

# Querying Semantically Enriched Sensor Observations: Short Paper

Jean-Paul Calbimonte<sup>1</sup>, Hoyoung Jeung<sup>2</sup>, and Oscar Corcho<sup>1</sup>

<sup>1</sup>Ontology Engineering Group, Universidad Politécnica de Madrid, Spain  
`jp.calbimonte@upm.es, ocorcho@fi.upm.es`

<sup>2</sup>Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland  
`hoyoung.jeung@epfl.ch`

**Abstract.** The increase that sensor network deployments are experiencing makes them one of the fastest growing sources of data nowadays. They allow live and in-situ data capture and provide high volumes of information about the real world that surrounds us. However their heterogeneity and error-proneness make them hard to use, as we normally require lots of details of each one before their measurements can be combined. We present an ontology-based approach for querying live sensor observations, enriched with semantic metadata. The metadata is used in two scenarios: (i) for adding contextual information to the sensor observations, such as which sensor recorded what, where, when, and in which conditions; and (ii) to help users filter sensor data in their queries. We provide automatic query translation and join of sensor data and metadata, using mappings of the underlying sensor schemas to semantically annotate sensor observations. With this approach we can express queries in terms of a high level sensor network ontology, using an extension of the SPARQL query language that considers window constructs.

## 1 Introduction

Sensor data applications are changing the way we conceive data acquisition and processing systems, and their use is extended to an increasing number of domains. However the great potential of these information data sources is sometimes lowered by the difficulty of managing and querying sensor data [1]. One of these difficulties is related to the lack of structure and coherence of data in large sensor network deployments. Systems require access to sensors from different vendors and with specific characteristics and data schemas, each of them producing different values, with different precision or accuracy, in different units of measurement. This heterogeneity complicates the task of querying and exploiting data from the sensor deployments coherently.

In this paper we address the problem of managing sensor data and metadata for distributed query processing, using semantic technologies. Using the SSN ontology<sup>1</sup> as a common model for the underlying heterogeneous sensor data

<sup>1</sup> <http://purl.oclc.org/NET/ssnx/ssn>, W3C Semantic Sensor Network Incubator Group (SSN XG) Ontology

sources, we expose a semantic query interface that can be used by users and applications, over both the sensor metadata and the actual observation values. This work has been applied to a real deployment, the Swiss Experiment (SwissEx)<sup>2</sup> sensor network platform running on the GSN<sup>3</sup> infrastructure[2], devised to support environmental research.

The paper is organized as follows: we describe in Section 2 the process of modeling the metadata using the SSN ontology, and the mappings from sensor data schemas. In Section 3 we introduce the ontology-based query translation approach we used. Finally we present our conclusions after describing some of the relevant related work.

## 2 Sensor ontology modeling

Sensors and their observations have been modeled with ontologies in the past (see Section 4), some of them focused on sensor descriptions, others in observations[3]. However most of these formalizations were project-specific, or discontinued, and some of them only covered a limited number of sensor characteristics. In order to overcome these issues the W3C SSN XG group<sup>4</sup> was created, with the goal of providing a generic and domain independent ontology, compatible with the OGC<sup>5</sup> standards, both at the Sensor and Observation levels.

For instance consider a Temperature Monitor sensor in a weather station deployed at the Wannengrat field site. The sensor is capable of measuring the air temperature on its location every 10 minutes. In terms of the SSN ontology the temperature measurements can be seen as observations, with a *feature of interest* (i.e. the air), and referring to an *observed property*, in this case the temperature (See Listing 1.1 for a sample representation). Notice that the SSN model does not define the possible types of observed properties, but these can be taken from a specialized vocabulary such as the NASA SWEET<sup>6</sup> ontology. The observation results are sensor outputs to which the actual data values can be attached, as instances of a data type.

```
swissex:WindSpeedObservation1
  rdf:type ssn:Observation ;
  ssn:featureOfInterest
    [ rdf:type sweet:Air ] ;
  ssn:observedProperty
    [ rdf:type sweetProp:Temperature ] .
ssn:observationResult
  [ rdf:type ssn:SensorOutput ;
    ssn:hasValue [qudt:numericValue "4.345"^^xsd:double]];
```

**Listing 1.1.** SSN representation of a Air Temperature observation in RDF

<sup>2</sup> <http://www.swiss-experiment.ch/> The Swiss Experiment Platform

<sup>3</sup> <http://sourceforge.net/apps/trac/gsn/> Global Sensor Networks

<sup>4</sup> <http://www.w3.org/2005/Incubator/ssn/> SSN Incubator Group

<sup>5</sup> <http://www.opengeospatial.org/> Open Geospatial Consortium

<sup>6</sup> <http://sweet.jpl.nasa.gov/> Semantic Web for Earth and Environmental Terminology

Although the observation model provides a semantically enriched representation of the data, sensors generally produce streams of raw data with very little structure and therefore there is a gap between the observation model and the original data. We take an ontology mapping-based approach to overcome this problem. We propose providing mappings from sensor data to the ontology using a declarative mapping language <sup>7</sup>. Basing our work on Ontology-based data access, we are able to express for instance, that for every sensor tuple of the `wan5` sensor, an instance of a SSN `ObservationValue` must be created, composing its URI according to a mapping rule (template), and the observation result value must be extracted from the `air_temperature` sensor field.

```

:Wan5Temp a rr:TriplesMapClass;
  rr:tableName "wan5";
  rr:subjectMap
    [ rr:template "http://swissex.ch/metadata#Wan5/Temperature/ObsValue/{timed}";
      rr:column "timed"; rr:class ssn:ObservationValue;
      rr:graph swissex:WannengratSensors.srdf ];
  rr:predicateObjectMap [ rr:predicateMap [ rr:predicate qudt:numericValue ];
                        rr:objectMap [ rr:column "air_temperature" ] ];
.

```

**Listing 1.2.** Mapping a sensor to a SSN `ObservationValue` in R2RML

By using the mappings and the SSN ontology, we are able to express the sensor metadata and observations data using a semantic model, even if the underlying data sources are relational streams. In the next section we provide details about the query processing and translation process that is carried out to query these sources through `SPARQLStream`.

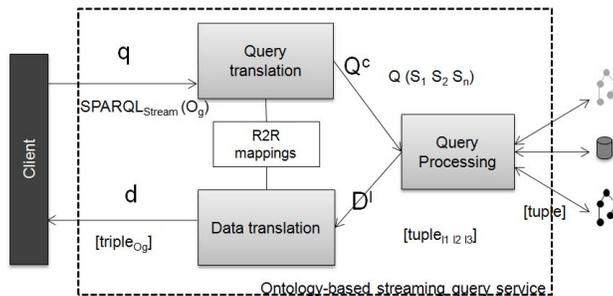
### 3 Ontology-based sensor data querying

The goal of ontology-based streaming data access is to generate semantic web content from existing streaming data source. Although previous efforts have been made in order to provide semantic content automatically from relational databases using mappings [4], only recently this idea has been extended to streams [5]. Our approach to enable ontology-based access to streaming data is depicted in Fig 1.

The service receives queries specified in terms of the classes and properties of the ontology using `SPARQLStream`, an extension of SPARQL that supports operators over RDF streams, which has been inspired by previous proposals such as C-SPARQL [6]. It improves and corrects window types and operations support with respect to previous extensions, in the sense that: (i) the result of a window operation is a window of triples, not a stream, over which traditional operators can be applied, (ii) we allow higher bounds for time windows, i.e. windows in the past, and (iii) we have adopted the SPARQL 1.1 <sup>8</sup> definition of aggregates.

<sup>7</sup> We use the W3C RDB2RDF Group, R2RML mapping language <http://www.w3.org/2001/sw/rdb2rdf/r2rml/> to represent the mappings

<sup>8</sup> <http://www.w3.org/TR/sparql11-query/> SPARQL 1.1 Query Language, working draft.



**Fig. 1.** Ontology-based sensor query service

In order to transform the  $\text{SPARQL}_{\text{Stream}}$  query, expressed in terms of the ontology, into queries in terms of the data sources, a set of mappings must be specified. These mappings are expressed in R2RML. The language is used to define declarative mappings from relational sources to datasets in RDF, as described in Section 2. This transformation process is called *query translation*, and the target is a continuous query language or a streaming query expression. After the continuous query has been generated, the query processing phase starts, and the execution is delegated to a sensor data processing engine, such as GSN. The transformation is based on the principle of transforming the incoming query into a relational algebra expression, extended with time window operators. As the window transforms streams into bounded sets of tuples, classical relational operators can be applied to them and finally a query expression can be generated from the algebra to a target language [5].

```

SELECT ?speed
FROM NAMED STREAM swissex:WannengratSensors.srdf [NOW - 5 HOUR ]
WHERE
{ ?obsresult a ssn:ObservationValue;
  qudt:numericValue ?speed. }

```

**Listing 1.3.**  $\text{SPARQL}_{\text{Stream}}$  query

For instance if the stream query engine is GSN, the algebra expression can be transformed to a GSN API<sup>9</sup> request. For example consider the  $\text{SPARQL}_{\text{Stream}}$  query in Listing 1.3, the query translation will produce a GSN request like the one in Listing 1.4, using the R2RML mapping definition that relates the `wan5` sensor and its `air_temperature` to an Observation Value. The time window in the GSN request is generated at compilation time according to the  $\text{SPARQL}_{\text{Stream}}$  query declaration.

```

http://montblanc.slf.ch:22001/multidata?vs[0]=wan5&field[0]=air_temperature&
from=15/05/2011+05:00:00&to=15/05/2011+10:00:00&

```

**Listing 1.4.** Generation of a GSN API URL

<sup>9</sup> <http://sourceforge.net/apps/trac/gsn/wiki/web-interfacev1-server>

## 4 Related Work

Several efforts in the past have addressed the task of representing sensor data and metadata using ontologies, and also providing semantic annotations and querying over these source. As recounted in [3], many proposals for representing sensor data through ontologies have been presented in the literature. However in many of the early approaches the focus was on sensor meta information, while the description of observation was overlooked. Besides most of these approaches lack the best practices of ontology reuse and alignment with standards. Recent proposals of ontologies use the concepts defined in the OGC SensorML<sup>10</sup> standard as a basis. Some of them, like [7] and [8], also consider the OGC Observations and Measurements (O&M) standard<sup>11</sup> to represent observations captured by sensor networks.

More recently, through the W3C SSN-XG group, the semantic web and sensor network communities have made an effort to provide a domain independent ontology, generic enough to adapt to different use-cases, and compatible with the OGC standards at the sensor level and observation level. Therefore it is more likely to be adopted in the community and become a standard.

Approaches providing search and query frameworks that leverage semantic annotations and metadata, have been presented in several past works. The architectures described in n [9] and [10], rely on bulk-import operations that transform the sensor data into an RDF representation that can be queried using SPARQL in memory, lacking scalability and real-time querying capabilities. In [11], the authors describe a metadata management framework based on Semantic Wiki technology to store distributed sensor metadata, which is joined with streaming data coming form GSN for query purposes. Approaches for annotating OGC-compliant sensor services have been studied in [12] and [13]. However none of the previous works considers exposing live sensor data through SPARQL queries with time windows in a distributed environment.

## 5 Conclusions

We presented an approach that enables ontology-based querying of sensor data in a federated environment, considering metadata and mappings to underlying stream processing engines. The presented work uses the SSN ontology for effectively modeling the sensor metadata such that users can pose queries that exploit their semantic relationships. By querying using these abstractions, users do not require any knowledge about sensor specific names or their attributes.

In order to achieve this, we have extended the use of the R2RML language specification to the case of streaming sensor data, declaring mapping definitions that allow dynamic translation of SPARQL<sub>Stream</sub> queries into algebra expressions that can be used to generate queries or data requests like the GSN API URLs.

<sup>10</sup> SensorML. <http://www.opengeospatial.org/standards/sensorml>

<sup>11</sup> OGC O & M: <http://www.opengeospatial.org/standards/om>

In this way we have enabled distributed processing of queries in a federated sensor network environment, through a centralized semantic sensor metadata processing service.

In the near future we will expand this work to enable the publication of observation data from these federated sensor sources as Linked Data. We are also planning to integrate this platform with external data sources that may provide additional information about the sensors, or locations, or features of interest that are referenced in the metadata.

## References

1. Corcho, O., García-Castro, R.: Five challenges for the Semantic Sensor Web. *Semantic Web* **1**(1) (2010) 121–125
2. Aberer, K., Hauswirth, M., Salehi, A.: A middleware for fast and flexible sensor network deployment. In: *Proceedings of the 32nd international conference on Very large data bases, VLDB Endowment* (2006) 1199–1202
3. Compton, M., Henson, C., Lefort, L., Neuhaus, H., Sheth, A.: A survey of the semantic specification of sensors. In: *Proc. Semantic Sensor Networks*. (2009) 17
4. Sahoo, S.S., Halb, W., Hellmann, S., Idehen, K., Jr, T.T., Auer, S., Sequeda, J., Ezzat, A.: A survey of current approaches for mapping of relational databases to RDF. *W3C* (January 2009)
5. Calbimonte, J., Corcho, O., Gray, A.: Enabling ontology-based access to streaming data sources. In: *Proceedings of the 9th international semantic web conference on The semantic web-Volume Part I, Springer-Verlag* (2010) 96–111
6. Barbieri, D.F., Braga, D., Ceri, S., Della Valle, E., Grossniklaus, M.: C-sparql: Sparql for continuous querying. In: *WWW '09: Proceedings of the 18th international conference on World wide web, New York, NY, USA, ACM* (2009) 1061–1062
7. Barnaghi, P., Meissner, S., Presser, M., Moessner, K.: Sense and sens ability: Semantic data modelling for sensor networks. In: *Proceedings of the ICT Mobile Summit*. (2009)
8. Compton, M., Neuhaus, H., Taylor, K., Tran, K.: Reasoning about sensors and compositions. In: *Proceedings of the 2nd International Workshop on Semantic Sensor Networks, SSN09. Volume 522*. (2009)
9. Lewis, M., Cameron, D., Xie, S., Arpinar, B.: ES3N: A semantic approach to data management in sensor networks. In: *A workshop of the 5th International Semantic Web Conference ISWC, Citeseer* (2006)
10. Huang, V., Javed, M.: Semantic sensor information description and processing. In: *The Second International Conference on Sensor Technologies and Applications, IEEE* (2008) 456–461
11. Jeung, H., Sarni, S., Paparrizos, I., Sathe, S., Aberer, K., Dawes, N., Papaioannou, T., Lehning, M.: Effective Metadata Management in Federated Sensor Networks. In: *2010 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, IEEE* (2010) 107–114
12. Babitski, G., Bergweiler, S., Hoffmann, J., Schön, D., Stasch, C., Walkowski, A.: Ontology-based integration of sensor web services in disaster management. *GeoSpatial Semantics* (2009) 103–121
13. Henson, C., Pschorr, J., Sheth, A., Thirunarayan, K.: SemSOS: Semantic sensor Observation Service. In: *Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems, IEEE Computer Society* (2009) 44–53