
An SDI for archaeological data with a RESTful interface to semantically modeled information

Tobias KOHR, Kai-Christian BRUHN, Ashish KARMACHARAYA, Christophe CRUZ,
and Frank BOOCHS

Abstract

The poster focuses on current activities carried out within a bi-national research project “HIGEOMES – Die historische Geographie Obermesopotamiens im 2. Jahrtausend v. Chr.“. Motivated by the increase of archaeological and epigraphic sources in recent decades the project aims at an integrated analysis and transparent access of these heterogeneous data. A designed application needs to consider that archaeological data contain absolute geography – the association with a specific geographic location, while epigraphic data imply relative geography – the topological relation of places. The solution is provided by a Web GIS interface supporting scientists to better understand interdependencies and contribute to a superior comprehension of past societies.

Archaeology comprises a vast amount of heterogeneous datasets with relevant spatial reference. Dissemination of these data is essential, in order to link knowledge of different research groups. Standards and technologies that are being developed with the advent of spatial data infrastructures appear suitable to achieve this objective. The European SDI initiative (INSPIRE) clearly targets environmental data, it is predominantly enforced in the administrative sector and only obliges to ensure access to data on legally protected sites (EU 2007). However, the initiative can be utilized as a driving force for the exchange of research data and is extensible to other domains. For the archaeological research community this presents an opportunity to facilitate transparent access on research data and thus enrich knowledge of individual scientists. This could be a motivation for the cultural heritage community to envision a broader approach and ensure data dissemination that goes beyond legally protected sites (MCKEAGUE et al. 2012). FERNÁNDEZ FREIRE et al. (2012) have proposed a Cultural Heritage Application Schema to realize this task.

The HIGEOMES portal uses an infrastructure based on OGC Web services to facilitate the exchange of collected archaeological information among project partners and to ensure external researchers an interoperable access to it. Archaeological find spots collected at the University of Mainz are shared via a Web Feature Service (WFS) along with Web Map Services (WMS) portraying relevant raster imagery. A Web GIS client acts as an entry point for visualization and simple analysis of the find spot locations.

Alongside the creation of Service Oriented Architectures, the Semantic Web is developing for the dissemination of data following the concept of Linked Data. The goal of the Semantic Web is to focus on meaning of words rather than the words themselves (MAZZOCCHI 2000) and thus provide smarter data. This promises the possibility to link highly heterogeneous data and to apply semantic analysis such as reasoning and inferring. Within HIGEOMES, semantic modelling is suitable for epigraphic data from the Archibab DB in Paris

(<http://www.archibab.fr/>). Toponyms and peoples that are mentioned in texts are interrelated among each other to a high extent and frequently miss a known location, which prevents absolute geographic modelling. Ontologies provide a fitting framework to describe this situation. However, an integration of the Semantic Web with the service oriented approach is still largely missing (PEDRINACI 2010). While the attempt to resolve this deficiency as a whole ranks among current most important research topics in geoinformatics (JANOWICZ et al. 2012), HIGEOMES strictly focuses on the integration of a knowledge base within an existing SDI. In a first approach we expose data from the knowledge base via a RESTful Web service. The Web service is able to query information to a specific toponym from the ontology and emit the revealing graph in JSON. This graph can be consumed and visualized by the Web client, which enables the user to browse through text data and explore relations between toponyms independent from geography. At the same time an overlap between the toponyms and the find spot location datasets exists. In order to connect both sources, the key identifier of the find spots is included in the semantic model of the toponyms. As a result the user can demand text information to a specific find spot and locate toponyms from a textual source on the map. An integrated visual exploration of geographically and semantically modelled information is possible.

The designed application is based on the present SDI and thus on well-established technologies. This ensures a mature infrastructure for the dissemination of archaeological data and allows presenting the potential to integrate text data. However the prototypical solution has several shortcomings that we need to address in future research. Both, places with multiple toponyms and multiple locations with identical names cause problems that are difficult to solve in the current model. Connecting the data sources on a semantic level could present a solution. Eventually ontologies are designed to resolve such issues. This would mean to expose the geographic information in a semantic way and to perform a mapping between the two data sources. In addition this approach would implicate the capacity to perform semantic analysis on both data sources. KARMACHARYA (2011) demonstrates the capabilities of geospatial inference through integrating geospatial semantics in existing knowledge schemas. Such analysis has the potential to locate toponyms from texts with unknown locations by considering facts from the archaeological data and vice versa.

In geospatial semantics, KUHN (2010) and JANOWICZ et al. (2012) distinguish between the two research areas of modelling and encoding. Both, semantic modelling and encoding of geospatial information are of special interest in a next step towards comprehensive semantic integration within HIGEOMES. In order to map the archaeological information to the current ontology of text data we will need to work on a fitting model and consider the extension of existing geo-ontologies. At the same time we will need to investigate how to encode information on the interfaces and how to expose semantic capabilities to the client towards the realization of a Semantic Web.

The poster will present the current state of development and invite the audience for further discussion.

Acknowledgements: HIGEOMES is a bi-national collaboration between the Universität Mainz, Fachhochschule Mainz/i3mainz, Freie Universität Berlin, Université de Bourgogne Dijon and CNRS Paris that is funded by DFG and ANR.

References

- EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION (2007), 'Directive 2007/2/EC (INSPIRE)', Strasbourg
- FERNÁNDEZ FREIRE, C.; PÉREZ ASENSIO, E.; DEL BOSQUE GONZÁLEZ, I.; URIARTE GONZÁLEZ, A.; VICENT GARCÍA, J.M.; FRAGUASBRAVO, A.; PARCERO-OUBIÑA, C.; FÁBREGA-ÁLVAREZ, P. (2012) 'Proposal for a Cultural Heritage Application Schema within the INSPIRE directive', Proceedings of the AGILE'2012 International Conference on Geographic Information Science, Avignon, April, 24-27, 2012
- JANOWICZ, K.; SCHEIDER, S.; PEHLE, T.; HART, G. (2012) 'Geospatial Semantics and Linked Spatiotemporal Data – Past, Present, and Future', Semantic Web Journal
- KARMACHARYA, A. (2011) 'Introduction of a spatial layer in the Semantic Web framework: a proposition through the Web platform ArchaeoKM', Ph.D thesis, Le2i, Université de Bourgogne
- KUHN, W. (2010) 'Modeling vs encoding for the semantic web' Semantic Web Journal, 1(1-2)
- MAZZOCCHI, S. (2000) 'Toward the Semantic Web: a view of XML from outer space' Proceeding: ApacheCON 2000, London
- MCKEAGUE, P.; CORNS, A.; SHAW, R. (2012) 'Developing a Spatial Data Infrastructure for Archaeological and Built Heritage', International Journal of Spatial Data Infrastructures Research, Vol.7, 38-65
- PEDRINACI, C.; DOMINGUE, J. (2010) 'Toward the Next Wave of Services: Linked Services for the Web of Data', Journal of Universal Computer Science, 16(3), pp. 1694–1719