ICT-MobileSummit 2009 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2009
ISBN: 978-1-905824-12-0

Sense and Sens'ability: Semantic Data Modelling for Sensor Networks

Payam Barnaghi, Stefan Meissner, Mirko Presser, and Klaus Moessner
Centre for Communication Systems Research (CCSR)
University of Surrey, GU2 7XH, United Kingdom
{p.barnaghi,s.meissner,m.presser,k.moessner}@surrey.ac.uk
http://www.ee.surrey.ac.uk/ccsr

Abstract: Sensor networks are used in various applications in several domains for measuring and determining physical phenomena and natural events. Sensors enable machines to capture and observe characteristics of physical objects and features of natural incidents. Sensor networks generate immense amount of data which requires advanced analytical processing and interpretation by machines. Most of the current efforts on sensor networks are focused on network technologies and service development for various applications, but less on processing the emerging data. Sensor data in a real world application will be an integration of various data obtained from different sensors such as temperature, pressure, and humidity. Processing and interpretation of huge amounts of heterogeneous sensor data and interoperability are important issues in designing a scalable sensor network architecture. This paper describes a semantic model for heterogeneous sensor data representation. We use common standards and logical description frameworks proposed by the Semantic Web community to create a sensor data description model. The work describes a sensor data ontology which is created based on the Sensor Web Enablement (SWE) and SensorML data component models. We describe how the semantic relationship and operational constraints are deployed in a uniform structure to describe the heterogeneous sensor data.

Keywords: Sensor Networks, Ontologies, Knowledge Modelling, SensorML

1. Introduction

The current Web is a document centric platform for exchanging data amongst the users. The Internet and Web applications in recent years have seen tremendous growth in facilitating data exchange for different applications and purposes. The current networks, however, are limited in sensing and measuring the physical world phenomena and employing them for observing and controlling real world incidents. Sensor networks provide a potential for Internet applications to acquire context data and observe and measure physical incidents. This will support constructing platforms which are aware of physical world incidents. This will enable construction of new services that remove the strict boundary between virtual and physical world. To achieve this, data collected from different types and levels of sensors and sensor networks will be used in different applications. Machines will need to collect and interpret the data provided by various types of sensor devices. This paper describes an ontology-based approach to structuring of data obtained from different types of sensors in a semantic framework. The rest of the paper is organised as follows. Section 2 describes background studies and related technologies. Section 3 discusses foundations of a sensor observation and measurement data model and describes the sensor data description ontology. Section 4 provides an evaluation of the work and section 5 concludes the paper and discusses the future work.

2. Background

The Open Geospatial Consortium (OGC)[1] has recently established a group which is called Sensor Web Enablement (SWE). This group is responsible for specifying interoperability interfaces and metadata encodings for integration of heterogeneous sensor data [1]. The main specifications defined by the group are described in the following.

- *Observations & Measurements (O&M)* which define standard models and XML Schema for encoding real-time and archived observations and measurements of sensor data.

- *Sensor Model Language (SensorML)* is a standard model to describe sensor systems and processes associated with sensor observations in an XML-based structure.

- *Transducer Model Language (TransducerML or TML)* provides a conceptual model to describe transducers and to support real-time data to and from sensor systems, sensors and actuators.

- *Sensor Observations Service (SOS)* is a standard Web service interface for requesting, filtering, and retrieving observations and sensor system information.

- *Sensor Planning Service (SPS)* is a standard Web service interface that acts as an intermediary between a client and a sensor collection management environment.

- *Sensor Alert Service (SAS)* is a standard Web service interface that enables publishing and subscribing to sensor alerts.

- *Web Notification Services (WNS)* enables asynchronous delivery of messages or alerts from SAS and SPS Web services and other elements of service workflows.

The models and interfaces provided by SWE describes a standard framework to deal with sensor data in heterogeneous sensor network applications. SensorML provides a description model for various attributes of sensor data [2]. Its primary representation is defined in XML schema form. Although XML provides a remarkable solution for heterogeneous data representation, there are significant limitations in semantic interoperability and describing the semantics and relationships between different data element using XML representations [3].

2.1 The Semantic Web Technologies

Semantic Web is an extension to the current Web in which the meaningful relationships between resources is represented in machine-processable and machine-interpretable forms [4]. The main idea in the Semantic Web is to provide well defined and machine accessible meanings to resources and to their relationships rather than simple links as they are offered by the simple hyperlink structure on the current Web. Ontologies are utilised by the Semantic Web applications to offer conceptualised representation of domains and to specify meaningful relationships between resources in a domain.

The primary technologies for the Semantic Web include the Resource Description Framework (RDF)[2], RDF Schema[3], and the Web Ontology Language (OWL)[4]. OWL is based on description logic and facilitates construction of ontologies for different domains. The OWL data can be accessed by software agents for reasoning and inferencing purposes and to enable systems to derive additional knowledge from the represented

[1] http://www.opengeospatial.org/
[2] http://www.w3.org/RDF/
[3] http://www.w3.org/TR/rdf-schema/
[4] http://www.w3.org/TR/owl-features/

data. There are also widely used software systems such as Jena [5] and Sesame [6] to deploy and manage the constructed ontologies.

2.2 Sensor Data Modelling

OntoSensor [7] is one of the initial efforts on constructing ontology-based descriptions for sensor data. OntoSensor adapts parts of SensorML descriptions and uses extensions to the IEEE Suggested Upper Merged Ontology (SUMO)[5] to describe sensor information and capabilities. The ontology is represented in OWL format and the authors have discussed the advantages of the proposed approach compared to SensorML and XML based solutions. The main enhancement is providing self-descriptive metadata for the transducer elements. The embedded semantics in the descriptions could be utilised in various sensor discovery and reasoning applications. OntoSensor illustrates a semantic approach to sensor description and provides an extensive knowledge model; however, there is no distinctive data description model to facilitate interoperable data representation for sensors observation and measurement data. A universal sensor observation and measurement data model in collaboration with a sensor specification model will support creating a semantic framework for sensor networks. The semantic sensor network will utilise Semantic Web technologies and reasoning mechanisms to interpret sensor data from physical devices that perform observations and measurements.

Ontology-based description of a service oriented sensor network is discussed in [8]. The SWE and Geography Markup Language (GML)[6] classes and properties in collaboration with SensorML, Suggested Upper Ontology (SUMO) and OntoSensor are employed to develop an ontology for sensor service description. The ontology consists of three main components *ServiceProperty*, *LocationProperty*, and *PhysicalProperty*. ServiceProperty explains the functionality of a service, and properties in the other two components describe contextual and physical characteristics of the sensor nodes in a wireless sensor network architecture. The ontology is represented in OWL form and some initial consistency checking and query results are provided to evaluate the validity of the proposed solution. The system, however, does not specify how complex sensor data will be described and interpreted in a sensor network application. The proposed framework concentrates on building a sensor description ontology for sensor discovery and description of sensor metadata in a heterogeneous environment.

A high level design for a universal ontology which consists of extension plug-in ontologies, sensor data ontology and sensor hierarchy ontology is described in [9]. The extension plug-in ontologies enable the developers to integrate domain specific ontologies into the main ontology. This functionality describes sensor network capabilities and provides relations between domain concepts and sensor functionalities. The sensor hierarchy ontology is a knowledge model for sensors and actuators and other physical devices in the network. It describes the features and capabilities of the elements and contains metadata related to devices such as measurement range, accuracy and calibration. The sensor data ontology describes the dynamic observational data for transducers. The ontology model describes context data with respect to the spatio-temporal attributes. The illustrated model does not specify details of sensor data specification and relationships between various types of complex sensor data. The taxonomy provided for the sensor hierarchy ontology specifies a set of primary numerical attributes for common

[5] http://www.ontologyportal.org/

[6] http://www.opengeospatial.org/standards/gml

types of sensors. In a practical scenario, sensor data will include more complex data types which are not addressed in the proposed model.

Sheth and Hanson [10] discuss a Semantic Sensor Web framework to provide enhanced descriptions for sensor data and to create situation awareness for sensor networks. The semantics data for sensor nodes is using temporal, spatial, and thematic data. Thematic data describes sensor node related information which can be derived by sensor data analysis. Thematic data also includes tags and textual descriptions [11]. A significant aspect of the Semantic Sensor Web architecture is employing a unified data model which provides universal interoperability and semantic description for sensor data. The semantic data also supports construction of content and context-aware sensor network applications.

Henson *et al.* [12] describe a prototype application for the Sensor Web by using annotated video data. The dataset contains Youtube[7] videos annotated with SensorML and XLINK[8] models with reference to a time ontology. Utilising the designated semantic enables the system to retrieve videos by specifying temporal concepts such as "*within*", "*contains*", or "*overlaps*" during a time interval query submission. The authors use keyword tagging and metadata description to provide references to temporal concepts and domain ontologies. An extension to this idea could be proposed as providing a universal metadata structure with a broaden scope to accommodate various sensor data types and domain knowledge.

3. Designing an Ontology for Sensor Data

The SWE common namespace defines several value types and data types for sensor measurement and observation data [2]. The data types fall into the following main categories.

- Primitive data types, which complement the data types defined in GML;
- General purpose aggregate data types such as records, arrays, vectors, and matrices;
- Aggregate data types with special semantics such as curve, and time aggregates;
- Standard encoding to include semantics, quality indications and constraints to primitive and aggregate types;
- Specialised components to support semantic definitions;
- A notation for the description of XML and non-XML array encoding.

The data types are represented in XML encoded form; however it is also possible to use other alternative encodings for the data. The primitive data types describe scalar values such as *Quantity*, *Count*, *Boolean*, *Category*, and *Time*. These data types provide primitives to define the sensor data. Figure 1(a) shows a model of the simple data types in SWE namespace.

A data component describes an object whose values can be defined as a set of simple data types. The simple data types contain properties that describe different attributes required for sensor data. The data types can be grouped together to construct an aggregate object [2]. The generic aggregate components are defined as *RecordTypes* and *ArrayTypes*. There are also derived aggregates such as *DataRecord*, S*impleDataRecord*, *DataArray*, *Vector*, *ConditionalValue*, and *Curve*. Figure 1(b) shows a UML model [13] for

[7]http://www.youtube.com/
[8]http://www.w3.org/TR/xlink/

the generic data aggregation models in SensorML which is based on SWE namespace. SensorML data types and SWE common namespace are described in detail in [2].

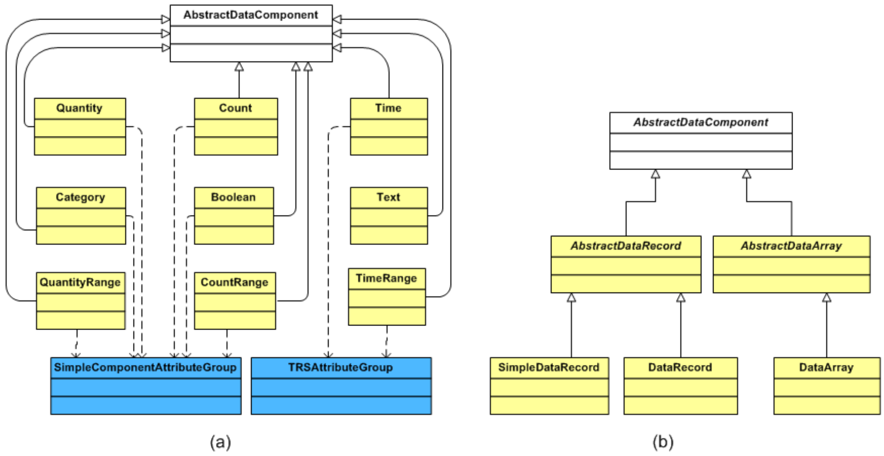

Figure 1: (a) The simple data types (b) The generic data aggregates

Although XML provides a flexible method to represent the data, it does not provide a full potential for the machines to acquire and interpret the emerging semantics from data. Extending the XML descriptions to ontology-based data representations enables advance analysis and enhanced data processing for heterogeneous sensor network applications. The SWE data model is utilised in the current work to construct an ontology for describing sensor observation and measurement data. We have used Protégé[9] an opensource ontology editor and knowledge acquisition system developed at the University of Stanford. The ontology also imports a part of the NASA's SWEET ontology [10] to specify measurement units. Figure 2 shows a fragment of the ontology which is called SensorData Ontology[11].

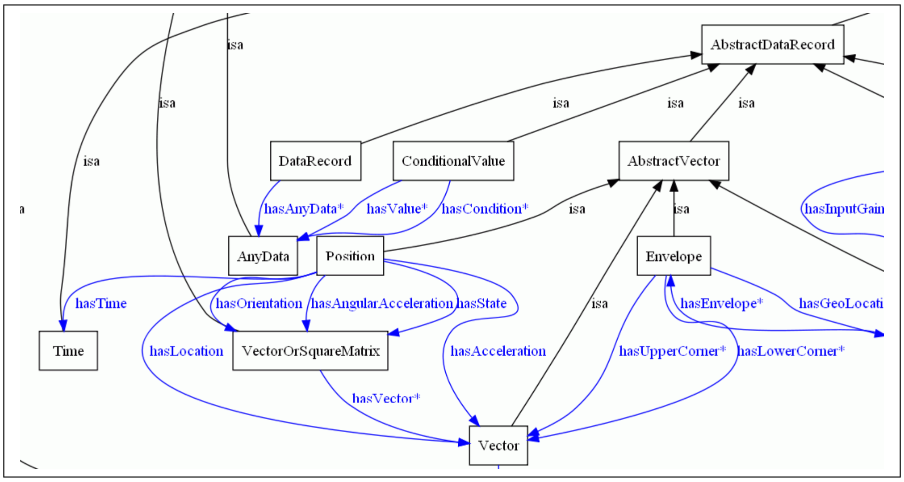

Figure 2: A snippet of the SensorData ontology

Employing ontology-based approach adds complexity to the data representation

[9] http://protege.stanford.edu/

[10] http://sweet.jpl.nasa.gov/ontology/

[11] An OWL version of the ontology is available at: http://purl.oclc.org/net/unis/ontology/sensordata.owl

structure. It requires extra information to describe the sensor data. Considering the fact that sensor nodes have limited process and memory capabilities, the data representation could appear as a bottleneck to the design. To address this issue, we propose a gateway component to wrap the observation data in the designated format. The data analysis and adding ontology-based descriptions to data will only occur in a gateway node which runs on machines with more powerful processing capabilities. The major cost of this method will be adding a extra gateway component which acts as an interface between sensor node and the network. Figure 3 demonstrates a sensor network architecture using gateways to mediate communication between sensor nodes and the network components.

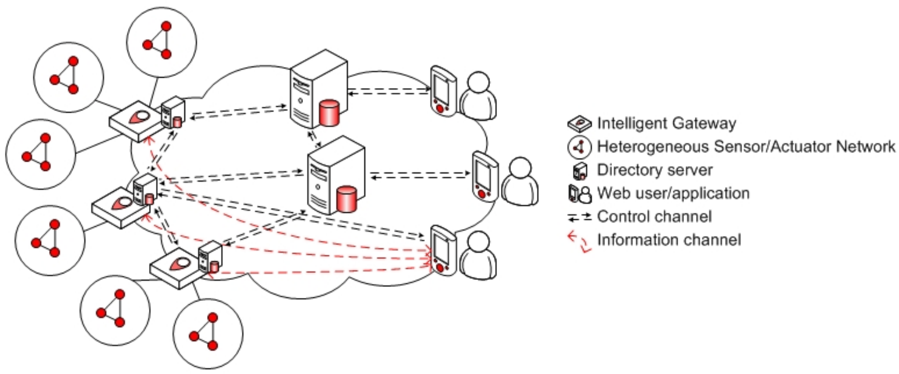

Figure 3: Using gateway nodes in sensor network architecture

4. Outlook and Discussion

To evaluate the proposed approach, we focus on expressibility and scalability of the representations for different types of sensor observation and measurement data. Using a composite data as an example, we demonstrate how the representations differ by employing pure XML serialisation as those suggested by SensorML and the proposed semantic model.

```
<swe:DataRecord definition="urn:ogc:def:property:OGC:atmosphericConditions">
    <swe:field name="AirTemperature">
        <swe:Quantity definition="urn:ogc:def:property:OGC:AirTemperature">
            <swe:uom code="Cel"/>
            <swe:value> 35.1 </swe:value>
        </swe:Quantity>
    </swe:field>
    <swe:field name="WindSpeed">
        <swe:Quantity definition="urn:ogc:def:property:OGC:WindSpeed">
            <swe:uom code="m/s"/>
            <swe:value> 6.5 </swe:value>
        </swe:Quantity>
    </swe:field>
</swe:DataRecord>
```

Figure 4: Sample sensor data in plain XML form

We also illustrate the RDFa[12] annotation of XML data which supports both legacy and the semantic data models [10]. Figure 4 shows a data snippet created in plain

[12]http://www.w3.org/TR/xhtml-rdfa-primer/

XML, and Figure 5 demonstrates the sample data description by employing RDFa annotations. Using RDFa provides compatibility to the legacy data models such as SensorML and at the same time semantic data can be included in the main structure. Figure 6 describes the sample data record according to the proposed ontology-based representation (represented in OWL form).

```xml
<swe:DataRecord definition="urn:ogc:def:property:OGC:atmosphericConditions">
    <swe:field swe-om:Quantity rdf:about="#AirTemperature" name="AirTemperature">
        <swe:Quantity definition="urn:ogc:def:property:OGC:AirTemperature">
            <swe:uom code="Cel" swe-om:hasUomIdentifier rdf:about=
                "http://sweet.jpl.nasa.gov/ontology/units.owl#degreeC"/>
            <swe:value swe-om:hasDoubleValue rdf:datatype="&xsd;double">35.1</swe:value>
        </swe:Quantity>
    </swe:field>
    <swe:field swe-om:Quantity rdf:about="#AirTemperature" name="WindSpeed">
        <swe:Quantity definition="urn:ogc:def:property:OGC:WindSpeed">
            <swe:uom swe-om:hasUomIdentifier rdf:about=
                "http://sweet.jpl.nasa.gov/ontology/units.owl#meter_perSecond"code="m/s"/>
            <swe:value swe-om:hasDoubleValue rdf:datatype="&xsd;double">6.5</swe:value>
        </swe:Quantity>
    </swe:field>
</swe:DataRecord>
```

Figure 5: Sample sensor data in XML + RDFa form

The sample data record is related to two quantities measuring two different phenomena of the physical world. All the concepts are defined externally and are therefore referenced in the descriptions. The overhead in OWL is significant compared to the pure XML specification. Sensor devices are typically constrained by the transmission power and processing capabilities. The machine interpretable representation for sensor data increases the amount of data to be transmitted to sensor network. If metadata is included in sensor node itself before transmission to the network, this will increase power consumption of sensor nodes.

```xml
<swe-om:Quantity rdf:ID="Quantity_AirTemperature">
    <swe-om:hasUomIdentifier rdf:resource=
        "http://sweet.jpl.nasa.gov/ontology/units.owl#degreeC"/>
    <swe-om:hasDoubleValue rdf:datatype="http://www.w3.org/2001/XMLSchema#double"
    >35.1</swe-om:hasDoubleValue>
    <swe-om:hasName xml:lang="en">air temperature</swe-om:hasName>
    <swe-om:hasDefinition rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI"
    >urn:ogc:def:property:OGC:AirTemperature</swe-om:hasDefinition>
</swe-om:Quantity>
<swe-om:Quantity rdf:ID="Quantity_WindSpeed">
    <swe-om:hasDefinition rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI"
    >urn:ogc:def:property:OGC:Windspeed</swe-om:hasDefinition>
    <swe-om:hasName xml:lang="en">wind speed</swe-om:hasName>
    <swe-om:hasUomIdentifier rdf:resource=
        "http://sweet.jpl.nasa.gov/ontology/units.owl#meter_perSecond"/>
    <swe-om:hasDoubleValue rdf:datatype="http://www.w3.org/2001/XMLSchema#double"
    >6.5</swe-om:hasDoubleValue>
</swe-om:Quantity>
<swe-om:DataRecord rdf:ID="DataRecord_AtmosphericConditions">
    <swe-om:hasField rdf:resource="#Quantity_AirTemperature"/>
    <swe-om:hasField rdf:resource="#Quantity_WindSpeed"/>
    <swe-om:hasDefinition rdf:datatype="http://www.w3.org/2001/XMLSchema#anyURI"
    >urn:ogc:def:property:OGC:atmosphericConditions</swe-om:hasDefinition>
</swe-om:DataRecord>
```

Figure 6: Sample sensor data in OWL form

Most of the overhead consists of self explanatory metadata that helps the receiver of the information to interpret the data. In many real world applications, the power saving and reducing the energy consumption is a high priority in designing sensor network applications. Increasing the power consumption means cutting the lifetime of a battery powered sensor node. Such a trade-off between lifetime and including machine interpretable data is a critical issue that needs to be addressed using other components in the sensor network architecture. By employing gateway nodes between sensor node and a sensor network, the gateway receives the binary data from sensor node and then applies the metadata template to the data to construct a semantic description for sensor observation and measurement data. This enables sensor nodes to operate in optimum mode where the gateway components will be responsible for constructing the semantic description of sensor measurement and observation data.

5. Conclusions

The current data exchange for sensor networks mostly relies on binary data models which do not provide machine interpretable meanings to the data. An ontology-based model not only will provide an sensor data description, it also enables machines to process and interpret the emerging semantics to create intelligent sensor network applications. This papers reports initial results of an ongoing research on creating a semantic data description framework and automated resource discovery for global sensor networks. The major challenge of introducing semantic data modelling to sensor networks (which are traditionally designed to be of low complexity) is the addition of metadata that needs to be exchanged alongside the measured data. There are however several deployment and operational mechanisms that can keep this added complexity at bay. Including the semantics can be achieved once the data has left the low complexity part of a sensor network, for instance the gateway or sink node can provide the additional processing; hence only keeping binary formats on the sensor network side. The meta data annotation will be assigned to a designated gateway which receives the raw data and wraps the value with annotations taken from a template (i.e. semantic model). The annotated data can then be transmitted to the information subscribers.

The future work will focus on the evaluation of the impact of adding metadata to the measurement data on the sensor side and using binary XML to keep the sensor network side lightweight. In addition, all other processing to integrate the sensor data into the semantic data model will be outsourced to the sink or gateway. The context modelling will be also another step in developing automated mechanisms for resources discovery, composition, and utilisation in a semantic-enabled sensor network architecture.

6. Acknowledgement

This paper describes work undertaken in the context of the SENSEI project, Integrating the Physical with the Digital World of the Network of the Future (www.sensei-project.eu). SENSEI is a Large Scale Collaborative Project supported by the European 7th Framework Programme, contract number: 215923.

References

[1] Botts, M. and Percivall, G., and Reed, C. and Davidson, J., Sensor Web Enablement: Overview and High Level Architecture, The Open Geospatial Consortium.

[2] Sensor Model Language (SensorML), The Open Geospatial Consortium, `http://www.opengeospatial.org/standards/sensorml/`.

[3] G. Antoniou and F. van Harmelen, *A Semantic Web Primer (Cooperative Information Systems)*. The MIT Press, April 2004.

[4] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific American*, vol. 284, pp. 28–37, May 2001.

[5] Jena: A Semantic Web Framework for Java, `http://jena.sourceforge.net/`.

[6] Sesame: RDF Schema Querying and Storage, `http://www.openrdf.org/`.

[7] D. Russomanno, C. Kothari, and O. Thomas, "Sensor ontologies: from shallow to deep models," *System Theory, 2005. SSST '05. Proceedings of the Thirty-Seventh Southeastern Symposium on*, pp. 107–112, March 2005.

[8] J.-H. Kim, H. Kwon, D.-H. Kim, H.-Y. Kwak, and S.-J. Lee, "Building a service-oriented ontology for wireless sensor networks," *Computer and Information Science, 2008. ICIS 08. Seventh IEEE/ACIS International Conference on*, pp. 649–654, May 2008.

[9] M. Eid, R. Liscano, and A. El Saddik, "A universal ontology for sensor networks data," *Computational Intelligence for Measurement Systems and Applications, 2007. CIMSA 2007. IEEE International Conference on*, pp. 59–62, June 2007.

[10] A. Sheth, C. Henson, and S. Sahoo, "Semantic sensor web," *Internet Computing, IEEE*, vol. 12, pp. 78–83, July-Aug. 2008.

[11] A. Sheth and M. Perry, "Traveling the semantic web through space, time, and theme," *IEEE Internet Computing*, vol. 12, no. 2, pp. 81–86, 2008.

[12] C. Henson, A. Sheth, P. Jain, and T. Rapoch, "video on the semantic sensor web," *W3C Video on the Web Workshop*, 2007. `http://www.w3.org/2007/08/video/papers.html`.

[13] I. Jacobson, G. Booch, and J. Rumbaugh, *The unified software development process*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999.