibility of free text or HTML editing. The text editor for composing posts, reformatted the entries on HTML level, which resulted in problems for embedding YouToube videos for example.

2.2 New Functionality, Features and Interfaces

Additional to solving prevalent problems of the system, the overhaul of the CRC806-Database also includes the implementation of several new functionality and feature requests from the project partners and from the Data Management Project itself, which were not planned from the original project proposal and were designed and implemented additionally.

Publication Database: In order to have a central 'point of entry', to avoid redundant entry of publication records for researchers and projects within the CRC 806, a publication database including interfaces to the three websites:

- CRC 806 Main Website, http://sfb806.de,
- CRC 806 IRTG, http://sfb806irtg.uni-koeln.de/,
- CRC806-Database, http://crc806db.uni-koeln.de,

of the CRC 806 was requested for implementation by the project partners, and because this appraoch was a success in the comparable TR32DB (http://tr32db.uni-koeln.de, Curdt et al. 2012) implementation for the CRC / Transregio 32 web presence (http://tr32.de).

Integration of Data Catalogue and SDI: Some published and / or archived datasets can contain spatial resources in form of geodata. We found, that it would be a good idea, to have these datasets in the CKAN based data repository linked with resources in the GeoNode based SDI and vice versa. This integration is facilitated by implementing a Typo3 Extbase (Lobacher 2014) based model (see section 3.2).

Improvement of Spatio-Temporal and Thematic Filters: One particular shortcoming of the previous data catalogue frontend was the temporal filter. There was none. The only possibility to discover data by its temporal location, was to tag them with a name of a time interval like Last Glacial Maximum or Mid-Holocene for example. The idea is to provide a thesaurus like list of pre-defined temporal intervals or events and place names, that translate into coordinates or bounding boxes. Additionally it is possible to define a time interval by its upper and lower temporal bound in yBP or an event by a date in yBP. In an comparable way to the implementation of the spatial metadata for bounding boxes or coordinates.

Automated Issuing of Digital Object Identifiers: Since 2012, the CRC806-Database has the capability to issue Digital Object Identifiers (DOI's) ISO2012 in two CRC 806 name spaces (10.5880/SFB806 and 10.5880/CRC806). Until now the issuing process is executed manually after a not formalized request from CRC806-Database dataset authors to the Z2 data management project. In the new version it will be possible to request DOI's through the web application, for datasets providing the minimum required metadata fields for a DOI.

3. Improving and Fixing the System!

For the improved version of the CRC806-Database, we decided to move from the RRZK maintained Typo3 instances to a self maintained, but RRZK web hosted Typo3 instance. With this change we decided also to switch from the previous Typo3 v4.5 to Typo3 v6. This has several advantages. The first advantage is that we are no more dependent on the RRZK for updating the Typo3 core. This also means, that we can choose to port to new releases or not, if compatibility between versions is broken, like it was between version v4.5 and v6. This is a crucial point for the long term availability of the system, that was not possible in the former RRZK hosted Typo3, where updates needed to be maintained regularly.

Typo3 v6 brings some very useful technical features, that were not available in v4.5. The Extension framework Extbase & Fluid (Rau et al. 2013) is now natively supported by v6, in v4.5 an extension was needed, to be able to run Extbase & Fluid based extensions. Extbase & Fluid extensions are developed in the Domain Driven Design (DDD) (Evans 2004) and Model View Controller (MVC) (Lindberg et al. 2002) pattern, which allows to implement complex tasks in a clean and relatively fast process. The downside is the relatively steep learning curve for programming with Typo3 and Extbase & Fluid applying the DDD and MVC patterns.

3.1 Data Catalogue Interface

An obvious improvement that needed to be implemented, was the switch from the AngularJS based to a server side rendered catalogue interface. The complete AngularJS based MVC model interfacing the CKAN action API, was redeveloped in Typo3 Extbase & Fluid technology. This has the advantage, that all DOM rendering takes place on the server side. JavaScript is only used for Styling and interactivity in web forms for processing and validating the entries, or in the Maps interface for displaying the interactive maps.

Another improvement, that the switch to Extbase & Fluid creates, is the possibility of a more streamlined development process of new features, because Extbase & Fluid can leverage also sophisticated persistence of the domain data, that is not stored in CKAN or GeoNode. This enables to develop relations between CKAN resources and GeoNode resources outside of both systems.

We developed two new main features for the new version of the catalogue. The first new feature is the BibTeX (Patashnik 1988) format export functionality. In the previous AngularJS based version of the catalogue, there was a field with BibTeX markup on the bottom end of each dataset page containing bibliographic information, in the new frontend, we provide a link to download the structured citation information of a dataset as a .bib BibTeX file, and further formats are possible to implement in the future.

The second new feature, a temporal filter, is a major improvement to the data catalogue. It is now possible to add temporal metadata, in the form of predefined events (points in time), or by predefined intervals (defined by a start and an end time), to a dataset and filter the datasets by defining temporal filter boundaries.

Also, the spatial filter was extended to filter by predefined regions and locations. In the near future the connection between the spatial filter of the GeoNode Maps and the CKAN resources will be integrated in the catalog interface.

3.2 GeoNode Based SDI and Extbase & Fluid maps Frontend

Similar to the CKAN interface, a frontend to interface GeoNode from within the CRC806-Database website, was developed. GeoNode (GeoNode Contributors 2014) is an integrated SDI geo-content publishing system that integrates a complete SDI stack based on GeoServer (GeoServer Contributors 2014), pyCSW (Kralidis et al. 2014), PostGIS (Ramsey et al. 2014), and GeoDjango (GeoDjango Contributors 2014).

The GeoNode functionality is integrated into the Typo3 frontend interfacing the REST API of GeoNode, available since the current version 2.1 (as of December 2014) of GeoNode. For the first release of the new maps frontend, the spatial datasets can be annotated with ISO 19115 metadata, additionally they can be categorized into predefined categories and tagged. The frontend has a full-text search, filters for the tags and categories, and also filters for the spatial data representation type (vector or raster).

A temporal, similar to and integrated with the catalogue filter will be developed in the near future, but is not included in this first release.

3.3 Publication Database

The publication database also uses the CKAN instance as the data backend, similar to the Data Catalogue in-

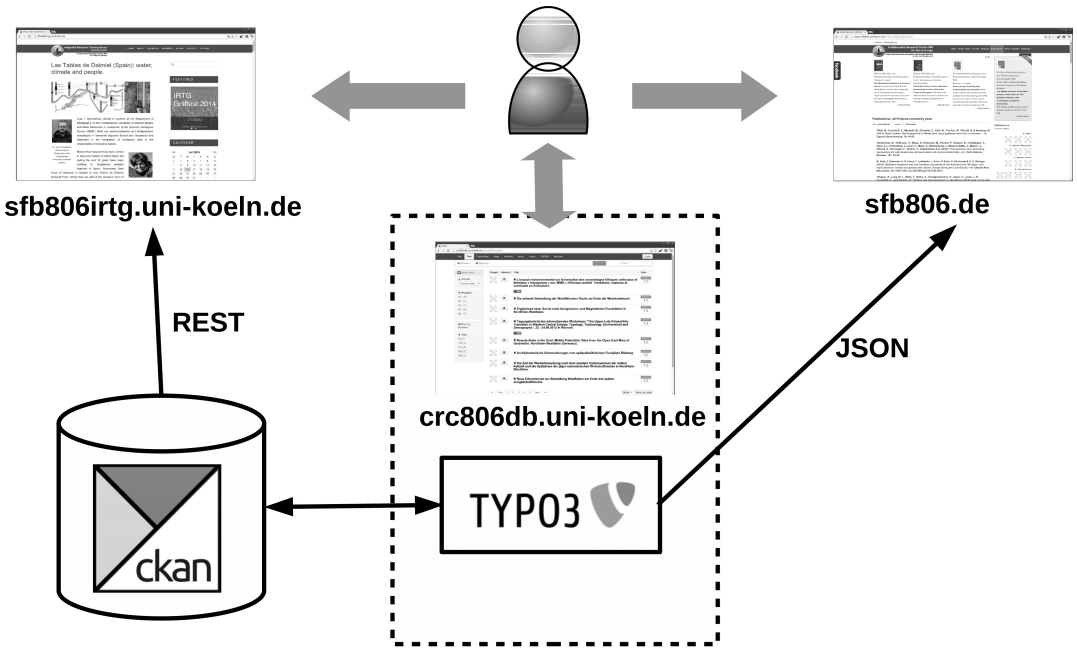

Fig. 3: CKAN based bibliographic database

terface. The publication records are stored as CKAN datasets, this datasets can have resources like links to external publisher websites or for example the PDF file, stored in the CRC806-Database backend, of the article and any other supplementary data. The publications can be related to datasets and geo-datasets of the CRC806-Database, this is implemented in the Extbase domain model of the data catalogue. It is also possible to tag publication records with keywords to allow filtering among them in the database. The augmentation with spatial and temporal metadata is possible as well, and allows spatio-temporal filtering of literature records.

For the main SFB / CRC 806 website (http://sfb806.de) and for the website of the Integrated Research Training Group (IRTG, http://sfb806irtg.uni-koeln.de) plug-ins to interface the CKAN action API were developed for the according Content Management Systems (CMS). For the Wordpress (Wordpress Contributors 2014) CMS based IRTG website, an according wordpress plugin interfacing the CKAN action API was developed. It is possible to display bibliographies by author or by SFB/CRC 806 cluster or project. The plug-in is mainly used on the IRTG alumni profile pages. Similarly, for the Joomla CMS (Joomla Contributors 2014) based SFB / CRC 806 main website, a plug-in to interface the CKAN based publication data base was implemented. This plug-in has some more capabilities compared to the Wordpress plug-in, because it also allows to show a list of all publication records in the data base additionally to display bibliographies by cluster, project and authors.

3.4 Members Directory

The new members directory is also an in-house developed Extbase & Fluid extension, building upon, and extending, the previously for the former members directory used fe_users (front-end users) extension of Typo3. This means, that the existing user data base can be re-used from the new implementation. And user data, registration and profile data base and functionality is interfaced from the fe_users extension. The frontend renderings, of the user directory and profile pages, as well as the filters of the user list and the according dataset lists on the user pages are implemented new from scratch using Extbase & Fluid technology. This approach was chosen, to fix the observed rendering problems mentioned in section 2.1, but it also enhances the web site with its additional functionality.

3.5 News Stream and Blog Post Extension

The new extension for news and blog post on the frontpage including the latest data, geodata and publication streams are also implemented in a custom developed Extbase & Fluid extension. This new extension solves the before mentioned main problems, that we observed using the Typo3 tx_news extension, which we used previously. The new extension is mainly fixing the problems with entering HTML markup into the posts, as described in section 2.1. The news streams for the latest data sets, publications and geo-data sets, the functionalites for accessing the according backends could be reused from the data catalogue, maps and publications implementations.

4. Results

The resulting user interfaces, of the previously in section 3 described implementations, are presented in the following of this section, after an overview of the new system architecture and its changes compared to the first version.

4.1 Overview

The system architecture of the CRC806-Database version 2 (see Fig. 4) changed in the Typo3 web application scope and in the SDI architecture, compared to the first version (see Fig. 1). Completely new is the publication database, which builds upon an existing part of the architecture, the CKAN backend.

As can be seen on the diagram (Fig. 4), the new glue that binds all parts of the CRC806-Database system together is the Typo3 Extbase & Fluid application framework. All five sub-models (News, data catalogue, publication database, the maps SDI interface and the members directory) of the web application, are now based on this technology, and integrate the four backends (CKAN, AFS, GeoNode and the Typo3 MySQL database) into the system. This redesign enables several positive synergy effects for further developments of the system. It is now possible to reuse functionalities, like filters, interface layouts, etc., build for one sub-model in an other sub-model. One example for such a reuse are the query and access functionalities implemented originally for the data catalogue CKAN backend, that could now be reused from the bibliography database interface to consume the publication data stored in the CKAN backend.

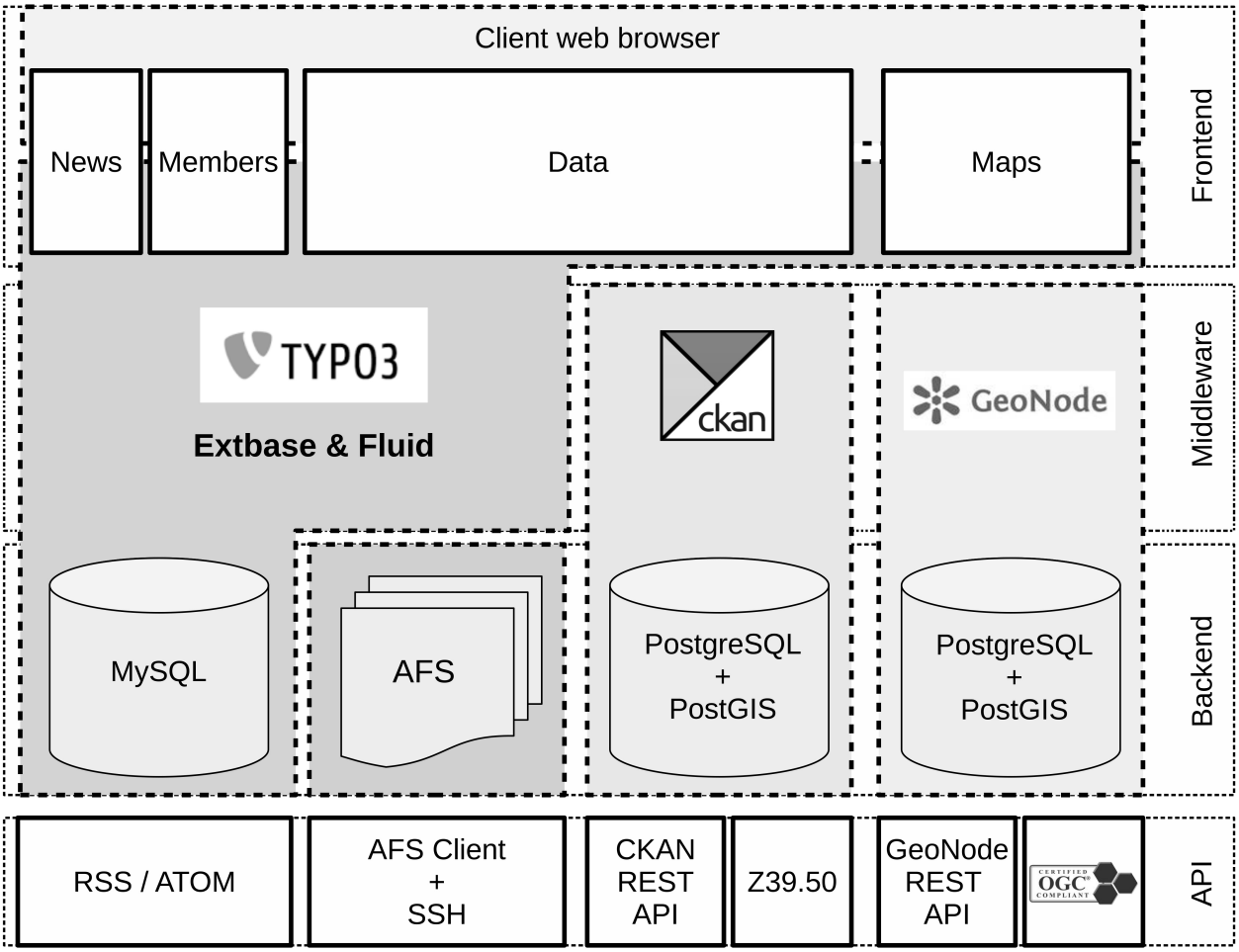

Fig. 4: Layered architecture diagram of the new CRC806-Database version

4.2 Data Catalogue

In Figure 5 the interface of the new Extbase & Fluid based data catalogue, including the in section 3.1 described new features is shown. The interface is divided into two rows. On the left side, the result set of data sets from the current query is displayed as a list. On the right side, the different filter, sorting and search interfaces are placed. This main design is the same as the design of the first version, but the filters are now on the right instead of the previously left side.

The new dataset list (see left screenshot in Fig. 5) has more information about each dataset displayed in the list view. Beside dataset type, title, creation date, CRC 806 project(s) and data format(s) of appended resources, the first 200 characters of the description or abstract and indicators for spatial and temporal metadata, as well as the dataset maintainer are displayed. The in section 3 described new filter interfaces for temporal data are included on the right side under the spatial filter (filter by location). An overview map of all spatially annotated data sets is available through a button on the top of the data set list view.

The dataset detail page (see right screenshot in Fig. 5) was also restructured. The main improvement is the technological switch to more stable server side rendering and avoiding the rendering issues of the first version (see Fig. 2 in section 2.1). Further advantages are features like the BibTeX export in a .bib file, to support the export of bibliographic information into most bibliographic data base applications. And the authors, who are members of the CRC 806 appear now with links to their data base profile pages, to access and review other contributions of an author to the data base.

4.3 Maps SDI Frontend

The new GeoNode based Maps interface is shown in Figure 6. For this first version of the GeoNode based Maps interface, a list view of the geo-datasets including several filters and a full-text search was implemented.

The filters are: 1.) a spatial filter, provided by a small map interface on the right side of the list; 2.) a filter for GIS data type (vector or raster); 3.) a filter for the main category of the data; and 4.) a filter for the key-

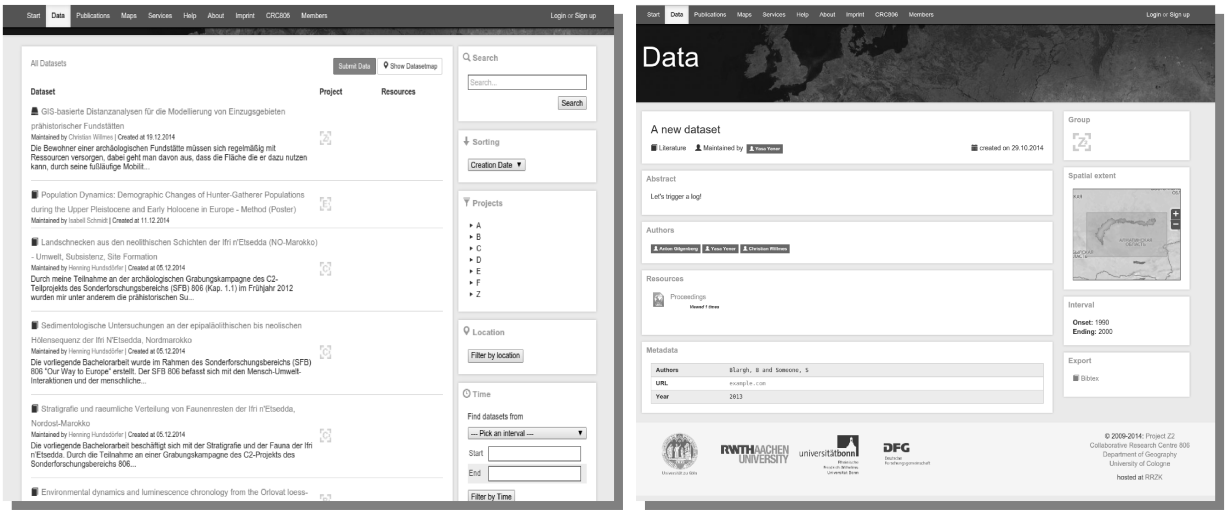

Fig. 5: Screenshots of the Data Catalogue interface, list view (left) and dataset page (right)

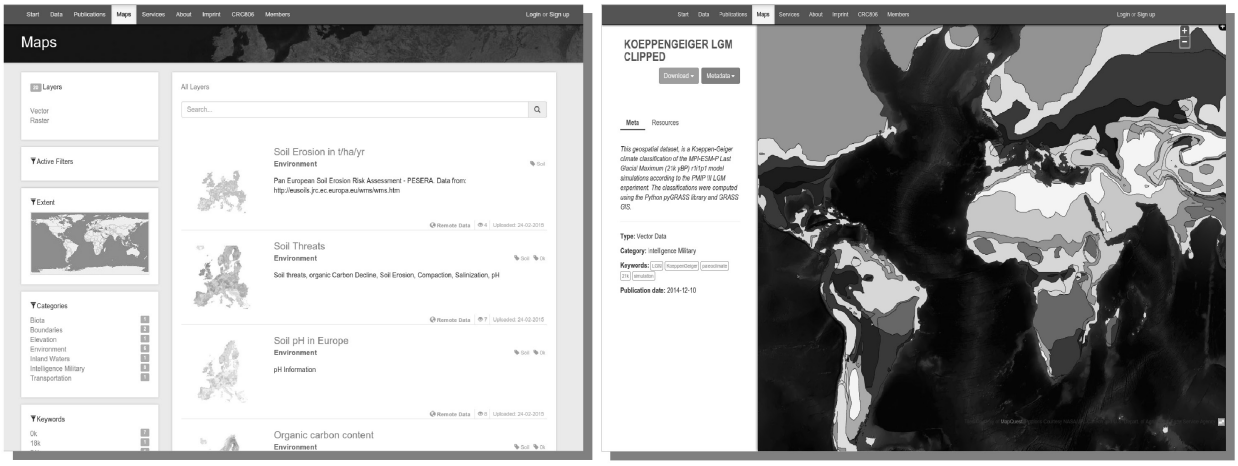

Fig. 6: Geospatial dataset list view (left) and detail view (right)

words or tags, the datasets are annotated with. A small box (named Äctive Filters) on the left indicates the user which filters and searches are currently applied to the geo-data directory, allowing the user to disable previously selected filters.

The detail view of a dataset, shows an interactive map with the currently selected dataset overlayed on a world map. On the bottom of the map, the detailed information or metadata view is shown. By hovering with the mouse pointer over the title of the dataset, the detail view can be expanded and offers the user to review the metadata information, as well as the download of the data in several well known GIS formats, such as ESRI Shapefile, GML, GeoTiff, to name a few options. The download of the metadata information in standard notations, such as ISO19115-1 (2014), Dublin Core Metadata Initiative (2014), ebRIM (2004), FGDC (1998) and more, is also offered through this interface.

Additionally it is possible to access the data via OGC web services, such as WMS, WFS or WCS for example using a desktop GIS over the network. The links to the GeoNode service endpoints are also provided on the detailed information view.

4.4 Bibliographic Database

The interface of the bibliographic database is a list similar to the data catalogue list, but only showing publication datasets, formatted according to uniform citation style. In Figure 3, some of the interfaces to the central CKAN based bibliographic data base are displayed.

The entry and editing of bibliographic information is facilitated through the data catalogue entry and editing interface. A bibliographic record is handled like any other resource in the CKAN backend, it has additionally to a free amount of resources and any kind of meta

information (including tags, temporal and spatial annotations) markup provided through the interface, also the possibility to assign a bibliographic publication type, defined by the BibTeX (Patashnik 1988) bibliographic information schema. According to the publication type, several informations like, Authors, Title, Year, etc. have to be provided. If all demanded informations are provided, the resource is included into the publication data base, and available through its interfaces, described in section 3.

4.5 Members Directory

The new members directory and its user interface (see Fig. 7) has also improvements compared to the previous implementation.

The main advantages are the additional filter capabilities of the members list and the enhanced user profile pages. In the new version, all datasets a user is related to (owner or author) is shown on his profile page. This information is also provided to the main CRC 806 website (www.sfb806.de) through interfacing the CKAN backend API from the Joomla system of the main website.

Another improvement is the more clear layout of the interface, compared to the previous implementation. The new layout is oriented along the look and feel of the data catalogue and the other interfaces (Maps, Publications), which should be more intuitive, compared to the old version, that was implemented in a jQuery accordion widget, where the user had to click on the tab of a cluster to see the projects and its according data base members. No further filtering was available in that implementation. The new implemented additional filter capabilities are a full-text search, a sorting by name or project and a filter by cluster and projects.

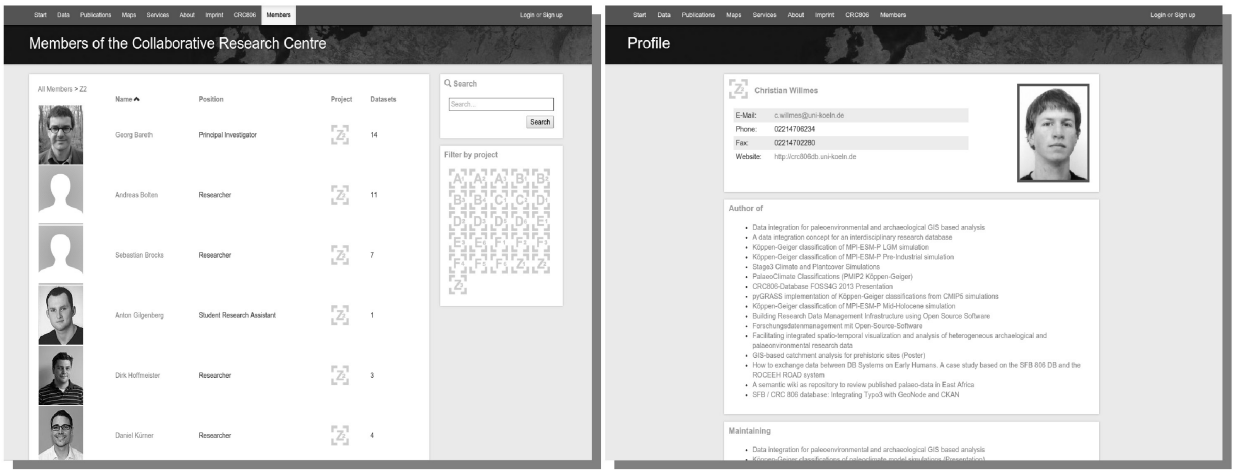

Fig. 7: The new members directory interface. Members list with filters (left) and Profile view with user details and datasets (right)

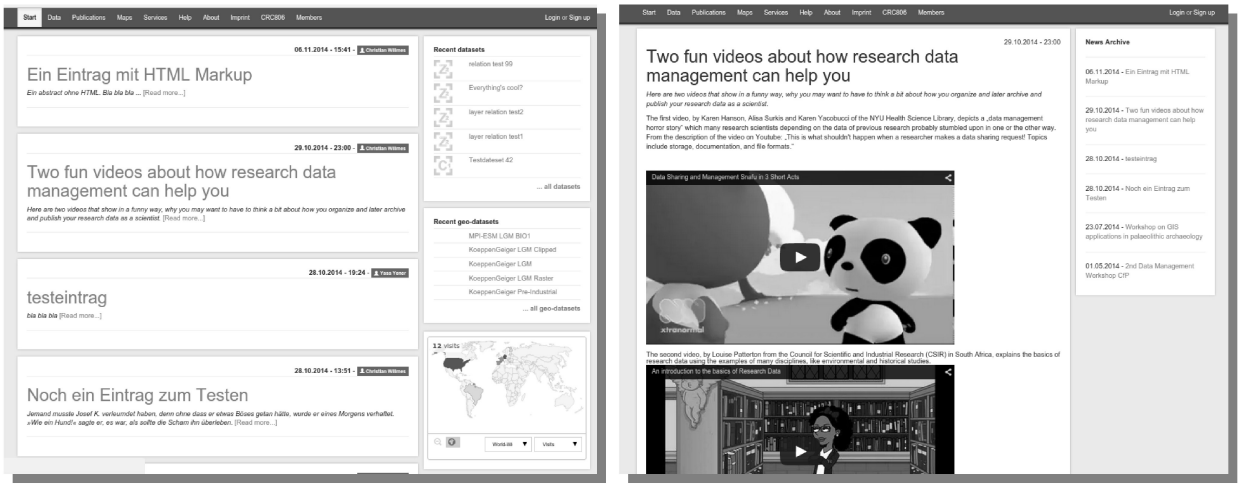

Fig. 8: The new frontpage news stream and blog posts interface

4.6 News and Blogs

As described above the news and blog interface, placed on the frontpage of the CRC806-Database website, was completely re-implemented in Extbase & Fluid technology. The most notable improvements compared to the previous tx_news based implementation, are the additional geo-data stream, the new web statistics user interface and the alignment of the layout according to the overall look and feel of the website.

A new feature of the implementation is, that it now provides RSS (2002) feeds of the CRC806-Database news, latest datasets, latest geo-datasets and bibliography records.

Another new feature of the interface is the PIWIK piwik2014 based visitor map widget, that replaces the old third-party hosted visitor map.

4.7 Web Site and Data Metrics

The redesign of the website enabled also to implement a comprehensive web site metrics application for access, usage and download tracking. The website statistics itself (access logs) are tracked using the open source PIWIK (Piwik Contributors 2014) system, this system is extended by an additional detailed tracking of data access and downloads. In the future we aim to provide these collected metrics to the database users and to the CRC 806 for internal evaluation and feedback in general. The web access statistics tracked by PIWIK are already shown to the users from the PIWIK map widget placed under the news streams on the CRC806-Database frontpage.

5. Discussion and Conclusions

Because of the disadvantages, shortcomings, bugs and problems of the previous version of the CRC806-Database, as described in section 2, it was decided to conduct a major redevelopment of the CRC806-Database web portal. This redevelopment was described in detail in the sections 3 and 4 of this paper. However, this major redesign did not affect the underlying CKAN based data catalogue and the AFS based long-term storage for the data files. The data archival and preservation aspects of the CRC806-Database were not affected by the changes to the overall system and remain as described in Willmes et al. (2014). Furthermore, it was taken care of preserving all URLs of dataset landing pages. This guarantees, that all links from other websites or even from already existing publications to content and especially data set landing pages remain unaffected. What actually did change, are the URLs of the OGC (WMS, WFS, WCS, etc.) SDI services. It is planned to publish a lot more geo-data resources via SDI services in the future, so that this particular change may be worth this backward compatibility break, that could not be avoided due to the transition from a custom SDI stack to a GeoNode implementation.

The approach to replace all functionality implemented in client side AngularJS technology with Extbase & Fluid Typo3 v6 core server side technology, results in more robust and flexible capabilities at the same time. The integration is tighter, because we could increase the maintainability and resilience of the system by avoiding a complete layer of complexity and possible failure, by building all MVC functionality directly using the Typo3 technology of Extbase & Fluid. The switch to Extbase & Fluid technology also provides better development capabilities for further maintainability of the system and its development. A rich set of features which can be taken from the framework that are not implemented yet and can be added in the future.

But not only the technological and user interface side of the CRC806-Database was improved, there were also some main new features developed. The feature, that provides possibly the most useful is the related feature (as described in section 3), that integrates the data catalogue, the publication database and the maps interface, by listing related entries for the resources in the according directories. Another useful new feature that needs to be mentioned, is the temporal filter. Although it is for this first release only implemented in the data catalogue, it provides a useful query interface to the data basis. In the future, this feature will be extended to the maps interface also.

The switch from the Computing Centre supported Typo3 instance, to aself supported and maintained instance is an important step to guarantee long term availability of the CRC806-Database, after funding for the data management project of the CRC 806 is terminated, which is the case at the latest in summer 2021, if a third phase of the CRC will be funded. The DFG demands in its funding guidelines, that data produced by CRC members have to be accessible at least 10 years after the funding terminated. And to meet this requirement, a system that can stay working without maintenance on the application side, must be implemented.

Summarizing the main lessons learned from this architecture improvements, are for one the User Interface

maintainability and its stability (rendering behaviour) considerations for choosing the right technology. For two, it is crucial to break as less as possible (functionality, URLs, known interfaces, etc.) to preserve as much backward compatibility as possible. And third, the most important lesson, is to only implement changes in the backend, if they are not avoidable apart from very good reason. In conclusion, we can say that the effort of redesigning the web portal on the basis of Extbase & Fluid technology was worth the effort.

References

Adida, B., Birbeck, M., McCarron, S., Pemberton S. (2008): RDFa in XHTML: Syntax and Processing. W3C. http://www.w3.org/TR/2008/REC-rdfa-syntax-20081014/.

Apple Inc. (2014): Safari – The best way to see the sites. (2014-11-06). Apple Inc. https://www.apple. com/safari/.

Brickley D., Guha, R. (2004): RDF Vocabulary Description Language 1.0: RDF Schema. W3C. http://www.w3.org/TR/2004/REC-rdf-schema-20040210/.

Carroll, J. J., Klyne, G., (2004): Resource Description Framework (RDF): Concepts and Abstract Syntax. Tech. rep. W3C. http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.

Curdt, C., Hoffmeister, D., Waldhoff, G., Jekel, C. & Bareth, G. (2012): Scientific Research Data Management for Soil-Vegetation-Atmosphere Data - The TR32DB. International Journal of Digital Curation, 7 (2), 68– 80.

DFG (1998): Proposals for Safeguarding Good Scientific Practice - Recommendations of the Commission on Professional Self Regulation in Science. Tech. rep. (http://doiop.com/empfehlung_wiss_praxis_0198.pdf). Deutsche Forschungsgemeinschaft, Weinheim, Germany.

DFG (2009): Recommendations for Secure Storage and Availability of Digital Primary Research Data. Committee on Scientific Library Services and Information Systems - Subcommittee on Information Management. Deutsche Forschungsgemeinschaft, 53170 Bonn,Wissenschaftliche Literaturversorgungs- und Informationssysteme (LIS), http:/ /www.dfg.de / download/pdf/ foerderung/programme/ lis/ua_inf_empfehlunge%20n_200901_en.pdf.

Dublin Core Metadata Initiative (2004): Dublin core metadata element set, version 1.1: Reference description. http://dublincore. org/documents/dces/.

ebRIM (2004): ISO/TS 15000-3:2004. International Organization for Standardization. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=39974.

Effertz, E. (2010): The Funder's Perspective: Data Management in coordinated programmes of the German Research Foundation (DFG) - In: Curdt C. & Bareth, G. (eds.): Proceedings of the Data Management Workshop, 29. 30.10.2010, University of Cologne, Germany. Kölner Geographische Arbeiten, 90: 35-38, Köln. doi: 10.5880/TR32DB.KGA90.7.

Evans, E. (2004): Domain-driven Design: Tackling Complexity in the Heart of Software. Addison-Wesley.

FGDC (1998): Content Standard for Digital Geospatial Metadata. Federal Geographic Data Committee. http://www.fgdc.gov/metadata/csdgm/.

Geodjango Contributors (2014): GeoDjango – A world-class geographic web framework. Django Software Foundation. http://geodjango.org. 2014-12-18.

Geoext Contributors (2014): GeoExt – JavaScript Toolkit for Rich Web Mapping Applications. GeoExt Community. http://geoext.org. 2014-12-18.

Goenode Contributors (2014): GeoNode–Open Source Geospatial Content Management System. http://geonode.org. 2014-12-18.

Geoserver Contributors (2014): GeoServer – open source server for sharing geospatial data. Open Source Geospatial Foundation. http://geoserver.org. 2014-12-18.

Google Inc. (2014a): AngularJS – Superheroic JavaScript MVW Framework. Google Inc. https://angularjs.org. 2014-11-06.

Google Inc. (2014b): Chrome – Fast, free web browser. Google Inc. https://www.google.co.uk/intl/en/chrome/browser. 2014-11-06.

ISO19115-1 (2014): Geographic information – Metadata – Part 1: Fundamentals. International Organization for Standardization. http://www.iso.org/iso/iso_catalogue/catalogue_tc/ catalogue_detail.htm?csnumber=53798.

ISO26324 (2012): Information and documentation — Digital object identifier system. International Organization for Standardization. https://www.iso.org/obp/ui/#iso:std:43506:en.

Joomla Contributors (2014): Joomla – The Platform Millions of Websites are built on. OpenSource Matters, Inc. http://www.joomla.org. 2014-12-18.

Kralidis, T., Tzotsos, A. (2014): pyCSW–Metadata Publishing Just Got Easier. http://pycsw.org.

Lindberg, H., Rydin, P. (2002): Model view controller. (US Patent App. 09/768,389). http://www.google.com/patents/US20020143800.

Lobacher, P. (2014): Typo3 Extbase - Moderne Extension-Entwicklung für Tyop3 CMS mit Extbase & Fluid. Open Source Press.

Mapserver Contributors (2014): MapServer – Open source web mapping. Open Source Geospatial Foundation. http://mapserver.org. 2014-12-18.

Microsoft Inc. (2014): InternetExplorer – Rethink the web. Microsoft Inc. http://windows. microsoft.com/en-gb/internet-explorer. 2014-12-18.

Mozilla Found. (2014): FireFox – Committed to you, your privacy and an open Web. Mozilla Foundation. https://www.mozilla.org/en-US/firefox. 2014-12-18

Open Knowledge Foundation (2014): CKAN – The open source data portal software. CKAN Association. http://ckan.org. 2014-10-30.

Patashnik, O. (1988): Designing BibTeX styles.

Piwik Contributors (2014): PIWIK - Open Analytics Platform. http://piwik.org. 2014-12-18.

Ramsey, P., Santilli S., Obe, R., Cave-Ayland, M. & Park, B. (2014): PostGIS – Spatial and Geographic objects for PostgreSQL. Open Source Geospatial Foundation (OSGeo). http://postgis.net. 2014-12-18.

Rau, J., Kurfürst, S. & Helmich, M. (2013): Zukunftssichere TYPO3-Extensions mit Extbase und Fluid. O'Reilly Verlag.

Richter, J., Melles, M. & Schäbitz, F. (2012): Temporal and spatial corridors of Homo sapiens sapiens population dynamics during the Late Pleistocene and early Holocene. Quaternary International, 274, 1–4. ISSN : 1040- 6182.

RSS (2002): Really Simple Syndication (RSS) 2.0. W3C. http://validator.w3.org/feed/docs/ rss2.html.

Tonnhofer, O. & Helle, D. (2014): MapProxy – Open source proxy for geospatial data. Omniscale GmbH & Co. KG. http://mapproxy.org. 2014-12-18.

Typo3 Contributors (2014): Typo3 – Open Source Enterprise CMS. TYPO3 Association. http://typo3.org. 2014-10-30.

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. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, I-2, 223–228.

Willmes, C., Kürner, D. & Bareth, G. (2014): Building Research Data Management Infrastructure using Open Source Software. Transactions in GIS, 18 (4), 496–509, ISSN : 1467-9671, doi: 10.1111/tgis.12060.

Wordpress Contributors (2014): Wordpress – Open Source Blog Engine. https://wordpress.org. 2014-11-09.

Contact information
Christian Willmes
University of Cologne
Institute of Geography
Albertus-Magnus-Platz
50923 Cologne
Germany
c.willmes@uni-koeln.de
+49 221 470 6234