

An OGC-based Service Oriented Architecture Approach for a Research-Data Infrastructure

Silke Boos, M.Eng.
Prof. Dr. Kai-Christian Bruhn
Thomas Engel, M.A.
Prof. Dr. Hartmut Müller

Research Group Geocycles

The Geocycles Earth System Research Centre consists of 30 research groups within different geoscientific realms. It is based on the cooperation of four institutions in Mainz: The Johannes Gutenberg University, the Max-Planck-Institute for Chemistry, the Roman-Germanic Central Museum and the Institute for Spatial Information and Surveying Technology i3mainz.

The Data Curation and Administration group (DCA), based at i3mainz, linking all four institutions involved in Geocycles, aims to support the research groups to address the multiple challenges of data treatment and archiving and supporting the members to contribute to openly available repositories through standards-based protocols and qualified meta data schemes.

DCA Survey

A survey was conducted to evaluate current data polices of the Geocycles research groups. Current and future status of data handling, data acquisition, data integration of external sources, use of free and open standards, existing mechanisms of data curation, the spatio-temporal dimension of data and the need for data sharing across different involved research disciplines were retrieved. Results are displayed for a representative sample of 15 research groups, clearly showing a demand for integrated approaches to various aspects of research data handling. Almost all data obtained by an earth system research centre has a decisive spatial and temporal aspect (see "Survey results") that is missing from existing solutions for research data structures.

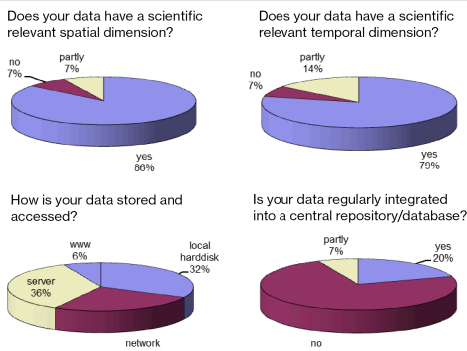


Fig. 1: Key findings of a comprehensive survey undertaken to evaluate the current data policy of the different research groups forming the earth system research center

Focus of research

The heterogeneous data created by the research groups can be divided into four realms: Hydrographic Data, Earth Data, Atmospheric Data and Human Data, covering earth system processes within the anthropocene (see Fig. 1). One of the key problems within the research groups is knowledge mobilization – transporting data out of the "data reservoirs" surrounding individual research groups, integrating it, and controlling its dissemination.

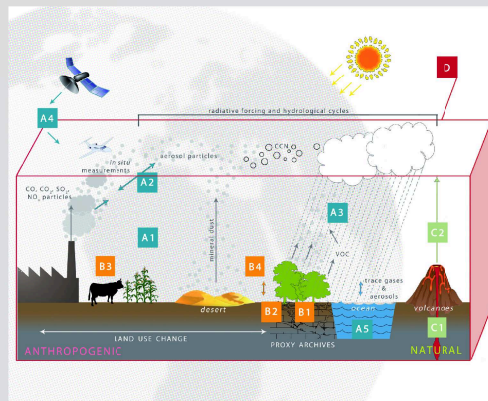


Fig. 2: Earth System processes covered by Geocycles research determining data domains

Data integration through space and time

As a demonstration of the possibilities of the approach, a web based data portal was developed, which integrates and visualizes heterogeneous spatial/temporal data/information based on archaeological and palaeoclimate archives. This tool – EifelGIS – uses OGC (Open Geospatial Consortium) compliant services, providing researchers with collecting, editing and viewing their own data corresponding to datasets from other research groups (Fig. 2).

Fundamental background of this approach is the use of a service-oriented architecture to make data/functionality available over standardized (Web-) services. Once the meta data of a dataset is published in a registry, it can be found by a client, which is then able to bind to the respective dataset or functionality (Publish-Find-Bind paradigm).

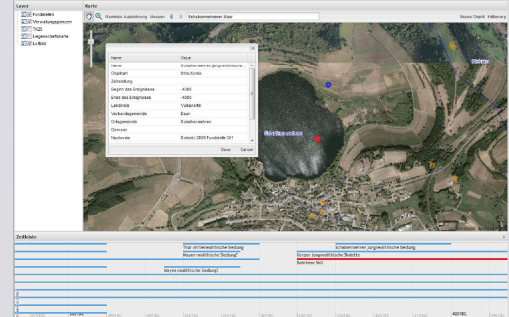


Fig. 3: EifelGIS – a webtool for the integration of archaeological/geological scientific data

The DCA-ERA-concept

The developed high-level layout of the research data infrastructure is divided into two major parts. The first part details the support provided by DCA, including legal aspects, meta data issues, data formats and models (Fig. 4 left). For this part, tools, applications and training are planned to be provided by DCA.

The second part (Fig. 4 right) depicts the technical design of the research data domain with controlled access to stored data in a central repository also allowing for long term access. The use of standardized web-services provides a single technological solution for communicating research data within Geocycles, as well as with the scientific and public domain.

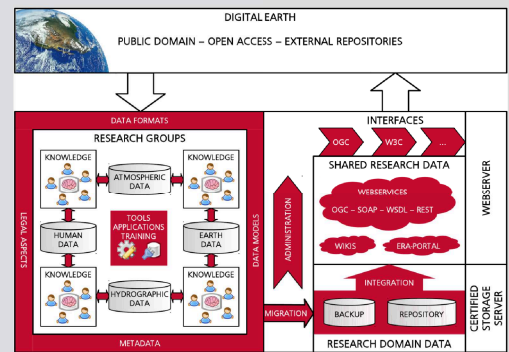


Fig. 4: Conceptual approach to the architecture of the research data infrastructure

