



Architectural Patterns for web-based information processing in scientific infrastructures

Rainer Haener (1), Henning Lorenz (2), Sylvain Grellet (3), Sebastien Hameau (3), Francois Robida (3), and Joachim Waechter (1)

(1) Helmholtz Centre Potsdam GFZ German Research Centre for Geosciences, CeGIT, Potsdam, Germany (rainer.haener@gfz-potsdam.de), (2) Uppsala University, Department of Earth Sciences, Uppsala, Sweden, (3) The French Geological Survey, Orleans, France

Current e-infrastructures in solid Earth sciences have a great diversity with regard to data volume, data types and formats, and the interfaces that are provided for accessing the content. Initiatives like the European Plate Observing System (EPOS) with the aim to provide trans-national access to multiple scientific domains expose the lack of standardisation concerning the encoding of data, metadata, and semantic information, partly due to a proliferation of domain and even application specific encodings. However, such international and multi-domain initiatives are also the primary driving force in amending this situation. Collaboration across domain boundaries does not only imply simple information exchange. Moreover, it requires that information can be further processed and thus, easily discovered, utilised by and integrated into existing workflows. Consequently, a common understanding of the exchanged information itself and its meaning must exist between all participants involved, which in turn requires a very high level and quality of standardisation regarding the specification of domain semantics and interfaces (Design by Contract).

To solve this problem, the patterns proposed here are based on a consequent application of the architectural paradigms for a functional integration of web resources, Representational State Transfer (REST) and Event-Driven Service-Oriented Architectures (SOA 2.0). Of greatest importance is the use of a small, reduced set of well-established and commonly accepted encodings and protocols that can be applied to any domain, regardless of its nature. This approach enables a seamless integration of both, web services and event processing as well as their automatized composition in processing workflows. Thus, it provides the means for a reproducible and fast and meaningful analysis of large data sets within responsive and resilient infrastructures.

Using examples from subsurface management, risk management, and tsunami early warning, this proposal illustrates the utilisation of standards-based information from different domains. Beyond discovery and retrieval of data we present the application of (web-based) processing and visualisation methods to them: Subsurface management for instance applies statistical analyses to geophysical and borehole measurements in order to approximate geological structures. Early warning systems on the other hand, use pre-processed numerical models like the simulation of a tsunami in order to predict (observe) time series of sea-level measurements at arbitrary virtual sensor locations and correlate them with real sensor observations.