

A RESTful Proxy and Data Model for Linked Sensor Data

Krzysztof Janowicz^{*1,3}, Arne Bröring^{2,3,4}, Christoph Stasch⁴, Sven Schade⁵, Thomas Everding⁴, and Alejandro Llaves⁴

¹University of California, Santa Barbara, USA

²ITC Faculty, University of Twente, The Netherlands

³52° North Initiative for Geospatial Open Source Software, Münster, Germany

⁴Institute for Geoinformatics, University of Münster, Germany

⁵Institute for Environment and Sustainability, Joint Research Centre, Ispra, Italy

Abstract

The vision of a Digital Earth calls for more dynamic information systems, new sources of information, and stronger capabilities for their integration. Sensor networks have been identified as a major information source for the Digital Earth, while Semantic Web technologies have been proposed to facilitate integration.. So far, sensor data is stored and published using the Observations & Measurements standard of the Open Geospatial Consortium (OGC) as data model. With the advent of Volunteered Geographic Information and the Semantic Sensor Web, work on an ontological model gained importance within Sensor Web Enablement (SWE). In contrast to data models, an ontological approach abstracts from implementation details by focusing on modeling the physical world from the perspective of a particular domain. Ontologies restrict the interpretation of vocabularies towards their intended meaning. The ongoing paradigm shift to Linked Sensor Data complements this attempt. Two questions have to be addressed: (i) how to refer to changing and frequently updated data sets using Uniform Resource Identifiers, and (ii) how to establish meaningful links between those data sets, i.e., observations, sensors, features of interest, and observed properties? In this paper, we present a Linked Data model and a RESTful proxy for OGC's Sensor Observation Service to improve integration and inter-linkage of observation data for the Digital Earth.

Keywords: Digital Earth, Geospatial data integration, Spatial Data Infrastructure, Earth Observation, Data exchange models, Linked Data, Semantic Enablement

*Corresponding author: Krzysztof Janowicz; email: jano@geog.ucsb.edu

1 Motivation

The initial vision of a Digital Earth was first formulated by former US Vice President Al Gore as a *multi-resolution, three-dimensional representation of the planet, into which we can embed vast quantities of geo-referenced data* (Gore, 1998). Ten years after this speech, Craglia et al. published a position paper to argue that this vision has not yet been achieved (Craglia et al., 2008). In parallel to the growing availability of information, the need to better understand the interplay of environmental and social phenomena has also increased, thus requiring more dynamic systems, new sources of information, and stronger capacities for their integration. The Sensor Web has been identified as a central building block to address these challenges (De Longueville et al., 2010). A *digital nervous system* for the globe has been suggested as a vibrant approach for the Digital Earth. An implementation based on Spatial Data Infrastructures (SDI), especially on the Sensor Web Enablement (SWE) standards of the Open Geospatial Consortium (OGC), has been proposed. These infrastructures do not only deliver data but also offer geospatial processing capabilities and the final rendering on a virtual globe. Grounded in spatial and temporal reference systems, the outcomes of a variety of services can be combined into a multi-layered representation of the Earth’s surface and help to answer scientific questions or assist in emergency situations. A rigid standardization process as well as conformance tests ensure that a multitude of services can be integrated into SDIs to realize advanced tasks such as predictions.

Yet, in contrast to other scientific domains, we cannot define a context-free and canonical representation of geographic features or even their corresponding types (Mark, 1993; Brodaric and Gahegan, 2007). For instance, there is no pre-given and common definition of *Forest*, *Mountain*, or *Lake* (Lund, 03/22/2010; Smith and Mark, 2003; Montello and Sutton, 2006). Nevertheless, a meaningful layering of geo-referenced data for the Digital Earth requires the integration of the thematic aspects as well, e.g., by grounding them using semantic reference systems (Kuhn, 2003; Scheider et al., 2009). This challenge, known as *semantic integration*, becomes even more urgent when taking Volunteered Geographic Information and the idea of citizens as sensors into account (Goodchild, 2007; Goodchild and Glennon, 2010). Every community has its own requirements and motivations for contributing data. This is reflected by differences in the local conceptualizations of geographic space – leading to semantic heterogeneity (Janowicz, 2010). Van Zyl et al., for instance, pointed out that the Sensor Web requires well defined semantics to make observation data discoverable and reusable (van Zyl et al., 2009).

The Semantic Web explicitly addresses the integration problem by (i) providing formal and machine-readable specifications for the conceptualizations used within different information communities, i.e., by creating ontologies, and by (ii)

using reasoning engines to discover implicit facts, relations, and contradictions. Up to now, SDIs and the Semantic Web are not connected and, therefore, cannot exchange data or combine their services. To address this shortcoming, we have specified and partially implemented a Semantic Enablement Layer (SEL) for SDIs (Janowicz *et al.*, 2010b). It encapsulates Semantic Web reasoners and repositories within OGC services and, thereby, enables a transparent and seamless integration of Semantic Web technologies with SDI. However, our work also focuses on the reverse direction – namely how to make spatial information available on the Semantic Web without changing existing SDI standards and implementations. Mazzetti *et al.* suggested that Representational State Transfer (REST) may provide an architectural solution and identified the connection to the Semantic Web as an item for future Digital Earth research (Mazzetti *et al.*, 2009).

In this paper, we follow this proposal and introduce a RESTful proxy for the OGC Sensor Observation Service (SOS) to assign meaningful identifiers to sensor data and to directly publish this raw data on the Web. The free and open source code as well as a demonstration are available at http://52north.org/RESTful_SOS. We present our research on developing a Linked Data model and argue why it is required in addition to classical data models and ontologies. Our software can be installed as a facade to a SOS without any modifications to the service interface or database, and hence follows the SEL methodology. The proxy provides an RDF representation of observation data, links between data sets, as well as URIs – all of these are fundamental building blocks of Linked Data. Our work goes beyond the initial proposal of Page *et al.* (Page *et al.*, 2009) and moreover builds up on top of the sensor ontologies developed by the W3C Semantic Sensor Network Incubator Group (W3C SSN-XG).

The remainder of this article is structured as follows. Based on our previous work (Janowicz *et al.*, 2010a; Schade and Cox, 2010; Janowicz and Compton, 2010; Keßler and Janowicz, 2010), we discuss the need for a Linked Data model and introduce all relevant parts in detail. Next, we describe the transparent and RESTful SOS proxy, offer conceptual as well as technical insights, and discuss potential applications. Finally, we conclude our paper by summarizing the achieved results and lessons learned, pointing out directions for further work, and discussing the future of SDI in relation to the Linked Data principles and the Digital Earth.

2 Background

This section introduces research on Spatial Data Infrastructure, Linked Data, and Resource-Oriented Architectures relevant for the presented work and also clarifies their interrelation.

2.1 Spatial Data Infrastructures and Sensor Web Enablement

Spatial Data Infrastructure refers to the specialization of information infrastructures for the geospatial sciences (Nebert, 2004). Dozens of SDIs have been developed across the globe, both on national as well as on international-level (Crompvoets and Bregt, 2007). Nowadays, SDIs define the service-oriented management, access, and processing of geospatial data and are implemented using Web Services. The demand for interoperability has boosted the development of standards and tools to facilitate data transformation and integration, mostly in terms of interfaces specified by the OGC.

An abstract structure for data modeling and encoding is provided in form of the Geographic Markup Language (GML). Specific data models and services for sensor networks are developed within OGC's Sensor Web Enablement (SWE) initiative (Botts *et al.*, 2008). SWE is responsible for the development of standards to make sensors and their observations accessible on the Web. The Observations & Measurements (O&M) and Sensor Model Language (SensorML) specifications define how to exchange data, while a number of additional Web Services are responsible for the storing, retrieving, and tasking of sensor-related data. One example is the Sensor Observation Service (SOS) which stores and gives pull-based access to observation data, while the upcoming Sensor Event Service (SES) aims at push-based notifications and also handles complex event processing.

2.2 The Semantic Web and Linked Data

While the OGC has been successful in addressing syntactic and structural heterogeneity by standardization, semantic heterogeneity remains a major challenge (Kuhn, 2009; Janowicz *et al.*, 2010b; Schade, 2010). For example, the missing capabilities for semantic matching are among the main obstacles towards a plug & play infrastructure for the Sensor Web (Bröring *et al.*, 2009a). The Semantic Web promises to address exactly these shortcomings. It reduces semantic interoperability problems by providing a family of formal and machine-readable knowledge representation languages such as the Web Ontology Language (OWL). Semantic Web technologies such as inference engines discover implicit relations and help to establish new facts. Web ontologies make semantic heterogeneity explicit by restricting domain vocabularies towards their intended interpretations (Guarino, 1998; Kuhn, 2009).

Nevertheless, SDI and the Semantic Web can only interact to a very limited degree. A common layer to integrate ontologies and reasoning services into SDI is missing. We have proposed such a Semantic Enablement Layer (SEL) in our previous work (Janowicz *et al.*, 2010b). Instead of defining novel services and

protocols, it encapsulates Semantic Web repositories and reasoners using established OGC service types such as the Web Processing Service (WPS). We have also introduced a similar approach for ontology repositories by encapsulating them by OGC catalog services. This layer provides a transparent and seamless integration of Semantic Web technologies into Spatial Data Infrastructures and does not require any modifications to existing services. So far, the suggested solution has been partially implemented and is available as free and open source software. A first release of the Web Reasoning Service (WRS) as part of the SEL is available at <https://svn.52north.org/svn/semantics/WRS>.

There is reason to assume that the development discussed previously for services will also lead away from top-down data models developed by authorities towards users becoming active knowledge engineers (Janowicz, 2010). This raises questions and problems that reach beyond the goals and capabilities of SDIs as well as the current state of the art of Semantic Web technologies. Flexible, minimalistic, and local vocabularies are required to interlink single, context-specific data fragments on the Web. Linked Data has been introduced to address some of these new requirements (Bizer *et al.*, 2009). In a nutshell, Linked Data describes a paradigm shift from a Web of linked documents towards a Web of linked data. In conjunction with ontologies, such *raw* data can be combined and reused on-the-fly. From a methodological point of view, Linked Data proposes unique identifiers for data, links between them, and relies on the Resource Description Framework (RDF) (Manola and Miller, 2004). The most common query language for RDF is SPARQL (Prud'hommeaux and Seaborne, 2008). SPARQL has similar capabilities as query languages for relational databases, but works by matching graph patterns and is optimized for RDF triple stores, such as Sesame or Virtuoso. In comparison to SDIs, the Linked Data paradigm is relatively simple and, therefore, can help to open up SDIs to casual users. Within the last years, Linked Data has become the most promising vision for the Future Internet and has been widely adopted by academia and industry. The *Linking Open Data cloud diagram* provides a good and up-to-date overview of the available data and the degree of interlinkage: <http://lod-cloud.net/>.

The combination of classical SWE approaches and work on the Semantic Sensor Web (Sheth *et al.*, 2008) with Linked Data has been recently proposed by various research groups (Phuoc and Hauswirth, 2009; Sequeda and Corcho, 2009; Page *et al.*, 2009; Janowicz *et al.*, 2010a; Patni *et al.*, 2010a,b; Schade and Cox, 2010). While they differ in application areas and the used methods, they all argue that sensors, features of interest, and observations should be identified using URIs, looked up by dereferencing these URIs over HTTP, encoded in machine-readable knowledge representation languages such as RDF, and interlinked with other resources. Linked Sensor Data can provide the resources and technologies required

to semantically integrate observation-centric data on the Digital Earth.

2.3 The RESTful Approach

SDI and Linked Data have been developed with fundamentally different and partially contradicting motivations in mind. While the OGC has standardized Web Services which follow a state-full request-response pattern and are designed based on the publish-find-bind paradigm of Service-Oriented Architectures (SOA) (Mazzetti *et al.*, 2009), Linked Data explicitly aims at breaking up such *data silos*. Instead, it proposes a Resource-Oriented Architecture (ROA) which is focused on distributed capabilities and services. ROA is an architectural style devoted to manage distributed, heterogeneous resources, e.g., features, in which client applications interact directly with the exposed resources. The main constraints behind ROA-based applications are the principles known as REpresentational State Transfer (REST) (Fielding, 2000):

1. Resources should be identified properly using URIs, i.e., each resource must be uniquely addressable;
2. Uniform interfaces should be provided through the use of HTTP as the unique application-level protocol;
3. Resources are manipulated through their representations, since clients and servers exchange self-descriptive messages with each another;
4. Interaction is stateless since servers only record and manage the state of the resources they expose, i.e., client sessions are not maintained on the server;
5. Hypermedia are the engine of application state, i.e., the application state is build following hyperlinks according to the navigation paradigm.

Providers that wish to publish their sensor data on the Linked Data cloud (and thereby follow a ROA style) can do so by converting their data to RDF, store it in a triple store, and make it accessible via SPARQL endpoints. Consequently, they do not need any OGC Web Service. This implies moving data out of the SDI. Instead, we propose an integrated approach following our notion of semantic enablement (Janowicz *et al.*, 2010b). Data governance and management remains with the SDI, while resources can still be exposed as Linked Data.

3 A Linked Data Model for Sensor Data

Exposing observations provided by a SOS as raw data to the Web, requires global identifiers for the different components of the O&M model (Sequeda and Corcho,

2009; Page *et al.*, 2009; Janowicz *et al.*, 2010a). In contrast to existing approaches, our work aims at introducing three of the core characteristics of Linked Data to OGC Web Services – namely, global instead of local identifiers, i.e, URIs, links between data, and an RDF serialization of O&M data.

While transforming GML-encoded geographic information to RDF is a major step towards making it accessible on the Linked Data Web, a purely automated mapping is of questionable value and does not add any semantics (Jain *et al.*, 2010; Schade and Cox, 2010). The problems of assigning meaningful URIs to data sets, the semantic annotation of data using ontologies, and how to establish links to external resources are still open research questions.

This section addresses each of these requirements. We will especially focus on URIs as sources of reference. Linked Data detaches information from its original creation context, such as documents, applications, or databases. While this eases accessibility and re-usability, it makes the interpretation of these data chunks more challenging (Janowicz, 2010).

3.1 A Linked Data Model for Sensor Observations

The O&M specification describes a conceptual data model for handling observations. It follows a classical object-oriented approach and provides an encoding of the model as XML schema. In contrast, work at the W3C Semantic Sensor Networks Incubator Group aims at developing an ontology for sensors and their observations that describes the physical processes of transforming stimuli to numeric values (Neuhaus and Compton, 2009; Compton *et al.*, 2009; Janowicz and Compton, 2010). Both approaches are not sufficient to address the needs of Linked Data. While O&M supports unique identifiers, it currently does neither prescribe the use of HTTP URI's, the persistence of identifiers, nor clear and flexible linking strategies between resources. Ontologies are an abstraction layer above data models and aim at describing the physical world. For instance, they introduce the notion of a stimulus that triggers the sensor and leads to the observation. The stimulus as such, however, cannot be stored in data repositories.

Therefore, we introduce an intermediate Linked Data model for sensor observations. It is derived from an ontology developed by the W3C SSN-XG, namely the Stimulus-Sensor-Observation (SSO) design pattern (Janowicz and Compton, 2010). The SSO pattern was developed as a flexible and extensible starting point for sensor ontologies and Linked Data vocabularies. The OWL ontologies and a detailed documentation are available at http://www.w3.org/2005/Incubator/ssn/wiki/Main_Page. Figure 1 shows the classes and relations adopted from the pattern together with new elements such as the ObservationCollection and the SamplingTime. In a nutshell, we use the following definitions:

- *FeatureOfInterest*: the entity that comprises observable properties; for example, a 3-dimensional body of air or a sampling point where measurements are taken.
- *ObservedProperty*: the property that inheres in a feature of interest; for example, the temperature of a 3-dimensional body of air.
- *ObservationCollection*: a set of observations that is grouped by a distinct criteria; for example, all observations performed by a particular sensor, or all observations of a particular observed property that have been performed within a particular time frame (sampling time).
- *Observation*: a (social) construct that connects observed properties with sensors, sensing results, and sampling times; for example, the connection between air temperature, a particular temperature sensor, 11.00am (as sampling time) and 23 (as result) degree Celsius.
- *SamplingTime*: The time instant or interval at which an observation was made; for example, 11.00am in the observation above.
- *Result*: a symbol representing an observed value; for example, 23 in the observation above.
- *Sensor*: an entity that performs observations and produces results in form of values; for example, a device that measures air temperature. Humans can also act as sensors.

[Figure 1 about here.]

The relations between the presented classes act as links in our model and define the multiple navigation paths and external references; see section 3.4. A link to additional sensor metadata modeled in SensorML can be established using the relation from an observation to the sensor and then link to additional sensor properties. Based on this model we define an URI schema that provides unique identifiers and at the same time acts as filter for querying.

3.2 A URI Schema for Linked Sensor Data

In order to provide observations to the Linked Data Web, URIs for the different components of the O&M model are required. The main O&M components associated with an *observation* are *features of interest*, *sensors (procedures)*, *sampling times*, *observed properties* and *results*. The URIs are assigned to these components by appending the component type to the URI which identifies the authority. For example, <http://my.authority.org/sensors> returns links to

all sensor descriptions. Consequently, <http://my.authority.org/sensors/thermometer1> provides the description of a certain thermometer and links to the observation collection containing the produced observations. An according scheme is defined for the other components of observations as described before.

To enable the RESTful access to sensor observations, the base URI scheme has the form: <http://my.authority.org/observations>. By following this base link an observation collection which contains all observations of a RESTful service are returned. Observations measured by specific sensors, gathered for particular observed properties, or from specific features can be retrieved by appending the component type in one segment followed by the identifiers of those resources to the base URI in the next segment. For example, the reference <http://my.authority.org/observations/sensors/thermometer1/featuresofinterest/measurementpoint23/observedproperties/temperature>, points to all observations gathered by the sensor *thermometer1* at the feature of interest *measurementpoint23* for the observed property *temperature*. The order of the component type segments can be chosen as desired by the client. This makes the new scheme more robust compared to our previous approach (Janowicz *et al.*, 2010a).

Another example is a reference to all temperature and windspeed observations which appends the segment [/observedproperties/temperature;windspeed](#) to the base URI. In general, several identifiers can be appended for certain resources. By adhering to the proposal of Richardson and Ruby (2007) for a sound URI scheme, these multiple identifiers are separated by semicolons, as their order does not matter. To refer to observations from a particular time instant or period, the *samplingtimes* token can be appended to the URI followed by two comma-separated time strings encoded according to the ISO 8601 specification (ISO, 2004). The temporal relationship between the time strings has to be specified, e.g. *ont:time:relation:between* defines that all observations between the first and second date should be returned. For instance, the URI <http://my.authority.org/observations/samplingtimes/ont:time:relation:between,2008-01-10T14:00,2008-01-12T16:00/sensors/thermometer1/observedproperties/temperature> points to the observation collection with all temperature observations from January 10th 2008 at 2pm until January 12th at 4pm made by *thermometer1*. In this case, the time strings are comma-separated as the order of the time strings is not arbitrary. The first time string represents the start date of the time period for which observation should be returned, the second indicates the end of the time period. The second time string can also be omitted when a link to observations of a particular time instant has to be specified, e.g. all temperature observations gathered at January the 9th are identified by the URI

<http://my.authority.org/observations/samplingTimes/ont:time:relation>equals,2008-01-09/observedProperties/temperature>.

The URIs for observations described before are used to provide links from sensor and feature descriptions to their related observations. For example, the sensor description at <http://my.authority.org/sensors/thermometer1> contains links to the observations produced by the sensor: <http://my.authority.org/observations/sensors/thermometer1>. This ability to collect observations by following links offered by the RESTful SOS proxy replaces the extensive filtering capabilities of the original SOS interface and reflects the interlinking paradigm of Linked Data.

As an additional spatial criteria, a bounding box can be appended to the URI. We use commas to separate the ordered parameters forming a bounding box. The first four values are the coordinates defining a two dimensional rectangle, while the fifth value is the identifier of their reference system: `<minCoord1>,<minCoord2>,<maxCoord1>,<maxCoord2>,<crsURI>`. Since observations do not necessarily have a location property, the spatial filter is applied to the position of the feature of interest associated with an observation. The SOS specification uses the 'propertyName' parameter to identify the property to which the spatial filter is applied. For the sake of simplicity we restrict this parameter to the position (shape) of the feature of interest. In contrast to the temporal filter, the coordinates are not in different URI segments, as leaving one of them aside would be worthless. An URI using a bounding box filter may be specified as follows: <http://my.authority.org/observations/boundingBox/3,6,23,36,urn:ogc:def:crs:EPSG:6.5:4326>.

Finally, using geographic coordinates in the URIs raises the question of how to encode geometries within the Linked Data model. The geometry of features can be encoded as plain-text lists of coordinates such as defined in GML or translated to an RDF representation of single points connected by RDF predicates. The latter solution leads to a substantial overhead as receiving the geometry of a feature of interest involves traversing all points defined within the RDF encoding. This is only feasible if introducing a labeled and directed graph adds further retrieval or reasoning capabilities. Existing solutions do not require nor benefit from such an RDF serializations, e.g., for computing topological relations. Consequently, we do not advocate to transform GML geometries from *well-known text* to RDF.

It is important to keep in mind that the proposed URI schema fulfills two requirements at the same time. First, it provides unique references for all resources stored in a SOS. Second, it defines a filtering language for querying the SOS based on a RESTful paradigm. Consequently, many URIs may be used to request the same data but a particular URI will always uniquely identify a data set. The vague and uncertain nature of features of interest as well as the dynamic nature of sensor

data in general make the definition of URIs and the creation of links between data sets challenging.

3.3 Establishing Meaningful URIs

One claim of the Linked Data initiative is to make *raw data* available on the Web and assign Uniform Resource Identifiers to each chunk of data. The notion of (and need for) *raw* data was introduced by Berners-Lee during a TED talk in 2009 calling for a direct and unfiltered access to data; see http://www.ted.com/talks/tim_berniers_lee_on_the_next_web.html. Taken seriously, this postulation leads to numerous problems ranging from object identity (Janowicz, 2010), over granularity, up to the question which processing steps are allowed before the data cannot be called raw anymore. In fact, most sensor data is pre-processed. e.g., by deleting or replacing measurement errors. In this work, we use OGC’s notion of a feature of interest to demonstrate some of these challenges.

Measurement is the process in which a sensors receives a stimulus and translates it into another, often digital, representation. However, we are not interested in such stimuli but in what they reveal about the properties of specific entities in the physical world – the features of interest. Such features and their corresponding types, however, do not exist a priori but are an artifact of human cognition and social convention (Brodaric and Gahegan, 2007; Mark, 1993; Lehar, 2003). The extraction of features from sensor observations requires several processing steps that are arbitrary to a certain degree. For example, the extraction of land-cover features such as forests from raster data varies among algorithms and applications. The classical *where is downtown* problem can be used for illustration (Montello *et al.*, 2003) – for instance, measuring the air temperature at the downtown of Santa Barbara, CA. One can determine the boundaries of such a (vague) region from human participants tests, by studying user assigned tags in Web 2.0 photo communities (Kebler *et al.*, 2009), and various other approaches. Each of these methods depends on additional factors, such as confidence values in case of the human participants tests. Hence, a URI for Santa Barbara’s downtown extracted using a yolk-egg model based on a 75% confidence may be constructed as http://my.authority.org/featuresofinterest/SantaBarbara_downtown/yolk-egg/C75/ and linked as a feature of interest to an air quality sensor system.

A meaningful URI scheme also requires a careful sequencing of the used segments. For example, omitting the */C75/* segment should still identify a resource – in this case, it should return all polygons based on the yolk-egg model available for downtown Santa Barbara. Further reducing the URI by removing the */yolk-egg/* segment, returns all potential downtown representations extracted by different approaches (and their parameters) from the data

stored at *my.authority.org*. One could argue that http://my.authority.org/featuresofinterest/SantaBarbara_downtown/ is the base URI identifying Santa Barbara’s downtown and providing all RDF-encoded information about this feature as well as the links to different geometries. While this is a feasible approach, it does not solve the problem of identity. This would require a context-free, i.e., global, notion of Santa Barbara or complex semantic mappings. It is worth mentioning that using *owl:sameAs* to link between different versions of the downtown URI would rather add to the confusion and not resolve the problem (Halpin and Hayes, 2010).

Finally, taking the call for *raw* data seriously, one could assign URIs to all unprocessed outputs of sensors. However, this turns out to be of questionable value for two reasons: (i) What is the appropriate granularity for such chunks of data? For example, remote sensing data from satellites is recorded based on the *swath width* that clearly has no reference to geographic features. Therefore, assigning URIs to such huge chunks of data would render them meaningless and moreover would fragment geographic features randomly over different data sets. Alternatively (and also pointless), one could create URIs for each single pixel representing a sensed value. (ii) A major reason for the limited re-usability of sensor observations is that by deploying a sensor we already make numerous assumptions about the studied phenomena. Summing up, there is hardly any context-free sensor data. In most cases raw data will be of limited value. The selection, processing, and publication of data using URIs already involves making certain decisions, e.g., about the features of interest.

3.4 Establishing Links to External Sources

While links are one of the crucial building blocks for Linked Data, Guéret et al. have shown that about 80% of all triples within the Linked Data cloud point either to URIs in the same namespace, blank nodes, or literals (Gueret *et al.*, 2010). In previous work, we described some of the difficulties involved in linking highly dynamic resources such as sensor data and argued for a curated approach (Keßler and Janowicz, 2010). For instance, links between sensors and features of interests may be of a temporal nature. An *observes* relation between a particular air quality monitoring station and a gas plume will only hold as long as the feature exists or until the station has been re-deployed at another location. For this reason, our Linked Data model does not offer direct links between sensors and features; see also figure 1.

So far, links expressed as RDF predicates do not support a temporal scoping. A partial solution are blank nodes or reifications in which the *observes* relation is represented as an RDF subject to which timestamps are linked. Such approaches,

however, complicate querying and require additional knowledge about the used predicates. To easy interlinking with external recourses, the proposed Linked Data model tries to keep the data identified by a URI as stable as possible. An observation, for instance, will only change due to manual modifications by the data provider and, hence, can be considered as stable. Therefore, Linked Data applications can use these URIs in combinations with *owl:SameAs* services such as sameas.org. In contrast, if we would provide a direct mapping from an ontology such as the Semantic Sensor Network Ontology (Neuhaus and Compton, 2009; Janowicz and Compton, 2010), data about sensors or features of interest identified by a particular URI would change between requests. This would result in misleading links and hamper meaningful information retrieval.

While this is essentially due to the nature of dynamic information, we reduce the resulting problems by introducing the *ObservationCollection* known from the original O&M specifications. As depicted in figure 1, sensors, features of interest, sampling times, results, and observed properties have a unidirectional link to the *ObservationCollection*. This collection contains a list of observations related to the subject of the query. Consequently, while the RDF triples describing a specific sensor will be relatively stable and may be used as basis for interlinking, the observation collection encapsulates the dynamic parts. This again, points out that a Linked Data model is required in addition to ontologies.

4 A Transparent RESTful SOS Proxy

This section introduces the RESTful SOS proxy, which can be installed as a software facade in front of any OGC conform SOS and offers the core functionality to make sensor data available to the Linked Data cloud. Based on the scheme introduced in section 3.2, the RESTful proxy extracts the user's query from the URI, encodes it into valid SOS queries, fetches the results from the underlying SOS, and converts them (after according content negotiation) to an RDF representation aligned with the Linked Data model described in section 3.1. Hence, the URI identifies a particular data set and at the same time encodes a query to the underlying Sensor Observation Service.

4.1 Design of the RESTful SOS Proxy

The RESTful SOS proxy (available online at http://52north.org/RESTful_SOS) is developed based on the *OGC Web Service Access Framework* (OX-*Framework*)(see <http://52north.org/oxf>), a software framework whose architecture can be used for an easy utilization of OGC Web Services, such as the

SOS. The OX-Framework (Bröring *et al.*, 2009b) offers developers a customizable and extendable system of cooperating classes which supply a reusable design applicable for the implementation of software clients. However, the framework does not only support access to data through the SOS. The OX-Framework has been developed by the 52° North Sensor Web community (see <http://52north.org/sensorWeb>) to provide an integrative view to access all kinds of OGC Web Services and enable the processing of the queried data.

The OX-Framework supports the access of various service interfaces by providing a generic architecture which includes a plug-in mechanism for service adapters as extension points of the framework. The architecture is structured into an *adapters* subsystem and a *core* subsystem as shown in figure 2.

[Figure 2 about here.]

The core subsystem incorporates a two-folded data model: the *common capabilities model* implements the OGC Web Service Common model (Whiteside, 2007) and introduces thereby the integrative view on service access to the framework. The internal *feature model* provides a basis for accessing and processing of features, including observations. It is based on OGC's feature model (Kottman and Reed, 2009), and, since the O&M design follows OGC's feature model, retrieved observations can be directly mapped. These two data models enable the communication between the various framework components.

The adapters subsystem contains realizations of three kinds of service adapters. The *service connectors* are used to trigger service operations and to instantiate the common capabilities model. Here, a specific service connector for the SOS is used, first, to call the *GetCapabilities* operation for retrieving metadata about the service and its contents, and second, to execute the *GetObservation*, *DescribeSensor* or *GetFeatureOfInterest* operations to retrieve data requested via the RESTful SOS interface. Once data has been retrieved from the service, *feature stores* provide the functionality to unmarshal received feature data into the core's feature model. In case of the RESTful SOS proxy, a specific feature store for O&M has been developed to instantiate the feature model from received observations. Finally, *data processors* can run on the instantiated feature model and are used to transform the feature data into other representations. Within this work, we developed a processor which converts observations into RDF-encoded Linked Data but we also support other representations such as KML or JPEG charts. The RESTful SOS proxy chooses the right data processor based on content negotiation and the *HTTP accept header*. The described interplay of different adapter components is illustrated in figure 3.

[Figure 3 about here.]

Service adapters are implemented for specific types of OGC Web Services, e.g., for the SOS. Implementations of service adapters can be incorporated into the framework in a dynamic way by means of a plugin mechanism which enables a flexible extension of the framework. Once plugged into the OX-Framework, the applications build on top of it can reuse the service adapter plugins. This flexibility also allows extending the RESTful proxy to work with other OGC Web Services for which service adapters already exist.

4.2 Discussion of the Proxy Approach

During the implementation of the proxy, several differences between the OGC approach for data provision services and a well-designed Linked Data model have become apparent. On the one hand, these differences increase the complexity of the implementation, on the other hand, these have implications on the performance of the proxy. Originally, we planned to directly map a SOS URI identifying a certain observation collection to a GetObservation request and translate the result to RDF using XSLT. However, this has not been possible for several reasons, which we discuss in the following as they are relevant for further work on semantic enablement of SDIs as well.

According to the current SOS specification (Na and Priest, 2007), every SOS provider has to group its observations in arbitrary groups called *observation offerings*. The specification does not further restrict this grouping which leads to very different interpretations of such offerings. Some providers group their observations by observed properties, others are using certain time periods, while observations are sometimes also grouped by spatial extent. Conforming to the current SOS specification, one GetObservation request can only query observations from one particular offering. Thus, metadata about all observation offerings need to be cached by the SOS proxy and kept updated for cases when new observation offerings are added to the SOS at the backend. The OGC is aware of these problems and the upcoming SOS specification version 2.0 (Bröring *et al.*, 2010) will only contain an optional observation offering parameter in an GetObservation request. Also, the observation offering is defined as a collection of observations produced by a certain sensor system.

In consequence, when a certain URI identifying an observation collection has to be resolved by the SOS proxy, the segments of the URI have to be mapped to one GetObservation request for each observation offering. Therefore, steps 2 and 3 in figure 3 have to be repeatedly executed if, for example, the feature of interest specified in the URI is registered at multiple observation offerings at the SOS. In case of large datasets, which are divided in several observation offerings at the SOS, this might lead to numerous requests that have to be parsed and then transformed to

RDF. To make the RESTful SOS proxy more performant and scalable, intelligent caching mechanism as well as parallel querying and response merging become necessary. Once the RESTful SOS proxy is deployed for a certain SOS, it should query all information required from the SOS and then only update the information, if the metadata of a particular offering (and thus its observations) have changed.

5 Application

This section illustrates how the RESTful SOS proxy can be used to access air quality observations made by a particular monitoring station. The RESTful SOS proxy is deployed at http://v-swe.uni-muenster.de:8080/52nRESTfulSOS/RESTful/sos/AirBase_SOS/. For the sake of readability, this endpoint URL is replaced with the abbreviation <http://myRESTfulSOS/> in the following. The deployed SOS offers data from the AirBase database provided by the European Environmental Agency. As an entry point to Linked Sensor Data, the user retrieves the RDF representation of a particular monitoring station by following the URI <http://myRESTfulSOS/sensors/HR:0002A>. The RDF serialization is shown below and contains links to the related observations. Following these links results in new query to the RESTful SOS.

```
<rdf:Description rdf:about="http://myRESTfulSOS/sensors/HR:0002A">
  <rdfs:label>HR:0002A</rdfs:label>
  <rdf:type rdf:resource="http://v-swe.uni-muenster.de:8080/
52nRESTfulSOS/RESTful/minionmld/Sensor"/>
  <minioamld:relatedObservations
    rdf:resource="http://myRESTfulSOS/observations/sensors
    .../HR:0002A"/>
</rdf:Description>
```

Following the link <http://myRESTfulSOS/observations/sensors/HR:0002A> to the related observations, the user receives an RDF representation of the observation collection listing all observations produced by the particular monitoring station. In case only observations from 2008 for nitrogen dioxide are required, the observation collection can be further restricted by calling the URI [http://myRESTfulSOS/observations/sensors/HR:0002A/samplingtimes/2008-01-01,2008-12-31/observedproperties/concentration\[NO2\]](http://myRESTfulSOS/observations/sensors/HR:0002A/samplingtimes/2008-01-01,2008-12-31/observedproperties/concentration[NO2]) according to the scheme defined in section 3.2. The resulting RDF serialization is shown below.

```
<rdf:Description rdf:about="http://myRESTfulSOS/observations
/sensors/HR:0002A/samplingTimes/2008-01-01,2008-12-31
/observedProperties/Concentration[NO2]">
  <rdfs:label>OC_NO2_HR:0002A_2008-01-01,2008-12-31</rdfs:label>
  <rdf:type rdf:resource="http://v-swe.uni-muenster.de:8080
```



```

</52nRESTfulSOS/RESTful/minionmld/ObservationCollection"/>
  <minioamld:hasObservation rdf:resource="http://myRESTfulSOS
</observations/ids/o_5633"/> [... shortened output]
</rdf:Description>

```

Implementing the browsing paradigm of Linked Data, users or applications can retrieve particular observations by following the `hasObservation` relationship; e.g. by following the URI http://myRESTfulSOS/observations/ids/o_5633 as shown below. Note that the URI to the `aboutProperty` is not pointing to a URI of the SOS, but to a Sensor Observable Registry which has been proposed as a registry mechanism for observable properties at OGC (Jirka and Bröring, 2009) and also serves a RESTful interface to the contained resources.

```

<rdf:Description
rdf:about="http://myRESTfulSOS/observations/ids/o_5633">
  <rdfs:label>o_5633</rdfs:label>
  <rdf:type rdf:resource="http://v-swe.uni-muenster.de:8080
</52nRESTfulSOS/RESTful/minionmld/Observation"/>
  <minioamld:aboutProperty rdf:resource="http://giv-genesis
...uni-muenster.de:8080/SOR/REST/phenomenon/OGC/Concentration[NO2]"/>
  <minioamld:samplingTime rdf:resource="http://myRESTfulSOS
</samplingTimes/2008-12-31T01:00:00.000+01:00"/>
  <minioamld:performedBy rdf:resource="http://myRESTfulSOS
</sensors/HR:0002A"/>
  <minioamld:hasResult rdf:resource="http://myRESTfulSOS
</results/5.48"/>
</rdf:Description>

```

With the RESTful SOS in place, unique identifiers can be used to persistently and globally refer to chunks of dynamically growing data sets. These identifiers are generated automatically by deploying the proxy on top of a classic SOS installation. Other resources become able to seamlessly connect to the data in a transparent and consistent manner. For example, reports on environmental conditions can directly refer to the earth observation data, which was used in the monitoring phase. In this way, we bridge the gap between SDI and the Semantic Web while keeping data management at the source. This proposed solution does not impose any change of technology which would result in changes in connected geospatial decision support systems. For SDI users and service providers, the proposed approach provides a major achievement opposed to the mirroring of geospatial data within triple stores. Others may still prefer to step outside of the SDI framework and publish their data directly to the Linked Data cloud.

Furthermore, spatial processing may be implemented inside a SOS, such that spatial reasoning of SDIs could be combined with logics-based reasoning on top of the RDF serialization. Following this endeavor, the relatively complex silos of current SDIs can be partially embedded into mainstream applications, which opens a

new dimension of geospatial data sharing for Digital Earth research. Raw measurement results may be equally integrated as value added information, such as spatial interpolations or the results of forecasting simulations. In the mid-term, the enabled flexible plug-and-play of observation data may lead to *micro-SDIs* (Janowicz *et al.*, 2010b), in which any SDI managed resource can be easily augmented with common Web pages and services. Such a step would not only allow for better analysis of the complex interplay between society and the environment, it would at the same time open up a new market place for innovative integrated applications.

6 Conclusions and Outlook

In this section, we summarize the presented work, discuss lessons learned, and point out directions for further work.

6.1 Summary

In this work, we introduced a Linked Data model to combine OGC’s Sensor Observation Service with the Linked Data cloud. We developed a URI scheme that offers unique global identifiers and build-in query filtering, and also discussed how to introduce links to external sources. We implemented a transparent and RESTful SOS proxy that can serve Linked Sensor Data without any modifications to existing services. Moreover, we gave insights into the problem of *identity* for the assignment of URIs. We argued that it stems from an entity-centric view taken by most Semantic Web and Linked Data research. It assumes that distinct entities can be identified and while their attributes may be different or vague, the existence of mind-independent entities is not questioned. In contrast, work on sensors and earth observations is focused on continuous fields and entities are introduced during analysis (if at all).

We decided to use a RESTful approach as it combines three key advantages. First, URIs are a fundamental building block of Linked Data. REST allows us to identify data and at the same time encode the query using our scheme. Second, a major requirement of our vision of Semantic Enablement is transparency. Using REST, Linked Data users and applications are not even aware that they are querying an underlying OGC service. Third, a RESTful approach is very flexible with respect to content negotiation. Our proxy does not only serve RDF but KML and other formats as well.

Summing up, the proposed approach provides an important step towards the semantic enablement of existing information systems and infrastructures, and thereby eases the integration of dynamic information sources such as sensor networks. De-

livering observations as Linked Data, connecting them with other data sources, and using ontologies and Semantic Web reasoners to improve retrieval, alignment, and matching are major building blocks for the implementation of a Digital Earth. We believe that this work can help to broaden the view of a Digital Earth as knowledge archive towards a *knowledge engine* vision similar to IBM’s Watson.

6.2 Lessons Learned

Linked Sensor Data is a radically new approach and differs from our previous work on Sensor Web Enablement and geospatial semantics in many ways. The work presented in this article points out several lessons to be learned.

First, just transforming data to RDF does not add any semantics. As pointed out by Jain *et al.* (2010), ontologies are a crucial part of the Linked Data vision. These ontologies may be lightweight or heavyweight, but they have to restrict the meaning of the used terms and relations towards their intended interpretation. While DBPedia is a great showcase for the idea of Linked Data, it also demonstrates the need for more expressive ontologies and reasoning. For instance, using the Santa Barbara example introduced before, there are at least six different relations in the DBPedia ontologies describing the fact that a company is located within a city: headquarters of, headquarter of, foundationPlace of, location of, city of, and locationCity. While the differences between some of them can be understood from their names, this is not possible for others. Moreover, semantics should not be encoded in literals that are not machine-understandable. Consequently, it is not clear which relations should be used or queried and users will get different results.

Developing ontologies for Linked Data is difficult. If the ontologies are too specific they will restrict the usage of the data and the possibility to interlink them with external sources; if they are not specific enough, data may be misinterpreted. With respect to the Digital Earth, data spanning different sources, topics, and perspectives should be integrated on-the-fly to answer complex scientific questions. As demonstrated by Probst and Lutz (2004), some mismatches cannot be discovered on the syntactic level and combining data from sources that seem to be compatible may lead to wrong results, e.g., when simulating the dispersion of a toxic gas plume using weather data from different Web services. To address this challenge, we use a slightly modified version of the Stimulus-Sensor-Observation ontology design pattern developed by the W3C SSN-XG (Janowicz and Compton, 2010). This pattern was developed with Linked Sensor Data in mind and represents the minimal ontological commitments for our RESTful SOS. The pattern comes with an optimal alignment to the DOLCE Ultra Light top-level ontology as well as the Semantic Sensor Network ontology. In this work, we present the Linked Data Model, URI schemes, and the implementation of the RESTful SOS – aligning more specific

ontologies to the pattern is up to the SOS providers and depends on their data and application areas.

Second, our URI scheme does not only identify resources but acts as a query filter at the same time. Each URI is decomposed into its segments and translated into calls for the Sensor Observation Service. While this approach is elegant to realize the transparent Semantic Enablement Layer, it does not cover the full functionality of a SOS. In the future, the URI scheme may be extended to mirror more filters. However, this makes a meaningful segmentation of the URI (following REST principles) difficult. It is important to notice that Linked Data allows to browse and navigate data sets by following relations to other internal or external sources. In our work, this browsing paradigm (known from the Document Web) replaces the extensive SOS filters. Finding the right balance between URI-based filters and following links needs to be determined based on feedback from users and SOS providers.

Third, as pointed out by [Schade and Cox \(2010\)](#) not all data stored in SDIs needs to be transformed to RDF. Essentially, Linked Data forms a global graph of interconnected resources. Some of these resources may be leaf nodes of the graph. It is not clear how to connect these data sets, e.g., binary files, and where to stop *triplifying* data. For instance, do we need an RDF serialization of shape files or geometries in general? Based on our work, we believe that these decisions need to be taken on a case-by-case basis depending on the added value. Before triplifying data, one should clarify which ontologies can be used, whether the data will benefit from internal and external links, and whether machine-readability and Semantic Web reasoning techniques are desired. For instance, we are skeptical about approaches that try to develop RDF representations for *well-known text* (WTK) encoded geometries. While this (and especially separating latitude and longitude) creates a substantial overcharge, the added value remains unclear.

6.3 Outlook

Further work will target the extension of the Linked Data model as introduced in Section 3. This particularly includes links between the sensor observations and external Linked Data sources. For reasons of simplicity, we have also not discussed the distinction between URIs referring to real world entities as proposed by the *Web of Things* ([Guinard and Trifa, 2009](#)) and URIs referring to data about these entities. Moreover, further OGC specifications will be included in the Linked Data model – especially SensorML. It is used to encode sensor descriptions and is therefore tightly coupled with the Sensor Observation Service and other sensor-related services. Extending the introduced model to resource types, such as coverages and maps, and RESTful proxies to other OGC services, e.g. the Web Coverage Service

and Web Mapping Service are also on the agenda of the 52° North Semantics Community. This is possible, as the presented implementation is based on the generic OX-Framework that can be extended through the development of according service adapters (see section 4). We also plan to implement semantic enablement strategies for push-based services such as the Sensor Event Service (Echterhoff and Everding, 2008).

Another important direction of further work is the ongoing development of a semantics-enabled Sensor *Plug&Play* infrastructure (Bröring *et al.*, 2011). Sensors can be automatically registered at a Sensor Observation Service and matched against the service profile. A sensor bus registers the sensors, while a mediator performs an ontology-based matching. Sensor *Plug&Play* is a pre-requisite for the vision of *smart dust* sensor networks and could extend the Digital Earth by large-scale, real-time observations in the future. Our current implementation is restricted to subsumption-based matching, but we plan to implement a rule-based system and integrate our SIM-DL similarity server in the near future.

Finally, we hope that realizing a micro-SDI (Janowicz *et al.*, 2010b) based on Linked Data and JavaScript will enable a ubiquitous Geospatial Web.

Acknowledgments

The presented work is developed within the 52° North semantics community, and partly funded by the European projects *UncertWeb* (FP7-248488), *ENVISION* (FP7-249170), as well as the *GENESIS* project (an Integrated Project, contract number 223996). We are thankful for discussions with members of the Münster Semantic Interoperability Lab (*MUSIL*), the Spatial Data Infrastructures Unit of the Joint Research Centre (JRC) of the European Commission, and many of our colleagues from the W3C Semantic Sensor Network Incubator Group (*SSN-XG*).

References

- Bizer, C., Heath, T., and Berners-Lee, T., 2009. Linked Data – The Story So Far, *International Journal on Semantic Web and Information Systems*, 5 (3), 1–22.
- Botts, M., Percivall, G., Reed, C., and Davidson, J., 2008. OGC Sensor Web Enablement: Overview and High Level Architecture, *Lecture Notes In Computer Science*, 4540, 175–190.
- Brodaric, B. and Gahegan, M., 2007. Experiments to Examine the Situated Nature of Geoscientific Concepts, *Spatial Cognition and Computation*, 7 (1), 61–95.

- Bröring, A., Janowicz, K., Stasch, C., and Kuhn, W., 2009a. Semantic challenges for sensor plug and play, *in*: J.D. Carswell, A.S. Fotheringham, and G. McArdle, eds., *Web and Wireless Geographical Information Systems, 9th International Symposium, W2GIS 2009*, Springer, *Lecture Notes in Computer Science*, vol. 5886, 72–86.
- Bröring, A., Jürrens, E.H., Jirka, S., and Stasch, C., 2009b. Development of Sensor Web Applications with Open Source Software, *in*: *First Open Source GIS UK Conference (OSGIS 2009), 22 June 2009, Nottingham, UK*.
- Bröring, A., Maué, P., Janowicz, K., Nüst, D., and Malewski, C., 2011. Semantically-enabled Sensor Plug & Play for the Sensor Web, *Sensors*; (*accepted for publication*).
- Bröring, A., Stasch, C., and Echterhoff, J., 2010. *OGC Interface Standard 10-037: SOS 2.0 Interface Standard [candidate standard]*, Open Geospatial Consortium.
- Compton, M., Henson, C., Neuhaus, H., Lefort, L., and Sheth, A., 2009. A survey of the semantic specification of sensors, *in*: T. K., A. Ayyagari, and D. De Roure, eds., *Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09)*, CEUR, vol. Vol-522, 17–32.
- Craglia, M., Goodchild, M., Annoni, A., Camara, G., Gould, M., Kuhn, W., Mark, D., Masser, I., Maguire, D., Liang, S., and Parsons, E., 2008. Next-generation digital earth: A position paper from the vespucci initiative for the advancement of geographic information science, *International Journal of Spatial Data Infrastructures Research*, 3, 146–167.
- Crompvoets, J. and Bregt, A., 2007. *Research and theory in advancing spatial data infrastructure*, ESRI Press, Redlands, USA, chap. National spatial data clearing-houses, 2000-2005, 133–146.
- De Longueville, B., Annoni, A., Schade, S., Ostländer, N., and Whitmore, C., 2010. Digital earth’s nervous system for crisis events: real-time sensor web enablement of volunteered geographic information, *International Journal of Digital Earth*, 3 (3).
- Echterhoff, J. and Everding, T., 2008. *OGC Discussion Paper 08-133: OpenGIS Sensor Event Service Interface Specification*, Open Geospatial Consortium.
- Fielding, R., 2000. *Architectural Styles and the Design of Network-based Software Architectures*, Ph.D. thesis, University of California, Irvine.

- Goodchild, M., 2007. Citizens as sensors: the world of volunteered geography, *GeoJournal*, 69 (4), 211–221.
- Goodchild, M.F. and Glennon, J.A., 2010. Crowdsourcing geographic information for disaster response: a research frontier, *International Journal of Digital Earth*, 3 (3), 231–241.
- Gore, A., 1998. The digital earth: Understanding our planet in the 21st century.
- Guarino, N., 1998. Formal Ontology and Information Systems, in: N. Guarino, ed., *International Conference on Formal Ontology in Information Systems (FOIS1998)*, Trento, Italy: IOS Press, 3–15.
- Gueret, C., Groth, P., van Harmelen, F., and Schlobach, S., 2010. Finding the achilles heel of the web of data: using network analysis for link-recommendation, in: *Proceedings of the Int. Semantic Web Conf 2010*.
- Guinard, D. and Trifa, V., 2009. Towards the Web of Things: Web Mashups for Embedded Devices, in: *International World Wide Web Conference*, Madrid, Spain.
- Halpin, H. and Hayes, P., 2010. When owl:sameas isn't the same: An analysis of identity links on the semantic web, in: *Linked Data on the Web (LDOW2010)*.
- ISO, 2004. 8601:2004(E) Data elements and interchange formats - Information interchange - Representation of dates and times.
- Jain, P., Hitzler, P., Yeh, P.Z., Verma, K., and Sheth, A.P., 2010. Linked Data is Merely More Data, in: *AAAI Spring Symposium 'Linked Data Meets Artificial Intelligence'*, AAAI Press, 82–86.
- Janowicz, K., 2010. The role of space and time for knowledge organization on the semantic web, *Semantic Web*, 1 (1-2), 25–32.
- Janowicz, K., Bröring, A., Stasch, C., and Everding, T., 2010a. Towards meaningful URIs for linked sensor data., *Towards Digital Earth: Search, Discover and Share Geospatial Data, Workshop at Future Internet Symposium, September 20th, 2010, Berlin, Germany.*, 640.
- Janowicz, K. and Compton, M., 2010. The stimulus-sensor-observation ontology design pattern and its integration into the semantic sensor network ontology, in: A.A.D.D.R. Kerry Taylor, ed., *Proceedings of the 3rd International workshop on Semantic Sensor Networks 2010 (SSN10) in conjunction with the 9th International Semantic Web Conference (ISWC 2010)*, CEUR-WS, vol. 668.

- Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maue, P., and Stasch, C., 2010b. Semantic enablement for spatial data infrastructures, *Transactions in GIS*, 14 (2), 111–129.
- Jirka, S. and Bröring, A., 2009. OGC Sensor Observable Registry Discussion Paper, OGC Discussion Paper 09-112, Open Geospatial Consortium.
- Keßler, C. and Janowicz, K., 2010. Linking sensor data – why, to what, and how?, in: A.A.D.D.R. Kerry Taylor, ed., *Proceedings of the 3rd International workshop on Semantic Sensor Networks 2010 (SSN10) in conjunction with the 9th International Semantic Web Conference (ISWC 2010)*, CEUR-WS, vol. 668.
- Keßler, C., Maué, P., Heuer, J.T., and Bartoschek, T., 2009. Bottom-up gazetteers: Learning from the implicit semantics of geotags, in: K. Janowicz, M. Raubal, and S. Levashkin, eds., *GeoSpatial Semantics, Third International Conference, GeoS 2009*, Springer, LNCS, vol. 5892, 83–102.
- Kottman, C. and Reed, C., 2009. *The OpenGIS Abstract Specification - Topic 5: Features*, Open Geospatial Consortium, Inc.
- Kuhn, W., 2003. Semantic reference systems (guest editorial), *International Journal of Geographical Information Science*, 17 (5), 405–409.
- Kuhn, W., 2009. Semantic Engineering, in: G. Navratil, ed., *Research Trends in Geographic Information Science*, Springer, Lecture Notes in Geoinformation and Cartography, 63–76.
- Lehar, S., 2003. *The World in Your Head: A Gestalt View of the Mechanism of Conscious Experience*, Lawrence Erlbaum.
- Lund, G., 03/22/2010. Definitions of forest, deforestation, afforestation, and reforestation, Tech. rep., Forest Information Services.
- Manola, F. and Miller, E., 2004. Rdf primer, w3c recommendation 10 february 2004., Tech. rep., W3C.
- Mark, D.M., 1993. Toward a theoretical framework for geographic entity types, in: *Spatial Information Theory: A Theoretical Basis for GIS, International Conference COSIT '93*, 270–283.
- Mazzetti, P., Nativi, S., and Caron, J., 2009. Restful implementation of geospatial services, *International Journal of Digital Earth*, 2 (1), 40–61.

- Montello, D.R., Goodchild, M.F., Gottsegen, J., and Fohl, P., 2003. Where's downtown?: Behavioral methods for determining referents of vague spatial queries, *Spatial Cognition & Computation*, 3 (2), 185–204.
- Montello, D.R. and Sutton, P.C., 2006. *An introduction to scientific research methods in geography*, Sage Publications.
- Na, A. and Priest, M., 2007. *OGC Implementation Specification 06-009r6: OpenGIS Sensor Observation Service (SOS)*, Open Geospatial Consortium.
- Nebert, D., 2004. *SDI Cookbook, Version 2.0*.
- Neuhaus, H. and Compton, M., 2009. The Semantic Sensor Network Ontology: A Generic Language to Describe Sensor Assets, in: *AGILE 2009 Pre-Conference Workshop Challenges in Geospatial Data Harmonisation, 02 June 2009, Hannover, Germany*.
- Page, K., De Roure, D., Martinez, K., Sadler, J., and Kit, O., 2009. Linked sensor data: Restfully serving rdf and gml, in: T. K., A. Ayyagari, and D. De Roure, eds., *Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09)*, CEUR, vol. Vol-522, 49–63.
- Patni, H., Henson, C., and Sheth, A., 2010a. Linked sensor data, in: *2010 International Symposium on Collaborative Technologies and Systems*, IEEE, 362–370.
- Patni, P., Sahoo, S., Henson, C., and Sheth, A., 2010b. Provenance aware linked sensor data, in: P. Kärger, D. Olmedilla, A. Passant, and A. Polleres, eds., *Proceedings of the Second Workshop on Trust and Privacy on the Social and Semantic Web*.
- Phuoc, D.L. and Hauswirth, M., 2009. Linked open data in sensor data mashups, in: A.A.D.D.R. Kerry Taylor, ed., *Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09)*, CEUR, vol. Vol-522, 1–16.
- Probst, F. and Lutz, M., 2004. Giving Meaning to GI Web Service Descriptions, in: *International Workshop on Web Services: Modeling, Architecture and Infrastructure (WSMAI 2004)*.
- Prud'hommeaux, E. and Seaborne, A., 2008. Sparql query language for rdf. w3c recommendation, <http://www.w3.org/tr/rdf-sparql-query/>, Tech. rep.
- Richardson, L. and Ruby, S., 2007. *RESTful Web Services*, O'Reilly Media, Inc.
- Schade, S., 2010. *Ontology-Driven Translation of Geospatial Data*, no. 001 in Dissertations in Geographic Information Science (GISDISS), IOS Press.

- Schade, S. and Cox, S., 2010. Linked data in sdi or how gml is not about trees, *in: Proceedings of the 13th AGILE International Conference on Geographic Information Science - Geospatial Thinking*.
- Scheider, S., Janowicz, K., and Kuhn, W., 2009. Grounding geographic categories in the meaningful environment., *in: K. Hornsby, C. Claramunt, M. Denis, and G. Ligozat, eds., Conference on Spatial Information Theory (COSIT 2009)*, Springer, LNCS, vol. 5756, 69–87.
- Sequeda, J. and Corcho, O., 2009. Linked stream data: A position paper, *in: Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09)*, CEUR-WS, vol. 522, 148–157.
- Sheth, A., Henson, C., and Sahoo, S., 2008. Semantic Sensor Web, *IEEE Internet Computing*, 78–83.
- Smith, B. and Mark, D., 2003. Do mountains exist? towards an ontology of land-forms, *Environment and Planning B*, 30 (3), 411 – 427.
- van Zyl, T., Simonis, I., and McFerren, G., 2009. The sensor web: systems of sensor systems, *International Journal of Digital Earth*, 2 (1), 16–30.
- Whiteside, A., 2007. *OGC Implementation Specification 06-121r3: OGC Web Services Common Specification*, Open Geospatial Consortium.

List of Figures

1	Concept map with the classes and relations of the Linked Sensor Data model.	28
2	Architecture of the RESTful SOS based on the OX-Framework . .	29
3	Resolving a URI by the RESTful SOS proxy	30

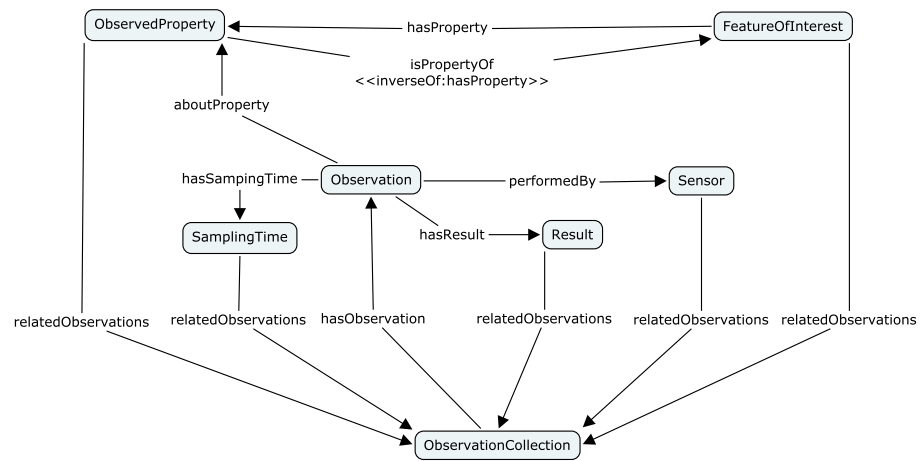


Figure 1: Concept map with the classes and relations of the Linked Sensor Data model.

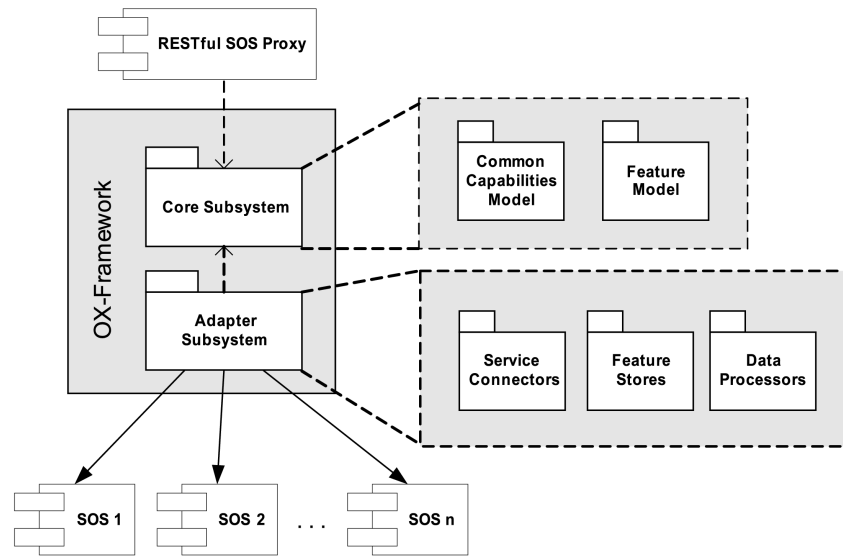


Figure 2: Architecture of the RESTful SOS based on the OX-Framework

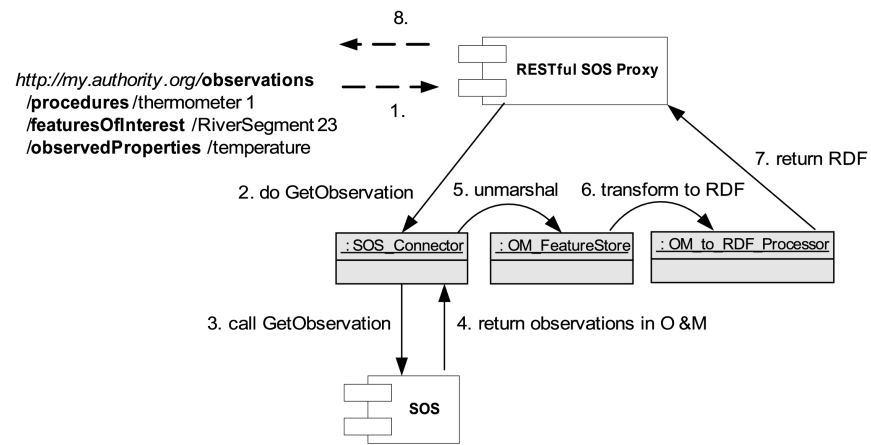


Figure 3: Resolving a URI by the RESTful SOS proxy