Semantics – Supportive Element for the Cooperative Evaluation of Geographical and Historical Information

Ashish Karmacharya, Tobias Kohr, Christophe Cruz, Kai-Christian Bruhn and Frank Boochs

Summary

The emergence of the Semantic Web and its underlying knowledge technologies has brought changes in data handling. Transferring expert knowledge to machines through knowledge formalization provides us the required support in managing huge datasets like the information in the World Wide Web. In the field of geospatial technology semantic technologies not only entail the capability to achieve higher degree of data integration but also infer semantics to discover new and hidden knowledge. This is of particular interest in the field of archaeology, where complex interrelations among heterogeneous datasets exist. Although researches on semantics are active areas in geospatial communities, their initial use is mainly for spatial data integration. This article tries to go one step further and imply semantics for spatial knowledge discovery through spatial built-ins within SWRL and SPARQL. The work resembles the approach of the Open Geospatial Consortium (OGC) to define standards for GeoSPARQL.

Zusammenfassung

Die Entstehung des Semantic Web und die damit verbundenen Wissenstechnologien haben Veränderungen in der Datenhandhabung mit sich gebracht. Wissensformalisierung, die es erlaubt, Fachwissen an Maschinen zu übermitteln, liefert die notwendige Unterstützung zur Verwaltung sehr großer Datensätze, wie die Informationen im World Wide Web. In der Geoinformatik sind semantische Technologien nicht nur dazu in der Lage einen höheren Grad an Datenintegration zu erreichen, sondern auch aus Rückschlüssen der Semantik neue Erkenntnisse zu gewinnen. Dies ist in der Archäologie, in der größtenteils heterogene Datensätze mit komplexen Verknüpfungen vorkommen, von besonderem Interesse. Wenngleich Semantik ein aktives Forschungsfeld in der Geoinformatik ist, konzentriert sich ihr Nutzen dort hauptsächlich auf die Integration räumlicher Daten. In diesem Artikel präsentieren wir erste Ansätze, Semantik zur raumbezogenen Wissensgenerierung einzubeziehen. Dies wird durch den Einbau räumlicher Parameter in SWRL und SPARQL erreicht. Die Arbeit ist mit dem Ansatz des Open Geospatial Consortium (OGC) vergleichbar, welches hierfür den GeoSPARQL spezifiziert hat.

Keywords: Knowledge Modeling, Spatial Integration, the Semantic Web, Archaeology, Philology

1 Introduction

Traditional geographic information systems (GIS) enable the capture, management, analysis and presentation of spatially referenced data. Fields of application are manifold, comprising regional planning, disaster prevention, traffic management, local administration and logistics, just to name a few. In times of spatial data infrastructures (SDI) and open accessible geodata over the Web, the relevance of GIS is steadily increasing. Besides the well-known application fields, GIS also serve a variety of marginal areas, such as archaeology. Here they provide the user distinct views on the data material, which sets the base for domain specific analysis. Considering the heterogeneity, variety and potential incompleteness of data, as well as the complexity of observation, attainable statements are confined however. Nevertheless GIS tools provide important assistance in this field that will further expand in the future.

A database structure that holds spatial and attributive feature information, as well as topological relations between features sets the base of the traditional GIS architecture. Based on these data, spatial analyses are performed, graphically visualized, evaluated by the user and (manually) used to infer further knowledge. Deficits of this approach are encountered at two major points: First, a once defined data structure is fixed and not adaptable in retrospect. Second, the information systems cannot contribute to the knowledge retrieval directly. This leads to drawbacks if the data structure is potentially subject to changes – as i.e. in the case of archaeological excavations, that often encounter unexpected facts – and as far as the interrelationships of interest are multilayered and to be developed in an explorative manner.

Broader flexibility in the creation and more possibilities in the supportive analysis would imply considerable benefits under such circumstances. New developments for the retrieval of the vast information of the Web are promising for this purpose. Currently emerging capabilities of knowledge based processing mechanisms possess significant benefits towards traditional approaches. Data are intelligently and flexibly managed and can therefore be dynamically adapted. This allows modification of the data model where required. Moreover the computer can directly contribute to further knowledge retrieval applying user-defined rules. In particular in complex situations, this can provide significant assistance.

In order to reach this objective, technologically unknown territory needs to be entered in many respects and it is still a long way to comprehensive alternatives of today's GIS. Yet without complete intelligent GIS, interesting capabilities to facilitate spatial issues are feasible. Geospatial communities have also started to give notice to the rising popularity of knowledge technologies within the Semantic Web framework. OGC work on standardizing spatial components in the Semantic Web technologies (Perry & Herring 2012) indicate in this direction and first work already presents the potential (Karmacharya 2011).

Geospatial disciplines require high level of semantics not only for integrating heterogeneous data sources but also allowing adapting user's cognitive abilities (Tanasescu 2007). Beyond that they allow spatial reasoning within GIS frameworks (Fonseca & Egenhofer 1999). Potentials of spatial semantics can be realized through associating spatial components into semantic technologies. *Räumliches Informationssystem zur Erfassung, Dokumentation und Analyse Industriearchälogischer Objekte* or the RIO project (Karmacharya et al. 2010) initiates this realization through proposing a geospatial layer into the Semantic Web framework. As a pioneer knowledge management project at i3Mainz, it was among the early works on spatial integration into the Semantic Web. It maintains its overall goal on data integration through semantically annotating data and documents. The key is to identify the objects in the excavation area and give them spatial signature and then to link up to relevant data and documents. This process of spatialization opens doors for spatial associations in the Semantic Web framework and paves the way for spatial knowledge processing. With spatial semantics, a level of spatial knowledge interpretation is possible that allows discovering new knowledge which was otherwise not possible. We take forward the advantages and experiences gained in RIO into HiGeoMes *(Die historische Geographie Obermesopotamiens)* (Kohr et al. 2013). The project provides an extra challenge of spatializing components described through relative geography as it incorporates other domains, which purely focus on textual interpretation like the domain of philology. It hence adds up an extra challenge in linking up such interpretations to their spatial semantics. The general approach is to spatialize non-spatial components by attaching to their relevant tangible objects having spatial identities (Cruz 2004, Cruz et al. 2004, Tanasescu et al. 2006).

The following chapter will give an introduction to knowledge management and semantic technologies. A background section addresses the need for the integration of semantic and spatial technologies. Subsequently the concept of spatial built-ins in the Semantic Web framework is presented. Finally the paper illustrates current and future capabilities emerging from this approach and terminates with a conclusion.

2 Knowledge Management and the Semantic Web

Knowledge Management in simplest term is the process of identifying, creating, distributing the experiences, expertise and insights possessed within an individual or group or even an organization. Knowledge is commonly distinguished from data and information (Zack 1999). Data are a representation of an observation or any singular fact kept out of context. Data are meaningless until they are put in context of space or an event. Additionally, unless the relationships between different pieces of data are defined, simply data do not have any significance. Once data are defined in terms of space or events and are defined through relationships, they become information. Information understands the nature of the data but does not provide the reasons behind the existence of data and is relatively static and linear by nature. Information is a relationship between data and, quite simply, is what it is, with great dependence on context for its meaning and with little implication for the future (Bellinger 2004). Beyond every relationship, arises a pattern which has capacity to embody completeness and consistency of the relations to an extent of creating its own context (Bateson 1979). Such patterns represent knowledge on the information and consequently on data. The term *Knowledge Management* has wide implications. However, very precisely Knowledge Management is about the capture and reuse of knowledge at different knowledge levels.

Knowledge Management re-evolved with the rise of the Semantic Web. The explosion of information in the World Wide Web (WWW) has led to the problem of managing it. It is generally perceived that this vast information cannot be managed through human effort only. In some form there should be interference from machines to assist humans manage the information. In order to have machines interfere and assist humans in managing information, it should understand the information first. This would require knowledge formalized from the information. In their paper Berners-Lee et al. (2001) have envisaged the next generation of the Web which they call "the Semantic Web". In this Web the information is given with well-defined meaning, better enabling computers and people to work in cooperation. Adding on, the Semantic Web aims at machine-processible information enabling intelligent services such as information brokers, search agents and information filters, which offer greater functionality and interoperability (Decker et al. 2000).

The association of knowledge with the Semantic Web has provided a scope for information management through knowledge management. Since both the technologies use ontologies to conceptualize the scenarios, Semantic Web technology could provide a platform for developments of knowledge management systems (Stojanovi & Handschuh 2002). The ontologies are core to both the technologies in whichever methods they are defined. The Semantic Web defines ontologies through XML based languages and with the advancements in these languages.

Semantic Web Architecture

The Semantic Web architecture comprises the technologies used within the Semantic Web framework. The technologies are grouped in hierarchical layers. Each layer takes advantage of the capabilities of the layer underneath. There have been numerous evolvements of the architecture since it was first described. However we will discuss it with respect to the architecture illustrated in Fig. 1.

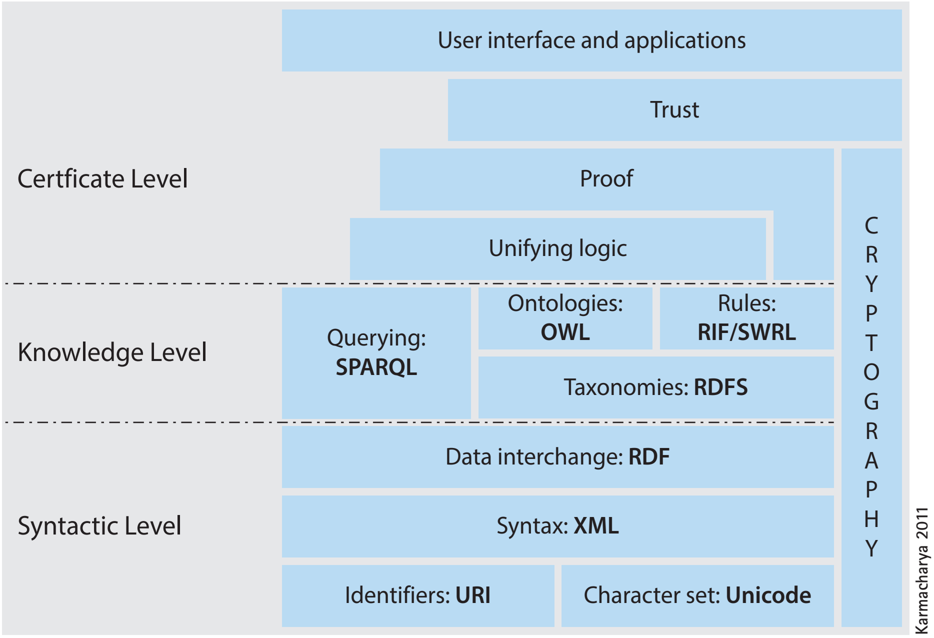

Fig. 1: The Semantic Web Architecture

Based on the character of the technologies, we have divided them into three different levels. The technologies at the bottom Syntactic level are designed to present information. The information presented can be non-structured, semi structured, or structured. The middle level is the Knowledge level. The technologies within this layer are designed to generate, represent, and manage generated knowledge. We focus our interest in this level and present our conceptual notes through technologies within this level. The technologies within this level are further discussed in the next sections. The top level is the Certificate level and is not yet fully conceptualized. The technologies within this level are meant for establishing trust on reliability of information. At present there is no technology recommended to support this layer but there is an attempt for developing a proof language called Proof Mark-up Language (PML) (da Silva et al. 2004, Al-Feel et al. 2009) by the knowledge systems laboratory at Stanford University.

Ontology Language

The term *Ontology* is being used for centuries to define an object philosophically. The core theme of the term remains the same in the domain of computer science; however the approach in defining it has been modified to adjust the domain. Within the computer science domain, ontology is a formal representation of the knowledge through the hierarchy of concepts and the relationships between those concepts. In theory ontology is a *formal, explicit specification of shared conceptualization* (Gruber 1993). In any case, ontology can be considered as formalization of knowledge representation and *Description Logics* (DLs) (Calvanese et al. 2001, Baader & Sattler 2000) provide logical formalization to the Ontologies (Baader et al. 2003). The World Wide Web Consortium (W3C) has standardized Web Ontology Language to model ontologies. OWL is actually a family of three language variants of increasing expressive power: OWL Lite, OWL DL, and OWL Full. The standardization of OWL has sparked off the development and/or adaption of a number of reasoners, including FacT++ (Tsarkov & Horroks 2006), Pellet (Nguyen & Nguyen 2010), RACER (Haarslev & Muller 2001) and HermiT (Shearer et al. 2008) and ontology editors, including Protégé (Protégé 2012) and Swoop (Kalyanpur et al. 2006). OWL 2 is a new version of OWL, the ontology language which considerably improves the datatype (Motik et al. 2009).

Query and Rule Languages

SPARQL is a query language for RDF triplets of which OWL is syntactically aligned. In this manner SPARQL queries the knowledge within OWL. As a query language, SPARQL is "data-oriented" in that it only queries the information held in the models; there is no inference in the query language itself. SPARQL is able to query OWL ontologies which use RDF graphs to structure it. SPARQL uses FILTERS to limit the solutions to only those which are returned true with the expression.

Semantic Web Rule Language (SWRL) is a logic programme which infers the knowledge base to derive a conclusion based on the observations and hypothesis. The SWRL as the form, antecedent --> consequent, where both antecedent and consequent are conjunctions of atoms written a1 ˆ … ˆ an. Variables are indicated by using the standard convention of prefixing them with a question mark (e.g., ?x). URI references (URIrefs) are used to identify ontology elements such as classes, individual-valued properties and data-valued properties. For instance, the following rule asserts that one's parents' brothers are one's uncles where parent, brother and uncle are all individual-valued properties.

hasParent(?x,?p)ˆhasBrother(?p,?u)→hasUncle(?x,?u)

3 Background

The basic tasks of a GIS can be broken down into four groups (4Ds): data acquisition, data management, data analysis and data visualization. In parallel, Bill (2010) refers to input, management, analysis and presentation (IMAP) among others. Most archaeological data such as artifacts, features, buildings, sites or landscapes, have spatial and aspatial attributes that can be explored by GIS. These attributes include the spatial location that informs about the local or global context concerning the pieces of information, and the morphology that defines the shape and the size of an object.

In recent years, the rapid growth in data acquisition techniques – that are applied in archaeology – has made some limitations visible in traditional standalone GIS. It is not only the volume of data any more. The diversity in data collected play equal role in current archaeological projects. Nevertheless, for many of the archaeological projects an information system is still either a GIS or a 3D modeling system. Applications like ArchaeoCAD from ArcTron and PointCloud from Kubit rely heavily on the geometry of the objects excavated. Correspondingly, research projects like 3D MURALE (Cosmas et al. 2001) and GIS DILAS (Wüst et al. 2004) lean heavily towards spatial data management focusing on extending GISs to fit in archaeological data management. One could thus argue spatial components play a major role in an archaeological project. The ever-growing sophistication in the data acquiring technology has exposed the incapability of current GIS as the data patterns get more diverse through the course of time and GIS lack behind accommodating them. Even data structures like the 3D point cloud which possess spatial dimension find itself extremely difficult to succor with existing GIS. This is even more prominent with aspatial data structures like multimedia datasets. Most of the archaeological projects today use 3D point clouds to document their findings for future reconstructions. Likewise, they use document files to document their processes of excavations and other multimedia to record their steps. All these need to be taken care of in order to have a comprehensive analysis process. 4Ds process within any GISs should thus need to take aspatial data structure into account. Data integration is an important issue within current information management systems in archaeological projects.

Besides, archaeology does not only produce very heterogeneous data within one project, but in numerous ventures. These data can have very diverse spatial, temporal and thematic scopes. At the same time very close interrelations among projects exist, making it essential to connect their information in order to gain new knowledge. In the field of geo-informatics, Spatial Data Infrastructures (SDIs) have emerged to enable the dissemination of spatial data over the Web. International initiatives like INSPIRE urge their introduction on all administrative levels. This trend acts as a driving force for the specification of relevant standards and the development of corresponding technologies. Becoming the leading standardization instance, the Open Geospatial Consortium (OGC) has defined a stack of Web service specifications generally complying with OGC Web service (OWS) Commons. These standards and technologies have taken the dissemination of data to a new level by providing syntactically interoperable interfaces. The actual meaning of the information remains unreadable to a machine however. This restrains discovery, integration and automated processing of the information (Janowicz et al. 2010). With the current trend of information explosion a next level of interoperability is aspired. Even with syntactical interoperable systems, the user needs to possess in-depth knowledge about the data structures of the data sources he or she is evoking. This is not always possible. In order to have higher degree of interoperability, data structures of one data source should understand diverse data structures of other data sources. This will require attaching meaningful semantics to each data source. Data sources in a distributed environment could be thus connected through a common semantic and syntactic model.

The importance of semantics in achieving higher level of interoperability is widely felt in the geo-spatial community. OGC is taking forward its efforts on existing web service standards for publishing geospatial data through attaching semantics to the services. Reports by Houbie et al. (2012) point out the importance of semantically annotating the OGC standards publishing geodata in web services like WFS. They identify three levels that are in need of semantic annotations: Service metadata, data models or process descriptions and data instances. Annotations enrich OWS data within the existing interface that can be used by semantically enabled and non-enabled consumers. The approach can enhance the aforementioned shortcomings in data discovery, integration and automated processing. Furthermost, the OGC is also moving towards standardizing spatial extensions to Semantic Web technologies. GeoSPARQL is the latest example of such activities. It defines vocabularies to represent geospatial data in RDF and defines spatial extensions to the SPARQL query language for processing geospatial RDF data (Perry & Herring 2012). Roth et al. (2013) pursue to provide linked data from OGC Web service output to harness capabilities of Semantic Web technologies. This has the potential to perform semantic analysis on the data and integrate them efficiently with non-spatial information. Moreover mapping schemas entail the capacity to link data sources with each other independent of their format.

Archaeology by this date remains a predominantly closed discipline where the archaeological records are often contested, and even more commonly restricted or altogether inaccessible to all but a few (Isaksen 2011). Hence, archaeology and semantic technologies seem to misfit (Isaksen 2011). However, numbers of research projects linking Semantic technologies to the archaeological domain have emerged in recent past. CIDOC CRM (CIDOC CRM 2013) provides formal definitions to define implicit and explicit concepts and relationships in cultural heritage. The model has extensive sections dedicated to archaeology and its findings. It has been used for data modeling of varied archaeological systems at English Heritage's Centre for Archaeology (Cripps & May 2004). STAR project (Binding et al. 2010) is one of such addition in using CIDOC CRM model along with W3C SKOS (SKOS 2012) to develop online terminologies and knowledge organization mechanism in archaeology. Pelagios (Simon et al. 2012) is a Linked Open Data hub that provides RDF serialization of the places in the Ancient World. It is a community driven project allowing independent groups to publish their works and link them to resources and works of other groups. Graph based data visualization techniques which it calls the Graph of Ancient World Data (GAWD) is implemented to visualize the data. STELLAR project provides a framework to extract archaeological datasets in RDF/XML representation conforming to CIDOC CRM (ADS 2013).

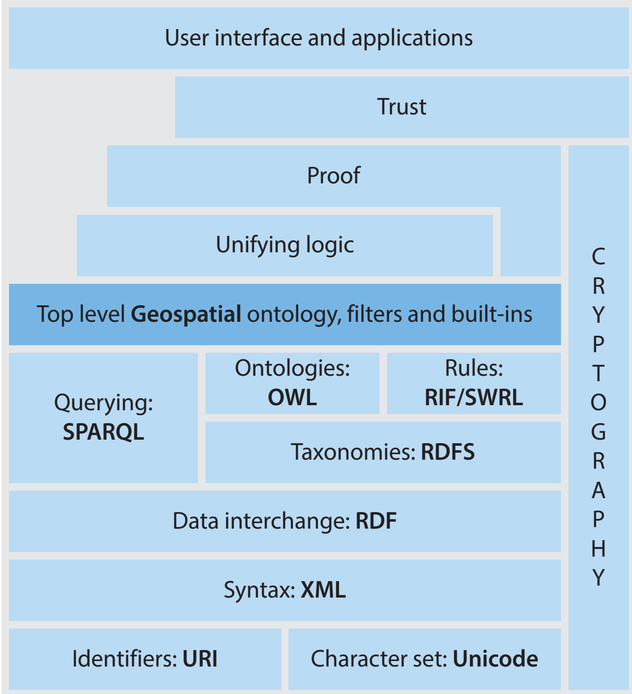

Fig. 2: Inclusion of geospatial layer within the Semantic Web architecture

4 Spatial semantics

In simplest terms, spatial semantics are realized through associating spatial components within semantic technologies. We keep abreast with this through a suggestion of a geospatial layer within the Semantic Web architecture that is demonstrated through the RIO project. RIO complements 4D principle of GIS through its 4K principle: knowledge acquisition, knowledge management, knowledge analysis and knowledge visualization. This is carried over to spatial knowledge through the geospatial layer.

The geospatial layer situates itself just over the Knowledge level intentionally not to contradict with the fact that the technologies within the Knowledge level are standardized by W3C. However, it associates the underlying spatial components to knowledge technologies (OWL Ontologies, Querying: SPARQL and Rules: RIF/SWRL) underneath. Fig. 2 illustrates the proposed geospatial layer within the Semantic Web architecture presented in section 2.

A spatial top level ontology expressed through OWL Ontology is designed that could be extended to domain ontologies. We call it Spatial Extension Ontology Schema (SEOS). Particular consideration is made so that every spatial operation can be performed through this schema. We have divided the spatial functions and operations into two major categories: geospatial processing functions and geospatial relationship functions. Geospatial processing functions return geometries while geospatial relationship functions return a Boolean value of their relationships. A specialized class is dedicated to store results from geospatial processing functions within SEOS whereas the geospatial relationship functions are reflected through object properties. We use four geospatial processing functions to demonstrate the effectiveness of the approach: buffer, union, intersection and difference. These four functions are presented through four generalized classes within SEOS. Likewise, SEOS contains generalized object properties to represent geospatial relationship functions like disjoint, touches, within, overlaps, equals, crosses, intersects and contains. Tab. 1 and 2 illustrate the geospatial processing and relationship functions and how they are integrated within SEOS.

The SEOS has laid a foundation to provide spatial signatures to the objects within any domain ontologies. It also provides background for performing spatial queries and spatial inference through a rule based inference mechanism on the knowledge base, which can be used alongside other filters in SPARQL or built-ins in SWRL. Tab. 3 demonstrates such filters and built-ins with their syntax of execution for geospatial processing and relationship functions. The table illustrates the most common syntax and thus presents those built-ins and filters that have unique executing characteristics. We have presented them in details in Karmacharya et al. (2011).

SEOS has thus provided a foundation where non-spatial data types can participate in spatial analysis or spatial queries. As a result, it is possible to retrieve a smart query through spatial SPARQL like "*SELECT the chimneys stored in CAD document xyz.dwg which are touching furnaces stored in point cloud file pc.xyz*". Likewise it is possible to infer domain rules to discover new knowledge through spatial built-ins like "*Buffer areas of Chimneys in point cloud file pc.xyz of 5 meters of furnaces mentioned in archaeological notes myNote1.pdf was restricted area*".

Tab. 1: The Geospatial processing functions and their integrations into SEOS

Functions	Concept	ObjectProperty	Execution Method
Buffer	sa:sp_Buffer	sa:hasBuffer(x,c)	$sa{:}sp\_Buffer \sqsubseteq \exists sa{:}hasBuffer.feat{:}Feature \sqcap sa{:}hasBufferDistance.\{c\}$ c is of float value providing the buffer distance
Union	sa:sp_Union	sa:hasUnion(x,c)	$sa{:}sp\_Union \sqsubseteq \exists sa{:}hasUnion.feat{:}Feature \sqcap (\geq 2\ hasUnion)$
Intersection	sa:sp_Intersection	sa:hasIntersection(x,c)	$sa{:}sp\_Intersection \sqsubseteq \exists sa{:}hasIntersection.feat{:}Feature \sqcap (\geq 2\ hasIntersection)$
Difference	sa:sp_Difference	sa:hasDifference(x,c)	$sa{:}sp\_Difference \sqsubseteq \exists sa{:}hasDifference.feat{:}Feature \sqcap (\geq 2\ hasDifference)$

Tab. 2: The Geospatial relationship functions and their integrations into SEOS

Functions	ObjectProperties	Characteristics
Disjoint	sa:hasDisjoint(x,y)	Symmetric
Touches	sa:hasTouch(x,y)	Symmetric
Within	sa:hasWithin(x,y)	Transitive
Overlaps	sa:hasOverlaps(x,y)	No characteristics
Equals	sa:hasEqual(x,y)	Symmetric, Transitive
Crosses	sa:hasCrosses(x,y)	Symmetric
Intersects	sa:hasIntersect(x,y)	Symmetric
Contains	sa:hasContain(x,y)	Transitive
A → B implies to B → A: Symmetric A → B, B → C implies to A → C: Transitive		

Tab. 3: The spatial SPARQL and SWRL syntax

Functions	Spatial SPARQL Syntax	Spatial SWRL	Remarks
Buffer	SPATIAL_FILTER [buffer (?x, b, ?y)]	spatialswrlb:Buffer(?x, b, ?y)	Result are first populated in the knowledge base as individuals of class `sa:sp_Buffer` before further processing.
Union	SPATIAL_FILTER [union (?x, ?y1,?y2)]	spatialswrlb:Union(?x,?y1,y2)	Result are first populated in the knowledge base as individuals of class `sa:sp_Union` before further processing.
Within	SPATIAL_FILTER [Within(?x, ?y)]	spatialswrlb:Within(?x, ?y)	

5 Outlook

Driven by the gained results in activities like RIO, we have started to apply spatial semantics as a supportive element in a variety of scenarios. Though the global efforts on semantic attached spatial data are driven through the motivation of data integration, such efforts have far more implications. One of the major potentials in our view is knowledge discovery. While RIO has demonstrated benefits of semantic data integration and harnessed first assets of knowledge management we want to demonstrate future potentials on another use case within the HiGeoMes project: The main goal of HiGeoMes is to connect archaeologically documented settlement sites with place names from epigraphic sources to better understand political, social and environmental developments in Upper Mesopotamia in the 2nd millennium BC.

Integration of the modeled archaeological and epigraphic data is not straightforward. We currently use a spatial data infrastructure (SDI) to facilitate archaeological data exchange between stakeholders and share it via a Web Feature Service. These data structures have absolute geography. While SDIs provide syntactic integration they lack semantic integration. SDIs thus need to be semantically enabled to provide high level semantic integration with epigraphic data (Janowicz et al. 2010). In contrast, the cuneiform records present fuzziness in their geographical locations. These locations are described through relative geography such as proximity to river or to other locations. In addition, most of these records are still in the process of proper evaluation and hence not documented. The records, which have been analysed, are transformed into an OWL knowledge base through certain underlying processes but the locations are still described through relative geography. The relative geography in the knowledge base is described through theorem and axioms representing expert knowledge within the knowledge schema. The challenge is to give spatial identities to these locations so to describe them through their absolute geography. The current HiGeoMes architecture (see Fig. 3) illustrates a mechanism that integrates the archaeological data shared through SDI and epigraphic data from the OWL ontology. Making the resources available via a web service interface, the approach establishes a connection on the visual level: A Web GIS client hosting the spatial dataset from the archaeological data is capable of making smart queries to related toponyms in the epigraphic OWL ontology. A graph visualization to represent evoked knowledge about the queried sites presents the semantic relation of the site with respect to other places, but independent to geolocalization of these places. In parallel, the user can locate places from texts on the map that have been associated with archaeological sites. An integrated visual exploration of geographically and semantically modeled information is possible.

We have established a platform where information can be integrated at higher semantic level in the future (see dotted line in Fig. 3). This has laid down a foundation for knowledge discovery. Problems still exist in geocoding each place in epigraphic data due to the lack of information (or better understanding of data). It is practically impossible to analyse each piece of information because of its volume and diversity. Research following current HiGeoMes efforts will bring on the hypothesis and facts from archaeologists and philologists and formulate individual knowledge models for the respective domain. This will extend present HiGeoMes knowledge schemas to fit in the formulated hypothesis. A semantic middleware will create the bridge between these knowledge schemas. In this context we will especially consider OGC standards and regard what role interfaces like WFS can play. They build a stable infrastructure for the syntactically inter-

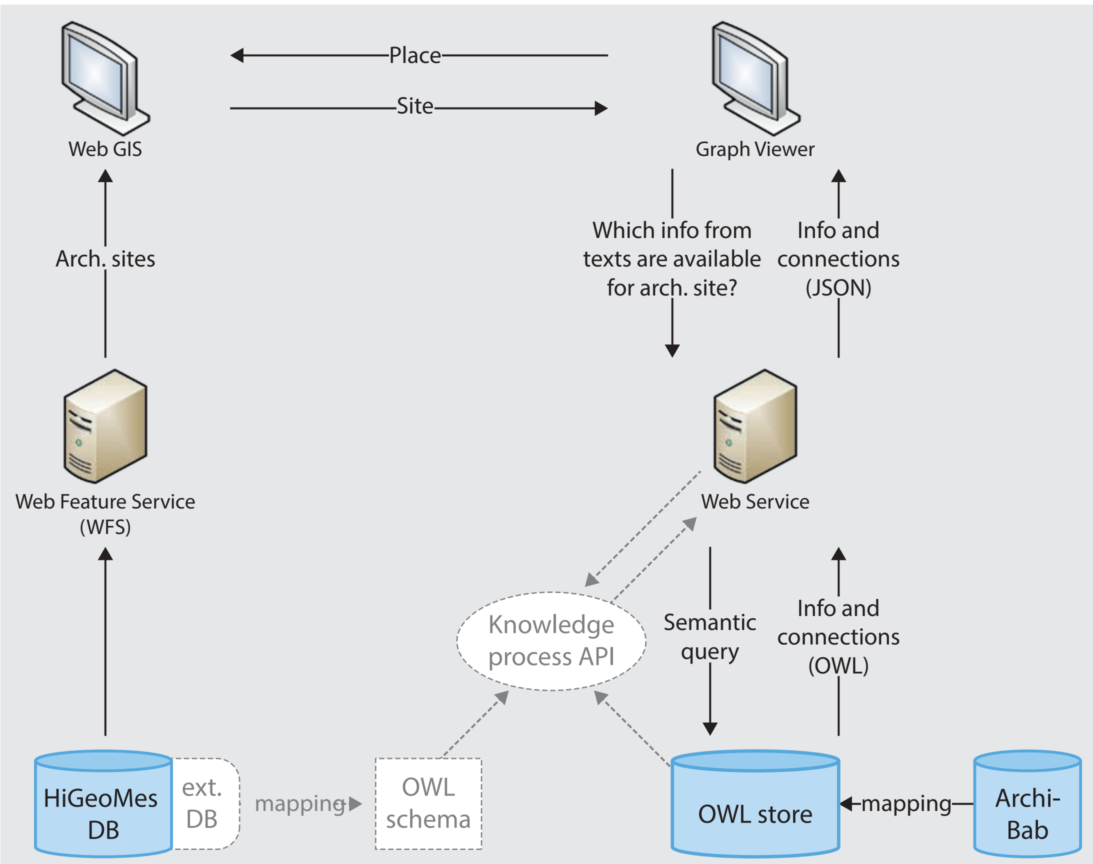

Fig. 3: HiGeoMes Architecture

operable dissemination of spatial HiGeoMes data that need semantic enablement. The semantic middleware not only acts as a facilitator for data integration but will also lay a foundation for knowledge discovery that can be based on the previously introduced spatial built-ins. We intend to use whatever limited semantics archaeological data possesses to relate them to semantics from the data of cuneiform records to discover new insight knowledge. In short, the hypothetical knowledge from the facts discovered in the excavated site will be related to see its relevance in the philological data from cuneiform records. This relevance will be deducted through inferring hypothetical rules against the facts in both datasets. It leads to discovering both spatial and non-spatial knowledge. An obvious benefit is that it will provide geocoded spatial signatures to places in cuneiform datasets which were not located previously.

6 Conclusions

In this paper we discussed the comprehension of spatial components within the Semantic Web framework. We reasoned for a layer in the Semantic Web architecture in order to enable geospatial analyses therein. The resulting benefits are twofold: First, semantic technologies are able to close the semantic gap of other technologies in data integration. Second, these technologies can manage large data volumes by automatizing analyses.

We presented our concept on examples from the RIO project. Spatial injections allow industrial archaeologists to manage their information through knowledge management. Such experts formulate domain rules to discover knowledge from the facts and hypothesis upon which these rules are built. In this process, the paper highlighted spatial extensions for SWRL, a rule language within the framework, to facilitate spatial inferences. The extension constitutes spatial built-ins which could be combined with other built-ins within the formulated rules. Likewise, it also presented spatial filters for the query language of SPARQL which were echoed through OGC standards for GeoSPARQL.

In a different scenario we demonstrate how the limitations of one dataset can be overcome by other datasets through semantically integrating those datasets. The HiGeoMes project requires the usage of semantics to suggest the relative geography of the places mentioned in epigraphic data through the absolute geography of the archaeological sites. Future activities in HiGeoMes will not only focus on data interoperability through semantics in OGC compliance standards but also concentrate on spatial knowledge discovery. This will bring the much needed reinforcement in this next dimension of knowledge management techniques where semantics could be implied to discover hidden knowledge from the existing facts and hypothesis.

References

ADS: Archaeology Data Service. Retrieved July 2, 2013, from http://archaeologydataservice.ac.uk.

Al-Feel, H., Koutb, M. A., Suoror, H.: Toward An Agreement on Semantic Web Architecture. World Academy of Science, Engineering and Technology (49), pp. 806–810, 2009.

Baader, F., Horrocks, I., Sattler, U.: Description Logics as Ontology Languages for the Semantic Web, 2003.

Baader, F., Sattler, U.: An overview of tableau algorithms for description logics. Scotland, LNAI, 2000.

Bateson, G.: Mind and Nature: A Necessary Unity (Advances in Systems Theory, Complexity, and the Human Sciences). s. l.: Hampton Press, 1979.

Bellinger, G.: Knowledge Management – Emerging Perspectives, www.systems-thinking.org/kmgmt/kmgmt.htm, 2004, last access 13/12/ 2010.

Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, May, pp. 34–43, 2001.

Binding, C., May, K., Souza, R., Tudhope, D., Vlachidis, A.: Semantic Technologies for Archaeology Resources: Results from the STAR Project. In Proc. of International Conference in Computer Applications & Quantitative Methods in Archaeology (pp. 555–563). Granada: Bar Internation Series 2494, 2010.

Bill, R.: Grundlagen der Geo-Informationssysteme, 5., völlig neu bearbeitete Auflage, Wichmann Berlin Offenbach, 2010.

Calvanese, D., Giacomo, G. D., Lenzerini, M., Nardi, D.: Reasoning in expressive description logics. In: Robinson, A., Voronkov, A., eds. Handbook of Automated Reasoning. s. l.: Elsevier Science Publishers, pp. 1581–1634, 2001.

CIDOC CRM: CIDOC CRM. Retrieved July 26, 2013, from www.cidoc-crm.org.

Cripps, P., May, K.: To OO or not to OO? Revelations from Ontology Modeling of and Archaeological Information System. International Conference in Computer Applications & Quantitative Methods in Archaeology. Prato: Conference Publication, 2004.

Cruz, I. F.: Geospatial Data Integration. ADVIS Lab, Department of Computer Science, University of Illinois, Chicago, 2004.

Cruz, I. F., Sunna, W., Chaudhry, A.: Semi-automatic ontology alignment for geospatial data integration. In: Geographic Information Science (pp. 51–66). Springer Berlin Heidelberg, 2004.

da Silva, P. P., McGuinness, D. L., Fikes, R. E.: A Proof Markup Language for Semantic Web Services. Stanford: Stanford University, 2004.

Decker, S., et al.: The semantic web: the roles of XML and RDF. IEEE Internet Computing, pp. 63–74, 2000.

Fonseca, F. T., Egenhofer, M. J.: Ontology-driven geographic information systems. In Proc. of the 7th ACM international symposium on Advances in geographic (pp. 14–19). New York: ACM Press, 1999.

Gruber, T. R.: A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2), pp. 199–220, 1999.

Haarslev, V., Muller, R.: RACER system description. Automated Reasoning, pp. 701–705, 2001.

Houbie, F., Duchesne, P., Maué, P.: Semantic annotations in OGC standards. OGC, 2012.

Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., Stasch, C: Semantic enablement for spatial data infrastructures. In: Transactions in GIS, 14(2), pp. 111–129, 2010.

Kalyanpur, Parsia, Sirin, Grau, Hendler: SWOOP: A web ontology editing browser. Web Semantics: Science, Services and Agents of World Wide Web, pp. 144–153, 2006.

Karmacharya, A.: Introduction to a spatial layer within the Semantic Web framework: A proposition through web application ArchaeoKM, Mainz: FH-MAINZ, 2011.

Karmacharya, A., Cruz, C., Boochs, F.: Spatialization of the Semantic Web. In: Semantics-Advances in Theories and Mathematical Models. s. l.: Intech, 2011.

Karmacharya, A., Cruz, C., Boochs, F., Marzani, F.: ArchaeoKM: Realizing Knowledge of the Archaeologists. Cotbus, Conference Publication, pp. 102–108, 2010.

Karmacharya, A., Cruz, C., Boochs, F., Marzani, F.: ArchaeoKM: Realizing Knowledge of the Archaeologists. Von Handaufmaß bis High-tech III (pp. 102–108). Cotbus: Conference Publication, 2010.

Karmacharya, A., Cruz, C., Boochs, F., Marzani, F.: Integration of Spatial Processing and Knowledge Processing Through the Semantic Web Stack. In Proc. of GeoS (pp. 200–216). Conference Publication, 2011.

Kohr, T., Bruhn, K.-C., Karmacharya, A., Cruz, C., Boochs, F.: An SDI for archaeological data with a RESTful interface to semantically modeled information. Geoinformatik 2013, Heidelberg, 2013.

Motik, Grau, Horrocks, Wu, Fokou, Lutz: OWL 2 Web Ontology Language Profiles, www.w3.org/TR/2009/REC-owl2-profiles-20091027, 2009, last accessed 28/10/2010.

Nguyen, D.T., Nguyen, T.P.-M.: Pellet reasoner based evaluation of OVL (Ontology for Vietnamese Language). s.l.: Conference Publications, pp. 180–184, 2010.

Isaksen, L.: Archaeology and the Semantic Web. Southampton: ePrints Soton, 2011.

Perry, M., Herring, J.: OGC GeoSPARQL – A Geographic Query Language for RDF Data, s.l.: Open Geospatial Consortium, 2011.

Perry, M., Herring, J.: OGC GeoSPARQL – A Geographic Query Language for RDF Data, s.l.: OGC, 2012.

Protégé: Protégé.: http://protege.stanford.edu last accessed 26/09/2012.

Roth, M., Bröring, A., Maué, P: Converting SDI Contents to Linked Data with Geographic Feature Pipes. Geoinformatik 2013, Heidelberg, 2013.

Shearer, R., Motik, B., Horrocks, I.: HermiT: A highly-efficient OWL reasoner. Karlsruhe, Conference Publication, 2008.

Simon, R., Barker, E., Isaksen, L.: Exploring Pelagios: a visual browser for geo-tagged datasets. In Proceeding of International Workshop on Supporting Users' Exploration of Digital Libraries (pp. 23–27). Paphos, Cyprus: Conference Publication, 2012.

Stojanovi, N., Handschuh, S.: A Framework for Knowledge Management on the Semantic Web. Hawaii, 2002.

SKOS: SKOS Simple Knowledge Organization System – Home Page. Retrieved July 26, 2013, from SKOS Simple Knowledge Organization System: www.w3.org/2004/02/skos.

Tanasescu, V.: Spatial Semantics in Difference Spaces. Spatial Information Theory, pp. 96–115, 2007.

Tanasescu, V., Gugliotta, A., Domingue, J., Gutiérrez-villarías, L., Davis, R., Rowlatt, M.: A Semantic Web Services GIS based Emergency Management Application. The 5th International Semantic Web Conference (ISWC06). Athens, Georgia: ISWC06, 2006.

Tsarkov, D., Horroks, I.: FaCT++ description logic reasoner: System description. s.l.: Springer-Verlag, pp. 292–297, 2006.

Wüst, T., Nebiker, S., Landolt, R.: Applying the 3D GIS DILAS to Archaeology and Cultural Heritage Projects Requirements and First Results. In: International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences. s.l.: Natural Resources Canada, pp. 407–412, 2004.

Zack, M.H.: Managing Codified Knowledge. Sloan Management Review, 40(4), pp. 45–58, 1999.

Authors' addresses

Ashish Karmacharya | Tobias Kohr | Kai-Christian Bruhn | Frank Boochs
Institut i3mainz, am Fachbereich 1 – Geoinformatik und Vermessung
Fachhochschule Mainz, Lucy-Hillebrand-Straße 2, 55128 Mainz
ashish@geoinform.fh-mainz.de
kohr@geoinform.fh-mainz.de
bruhn@geoinform.fh-mainz.de
boochs@geoinform.fh-mainz.de

Christophe Cruz
Laboratoire Le2i, UFR Sciences et Techniques
Université de Bourgogne
B.P. 47870, 21078 Dijon Cedex, France
christophe.cruz@u-bourgogne.fr