

Research Article

Semantic Enablement for Spatial Data Infrastructures

Krzysztof Janowicz

*Department of Geography
The Pennsylvania State University*

Sven Schade

*Institute for Environment and
Sustainability
European Commission – Joint
Research Centre*

Arne Bröring

*52° North Initiative for Geospatial
Open Source Software*

Carsten Keßler

*Institute for Geoinformatics
University of Münster*

Patrick Maué

*Institute for Geoinformatics
University of Münster*

Christoph Stasch

*Institute for Geoinformatics
University of Münster*

Abstract

Building on abstract reference models, the Open Geospatial Consortium (OGC) has established standards for storing, discovering, and processing geographical information. These standards act as a basis for the implementation of specific services and Spatial Data Infrastructures (SDI). Research on geo-semantics plays an increasing role to support complex queries and retrieval across heterogeneous information sources, as well as for service orchestration, semantic translation, and on-the-fly integration. So far, this research targets individual solutions or focuses on the Semantic Web, leaving the integration into SDI aside. What is missing is a shared and transparent Semantic Enablement Layer for SDI which also integrates reasoning services known from the Semantic Web. Instead of developing new semantically enabled services from scratch, we propose to create profiles of existing services that implement a transparent mapping between the OGC and the Semantic Web world. Finally, we point out how to combine SDI with linked data.

Address for correspondence: Krzysztof Janowicz, Department of Geography, 302 Walker Building, The Pennsylvania State University, University Park, PA 16802, USA. E-mail: jano@psu.edu

1 Motivation

Developing and deploying Spatial Data Infrastructures (SDIs) based on OGC services is attractive for two reasons. First, these services are well standardized and their implementations can be tested for conformity. Second, the OGC has defined a top-level interface standard called OWS Common (Whiteside 2007) defining main aspects that are shared by most OGC Web services. Frequent testbeds investigate, report on, and discuss the interoperability between specific services. Both aspects ease the integration of services into SDIs, make them adaptable, and form the basis for their orchestration (Weiser and Zipf 2007).

Services, however, are not built for their own sake but to encapsulate data or processing models. To exchange data between services, i.e. to make them interoperable, they have to share common schemas or translate between them. For example, if one processing service requires a string representing wind direction as input and was developed with a *wind blows from* conceptualization in mind, a second service offering wind direction observations as strings, but based on a *wind blows to* conceptualization, can still act as an input source (Probst and Lutz 2004). The OGC standards guarantee interoperability on a syntactic level. Services can exchange data if they agree on names and types for their inputs, outputs, and operations. Whether data exchanged between services can be interpreted in a meaningful way is not covered by the specifications. For example, a Web Processing Service (WPS) (Schut 2007) can be used to compute the dispersion of a gas plume caused by a factory fire based on wind direction observations delivered by a Sensor Observation Service (SOS) (Na and Priest 2007). Both services need to share a common understanding of wind direction to compute meaningful results (Probst and Lutz 2004, Bröring et al. 2009); otherwise, the simulated dispersion plume would point in the opposite direction. Hence, the challenge is to establish semantic interoperability, i.e. the ability of services to exchange data in a meaningful way and with a minimum of human intervention (Harvey et al. 1999, Manso and Wachowicz 2009). In this article, we propose a transparent vertical and horizontal Semantic Enablement Layer (Janowicz et al. 2009) for spatial data infrastructures that supports the required functionality.

The remainder of the article is structured as follows. First, we introduce previous work on geo-semantics related to OGC services. We then discuss semantic challenges for SDI in general. Next, we introduce the idea of a transparent horizontal and vertical Semantic Enablement Layer and its integration with OGC services. We stick to the gas plume dispersion example throughout the article as a running scenario. We conclude our work by summarizing the proposed approach and pointing to further work, such as the idea of a micro-SDI for linked spatiotemporal data.

2 Related Work

Over the last years, work on semantics (Kuhn 2005) and geo-ontologies has focused on semantic interoperability between OGC services. This includes work on the role of ontology for spatio-temporal databases (Frank 2003), the notion of semantic reference systems and the grounding of geographical categories (Kuhn 2003, Probst 2007, Scheider et al. 2009), semantics-based and context-aware retrieval of geographic information (Janowicz et al. 2007, Lutz and Klien 2006, Keßler et al. 2009a, Schade et al. 2008,

Maué and Schade 2009), ontology alignment (Cruz and Sunna 2008), as well as work on Semantic Geospatial Web services (Roman and Klien 2007) and their chaining (Lemmens et al. 2006, Fitzner et al. 2010). This research has led to tools such as ConceptVISTA (<http://www.geovista.psu.edu/ConceptVISTA/>) for ontology creation and visualization, the Concept Repository (<http://purl.org/net/concepts/>), the SIM-DL similarity server and Protégé plug-in (<http://sim-dl.sourceforge.net/>), the semantically-enabled Sensor Observation Service SemSOS (Henson et al. 2009), the sensor observable registry (Jirka and Bröring 2009), or the OWL application profile for the OGC Web Catalogue (CSW) (Stock et al. 2009).

Opposed to work on SDIs, these services do not share common interfaces. They are isolated solutions which lack a binding to each other and partially to OGC Web services. For instance, the SIM-DL server computes the similarity of geographic feature types. It depends, however, on an extended version of the Description Logics Interface Group (DIG) protocol for communication and the Web Ontology Language (OWL) for knowledge representation². OGC-compliant Web services such as the Web Feature Service (WFS) use GetCapabilities requests and the Geographic Markup Language (GML).

Recent approaches to enrich SDIs with semantics are coupled to a specific technology. An OWL-Profile for CSW suggested by Stock et al. (2009) depends on implementations in eBRIM and is restricted to the ontology language OWL. The registry proposal of Jirka and Bröring (2009) is even more restricted, namely to features observable by sensors. In contrast, we propose a transparent approach which abstracts from a particular inference engine and ontology language such as OWL, OWL 2.0, Web Service Modeling Language (WSML), or Topic Maps.

3 Semantic Challenges for Spatial Data Infrastructures

Misunderstanding and incorrectly using geographic data can be usually traced back to missing or unclear descriptions of their intended interpretation (Guarino 1998). Interest in semantics and reasoning for complex tasks such as geospatial decision making (Maué and Schade 2009) or retrieval (Janowicz and Wilkes 2009) is growing. Semantics can support decision makers to identify potential solutions and alternative paths. Reasoners embedded in workflow engines automatically select and process potentially relevant data to finally represent the results on a decision-support map.

SDIs are designed as service oriented architectures. Within such infrastructures functionalities such as storage and retrieval are realized by Web services. Complex workflows can be established by coupling such services. A typical compound activity includes the discovery and download of relevant geospatial data, applying pre-processing and appropriate analysis methods, and finally rendering the results on a map. Catalogues can be used to discover resources published in an SDI according to the CSW standard (Nebert et al. 2007). The access to geospatial data depends on the underlying format. Coverages (multi-dimensional fields modelling one attribute's variation over space and time) are provided by OGC's Web Coverage Service (WCS) (Whiteside and Evans 2008). Datasets comprising features with an open range of attributes are managed and offered by the OGC's Web Feature Service (WFS) (Vretanos 2005). In this sense, the Sensor Observation Service (SOS) can be considered as a specialization of the WFS restricting the served features to sensor observations. Processing data, e.g. running a spatial analysis or interpolation algorithm, is accomplished by Web Processing Services (WPS) (Schut

2007). The result of the process should³ still either be a feature set or a coverage, formatted according to a standardized data encoding. The Web Mapping Service (WMS) (de la Beaujardiere 2003) loads the data coming either directly from a data service (e.g. WFS, WCS or SOS) or a processing service. It renders a visualization of the data as a map as well as a corresponding legend and finally returns it in a common image format.

The different Web services have individual semantic challenges (identified below), but most semantic problems arise due to the lack of meaningful descriptions of the actual content. It has to be discovered, downloaded, processed and visualized: each Web service interacts in some way with data. Most semantic conflicts during a workflow appear if source data not been sufficiently specified in the beginning and the arising ambiguities are propagated through the whole workflow. Feature-based content is encoded in the XML-based OGC Geography Markup Language (GML) (Portele 2007) or, for simple cases, OGC KML (Wilson 2008). Specific GML Profiles like SensorML (Botts 2007) or Observation and Measurements (Cox 2007) extend GML with application-specific details. OGC standards do not restrict the WCS or WMS interface to certain data formats. However, OGC recommends common formats for coverages and maps, but developers are free to adapt them. The semantics of geospatial data do not depend on their format. Semantic descriptions are needed for all types of geospatial data to ensure their correct interpretation. Consequently, there is a need for techniques to propagate semantics through workflows. For instance, a WMS located at the end of a workflow chain has to be able to correctly interpret and visualize the results according to the semantic descriptions of the underlying data.

3.1 Semantic Challenges for Geospatial Data

Feature-based geospatial data are typically stored in spatially-enabled object relational databases and provided via the WFS interface. The data model reflects the various feature attributes and topological relations. The different entities can functionally depend on each other (the value of one attribute depends on the value of another). Explicit descriptions of such inner relationships or the intended use of the data may help to avoid semantic conflicts. However, the application schema alone is not sufficient to grasp the meaning of the underlying data model. The labels identifying the different data entities are often ambiguous; application-specific knowledge and semantic heterogeneities impair their correct interpretation (Maué and Schade 2009). Semantic annotations linking the feature types or instances to explicit and shared conceptualizations support the clarification of such ambiguities. Section 4.1 contains an example of an observation result, which is semantically annotated to clarify the provider's understanding of wind direction.

Geographic information represents geographic space, ontologies are conceptualizations of our common understanding of this space. To describe and understand this relation, it is necessary to understand the ontologies to which the data have been linked. Wind direction can be modelled and offered to clients in various ways: as near real-time observations coming from an SOS, as an attribute of a weather station feature hosted by a WFS, or as phenomenon varying over space and time modelled as a coverage provided by a WCS. Our understanding of wind direction is independent from the representation of the data. Hence, the challenges identified above for feature-based geospatial data also apply for coverages. As long as well-defined anchors within the metadata can be used to inject semantic annotations, different representations of geographic information can be

semantically enabled. Retrieval of individual features can also be supported by semantic annotations on the instance level. The actual features, or the placemarks in OGC KML files, can then be linked to instances in shared ontologies.

3.2 Semantic Challenges for Geospatial Activities

Above we identified the five most typical activities performed within SDIs: finding, accessing, updating, processing, and visualizing geospatial data. Semantics usually refer to the actual data, which represents real world entities and phenomena. In the following, the core semantic challenges and core benefits are discussed in more detail for each activity. Note that the different activities cannot be isolated from each other. Semantic conflicts arise during the combination of the various workflow elements, semantic propagation can ensure that changes to the data set's original intended meaning, e.g. by a WPS, are also forwarded and communicated to the end of the workflow. We consider the last step – usually the WMS responsible for rendering the data – as a sink where semantics and its changes of all input sources have to be aggregated, interpreted, and visualized in a meaningful way.

Discovery of geospatial data in an SDI is usually managed by catalogues, which enable the registration and discovery of data, Web services and other relevant documents. The retrieval of information is a multi-step process, starting with the user's task to formulate her information need as a query, processing of the query, finding and returning matching metadata in the repository, and finally evaluating the results. Semantics can help in each of these steps. Information Retrieval (IR) systems can support users to formulate queries by recommending appropriate concepts from ontologies after analyzing the already typed in query (more sophisticated techniques may analyze the user's context such as her current location (Keßler et al. 2009b)). Free text queries do not necessarily depend on semantic annotations; query expansion techniques can, for example, add other suitable search terms such as synonyms to improve the potential recall. Semantic queries are directly formulated in formal languages and forwarded to reasoning engines which then return matching records according to the semantic annotations. Semantic queries encoded, for example, in the Semantic Web Rule Language (SWRL) have to be combined within traditional catalog queries based on OGC Filter Encodings (ISO/TC211, 2009). These rules can more precisely represent the user's information needs. The chaining of Web services, e.g. a WPS expecting sensor observations of wind direction from an SOS to compute gas dispersion, also relies on semantically supported discovery. An approach for ontology-based descriptions of geoprocesses is presented by Lutz and Klien (2006). The expected input of the WPS can be regarded as the user's goal which has to be compared with the outputs of registered Web services. Again, reasoners compare semantic annotations of the outputs of registered Web services with the goal of finding matching candidates (Fitzner et al. 2010).

Access to geographic information within an SDI is managed by Web services such as the WCS, WFS, or SOS. They provide effective filtering and management techniques for the served data. However, they do not have an effect on the semantics of the data but simply return content matching the request parameter (such as the spatial or temporal extent). The Web services have to ensure semantic propagation; the results coming from a database have to be extended with semantic annotations to ensure that subsequent activities can benefit from semantics. In the end, it should not make a difference whether

the processed data has originally been downloaded as a file or was retrieved from a Web service.

Registration of geospatial datasets, features, coverages, sensor observations and so forth, can be supported by specific Web service operations. A Web service has to preserve certain aspects of data quality, such as logical and conceptual consistency (ISO/TC211 2002) of the registered data with respect to the represented real world entities. If a new street feature is added to a WFS-T (transactional WFS) offering a street network, the Web service has to make sure that the structure and semantics of the new feature matches the present feature set. The new feature may be marked as Highway, but the values of the attributes, e.g. number of lanes, may contradict with rules in the ontologies. Semantically supported integrity checks applied during the registration of new data can test for such heterogeneities, and either reject or automatically transform these datasets. Registering new sensor descriptions into an SOS raises other semantic challenges which have been recently discussed by Bröring et al. (2009). Here, the main difficulty lies in the dependencies (and accordingly inconsistencies) between sensor, observation, and the feature of interest. The relation between the real world entity and its computational artefact, the so-called feature of interest, can be inconsistent. If two sensors of different type deliver observations assigned to a particular feature of interest in an SOS, do they both refer to the same real world entity? The origin of this challenge lies in the symbol grounding problem discussed by Harnad (1990). The second challenge arises during the selection of an appropriate sensor. Its purpose is to observe certain characteristics of a real world phenomenon; the sensor inputs have to match these characteristics. Third, the sensor output has to comply with the property of the feature of interest stored in the SOS. A purely syntactic matching is not sufficient to avoid such inconsistencies (Bröring et al. 2009).

Processing of geographic information is managed by Web services compliant with the OGC Web Processing Service interface. An atomic process can be understood as a transformation of geospatial data based on well-defined functions. Such processes may change the form of representation: an interpolation of point-features such as sensor observations produces a continuous coverage. Classification of raster-based data such as satellite images can result in features, e.g. polygons sharing a common property. Processes based on either Tomlin's map algebra (Tomlin 1990) for continuous raster-based data or traditional geometrical functions for features such as merging or intersecting, combine datasets. A process understood as a mathematical function applied to geometries and attributes should not be confused with real processes in geographic space; it is purely syntactic, mapping an input to an output according to certain rules. The semantic challenge here is therefore not to describe what the process means, but to understand how the intended meaning of the output compares to the semantics of the input. A WPS computing a risk map based on wind directions clearly changes the semantics of the served geographic information. Fitzner et al. (2010) make use of functional descriptions based on Datalog to represent this relation between input and output.

Visualization or rendering of geospatial data is traditionally the last activity in complex workflows. However, a simple visualization of geodata in a WMS also unveils semantic challenges. The map layers served by the WMS contain the whole range of cartographic symbolization as well as more sophisticated 2D geo-visualizations. In association with those visualizations, the WMS is capable of offering explanatory legends for each layer and feature instances via its *GetFeatureInfo* operation. If a WMS is set up to generate maps based on data coming from a WFS or WCS it has to be aware of the

semantics of the received data to render meaningful visualizations. Additionally, it has to be aware of the semantics of the styles and symbols it can apply to such data. For example, a WMS should not draw cross icons into a map for all kinds of religious facilities, such as churches, synagogues, or mosques. Mapping services allow users to define the symbolization of the geodata. The Styled Layer Descriptor (SLD) (Lupp 2007) standard has been defined for this purpose. A semantic annotation of SLDs could clarify the meaning of styles offered to a user. An application can make use of such annotated SLDs to recommend specific styles for particular feature types or applications.

4 Towards a Transparent Semantic Enablement Layer

To integrate Semantic Web services into SDI we propose a transparent Semantic Enablement Layer (SEL) for OGC services. It resides on top of recent standards and considers the following three challenges: (1) How to link data encodings and service protocols to formal specifications stored within ontologies?; (2) How to manage and maintain these ontologies?; and (3) How to incorporate reasoning services known from the Semantic Web?

Based on these challenges we can derive functionalities, which should be provided by the SEL (see Table 1). For further structuring, we categorized atomic functionalities (not to be confused with normative service operations) into four conformance classes. *Storage* groups functionalities which are required for ontology storage, evolution, and access. The functionality to connect elements of a specific resource, e.g. a GML or RDF data model, with concepts or instances from an ontology is provided by the *Lookup and Retrieval* conformance class. *Reasoning* groups operations about inferring hidden facts as well as adding new ones, while the *Deployment* functionality supports the deployment of OGC services if their data models have been encoded in ontologies. Such deployment includes the generation of a content description for an OGC Web service which is advertised in its capabilities document, as well as an automated creation of descriptions of resources such as feature type serializations using XSD. This functionality ensures explicit linkages between services and content descriptions.

We propose to group the functionalities of the conformance classes in two services, the Web Ontology Service (WOS) for managing and accessing ontologies and the Web Reasoning Service (WRS) for providing reasoning functionality within SDIs. Instead of creating new services from scratch, the WOS is defined as a profile of the Web Catalog Service and the WRS as a profile of the Web Processing Service. This facilitates the integration with existing SDI technologies and simplifies the service orchestration. As WRS and WOS have to follow the OWS Common specification, a major challenge is the mapping between the protocols and representation languages used on the Semantic Web and in the OGC world. Note that we do not propose to develop separate reasoners or ontology repositories for SDI but to transparently encapsulate existing Semantic Web solutions by the WOS and WRS.

Components for authoring semantic annotations, i.e. which support users to link elements from data model to concepts in an ontology, are not considered here. Examples of annotation and authoring tools are given in Grcar (2008). The theory of semantic annotations of OGC-compliant content is explained in more detail in the following subsections. We illustrate the integration of the proposed services into SDI using the gas plume example. We assume the emissions resulting from a factory fire endanger an

Table 1 Overview of SEL functionalities

Conformance Classes	Operation
Storage	<i>create ontology</i> : creates/uploads a new ontology with all its classes and relations inside the repository <i>update ontology</i> : registers a new version of an ontology to the repository <i>get concept, get relation, get ontology</i> : returns different types of elements from a registered ontology
Lookup and Retrieval	<i>get model reference</i> : returns the appropriate ontology element ID for a given resource ID, e.g. GML Feature ID <i>retrieve</i> : executes semantic matchmaking between a goal/query and (1) available Web service advertisements and (2) feature type definitions
Reasoning	<i>load ontology</i> : loads a specific ontology into the reasoner. <i>release Ontology</i> : removes a specific ontology from the reasoner <i>tells</i> : inserts a new fact into the knowledge base <i>asks</i> : returns facts from the knowledge base
Deployment	<i>create capabilities</i> : creates content-specific section of an OGC Capabilities Document <i>create feature type description</i> : creates a GML feature type in XSD format (created file may contain annotations)

important European bird sanctuary, the so-called Rieselfelder in Münster (Germany), as well as the surrounding natural reserves. For reasons of simplification, we further assume that a local Sensor Web is already set up and used by a disaster relief organization, i.e. mobile sensors are deployed to monitor air pollutants, wind speed, and wind direction.

4.1 Semantic Annotation

Application-specific data models describe dependencies between data entities. A relational data model is an appropriate choice for local access and storage of data. Applications bundled with data do not need descriptions of how and why to use the data. Applying methods to ensure interoperability only makes sense if sharing data across applications is desired. Standards like the XML-based GML for feature-based geospatial data enable syntactic and structural interoperability between different applications, they are not meant to be used within the applications. The same is valid for semantic interoperability. The description of the semantics is not an intrinsic feature of geospatial data; the references to concepts from external vocabularies are not part of the features in the database. The idea of semantic annotations preserves this clear separation between real world semantics and application-specific data models. Figure 1 illustrates how metadata for geographic information served by an SOS can be extended with a reference to external domain ontologies.

Building on the work of Klien (2007) and Verma and Sheth (2007), Maué et al. (2009) proposed a methodology for the annotation of OGC-compliant Web services. Metadata for geographic information served by Web services exists on multiple levels: on the first (and most generic) level, references are added to descriptions valid for the whole data set or Web service, for example by adding them to the keywords section in the OGC Capabilities document. The second level covers the data model, with the goal to rebuild and explain the inner relationships and dependencies between different aspects of a feature. Figure 1 illustrates different options to semantically annotate an SOS serving current values for wind direction. As mentioned in Section 3, semantics can help to clarify (and communicate) the dependency between the observation procedure and the observed property. The OGC Standard for Observation and Measurements (Cox 2007) defines how to encode observations, for which Figure 2 serves as an example.

Figure 2 shows an observation delivered by an SOS serving wind direction values. The procedure is referenced to the instance *Anemometer01* within an information source ontology. Application-specific details such as the data provider’s perception of wind

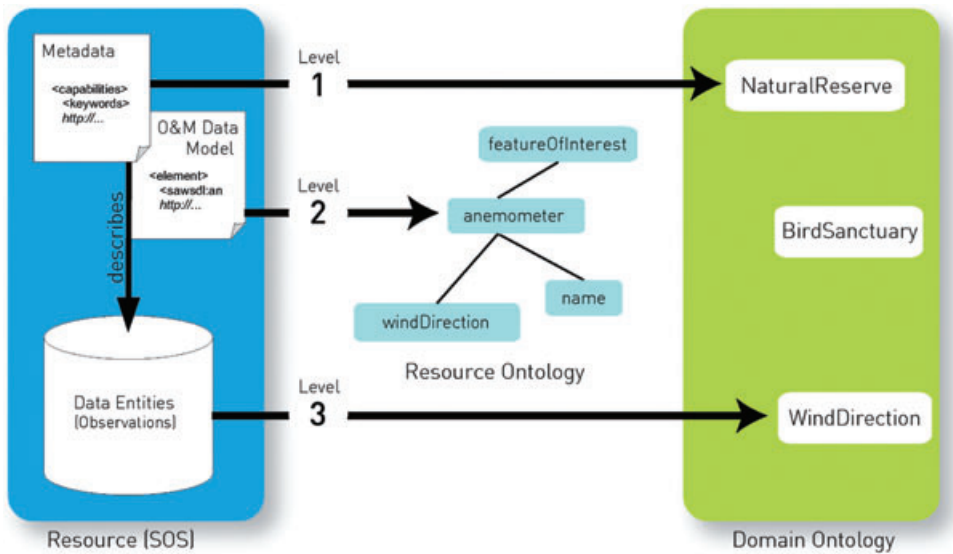


Figure 1 Levels of annotations; adopted from Maué et al. (2009)

```
<om:Observation xmlns:xlink="http://www.w3.org/1999/xlink">
  <om:procedure
    xlink:href="http://myWOS.com#Anemometer01"
    xlink:role="http://www.w3.org/ns/sawsdl#modelReference"/>
  <om:observedProperty
    xlink:href="urn:ogc:def:phenomenon:WindDirection"/>
  <om:featureOfInterest
    xlink:href="http://myWFS.com/request=getFeature&
    featureId=f23"/>
  <om:result xsi:type="gml:MeasureType"
    uom="urn:ogc:def:uom:OGC:deg"> 231 </om:result>
</om:Observation>
```

Figure 2 Example annotation of an SOS GetObservation request

```

<element name="resFeature" type="resType"
  sawsdl:modelReference="http://myWOS.com#NatureReserve">
<complexType name="resType">
  ...
  <element name="resName" type="string"
    sawsdl:modelReference="http://myWOS.com#Identifier">

```

Figure 3 Adding semantic annotations to a simple feature type

direction are explicitly described here. References within this local ontology point to a common OGC vocabulary of phenomena, which is also used to identify the observed property. The real-world observation target is represented as the feature of interest, served by a WFS as GML. Figure 3 contains the potential and simplified GML application schema for this particular WFS and illustrates how feature types are annotated. In this case, features of this type represent natural reserves, and have an identifying name.

Both data such as observation results and schemata such as a feature type description are dynamically generated. Semantic annotations are not an intrinsic part of the data. The actual references to the ontologies are stored at a different location. The software component responsible for creating OGC-compliant metadata documents has to dynamically *inject* the links during the serialization. An external lookup component maps unique identifiers of XML elements to an URI pointing to terms in shared vocabularies. The element identifiers comprise the URL of the resource serving the XML document and the XPath expression to identify the element within the document. The type of annotation depends on the document format. Elements in XML schema such as in Figure 3 can be referenced using the SA-WSDL standard. XML dialects such as GML or Observations & Measurements (O&M) (see Figure 2) usually have predefined extension points where links to external metadata documents are allowed. Maué et al. (2009) discuss the various extension points in existing OGC standards.

Pushing the links into the metadata can be complex. *Sapience*, the open source API for semantic annotations (<http://purl.org/net/sapience/docs/>) comprises a set of Java libraries which manage the lookup and injection of references into known metadata documents. The Web service developer simply forwards the serialized XML document to *sapience*, which adds the links to the appropriate locations within the document and returns the updated document to the source application. The annotations have been looked up in a database. *Sapience* does not support the authoring of annotations. External editors, potentially supported by data mining techniques, let data providers specify the annotations. As an example, imagine a user just set up a new WFS using a generic software package (supporting *sapience*). He or she configures one feature type, using a PostGIS database as source. The WFS automatically generates the feature type schema requested by the user. References to an external domain ontology can be added using an external editing tool. By uploading this document to *Sapience* and pushing it into its lookup database, the WFS connected to *Sapience* will add semantic annotations to its metadata. In this case, a simple call of the *Sapience* API activates the Semantic Enablement of the Web service.

4.2 Web Ontology Service: Managing and Accessing Ontologies

Annotations link elements within data or service models to concepts, individuals, and relations in ontologies. Such ontologies are typically stored in repositories. Existing

repositories⁴ provide auxiliary capabilities such as fine-grained access through Web services, querying, visualizing, versioning, editing, and even reasoning. Technologies used for reading and querying ontologies are largely based on established W3C standards for the Semantic Web. With regard to SDIs built on OGC services, a decomposition of the functionality into separate Web services is more appropriate (comparable to the separation of WFS and WMS). Embedding existing ontology repositories such as the concept repository CORE into SDIs requires a transparent solution which serves as a proxy between Semantic Web interfaces and those used in the OGC world.

The proposed OGC-compliant Web Ontology Service provides access, lookup, and retrieval functionalities. It encapsulates existing ontology repositories (or even simple text files containing the ontology definitions). A WOS can serve ontology definitions for different types of geographic features, processes, observations, and sensors. In the case of the gas plume example, a WOS contains feature types such as *Factory*, *NaturalReserve*, and *InhabitedPlace*, as well as sensor types such as *Anemometer*. The WOS serves the formal specifications for semantic annotations. Coupled with a Web Reasoning Service, a WOS can support semantics-based discovery of resources. In this sense, a WOS is a semantically-enabled catalogue supporting information retrieval beyond simple keyword search (Lutz and Klien 2006, Janowicz et al. 2007). Therefore and in conformity with Lieberman et al. (2006), we argue that a WOS should be designed as a profile of the OGC Catalogue Service (CSW) (Nebert et al. 2007). Thereby, it abstracts from spatial and temporal search, while focusing on thematic aspects. As ontologies need specific querying languages, the filter encoding standard (ISO/TC211 2009) requires an additional profile.

Using the gas plume scenario, Figure 4 illustrates how the WOS can support the transparent gathering of relevant data, e.g. sensor observations. A WOS is queried for all subtypes of *NaturalReserve* which are located within a particular bounding box, e.g. the greater Münster area. To process such a query, the WOS utilizes an associated Web Reasoning Service. The WOS response contains all feature types satisfying the input query, e.g. *Bird Sanctuary*. These types can be used in further discovery tasks to find features of interest affected by the fire and gas plume.

We identified different options for WOS development, a thick and a thin version. The thick WOS enables retrieval of resource descriptions and at the same time semantic matchmaking for data and services. If matchmaking should be performed in addition, an extended form of filter encoding has to be used as catalogue input in conjunction with a Web Reasoning Service. It is up to the implementation, whether the components are tightly- or loosely-coupled. Figure 5 illustrates the transparent encapsulation of an RDF repository.

4.3 Web Reasoning Service: Bringing Reasoning to SDI

While the Web Ontology Service encapsulates ontologies, a second service has to encapsulate the functionalities defined in the reasoning conformance class. This service has to bridge between the inference engines as key components of the Semantic (Geospatial) Web and the OGC world. Reasoners are not restricted to subsumption reasoning, but include non-standard inference such as finding the most specific concept, least common subsumer, similarity reasoning (Janowicz et al. 2007), as well as context-aware instantiation based on SWRL rules and built-ins (Keßler et al. 2009a). We argue that such a Web Reasoning Service should be developed as a profile of the Web Processing Service specification (Schut 2007). Since the WRS should encapsulate Semantic Web reasoners

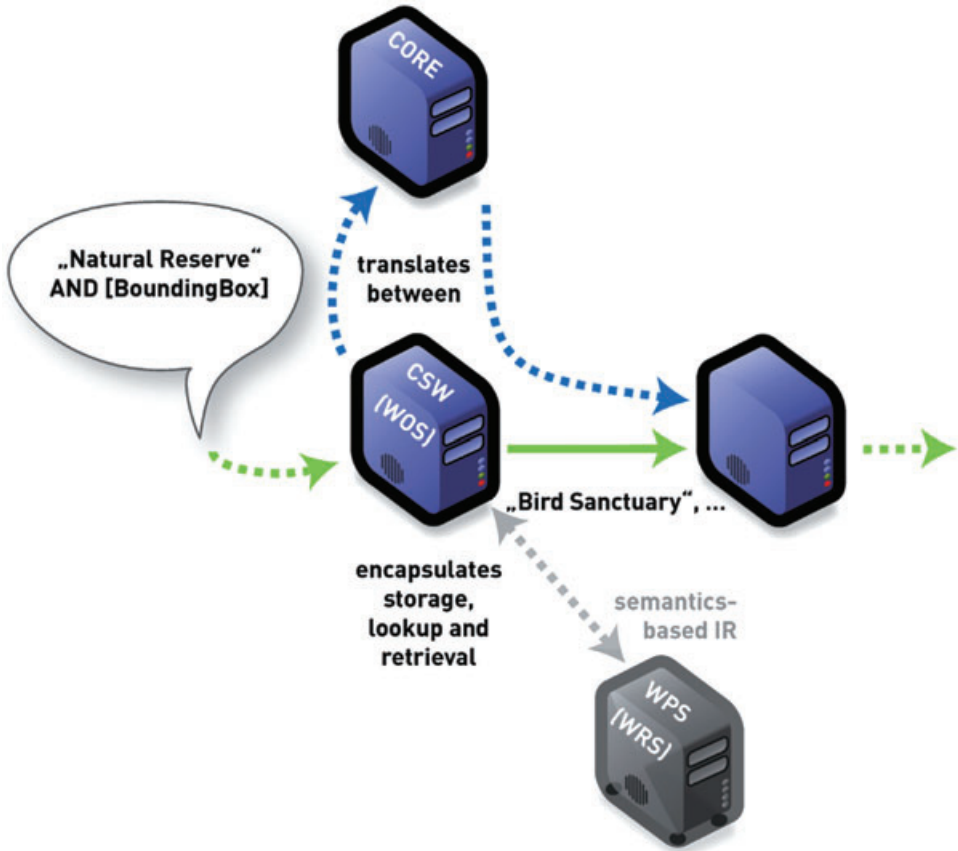


Figure 4 Transparent integration of existing ontology repositories into an SDI to support semantics-based lookup and retrieval using the Web Ontology Service as a CSW profile

and make them accessible for SDIs, it has to map in both directions between *DIG tells and asks* calls on the one side and *GetCapabilities* request and GML on the other side.⁵

With respect to Sensor Web Enablement (SWE), a WRS could be used to discover appropriate sensors using a feature of interest as query (Bröring et al. 2009). For instance, a semantically-enabled SDI could automatically choose and register sonic anemometers if the user is interested in data on the dispersion of a gas plume. In the case of semantics-based retrieval of feature types (Lutz and Klien 2006, Janowicz et al. 2007) as depicted in Figure 4, the WRS provides the necessary reasoning power for the WOS.

Figure 6 illustrates how the WRS can be used to incorporate reasoning services into an OGC service chain. With respect to the gas plume example, the WRS encapsulated the SIM-DL similarity server (Janowicz et al. 2007) to select features similar to the *Rieselfelder* in the greater Münster area, e.g. the *Wienburgpark*. Next, an SOS is used to access sensor observations about the potential pollution of these features. Finally, a Web processing service delivers a risk analysis.

The WRS provides encapsulation in the sense that GML instances can be used as input (for example a data entry representing the *Rieselfelder*) for the various supported

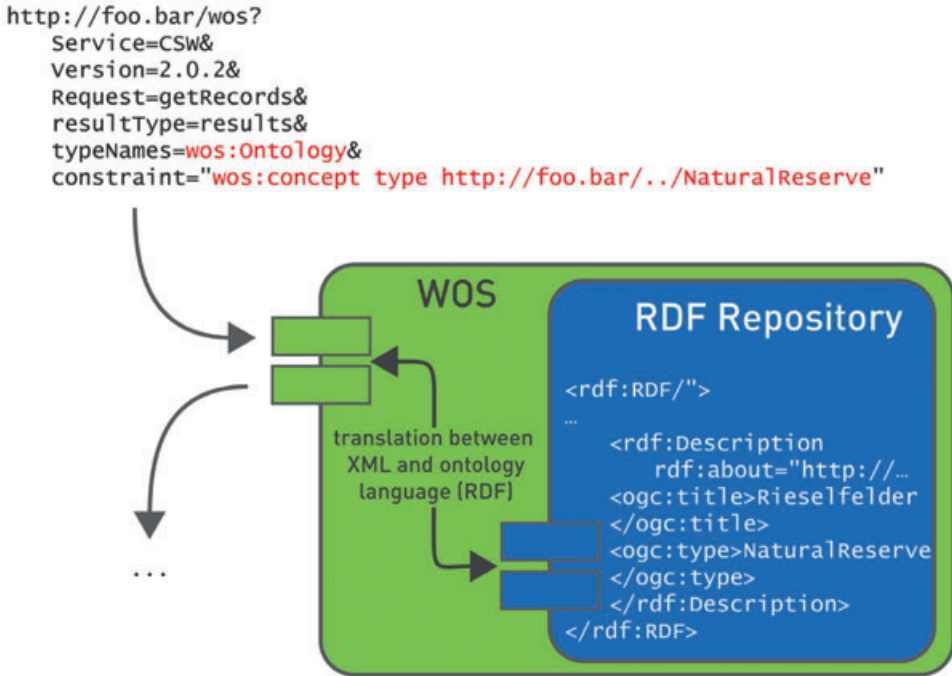


Figure 5 Encapsulation of an RDF repository in a Web Ontology Service. The WOS receives CSW-compliant requests and translates them to the language supported by the encapsulated repository – in this case RDF or SPARQL, respectively

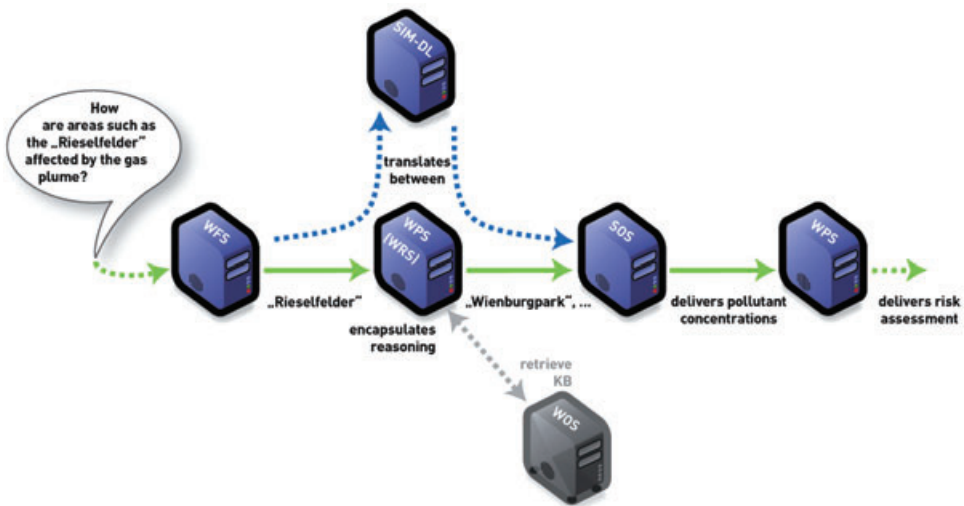


Figure 6 Transparent integration of existing reasoners into an SDI using the Web Reasoning Service as a WPS profile

reasoning mechanisms. Each mechanism is provided as an executable WPS process. A concrete setup for an encapsulation of a reasoning service as a profile of a WPS is shown in Figure 7. The process offered in this example is the retrieval of *similar* features, i.e. GML features annotated with similar types as the input feature. The WPS compliant request is shown at the top of the figure, with the relevant parts highlighted in red: The identifier for the triggered process is *Similarity*, the input GML file is referred to via *href*. As an additional input, we specify the context in which similarity is to be computed, which is in this case reduced to *EnvironmentalFeature*; see Janowicz et al. (2007) for details. The WRS' functionality consists in translating this request to DIG compliant calls that are forwarded to the SIM-DL similarity server. The SIM-DL server's interface extends the DIG interface with functions for similarity reasoning (Janowicz et al. 2007). As a first step, the SIM-DL server retrieves the input feature's type – in this case, *NaturalReserve* – using the feature type ontology provided by the WOS. The `<ccsimilarity>` request then takes this as the source concept and calculates a list of target concepts with similarity values. This query is restricted to subconcepts of *EnvironmentalFeature*, as specified in the WRS request. Finally, the SIM-DL server retrieves all instances of the similar concepts computed in the previous step. This DIG compliant list

```

http://foo.bar/wrs?
  Service=WPS&
  Version=1.0.0&
  Request=Execute&
  Identifier=Similarity&
  DataInputs=[InputPolygon=@xlink:href=http://foo.bar/Rieselfelder.xml;
    Context=EnvironmentalFeature]&
  RawDataOutput=[SimilarFeatures]
  
```

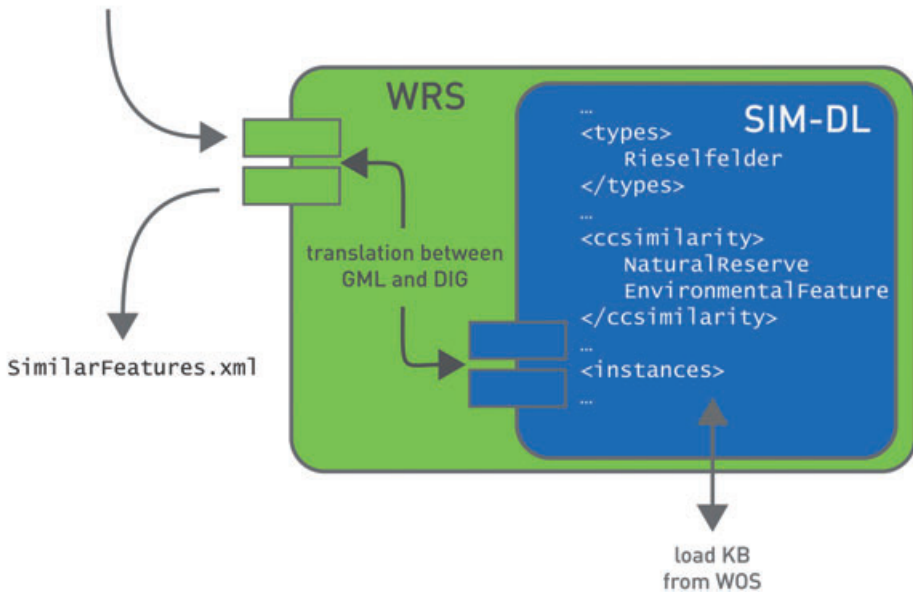


Figure 7 Encapsulation of the SIM-DL similarity server in a Web Reasoning Service. The WRS as a profile of the OGC Web Processing Service receives WPS-compliant requests and translates them to the DIG interface provided by the SIM-DL server

of instances is then translated to GML features by the WRS again and returned to the client in the file *SimilarFeatures.xml*, as specified in the WRS request.

5 Linked Data and Micro-SDI

An interesting question is whether and to what degree OGC services will co-exist with upcoming Semantic Web technologies and especially with linked data infrastructures (Bizer et al. 2010). While this is difficult to predict, we assume that both approaches do not exclude each other. For instance, one could think of a micro-SDI (μ SDI) for light-weight linked spatiotemporal data⁶ applications and still keep the established OGC services for more complex applications. Note that this is not a technical discussion about the long lasting conflict between the two camps supporting either the WS-* technology stack (SOAP, WSDL, WS-Addressing, WS-Security, etc.) or RESTful Web services. It is about a general paradigm shift. A μ SDI should consist of simplified and lightweight OGC services which can be directly embedded into Web pages and applications. In most cases one may think of the micro services as simplified 1 : 1 correspondences of classical OGC services; however, some of them will probably have to be split up in multiple other services. Some OGC developments such as the decomposition of the Sensor Alert Service and Web Notification Service already point in this direction.

Examples towards establishing such a μ SDI include recent work on next generation gazetteers (Kefler et al., 2009a), a linked data serialization of OpenStreetMap (Auer et al. 2009), or JavaScript reasoners such as JSExplicit (<http://jsexplicit.sourceforge.net/>) which can be directly embedded into Web pages to generate context and user-aware information from RDF or OWL data on-the-fly. For instance, instead of adding a static 'How to reach us'-page to a hotel description, one could integrate JavaScript calls to μ SDI services which use OpenStreetMap and public transport data. The website could then provide its users with a list of up-to-date public transportation opportunities by combining a query for all instances of subtypes of *Transportation* within a particular distance from the hotel. Note that the query is directly embedded into the HTML code of the webpage and executed in the user's browser – making it context aware (e.g. using the geolocation API in Firefox, language settings, and so forth).

The transparent encapsulation services proposed in this work can also act as proxies between SDIs and linked data. For instance, one could create linked data on-the-fly from exiting SWE services such as Sensor Observation Services or Sensor Alert Services.

6 Conclusions and Further Work

We outlined the need for a Semantic Enablement Layer for OGC Web services. We argue that it is a prerequisite for a semantically supported discovery of geospatial content tailored to the user's context, semantic translation, dynamic orchestration of sensors and Web services, and eventually semantic interoperability. We introduced four conformance classes: storage, look-up and retrieval, reasoning, and deployment. Two new Web service interfaces – developed as profiles of existing standards – implement a Semantic Enablement Layer for OGC services. Our proposal is considerably different from previous suggestions, especially with respect to the following:

- Unlike the SemSOS (Henson et al. 2009) and the registry for sensor observables (Jirka and Bröring 2009), the WOS and WRS concepts are applicable to any kind of OGC service and content.
- In contrast to previous approaches (Roman and Klien 2007, Stock et al. 2009), the proposed solution does not rely on a specific ontology language.
- Unlike the solution proposed by Roman and Klien (2007), the WRS provides a means to the complete encapsulation of reasoning. This is an important benefit considering the variety of application requirements.
- We do not rely on a specific reasoning engine or technology for ontology repositories, as for example in the approach introduced by Stock et al. (2009).

Three steps towards establishing an SEL have been identified. First, data encodings and service protocols have to be linked to formal specifications stored in ontologies using semantic annotations. Second, a service has to be established for managing and maintaining these ontologies. Third, Semantic Web reasoners have to be encapsulated to integrate them into SDIs. Supporting services, such as the WOS and WRS can be integrated into SDIs without changing existing clients. The proposed approach generalizes over previously suggested solutions and provides a tight (and transparent) integration into recent SDI developments. We also clearly separate data models (in any encoding) from domain ontologies. This separation acknowledges the distinction between information items and the real world.

While we focused on introducing the need for and components of the Semantic Enablement Layer, the reference implementation of the WOS and WRS is part of the 52°North semantics community (<http://www.52north.org/semantics/>). Currently, our work on the WRS focuses on the encapsulation of the SIM-DL similarity server and Pellet reasoner to make them accessible for OGC services such as the SOS and WFS (see <http://52north.org/svn/semantics/WRS/> for additional details). A semantic annotation API for the lookup and injection functionality is developed in the *sapience* project. Part of this project is also the CORE concept repository, which is planned to be encapsulated in the WOS. Adding annotations on-the-fly to existing OGC metadata is required as long as the data models are not represented as ontologies within the WOS. In the long-term, the functionality described by the deployment conformance class will enable the creation of parts of the Web service capabilities. Evaluation of thick versus thin WOS implementations is a subsequent step.

Finally, with the increasing popularity of linked data, the development of a common (and minimalistic) geo-vocabulary focusing on more than just topological relations becomes even more important. Instead of trying to agree on a common conceptualization of geographic features, the aim should be to develop an affordance/action-based domain level which allows mappings between local vocabularies (Janowicz and Keßler 2008).

Acknowledgements

The presented work is funded by the International Research Training Group on *Semantic Integration of Geospatial Information* (DFG GRK 1498), the DFG *SimCat II* project (DFG Ja1709/2-2), the BMBF *GDI-Grid* project (BMBF 01IG07012), the EC-funded projects *SWING* (FP6-026514) and *GENESIS* (FP7-223996), as well as the 52°North semantics community which aims at establishing a Semantic Enablement Layer for OGC services.

Endnotes

- 1 This is still a working definition as it does not define when a combination of data is considered to be *meaningful*.
- 2 This is not only the case for SIM-DL but most reasoners on the Semantic Web.
- 3 The WPS standard is very generic to allow various kinds of processing procedures, application specific profiles are necessary to ensure syntactic and structural interoperability here.
- 4 Examples of repositories and collaborative tools include work by the Open Ontology Repository Initiative, the NeON Cupboard, OwlSight, Web Protégé, or OWLDiff.
- 5 If the WRS should also encapsulate other ontology languages and their reasoning services, such as WSMML and IRIS, it has to implement additional mappings.
- 6 We propose the term *linked spatiotemporal data* instead of *linked geo data* as it is broader, does not limit the notion of space to geo-space, and also includes the temporal dimension which is important for work on cultural heritage.

References

- Auer S, Lehmann J, and Hellmann S 2009 LinkedGeoData: Adding a spatial dimension to the Web of Data. In *Proceedings of the International Semantic Web Conference (ISCW 2009)*, Washington, DC
- Bizer C, Heath T, and Berners-Lee T 2010 Linked data: The story so far. *Journal on Semantic Web and Information Systems* 6: in press
- Botts M (ed) 2007 *OGC Implementation Specification 07-000: OpenGIS Sensor Model Language (SensorML)*. Wayland, MA, Open Geospatial Consortium Technical Report
- Bröring A, Janowicz K, Stasch C, and Kuhn W 2009 Semantic challenges for sensor plug and play. In Carswell J D, Fotheringham A S, and McArdle G (eds) *Web and Wireless Geographical Information Systems*. Berlin, Springer Lecture Notes in Computer Science Vol. 5886: 72–86
- Cox S 2007 *OGC Implementation Specification 07-022r1: Observations and Measurements, Part 1 – Observation Schema*. Wayland, MA, Open Geospatial Consortium Technical Report
- Cruz I F and Sunna W 2008 Structural alignment methods with applications to geospatial ontologies. *Transactions in GIS* 12: 683–711
- de la Beaujardiere J 2003 *OGC Implementation Specification 03-109r1: OGC Web Map Service Interface*. Wayland, MA, Open Geospatial Consortium Technical Report
- Fitzner D, Hoffmann J, and Klien E 2010 Functional description of geoprocessing services as conjunctive datalog queries. *GeoInformatica* 14: in press
- Frank A 2003 Ontology for spatio-temporal databases. In Sellis T, Koubarakis M, Frank A, Grumbach S, Güting R H, Jensen C, Lorentzos N, Manolopoulos Y, Nardelli E, Pernici B, Theodoulidis B, Nectaria Tryfona N, Schek H-J, and Scholl M (eds) *Spatio-Temporal Databases*. Berlin, Springer Lecture Notes in Computer Science Vol. 2520: 9–77
- Grcar M 2008 D4.5: Software Module for Semantic Annotation of a Web Service. WWW document, <http://swing-project.org/deliverables/>
- Guarino N 1998 Formal ontology and information systems. In Guarino N (ed) *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS 1998)*. Trento, Italy, IOS Press: 3–15
- Harnad S 1990 The symbol grounding problem. *Physical Review D* 42: 335–46
- Harvey F, Kuhn W, Pundt H, Bisher Y, and Riedemann C 1999 Semantic interoperability: A central issue for sharing geographic information. *Annals of Regional Science* 33: 213–32
- Henson C A, Pschorr J K, Sheth A P, and Thirunarayan K 2009 SemSOS: Semantic sensor observation service. In *Proceedings of the International Symposium on Collaborative Technologies and Systems (CTS 2009)*, Baltimore, Maryland
- ISO/TC211 2002 *ISO/FDIS 19113:2002: Geographic Information – Quality Principles Schema*. Geneva, International Standards Organization
- ISO/TC211 2009 *ISO/DIS 19143: Geographic Information – Filter Encoding*. Geneva, International Standards Organization

- Janowicz K and Keßler C 2008 The role of ontology in improving gazetteer interaction. *International Journal of Geographical Information Science* 22: 1129–57
- Janowicz K and Wilkes M 2009 SIM-DLA: A novel semantic similarity measure for description logics reducing inter-concept to inter-instance similarity. In Aroyo L, Traverso P, Ciravegna F, Cimiano P, Heath T, Hyvoenen E, Mizoguchi R, Oren E, Sabou M, and Simperl E (eds) *Proceedings of the Sixth Annual European Semantic Web Conference (ESWC 2009)*. Berlin, Springer Lecture Notes in Computer Science Vol. 5554: 353–67
- Janowicz K, Keßler C, Schwarz M, Wilkes M, Panov I, Espeter M, and Baeumer B 2007 Algorithm, implementation and application of the SIM-DL similarity server. In Fonseca F T, Rodriguez A, and Levashkin S (eds) *Proceedings of the Second International Conference on GeoSpatial Semantics (GeoS 2007)*. Berlin, Springer Lecture Notes in Computer Science Vol. 4853: 128–45
- Janowicz K, Schade S, Bröring A, Keßler C, Stasch C, Maué P, and Diekhof T 2009 A transparent semantic enablement layer for the geospatial web. In *Proceedings of the Terra Cognita 2009 Workshop, held in conjunction with the Eighth International Semantic Web Conference (ISWC 2009)*, Washington, DC
- Jirka S and Bröring A H 2009 *OGC Sensor Observable Registry*. Wayland, MA, Open Geospatial Consortium Discussion Paper No. 09-112
- Keßler C, Janowicz K, and Bishr M 2009a An agenda for the next generation gazetteer: Geographic information contribution and retrieval. In *Proceedings of the ACM International Conference on Advances in Geographic Information Systems*, Seattle, Washington
- Keßler C, Raubal M, and Wosniok C 2009b Semantic rules for context-aware geographical information retrieval. In Barnaghi P (ed) *Proceedings of the European Conference on Smart Sensing and Context (EuroSSC 2009)*. Berlin, Springer Lecture Notes in Computer Science No. 5741: 77–92
- Klien E 2007 A rule-based strategy for the semantic annotation of geodata. *Transactions in GIS* 11: 437–52
- Kuhn W 2003 Semantic reference systems. *International Journal of Geographical Information Science* 17: 405–9
- Kuhn W 2005 Geospatial semantics: Why, of what, and how? *Journal on Data Semantics* 3: 1–24
- Lemmens R, Wytzisk A, de By R, Granell C, Gould M, and van Oosterom P 2006 Integrating semantic and syntactic descriptions to chain geographic services. *IEEE Internet Computing* 10(5): 42–52
- Lieberman J, Pehle T, Morris C, Kolas D, Dean M, Lutz M, Probst F, and Klien E 2006 *Geospatial Semantic Web Interoperability Experiment Report*. Wayland, MA, Open Geospatial Consortium Technical Report
- Lupp M 2007 *OGC Implementation Specification 05-078r4: Styled Layer Descriptor Profile of the Web Map Service Implementation Specification*. Wayland, MA, Open Geospatial Consortium Technical Report
- Lutz M and Klien E 2006 Ontology-based retrieval of geographic information. *International Journal of Geographical Information Science* 20: 233–60
- Manso M and Wachowicz M 2009 GIS design: A review of current issues in interoperability. *Geography Compass* 3: 1105–24
- Maué P and Schade S 2009 Data integration in the geospatial semantic web. In Kalfoglou Y (ed) *Cases on Semantic Interoperability for Information Systems Integration: Practices and Applications*. Hershey, PA, IGI Global: 100–22
- Maué P, Schade S, and Duchesne P 2009 *Semantic Annotations in OGC Standards*. Wayland, MA, Open Geospatial Consortium Discussion Paper No. 08-167r1
- Na A and Priest M 2007 *OGC Implementation Specification 06-009r6: OpenGIS Sensor Observation Service (SOS)*. Wayland, MA, Open Geospatial Consortium Technical Report
- Nebert D, Whiteside A, and Vretanos P 2007 *OGC Implementation Specification 07-006r1: OpenGIS Catalogue Services Specification*. Wayland, MA, Open Geospatial Consortium Technical Report
- Portele C 2007 *OGC Implementation Specification 07-036: OpenGIS Geography Markup Language (GML) Encoding Standard*. Wayland, MA, Open Geospatial Consortium Technical Report

- Probst F 2007 Semantic Reference Systems for Observations and Measurements. Unpublished PhD Dissertation, Institute for Geoinformatics, University of Münster
- Probst F and Lutz M 2004 Giving meaning to GI web service descriptions. In *Proceedings of the Second International Workshop on Web Services: Modeling, Architecture and Infrastructure (WSMAI 2004)*, Porto, Portugal
- Roman D and Klien E 2007 SWING: A semantic framework for geospatial services. In Arno Scharl K T (ed) *The Geospatial Web*. Berlin, Springer: 227–37
- Schade S, Klien E, Maué P, Fitzner D, and Kuhn W 2008 Report on Modelling Approach and Guideline. WWW document, <http://swing-project.org/deliverables/>
- Scheider S, Janowicz K, and Kuhn W 2009 Grounding geographic categories in the meaningful environment. In Hornsby K S, Claramunt C, Denis M, and Ligozat G (eds) *Conference on Spatial Information Theory (COSIT 2009)*. Berlin, Springer Lecture Notes in Computer Science Vol. 5756: 69–87
- Schut P 2007 *OGC Implementation Specification 05-007r7: OpenGIS Web Processing Service*. Wayland, MA, Open Geospatial Consortium Technical Report
- Stock K, Small M, Ou Y, and Reitsma F 2009 *OWL Application Profile of CSW*. Wayland, MA, Open Geospatial Consortium Discussion Paper No. 09-010
- Tomlin C D 1990 *Geographic Information Systems and Cartographic Modeling*. Englewood Cliffs, NJ, Prentice Hall
- Verma K and Sheth A 2007 Semantically annotating a web service. *IEEE Internet Computing* 11(2): 83–5
- Vretanos P 2005 *OGC Implementation Specification 04-094: Web Feature Service Implementation Specification*. Wayland, MA, Open Geospatial Consortium Technical Report
- Weiser A and Zipf A 2007 Web Service Orchestration of OGC Web Services for Disaster Management. In Li J, Zlatanova S, and Fabbri A (eds) *Geomatics Solutions for Disaster Management*. Berlin, Springer Lecture Notes in Geoinformation and Cartography: 239–54
- Whiteside A 2007 *OGC Implementation Specification 06-121r3: OGC Web Services Common Specification*. Wayland, MA, Open Geospatial Consortium Technical Report
- Whiteside A and Evans J 2008 *OGC Implementation Specification 07-067r5: Web Coverage Service (WCS) Implementation Standard*. Wayland, MA, Open Geospatial Consortium Technical Report
- Wilson T 2008 *OGC Implementation Specification 07-147r2: OGC KML*. Wayland, MA, Open Geospatial Consortium Technical Report