Geophysical Research Abstracts
Vol. 20, EGU2018-10307, 2018
EGU General Assembly 2018

A Research Data Management Platform to Support the Daily Tasks of Environmental Researchers

Stephan Mäs (1), Daniel Henzen (1), Simon Jirka (2), and Lars Bernard (1)
(1) Technische Universität Dresden, Dresden, Germany, (2) 52°North GmbH, Münster, Germany (jirka@52north.org)

Most of the available research data infrastructures focus on the final publication of research data, which is often the very last step of the research workflow. Usually, data management functionalities to support the daily work of the researchers are hardly considered. Consequently, the various technical components for data acquisition, transformation, integration, and management are often not very well connected to data publishing and provisioning services. This has consequences for acquisition of metadata and in particular data provenance, which most researchers experience as a tedious task during the data publication process. This submission proposes an integrated software architecture to support research data management over the complete data live cycle. The architecture is a result of the research project COLABIS (Collaborative Early Warning Information Systems for Urban Infrastructures, https://colabis.de).

The presented research data management architecture and its pilot implementations showcase a seamless, simple to use and secure solution, trying to get as close as possible to data creation and easing and minimizing manual metadata acquisition wherever possible. This starts with capturing information about the specific research workflows within the respective project and the steps in which data are collected, generated or changed. This information is used to automatically enrich the data descriptions with provenance information, such as references to source or input data for environmental simulations or versioning information of a data set. Further, the data management platform generates data folders for each of the workflow steps for the data storage and filing for the researcher's daily tasks. The access to the data can be restricted to specific team members. To support the collaboration among the involved researchers, a messaging system informs, when new data has been inserted to the platform. The data management platform also allows to integrate previous data from distributed sources and simplifies the inclusion of external data sources within projects. All metadata are collected once and continuously updated during the data life cycle. This integrated metadata enrichment process eases the follow-up publishing process.

Sufficiently qualified data is published in a CKAN (http://ckan.org) catalogue, an open source software solution for open data management and dissemination. Within COLABIS CKAN has been extended with data offerings as interoperable, OGC-compliant services to provide spatial data and map visualisations as well as interactive provenance visualizations that illustrate the data history. To allow the development of different information applications the platform also provides lightweight interfaces, for example for graph visualization of time series data or maps for decision support and early warning. Within COLABIS, a specific focus is laid on the use of standardized interfaces and formats for the exchange of environmental data.