Linking Sensor Data – Why, to What, and How?

Carsten Keßler*, Krzysztof Janowicz†

*Institute for Geoinformatics, University of Münster, Germany
carsten.kessler@uni-muenster.de
†GeoVISTA Center, Pennsylvania State University, USA
jano@psu.edu

Abstract. The Sensor Web provides access to observations and measurements through standardized interfaces defined by the Open Geospatial Consortium's Sensor Web Enablement (SWE) initiative. While clients compliant to these standards have access to the generated sensor data, it remains partially hidden from other knowledge infrastructures building on higher-level W3C standards. To overcome this problem, it has been proposed to make sensor data accessible using Linked Data principles and RESTful services. This position paper discusses the embedding of such data into the Linked Data cloud with a focus on the outgoing links that hook them up with other data sources. We outline how such links can be generated in a semi-automatic way, and argue why curation of the links is required. Finally, we point to the query potential of such an additional interface to observation data, and outline the requirements for SPARQL endpoints.

1 Introduction and Motivation

Blending the Sensor Web with Semantic Web technologies is attractive for several reasons [1]. From a Sensor Web perspective, it is desirable to enrich the SWE services by semantic annotations[1] [2] and ontologies to reduce ambiguity and to provide machine-readable descriptions of the provided data, the underlying processes, as well as the relevant instruments. This would improve interoperability for automatic service chaining, enable search beyond code lists, and support the reuse of sensor data [3]. From a Semantic Web and Linked Data perspective, semantically enabled SWE services such as the Sensor Observation Service (SOS) are rich sources of data to augment static knowledge about the world with dynamic sensor observations. Stream Reasoning engines and applications can be used to mine for specific patterns and to detect change. Recent examples include weather warning systems [1,4] or context-aware mobile decision support systems [5]. Finally, many applications such as Web mashups benefit from a URI-based access to sensor observations encoded using accepted standards such as the Resource Description Framework (RDF) [6]. The Sensor Web community has recently adopted the Linked Data principles [7,8] to use URIs as reference and for look-up as well as RDF and SPARQL for storage, access, and querying.

---

[1] See, for example, `http://my-trac.assembla.com/sapience/`.

First attempts have been made to provide Linked Sensor Data [9] and to serve them via RESTful interfaces[2] as transparent encapsulations of existing SWE services [10,3].

While providing observations or sensor meta-data as Linked Data is desirable to make them accessible for a broad audience, an in-depth discussion of the challenges and benefits is still missing. The exact motivations to provide Linked Sensor Data, the ontological and technological implications, and the potential of Linked Sensor Data remain partly ambiguous. Moreover, it is hard to measure how successful a Linked Sensor Data initiative is because of a lack of benchmarks – just counting triples for a sensor service that potentially produces hundreds of measurements per minute will not be sufficient. These questions have to be carefully considered to ensure that Linked Sensor Data does not remain on the level of yet another data encoding [11].

To foster discussion, this paper outlines these questions both from a data provider's and from a data consumer's perspective. The discussion is structured by the three questions indicated in the title and inspired by previous work from Kuhn [12]. *Why* is concerned with the motivations, potentials, and benchmarks for Linked Sensor Data. We analyze the different requirements for a Linked Sensor Data service and outline how providers can measure the success of their service. *To what* discusses how to hook Linked Sensor Data into the Linked Data cloud so that useful additional resources can be discovered. We argue for a curated approach, where potentially useful out-links are recommended by the system, but need to be verified to assure correctness. *How* sheds light upon the technical aspects, specifically the question how to identify and store specific events [13] in the long term to reduce the amount of stored data. We also outline how the conversion process that has already been demonstrated for static datasets can be implemented to generate Linked Data on the fly. As recent research has shown, 80% of all triples in the Linked Data cloud point to URIs within the same namespace, literals, or blank nodes [14]. The linking aspect thus needs to be taken seriously, both to provide context for the provided data, and to make the Web of Data less sensitive to outages of single hubs. Therefore, while the paper is intended to provide a general overview, we will especially focus on outgoing links and the the interplay of the spatial, temporal and thematic dimensions of Linked Sensor Data.

The Deepwater Horizon oil spill in the Gulf of Mexico serves as a running example to demonstrate general challenges and to point towards specific solutions. To do so, we have updated the Freebase page on the oil spill to provide up-to-date information. The oil spill is an interesting use case as there are numerous different sources for data on the oil spill, a number of affected parties and interest groups, and there is ample public interest in independent data. Moreover, the example demonstrates the value of raw data and freedom of interpretation, since many documents that were released by official sources were biased into a certain direction.

---

[2] See `http://52north.org/SensorWeb/clients/OX_RESTful_SOS/index.html`.

The remainder of this paper is organized as follows. In the next section, we give an overview of relevant related work from the areas of Linked Data and the Semantic Sensor Web. Section 3 illustrates the different motivations for providing observations and measurements as Linked Data. Section 4 discusses the question how sensor data can be turned into Linked Data. The technical requirements and approaches for implementation are discussed in Section 5, followed by conclusions and an outlook on future work in Section 6.

2 Related Work

The Semantic Sensor Web [1] is essentially a fusion of technologies of the Sensor Web on the one hand and the Semantic Web on the other. The Sensor Web builds on standards for services such as the Sensor Observation Service (SOS) and the Sensor Planning Service (SPS), as well as on data models and encodings such as Observations and Measurements (O&M) or the Sensor Model Language (SensorML). These standards are developed unter the umbrella of the Sensor Web Enablement (SWE) initiative[3]. While these specifications are adopted by an ever-growing number of sensor data providers, they target syntactic rather than semantic interoperability. The semantics of the provided observations, procedures, and observed properties remains ambiguous to a certain degree. This is especially relevant for the discovery and retrieval of sensor data [15]. The Semantic Web [16] targets these problems for arbitrary data using ontologies, semantic annotations, as well as deductive and inductive reasoning.

Combinations of these two infrastructures have been proposed in a number of different flavors. Sheth et al. [1] proposed a metadata approach for SWE services using RDFa[4] (RDF in attributes). Based on a Semantic SOS [4], rule-based reasoning on data from different sensors has been demonstrated in an application that identifies potentially dangerous weather conditions. Neuhaus and Compton [17] introduce an ontology for sensor descriptions that links a sensor to its measurement process, the physical feature for which a certain value is observed, and the corresponding domain of discourse. Devaraju et al. [18] combine this approach with a generic process ontology to facilitate sensor data retrieval. In a previous paper, we have outlined a semantic enablement approach for spatial data infrastructures that enables reasoning on spatial – and specifically on sensor – data that does not require a modification of established OGC standards [3]. It can thus be implemented without blocking access for 'non-semantic' clients.

These approaches to the Semantic Sensor Web all rely on the Semantic Web layer cake and especially on the Web Ontology Language (OWL) or the Semantic Web Rule Language (SWRL), which implies a level of complexity that is often not required and leads to new problems instead of solving them. The latest incarnation of the Semantic Sensor Web thus takes a more light-weight approach based on Linked Data. Providing sensor data in RDF format has been proposed by different researchers [19,20,9,10], as this format exposes observation data to

---

[3] See `http://opengeospatial.org/ogc/markets-technologies/swe` for an overview.

[4] See `http://www.w3.org/2006/07/SWD/RDFa/`.

a large number of clients and users that are often not aware of the geospatial services defined by the OGC. Moreover, this approach allows for easy integration with other sources in the Linked Data cloud[5]. Existing implementations range from static, converted data sets[6] to tools for on-the-fly translation between OGC services and RDF [10]. The pure mapping of encodings between the GML-based OGC standards and RDF is straight-forward as they are isomorphic [21].

Patni et al. [22] discuss the challenge of provenance in Linked Sensor Data, which is especially challenging for phenomena that are observed by a number of different sensors. The paper applies a provenance ontology to solve this problem that establishes an explicit link between the observed phenomenon and all involved sensors. Janowicz et al. [10] illustrate the need for a Linked Data Model in addition to classical data and conceptual models and discuss the challenge of assigning meaningful URIs [20] for highly dynamic information derived from sensor data.

3 Motivations: Why?

This section discusses the motivations for making sensor data available using Linked Data principles [7] and what makes Linked Data more than just another encoding. It also discusses some first ideas on how to benchmark the success of a Linked Sensor Data project.

3.1 Motivations

The prime motivation to publish data about sensors and their observations as Linked Data is to make them available outside of Spatial Data Infrastructures, provide unique HTTP-resolvable identifiers using URIs, and hence ease the access and re-usability of sensor data as well as support their integration and fusing [23]. While this motivation highlights the role of sharing data, sensor data providers need to take into account several other aspects and understand their implications:

- *Why Linked Sensor Data instead of classical SDIs?* Besides increasing the number of potential clients and thus the usage of the service, the integration with external (non-OGC) data sources is a classical task that can be addressed by using RDF as common data encoding. Additionally, at least for government data, it may turn out that open exchange formats for raw data become a legal requirements in the near future[7]. Such a legislative initiative would be especially desirable for natural disasters, as it would allow for informed, independent evaluations, and increase the accessibility of relevant data sets for local interest groups. In case of the oil spill example, observation

---

[5] See `http://richard.cyganiak.de/2007/10/lod/` for the current version.

[6] See, for example, `http://wiki.knoesis.org/index.php/SSW_Datasets`.

[7] See `http://data.gov.uk/` and `http://www.data.gov/`, for example.

data on the position of underwater oil plumes could be compared against different spreading models, or the data could be combined with marine data on fish habitats, for example. Finally, the close relationship between Linked Data and ontologies as conceptual reference models offers a promising alternative to one of the Achilles' heels of SDIs – namely, catalog services and code lists. An impressive example, demonstrating the interplay of Semantic Web technologies, ontologies, and Linked Spatiotemporal Data was recently discussed by Vilches-Blazquez et al. [24] for hydrographical data.

- *When we say Linked Data, do we mean it?* Linked Sensor Data can only unfold its full potential if the linking part is taken seriously [14]. Accordingly, it is crucial to identify other sources in the Linked Data cloud that could be linked in a meaningful way; see Section 4 for details. While links between data differ from the classical inter-document links of the Web, they are still created for some purpose and to express relatedness. However, providers of sensor data are interested in keeping their repositories as free of a particular interpretation as possible. It is unlikely that a provider of sensor data about water quality will link certain data sets to a DBpedia entry about the Deepwater Horizon oil spill. Therefore, in most cases links will be created on-the-fly by users as knowledge engineers [25]. Based on the experience with documents on the Web so far, incoming and outgoing links may therefore become an issue at law. This is especially interesting as, in contrast to classical Web links, the *owl:sameAs* construct is bidirectional. With a growing interest in Linked Data, link hubs such as sameAs.org will need to find a solution to the curation and ownership of links.

While an RDF representation supports the integration of O&M data, sensor metadata and ontologies, it also reduces the spatial querying capabilities. While GML and RDF are isomorphic, complex spatial queries, are only standardized for GML so far. Most Semantic Web applications and reasoners still reduce spatial queries to simple *containment* based on bounding boxes or *nearby* with respect to point data [26]. This situation will likely change in the future, as GeoSPARQL – a GML-compliant spatial extension to SPARQL – is currently under development in a special interest group at the OGC. In case of the oil spill scenario, many interesting queries require rather complex spatial operators such as buffers or overlaps. The following GeoSPARQL example (adapted from the OGC working group) queries for turtle habitats that have been reached by the oil:

```
PREFIX ogc: <http://www.opengis.net/rdf#>
PREFIX ex: <http://www.example.com#>
SELECT ?habitat
WHERE { ?habitat ogc:overlaps ex:oilSpill }
```

The underlying geometries must be defined as literals in well-known text (WKT) format, as in the following example showing an extract from the specification of a habitat of the endangered Green Turtle:

```
ex:greenTurtleHabitat ex:hasWKTSerialization
 ‘‘Polygon((28.7366 -88.3659, ...))’’ .
```

It is worth mentioning that transforming such geometries to an RDF-based representation is not necessarily useful, but it always adds complexity. Therefore, providers need to decide from case to case whether such a transformation is reasonable, i.e., whether it adds additional retrieval, data mining, or reasoning capabilities [10].

3.2 Benchmarks

While the success of a certain Web page or application can be measured in terms of its search engine ranking, incoming links, user counts, and so forth, measuring the success of Linked Sensor Data will require other or additional criteria. For instance, while the number of incoming links can serve as an estimation of the degree of embeddedness within the Linked Data cloud, the type of links play a role as well. If most of these links are owl:sameAs links, they could either connect different information about the same entity and, hence, enrich both data sets, but they could also indicate that the provided Linked Data are rather needless or regarded as redundant. A typed version of an algorithm like PageRank [27] could be used to rate the different sources on the Linked Data cloud. A weather service with thousands of users per day will typically be rated more important than a prototype platform in its early stages. Likewise, important hubs of the Linked Data Cloud as sources of incoming links will receive higher weights than small, specialized datasets. Semantic ping services such as `http://pingthesemanticweb.com/` are another useful way to keep track of incoming links. In contrast, the number of generated triples should not be used as benchmark as even single sensors may produce hundreds of new observation triples per minute.

4 Outgoing Links: To What?

This section introduces relevant datasets to which Linked Sensor Data can refer to. Moreover, we outline how to identify potential outgoing links depending on the corresponding application and argue why these links have to be curated.

4.1 Outgoing Links

Links are the glue of the Web of data and the connections that turn single, isolated information silos into one global graph. They enable cross-dataset queries, such as the comparison of the current state of a feature of interest with historical data [28] in the first place. Nonetheless, the generation of Linked Data is often reduced to the conversion of an existing data source to RDF – without discussing how to *link* the data to other sources in the Linked Data cloud. In order to properly embed Linked Sensor Data in the Linked Data cloud and increase findability, it is hence crucial to identify useful sources for further information that are related to a specific service, the sensors it offers, the observed phenomena, or the features of interest. Moreover, picking the appropriate vocabularies is an essential step to support retrieval and reuse of data. Besides the omnipresent `rdf:about` and `owl:sameAs`, popular vocabularies include Dublin Core (DC) for metadata, Friend Of A Friend (FOAF) for relationships among people, or the Simple Knowledge Organisation System (SKOS) for categorizations and concept maps. A vocabulary for sensors and observations is currently developed by the W3C Semantic Sensor Network Incubator Group[8].

The following example shows outgoing links for a sensor observing the Deepwater Horizon oil spill, using the Dublin Core, Semantic Sensor Web, and Marine Metadata Interoperability Platforms ontologies. Outgoing links point to a FOAF profile, the freebase entry on the Deepwater horizon explosion, to the Geonames entry on the Gulf of Mexico, and to the New York Times data entry about BP.

```
...
@prefix dc:  <http://purl.org/dc/elements/1.1/> .
@prefix obs: <http://knoesis.wright.edu/ssw/ont/sensor-observation.owl> .
@prefix mmi: <http://mmisw.org/ont/mmi/platform> .
@prefix son: <http://www.csiro.au/Ontologies/2009/SensorOntology.owl> .
...
:dhSOS
   dc:description "Deepwater Horizon Observation" ;
   dc:creator <http://ifgi.uni-muenster.de/~kessler/foaf.rdf> ;
   dc:subject <http://rdf.freebase.com/rdf/en.deepwater_horizon_drilling_rig_explosion> ;
   dc:coverage <http://sws.geonames.org/3523271/about.rdf> ;
   dc:relation <http://data.nytimes.com/63774392544048824322> ;
   obs:observedProperty :oilConcentration ;
   dc:relation :dhBuoy .
...
:oilConcentration rdf:type obs:PropertyType .
:dhBuoy rdf:type mmi:NeutrallyBuoyantFloat .
:sensor1 rdf:type son:Sensor .
...
```

4.2 Link Recommendations and Curation

Since links are a core component of Linked Data, their quality and accuracy is key to meaningful retrieval and reasoning. Undefined vocabulary terms, mismatched semantics, and unintended inferences can reduce the quality of linked data or even render them useless[9]. Consequently, a fully automated generation of links and especially of `owl:sameAs` relations to existing sources may rather do harm than add semantics. We therefore propose a curated approach where relations are recommended to the service administrator based on a previously selected set of Linked Data sources and vocabularies for relations. Adding semantic annotations

---

[8] See `http://www.w3.org/2005/Incubator/ssn/`.

[9] See `http://pedantic-web.org/` for a detailed discussion.

based on automatically generated recommendations has already been proposed for Volunteered Geographic Information [29]; we adopt this approach here for Linked Sensor Data. Figure 1 shows the workflow for link recommendation and curation:

1. Sensor metadata (as defined in the service capabilities) and O&M data are converted to RDF documents [20].
2. Keywords are extracted that match entities on the linked data cloud.
3. These keywords and their specifications and sources are presented to the user, who can then establish the first `owl:sameAs` relations.
4. Further potential thematic matches from the Linked Data cloud are computed based on similarity [30]. Potential spatial matches are computed based on co-location and containment.
5. These potential matches are presented to the user with an indication of the degree of similarity. The user can then curate these recommendations and establish the relations with the appropriate vocabularies.

The huge amount of sensor data will require the definition of templates for O&M data that are applied to new observations as they come in. The overall linking approach follows the idea of *bootstrapping*: As the amount of Linked Sensor Data increases, the linking opportunities increase as well. Therefore, new potential links can eventually be discovered and recommended after every iteration by inspection of the outgoing link selected in the previous step. Note, however, that there is no cold start problem, as newly created Linked Sensor Data can link to other parts of the Linked Data cloud as long as no related Linked Sensor Data are available. A more connected Linked Data cloud could be generated if the underlying ontologies specifying the used vocabularies would be linked to each other [31]; however, this is out of scope for this research. Figure 2 shows a conceptual design for are curation interface that implements this functionality for service administrators.

In the mapping process, the recommender service automatically replaces any concrete values by variables, so that later measurements following the same scheme can be converted on the fly. The following code shows a sample extract from a template for oil concentration measurements.

```
...
@prefix obs: <http://knoesis.wright.edu/ssw/ont/sensor-observation.owl> .
...
:observation$id obs:result :result$id ;
:result$id rdf:value ?//om:result/swe:DataArray/swe:values
...
```

The `$id` variable in templates is replaced by unique IDs upon conversion, so that every literal (e.g., every observation) within the given name space is distinct. Fragments starting with `?` contain an XPATH query for the node in the input O&M document whose value is to be integrated here. These queries are generated by the mapping engine on the fly when the administrator creates a sample mapping for a single O&M document.

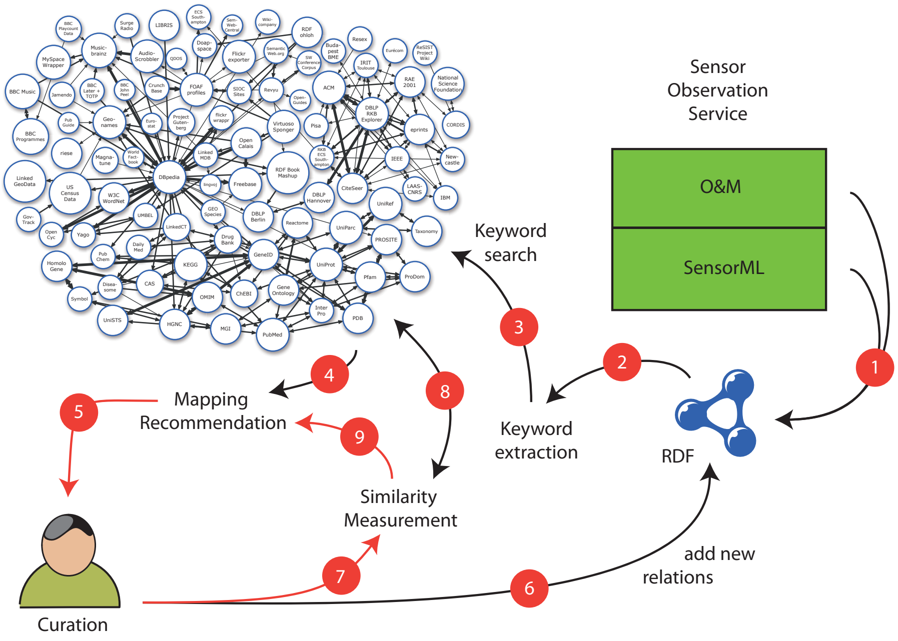


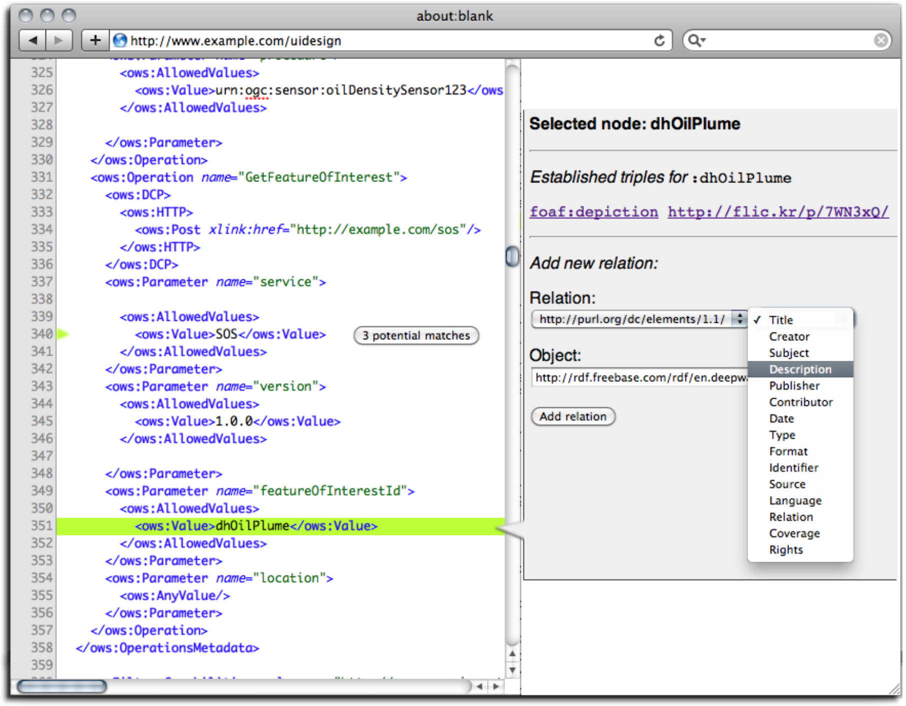
Fig. 1. Workflow for link curation. After converting O&M and SensorML data to RDF (1), keywords are extracted (2) and matched against previously selected sources in the Linked Data cloud (3). Matches are recommended to (4) and curated by (5) the user, who adds new relations to the RDF documents and templates this way (6). These relations can then iteratively be used as input for a similarity measurement (7) that finds similar concepts and entities in the Linked Data cloud (8), which serve again as input for the recommendation (9). [Linked Data cloud figure by Richard Cyganiak.]

5 Implementation: How?

The section discusses the technical aspects of publishing Linked Sensor Data. We focus on the peculiarities caused by the spatio-temporal dynamics of sensor data and the challenges they cause for representation, reasoning and provenance.

5.1 Handling the Dynamics of Linked Sensor Data

Any kind of sensor-related data is inherently dynamic. While this is obvious for observation data whose sole intention is to keep track of the dynamics of the real world, it also applies to meta-data about sensors. Sensors can be relocated, their *feature of interest* can change, and the actual instruments can be replaced. Phenomena such as the Deepwater Horizon oil spill are characterized by their three-dimensional spatial distribution. Moreover, they also involve a temporal



```
325         <ows:AllowedValues>
326           <ows:Value>urn:ogc:sensor:oilDensitySensor123</ows
327         </ows:AllowedValues>
328 
329       </ows:Parameter>
330     </ows:Operation>
331     <ows:Operation name="GetFeatureOfInterest">
332       <ows:DCP>
333         <ows:HTTP>
334           <ows:Post xlink:href="http://example.com/sos"/>
335         </ows:HTTP>
336       </ows:DCP>
337       <ows:Parameter name="service">
338 
339         <ows:AllowedValues>
340           <ows:Value>SOS</ows:Value>
341         </ows:AllowedValues>
342       </ows:Parameter>
343       <ows:Parameter name="version">
344         <ows:AllowedValues>
345           <ows:Value>1.0.0</ows:Value>
346         </ows:AllowedValues>
347 
348       </ows:Parameter>
349       <ows:Parameter name="featureOfInterestId">
350         <ows:AllowedValues>
351           <ows:Value>dhOilPlume</ows:Value>
352         </ows:AllowedValues>
353       </ows:Parameter>
354       <ows:Parameter name="location">
355         <ows:AnyValue/>
356       </ows:Parameter>
357     </ows:Operation>
358   </ows:OperationsMetadata>
359 
```

Selected node: dhOilPlume

Established triples for `:dhOilPlume`

`foaf:depiction` `http://flic.kr/p/7WN3xQ/`

Add new relation:

Relation: http://purl.org/dc/elements/1.1/ Title

Object: http://rdf.freebase.com/rdf/en.deepw

Fig. 2. Conceptual design for curation user interface. After selecting a node in the service output (a capabilities document, in this case), the service offers mappings and allows the user to insert link targets. Potential matches identified by the keyword mapping and similarity reasoners are highlighted and can be curated by clicking the corresponding button shown in line 340 of the capabilities document.

dimension as well as various attributive aspects. As discussed in previous work [10,25], assigning URIs to properties and features of interest is not straightforward. Similarly, difficulties may arise when using the URIs provided for RESTful access to sensor data stored in SOS as references. For instance, constructing a triple *<#OilSpill> <#observedBy> <#Sensor1>* bears the danger that the sensor will be deployed in a different environment in the future. Additionally, a URI such as `http://.../sos/observations/Sensor1/SpillRegion42/WindDirection` refers to all wind direction observations of Sensor1 for the region 42 [10]. As new records are added over time, the URI does not provide a unique reference for a specific set of observations. Some of these difficulties go beyond certain characteristics of the Sensor Web, as they are essential design issues of RDF; see also Hayes' notion of surfaces[10].

---

[10] See `http://www.ihmc.us/users/phayes/RDFGraphSyntax.html`.

The URI encoding for dynamic information discussed above makes information on timestamps available for clients working directly on one these service URIs. However, once the delivered data is cached or passed on for further processing, information on *when* the represented facts were valid is lost. This problem applies to all kinds of data delivered via such URIs, which can come in different forms based on content negotiation between client and server [32]. For example, an RDF triple – once published – does not bear any information about whether the encoded fact is still valid or not. The temporal dimension therefore also needs to be covered *within* the dataset. The following example demonstrates how this information can be attached using a named graph [33]:

```
...

:G1 { ex:sensor128 swe:hasFOI ex:foiPlume .
      :G1 dc:date "2010-08-30T11:00:00-5:00" }
...
```

It is thus made explicit that `sensor 128` is observing a feature of interest called `foiPlume` on August 30, 2010 at 11:00 AM local time.

5.2 Provenance and Persistence

The curation approach outlined in Section 4.2 is based on the assumption that a human user is required to confirm (or reject) proposed links. The data provider, for example, has additional contextual information and can therefore make sense of literals in RDF triples. He can hence make a more informed decision about the relations to establish and the appropriate vocabularies for that. In order to allow users of the data generated this way to assess their quality, however, meta-data about the relations are required. This applies especially in the open government data case, where legal liabilities may be implied. Particularly information about the creation timestamp of a new relation, the recommender engine (if applicable), as well as the curator provide useful information on the provenance of a dataset. Such information may even be used in clients to process only triples confirmed by trusted curators, for example. These meta-data can also be attached using named graphs, as demonstrated in Section 5.1.

An equally important issue to make Linked Sensor Data successful in the long run is the question how to store observation data persistently. Data on natural phenomena are potentially useful to lern about processes such as climate change. Simply storing *all* collected observations, however, does not seem feasible as we have already reached a point where more data than storage (hard disks etc.) is being produced at any time [34]. This trend is likely to continue, as more and more sensors are being deployed, and humans as sensors [35] produce even more potentially useful data. While a detailed discussion of this issue is out of scope for this paper, the Linked Data cloud bears potential to overcome this problem. As information from numerous different sources is available in the cloud, it should be possible to thin out the comprehensive data collection so that only relevant data are kept. In addition to periodic observation data, more dense data should be kept for specific events such as natural disasters [13]. The automatic identification of such phenomena based on different sources on the Linked Data cloud will require an annotation of the observation data that goes beyond RDF, as it has been indicated in previous research [1,4,3].

6 Conclusions and Future Work

Linked Sensor Data is a promising approach to make observation data available for clients that are not compliant to OGC SWE standards, and to facilitate data integration with other sources on the Semantic Web. In this position paper, we have discussed the motivations and challenges for publishing Linked Sensor Data. We have stressed the importance of embedding Linked Sensor Data properly into the Linked Data cloud. Different target data sets and vocabularies have been discussed. In order to facilitate the linking process, we have outlined a semi-automatic approach based on recommendation and curation that helps service providers to establish the required mappings. The challenges that arise from the spatio-temporal dynamics of sensors and the corresponding observation data have been pointed out, especially with respect to finding appropriate representations in RDF. We have proposed to turn triples describing spatio-temporally dependent properties into named graphs to enable the annotation with time stamps and locations. While this approach makes querying Linked Sensor Data more complex, it makes explicit what the meta-data refer to. The same applies for provenance data, which are required to fully document the lineage of both the observation data and their conversion into RDF. Finally, we have touched upon the persistence of Linked Sensor Data, which is especially challenging given the huge volumes of data that are produced.

The idea of Linked Sensor Data is a new approach that is just about to be put into practice, with first data sets and publishing tools available. Some of the ideas outlined in this paper will hence only be realizable in the medium term, when a reasonable number of Linked Sensor Data services are available. A first step in this direction is the further development of the RESTful Sensor Observation Service [10]. The current version only supports `GET` requests. An integration of the Sensor Planning Service within the same URI scheme could also make use of the `PUSH`, `POST` and `DELETE` requests to task a sensor. Moreover, the conceptual design for the relation recommender needs to be turned into an actual implementation. This requires an integration of the SIM-DL similarity server [36] which is currently being updated to a new version.

Acknowledgments

This work has been partly funded by the German Research Foundation's *SimCat* project (DFG Ra1062/2-1 and Ja1709/2-2; see `http://sim-dl.sf.net`) and is part of the 52°*North* semantics community; see `http://52north.org/semantics`.

References

1. Sheth, A., Henson, C., Sahoo, S.: Semantic Sensor Web. Internet Computing, IEEE **12**(4) (2008) 78–83
2. Maué, P., Schade, S., Duchesne, P.: Semantic annotations in OGC standards. Discussion paper 08-167r1, available from `http://portal.opengeospatial.org/files/?artifact_id=34916`, OGC (2009)
3. Janowicz, K., Schade, S., Bröring, A., Keßler, C., Maué, P., Stasch, C.: Semantic Enablement for Spatial Data Infrastructures. Transactions in GIS **14**(2) (2010) 111–129
4. Henson, C.A., Pschorr, J.K., Sheth, A.P., Thirunarayan, K.: SemSOS: Semantic Sensor Observation Service. In: Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems (CTS 2009), Baltimore, MD. (2009)
5. Keßler, C., Raubal, M., Wosniok, C.: Semantic Rules for Context-Aware Geographical Information Retrieval. In Barnaghi, P., ed.: 4th European Conference on Smart Sensing and Context, EuroSSC 2009. Volume 5741 of Lecture Notes in Computer Science., Berlin, University of Surrey, Springer (2009) 77–92
6. Phuoc, D.L., Hauswirth, M.: Linked Open Data in Sensor Data Mashups. In Kerry Taylor, Arun Ayyagari, D.D.R., ed.: Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09) in conjunction with ISWC 2009. Volume Vol-522., CEUR (2009)
7. Berners-Lee, T.: Linked Data. Personal view available from `http://www.w3.org/DesignIssues/LinkedData.html` (2009)
8. Bizer, C., Heath, T., Berners-Lee, T.: Linked Data – The Story So Far. International Journal on Semantic Web and Information Systems **5**(3) (2009) 1–22
9. Patni, H., Henson, C., Sheth, A.: Linked sensor data. In: 2010 International Symposium on Collaborative Technologies and Systems, IEEE (May 2010) 362–370
10. Janowicz, K., Broering, A., Stasch, C., Everding, T.: Towards Meaningful URIs for Linked Sensor Data. In Devaraju, A., Llaves, A., Maué, P., Keßler, C., eds.: Towards Digital Earth: Search, Discover and Share Geospatial Data. Workshop at Future Internet Symposium, September 20th, 2010, Berlin, Germany. (forthcoming 2010)
11. Jain, P., Hitzler, P., Yeh, P.Z., Verma, K., Sheth, A.P.: Linked Data is Merely More Data. In: AAAI Spring Symposium 'Linked Data Meets Artificial Intelligence', AAAI Press (2010) 82–86
12. Kuhn, W.: Geospatial Semantics: Why, of What, and How. Journal on Data Semantics III (Special Issue on Semantics-based Geographical Information Systems) (2005) 1–24
13. Winter, S., Dupke, S.: Queries for historic events in geosensor networks. Journal of Location Based Services **2**(3) (2008) 177–193
14. Guéret, C., Groth, P., van Harmelen, F., Schlobach, S.: Finding the achilles heel of the web of data: using network analysis for link-recommendation. In: International Semantic Web Conference 2010. (forthcoming 2010)
15. Jirka, S., Bröring, A., Stasch, C.: Discovery mechanisms for the sensor web. Sensors **9**(4) (2009) 2661–2681
16. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American **284**(5) (May) 34–43

17. Neuhaus, H., Compton, M.: 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. (2009)
18. Devaraju, A., Neuhaus, H., Janowicz, K., Compton, M.: Combining Process and Sensor Ontologies to Support Geo-Sensor Data Retrieval. In: 6th International Conference on Geographic Information Science (GIScience 2010), Zurich, CH 14-17th September, 2010. (forthcoming 2010)
19. Balazinska, M., Deshpande, A., Franklin, M.J., Gibbons, P.B., Gray, J., Hansen, M., Liebhold, M., Nath, S., Szalay, A., Tao, V.: Data Management in the Worldwide Sensor Web. IEEE Pervasive Computing **6**(2) (2007) 30–40
20. Page, K., Roure, D.D., Martinez, K., Sadler, J., Kit, O.: Linked Sensor Data: RESTfully serving RDF and GML. In Taylor, K., Ayyagari, A., Roure, D.D., eds.: Proceedings of the 2nd International Workshop on Semantic Sensor Networks (SSN09), collocated with the 8th International Semantic Web Conference (ISWC-2009), Washington DC, USA, October 26, 2009. (2009) 49–63
21. Schade, S., Cox, S.: Linked Data in SDI or How GML is not about Trees. In Painho, M., Santos, M.Y., Pundt, H., eds.: Proceedings of the 13th AGILE International Conference on Geographic Information Science – Geospatial Thinking, Association of Geographic Information Laboratories for Europe (AGILE) (2010)
22. Patni, H., Sahoo, S.S., Henson, C., Sheth, A.: Provenance Aware Linked Sensor Data. In Kärger, P., Olmedilla, D., Passant, A., Polleres, A., eds.: Proceedings of the Second Workshop on Trust and Privacy on the Social and Semantic Web, Heraklion, Greece, May 31, 2010. (2010)
23. Corcho, O., García-Castro, R.: Five Challenges for the Semantic Sensor Web. Accepted for publication in Semantic Web – Interoperability, Usability, Applicability; see `http://www.semantic-web-journal.net/content/new-submission-five-challenges-semantic-sensor-web` (2010)
24. Vilches-Blazquez, L.M., Villazon-Terrazas, B., Leon, A.D., Priyatna, F., Corcho, O.: An Approach to Publish Spatial Data on the Web: The GeoLinked Data Use Case. In Janowicz, K., Pehle, T., Hart, G., Maué, P., eds.: Workshop On Linked Spatiotemporal Data, in conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010). Zurich, 14th September. (forthcoming 2010)
25. Janowicz, K.: The Role of Space and Time For Knowledge Organization on the Semantic Web. Semantic Web – Interoperability, Usability, Applicability (2010; to appear)
26. Auer, S., Lehmann, J., Hellmann, S.: LinkedGeoData: Adding a Spatial Dimension to the Web of Data. In: ISWC '09: Proceedings of the 8th International Semantic Web Conference, Berlin, Heidelberg, Springer-Verlag (2009) 731–746
27. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank Citation Ranking: Bringing Order to the Web. Technical report, Stanford Digital Library Technologies Project (1998)
28. Simon, R., Sadilek, C., Korb, J., Baldauf, M., Haslhofer, B.: Tag Clouds and Old Maps: Annotations as Linked Spatiotemporal Data in the Cultural Heritage Domain. In Janowicz, K., Pehle, T., Hart, G., Maué, P., eds.: Workshop On Linked Spatiotemporal Data, in conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010). Zurich, 14th September. (forthcoming 2010)

29. Westermann, A.: Building a REST based Annotation Service for Volunteered Geographic Information. Bachelor's thesis, Institute for Geoinformatics, University of Münster, Germany (2010)
30. Janowicz, K., Wilkes, M., Lutz, M.: Similarity-based Information Retrieval and its Role within Spatial Data Infrastructures. In Cova, T.J., Miller, H.J., Beard, K., Frank, A.U., Goodchild, M.F., eds.: 5th International Conference on Geographic Information Science. (2008)
31. Parundekar, R., Knoblock, C.A., Ambite, J.L.: Aligning Ontologies of Geospatial Linked Data. In Janowicz, K., Pehle, T., Hart, G., Maué, P., eds.: Workshop On Linked Spatiotemporal Data, in conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010). Zurich, 14th September. (forthcoming 2010)
32. Schade, S., Granell, C., Diaz, L.: Augmenting SDI with Linked Data. In Janowicz, K., Pehle, T., Hart, G., Maué, P., eds.: Workshop On Linked Spatiotemporal Data, in conjunction with the 6th International Conference on Geographic Information Science (GIScience 2010). Zurich, 14th September. (forthcoming 2010)
33. Carroll, J.J., Bizer, C., Hayes, P., Stickler, P.: Named graphs, provenance and trust. In: WWW '05: Proceedings of the 14th international conference on World Wide Web, ACM (2005) 613–622
34. The Ecomomist: Data, data everywhere. Special report in issue from February 25 (2010)
35. Goodchild, M.: Citizens as sensors: the world of volunteered geography. GeoJournal **69**(4) (2007) 211–221
36. Janowicz, K., Keßler, C., Schwarz, M., Wilkes, M., Panov, I., Espeter, M., Bäumer, B.: Algorithm, Implementation and Application of the SIM-DL Similarity Server. In Fonseca, F., Rodriguez, M.A., eds.: Second International Conference on GeoSpatial Semantics, GeoS 2007. Lecture Notes in Computer Science 4853, Springer-Verlag Berlin Heidelberg (2007) 128–145