OVERVIEW ARTICLE

Semantics for the Internet of Things: Early Progress and Back to the Future

Payam Barnaghi, University of Surrey, UK

Wei Wang, University of Surrey, UK

Cory Henson, Wright State University, USA

Kerry Taylor, Commonwealth Scientific and Industrial Research Organisation, ICT Centre, Australia

ABSTRACT

The Internet of Things (IoT) has recently received considerable interest from both academia and industry that are working on technologies to develop the future Internet. It is a joint and complex discipline that requires synergetic efforts from several communities such as telecommunication industry, device manufacturers, semantic Web, and informatics and engineering. Much of the IoT initiative is supported by the capabilities of manufacturing low-cost and energy-efficient hardware for devices with communication capacities, the maturity of wireless sensor network technologies, and the interests in integrating the physical and cyber worlds. However, the heterogeneity of the "Things" makes interoperability among them a challenging problem, which prevents generic solutions from being adopted on a global scale. Furthermore, the volume, velocity and volatility of the IoT data impose significant challenges to existing information systems. Semantic technologies based on machine-interpretable representation formalism have shown promise for describing objects, sharing and integrating information, and inferring new knowledge together with other intelligent processing techniques. However, the dynamic and resource-constrained nature of the IoT requires special design considerations to be taken into account to effectively apply the semantic technologies on the real world data. In this article the authors review some of the recent developments on applying the semantic technologies to IoT.

Keywords: *Cyber Worlds, Future Internet, Information Modeling, Internet of Things, Web Community*

1. INTRODUCTION

Extending the current Internet with interconnected physical objects and devices (or referred to as "*Things*") and their virtual representation has been a growing trend in recent years. This will create a range of potentially new products and services in many different domains, such as smart homes, e-health, automotive, transport and logistics, and environmental monitoring (Kranenburg et al., 2011). The research in this area has recently gained momentum and is supported by the collaborative efforts from

DOI: 10.4018/jswis.2012010101

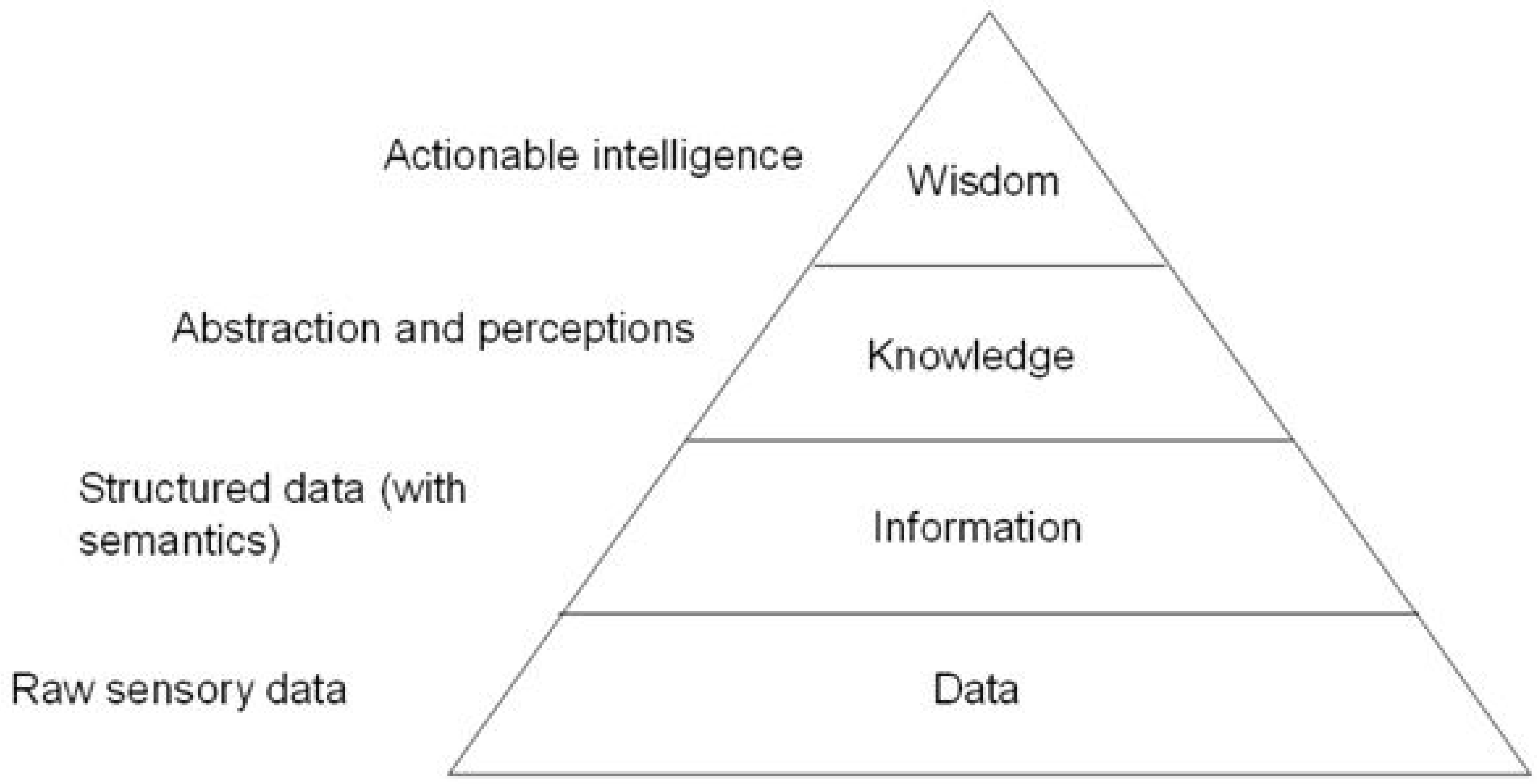
academia, industry, and standardization bodies in several communities such as telecommunication, semantic Web, and informatics. For example, we have seen that new protocols and standards for low-level device communications in resource-constrained environments have been developed (Bormann, Castellani, & Shelby 2012). While for many years legacy systems have been primarily designed for specific purposes with limited flexibility, the current initiative on building the IoT (or more general, the future Internet) demands application and service platforms which can capture, communicate, store, access and share data from the physical world. This will create new opportunities in a long list of domains such as e-health, retail, green energy, manufacturing, smart cities/houses and also personalized end-user applications.

A primary goal of interconnecting devices (e.g., sensors) and collecting/processing data from them is to create situation awareness and enable applications, machines, and human users to better understand their surrounding environments. The understanding of a situation, or context, potentially enables services and applications to make intelligent decisions and to respond to the dynamics of their environments. Data collected by different sensors and devices is usually multi-modal (temperature, light, sound, video, etc.) and diverse in nature (quality of data can vary with different devices through time and it is mostly location and time dependent). The diversity, volatility, and ubiquity make the task of processing, integrating, and interpreting the real world data a challenging task. The volume of data on the Internet and the Web has already been overwhelming and is still growing at stunning pace: everyday around 2.5 quintillion bytes of data is created and it is estimated that 90% of the data today was generated in the past two years (IBM, 2012). Sensory data (including the citizen sensors) (Sheth, 2009a) related to different events and occurrences can be analyzed and turned into actionable knowledge to give us better understanding about our physical world and to create more value-added products and services, for example, readings from meters can be used to better predict and balance power consumption in smart grids; analyzing combination of traffic, pollution, weather and congestion sensory data records can provide better traffic and city management; monitoring and processing sensory devices attached to patients or elderly can provide better remote healthcare. This data transformation process can be better illustrated using the well known "knowledge hierarchy" (Rowley, 2007). We adapt the meanings of the layers to the context of IoT and semantics (Figure 1).

The lower layer refers to large amount of data produced by the IoT resources and devices. The layer helps create structured and machine-readable information from the raw data of various forms to enhance interoperability. However, what *is* required by humans and high-level applications and services often is not the information, but high-level abstractions and perceptions that provide human and machine-understandable meanings and insights of the underlying data. The high-level abstractions and perceptions then can be transformed to actionable intelligence (wisdom) with domain and background knowledge to exploit the full potential of IoT and create end-to-end solutions.

The "big data" solutions and cloud platforms can provide infrastructure and tools for handling, processing and analyzing deluge of the IoT data. However, we still need efficient methods and solutions that can structure, annotate, share and make sense of the IoT data and facilitate transforming it to actionable knowledge and intelligence in different application domains. Since many of the devices and resources in IoT are highly distributed, heterogeneous, and resource-constrained (e.g., battery powered devices, nodes with limited processing and memory capabilities), the requirements for designing services and applications in IoT are different from those currently used on the Internet and the Web (specifically in terms of interoperability, scalability, reliability, autonomy, security and privacy). This is reflected in the recent architecture design and development efforts for the Future Internet and Web (Zorzi et al., 2010).

Figure 1. "Knowledge Hierarchy" in the context of IoT



Issues related to interoperability, automation, and data analytics naturally lead to a semantic-oriented perspective towards IoT (Atzori, Iera, & Morabito, 2010). Applying semantic technologies to IoT promotes interoperability among IoT resources, information models, data providers and consumers (Selvage et al., 2006), and facilitates effective data access and integration, resource discovery, semantic reasoning, and knowledge extraction. In this article, we provide an overview of the recent developments in applying semantic technologies in various aspects of the IoT. We emphasize that the use of semantic technologies should take the dynamicity and constraints of the IoT domain into consideration. We extend the discussion on the semantic Sensor Web (Sheth et al., 2008) and quality of sensor data on sensor Web (e.g., Corcho & Castro, 2010) to IoT and provide an analysis of the major research issues. We describe some of the initial progress and developments that have been made in the past few years in using the semantic technologies in IoT and discuss the future prospects and challenges of developing efficient semantic-enabled IoT systems. The rest of the paper is organized as follows. Section 2 discusses why semantics play such a significant role in the current development of IoT. Section 3 describes the experiences gained from the existing works that apply semantic technologies to IoT. Section 4 reviews the recent developments in this field; in particular, discusses resource and information modeling, linked sensor data, sensor data abstraction and perception, and the supporting tools for IoT data query and processing. In Section 5, we look at the potential research areas where semantic technologies can be further exploited and discuss the associated challenges. Section 6 concludes the paper.

2. WHY SEMANTICS ARE IMPORTANT?

It is estimated that there will be around 25 billion devices connected to the Internet by 2015 and 50 billion by 2020 (Evans, 2011). Such a stunning number of highly distributed and heterogeneous devices will need to be interconnected and communicate in different scenarios autonomously. This implies that providing interoperability among the "*Things*" on the IoT is one of the most fundamental requirements to support object addressing, tracking, and discovery as well as information representation, storage, and exchange. The suite of technologies developed in the Semantic Web (Berners-Lee, Hendler, & Lassila, 2001), such as ontologies, semantic annotation, Linked Data (Berners-Lee, 2006) and semantic Web services (McIlraith, Son, & Zeng, 2001), can be uses as principal solutions for the purpose of realizing the IoT. In what follows, we review different scenarios that demonstrate the importance of semantics to the research and development of IoT.

2.1. Semantics for Interoperability

Semantic interoperability means that different stakeholders can access and interpret the data unambiguously. "*Things*" on the IoT need to exchange data among each other and with other users on the Internet. Providing unambiguous data descriptions in a way that can be processed and interpreted by machines and software agents is a key enabler of automated information communications and interactions in IoT. Semantic annotation of the data (for example, with domain knowledge) can provide machine-interpretable descriptions on what the data represents, where it originates from, how it can be related to its surroundings, who is providing it, and what are the quality, technical, and non-technical attributes.

2.2. IoT Data Integration

IoT data usually originates from a device or a human, and refers to attributes of a phenomenon or an entity in the physical world. The data can be combined with other data to create different abstractions of the environment, or it can be integrated to the data processing chain in an existing application to support context and situation awareness. In all these cases, it is important that heterogeneous data can be seamlessly integrated or one type of data can be combined with other cyber, social, or physical world data (Sheth, 2011). Semantic descriptions can support this integration by enabling interoperability between different sources; however, analysis and mapping between different semantic description models is still required to facilitate the IoT data integration with other existing domain knowledge.

2.3. IoT Data Abstraction and Access

Data abstraction in IoT is concerned with the ways that the physical world data is represented and managed. The current research has mainly focused on representing the observation and measurement data from sensor networks according to the OGC[1] (the Open Geographical Consortium) model. More recently, ontologies such as the W3C's SSN ontology (Barnaghi et al., 2011; Compton et al., in press) have been developed, which provide a number of constructs to formally describe not only the sensor resources but also the sensor observation and measurement data. With the semantic descriptions, the sensor data, or more generally, IoT data, can also be characterized on different abstraction levels. This is accomplished with semantic reasoning offered by semantic query languages (e.g., derived data on accuracy or average (Corcho & Castro, 2010).

Data access in IoT can be implemented at low-levels (e.g., device or network levels) by the use of low-level programming languages and operating systems (Corcho & Castro, 2010). Obviously, heterogeneity of the devices and (sensor) networks makes data access across the networks a difficult task. Service oriented principles, which allow complex software systems to be decomposed into smaller sub-systems or services have been used to integrate the IoT data with enterprise services (Spiess et al., 2009). The idea of "sensing as a service" represents a scalable way to access the sensor data through standard service technologies and has received consensus from the community. For example, a recent work by De et al. (2011) proposes a semantic description model for services exposed by the IoT resources.

2.4. Resource/Service Search and Discovery

In IoT, a resource is referred to as a device or entity that can provide data or perform actuation (e.g., a sensor or an actuator), and a service is a software entity that exposes the functionality of its corresponding resource (De et al., 2011). The search and discovery mechanisms allow locating resources or services that provide data related to an entity of interest in the physical world. Search and discovery are among the most important functionalities that are required in

IoT. Semantic annotation of the IoT resources and services, and processing and analyzing the semantically annotated data are essential elements to support the search and discovery methods for resources, services, and real physical world entities with different attributes and functionalities. With the dynamicity of IoT and the resource-constrained nature of the many IoT resources, energy efficiency considerations for discovery (e.g., sending requests to the resource itself only when it is needed) or compensation mechanisms (e.g., when a resources becomes unavailable because of running out of power or network loss) are often needed. An interesting work in this regard involves the selection and ranking of service instances according to contextual information (Guinard et al., 2010). The idea of the linked sensor data (see Section 4.3 for details) that enables publishing and use of sensor data using the linked data principles (Berners-Lee, 2006) can also be applied to support discovery and search of resources and services. In this case, resources are linked to each other as well as to other types of virtual and/or real world objects through the semantic links.

2.5. Semantic Reasoning and Interpretation

The knowledge representation formalism used in the suite of semantic Web technologies allows logical reasoning that is able to infer new information or knowledge from existing assertions and rules. Semantic reasoning is an important instrument in the domain of IoT for various purposes such as resource discovery, data abstraction, and knowledge extraction. The actual inference algorithms are usually implemented within available reasoners (e.g., FACT++[1] and Jena[2]) so IoT developers do not need to be concerned with the complexities of the reasoning process itself. The SPARQL query language can be also used to construct queries to explore the semantic descriptions. Some examples using the SPARQL language to discover IoT resource in the linked data are presented in De et al. (2012) and Pschorr et al. (2010).

3. SEMANTICS ALONE ARE NOT ENOUGH

It is important to note that providing semantic descriptions alone does not provide semantic interoperability and will not solve all the issues regarding discovery, management of data, and supporting autonomous interactions. The semantic description still needs to be shared, processed, and interpreted by various methods and services across different domains. The following highlights some of the practical issues that need to be considered in applying semantic technologies to the IoT domain.

3.1. Ontologies do not make Data Interoperable at a Global Scale

Defining an ontology and using semantic descriptions for data will make it interoperable for users and stakeholders that share and use the same ontology. In the IoT domain different stakeholders need to have a common agreement on ontological definitions. Most of the current ontologies and semantic description frameworks in the IoT domain are defined in the context of different projects and applications or they are currently at an early stage. To achieve global scale semantic interoperability, common semantic annotation frameworks, ontology definitions, and adaptation are key issues. Recent efforts, such as the W3C SSN ontology, are effective steps towards achieving this goal. For the current and existing applications, it is also important that their ontologies and knowledge base can be accessed and reused by large groups of potential consumers. Developing and sharing ontologies and contributing towards description and annotation frameworks that can support legacy applications are effective steps in achieving semantic interoperability on a large scale. Other solutions, such as ontology mapping and matching (i.e., manual, semi-automated, or automated) can help link the resources described using different semantic annotation models. The ontology designers can also reference existing common ontologies and provide links to other upper-level ontologies to support interoperability between different semantic descriptions in the IoT domain.

3.2. Semantic Annotations Need To Be Processed and Analyzed

Using semantic annotations in the IoT domain provides machine-readable and machine-interpretable metadata to describe the IoT resources and data. However, a key issue that needs to be considered is that machine-interpretable data is still not necessarily machine-understandable data. The semantic Web technologies include well-defined standards and description frameworks (e.g., RDF, OWL, SPARQL) and a variety of open-source and commercial tools for creating, managing, querying, and accessing semantic data. However, this still does not eliminate the key role of information analytics and intelligent methods, which can process and interpret the data and create meaningful abstractions. The semantic annotations can support more effective mechanisms to be designed to utilise and integrate the IoT data, but autonomous and seamless integration of the data still requires effective reasoning and processing mechanisms (to be further discussed in Sections 4 and 5). The ontologies and semantic models need to be simple and light weight to make them suitable for the resource constrained environments. Accessing to the IoT data and semantic descriptions, and management of the resources can be also supported using service oriented solutions.

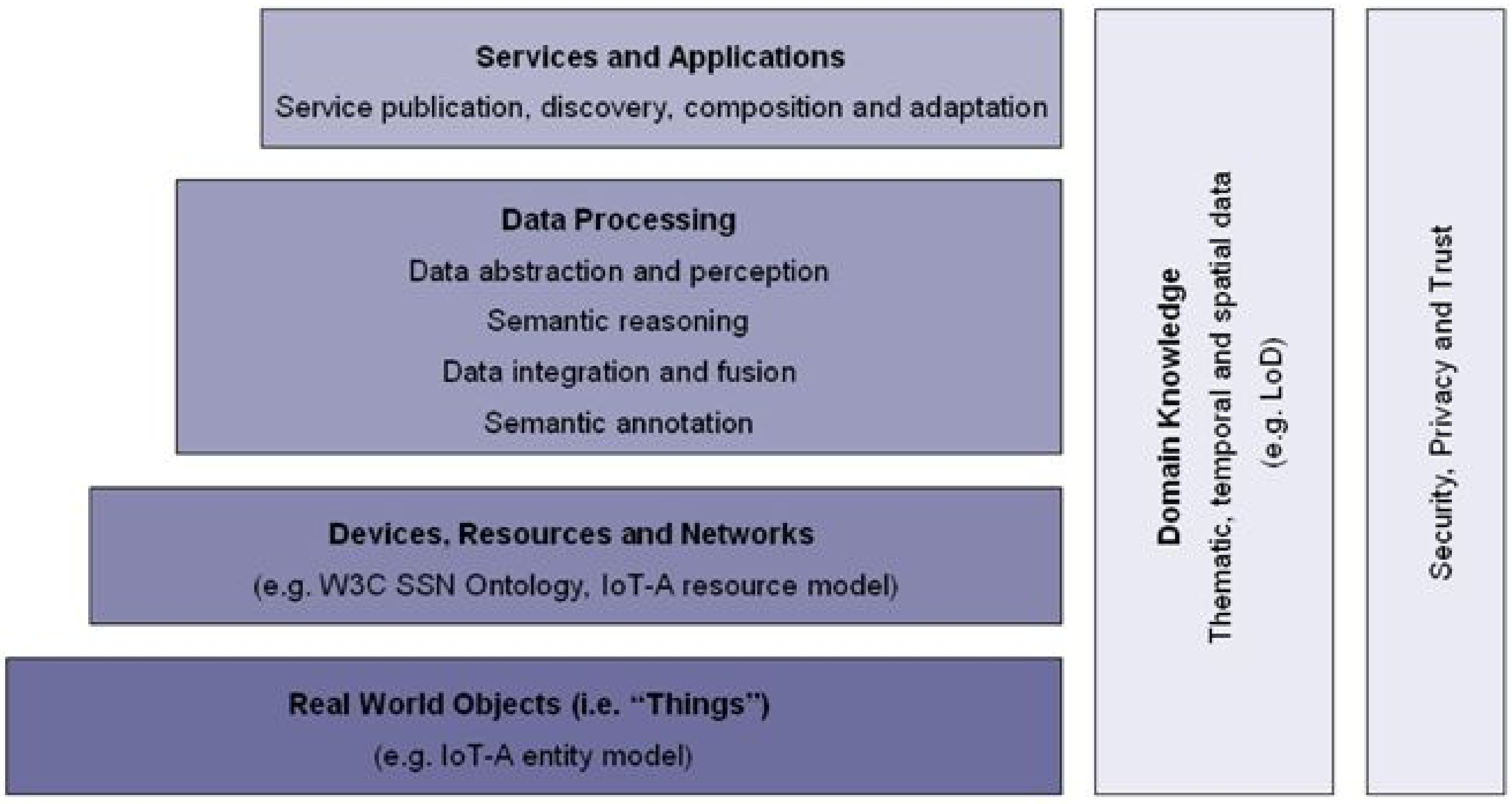
3.3. Semantic Technologies are not Just Hype!

Semantic technologies have matured over the years, and there are a number of existing tools and solutions to publish, annotate, query, search, and discover the semantic data. Semantic technologies have also been applied to service oriented technologies to provide interoperable interface, process, and service descriptions. Ontologies and semantic description frameworks provide an effective way to share and agree on a common vocabulary and knowledge model for describing the data, which can be machine-interpretable and represented in interoperable and re-usable forms. However, the IoT resources can be constrained devices that operate in dynamic environments. Therefore, it can be argued that introducing semantic annotations and metadata hinders effective utilization of resources and that they are not suitable for use in networks and devices with limited memory, process and energy resources (Preuveneers, 2008). Fortunately, some complex Internet and Web technologies have already been customized and applied to the resource constrained environments. 6LowPAN (Shelby & Bormann, 2009) and CoAp (Bormann, Castellani, & Shelby 2012) are examples of recent technologies that have been developed to address the limitations of applying Internet and Web –based solutions to the IoT domain. 6LowPAN provides IP-based solutions and CoAp provides a transfer protocol for constrained environments. Similarly, lightweight semantic models can also be introduced for the IoT domain. Compression mechanisms, similar to those used in 6LowPAN and CoAp, can be used to create and communicate small size semantic descriptions. Another key aspect is that the semantic annotations can be added to the data at different stages (e.g., when the data arrives at a node with more powerful resources, such as a gateway). Figure 2 shows a view on how semantics can be used at different levels in IoT. For example, in Ganz et al. (2011) a resource annotation and sensor device description based on W3C SSN ontology is provided when the nodes are connected to a gateway. Designing lightweight semantic description models (Guinard et al., 2010) and effective representation frameworks such as Binary RDF Representation (Fernández et al., 2011) are some of the recent works that can provide effective semantic data representations for the IoT domain.

4. RELATED DEVELOPMENTS

To help solve problems of interoperability among IoT systems, caused by the heterogeneous and distributed nature of the "Things," the IoT community has begun to adopt semantic

Figure 2. Semantics at different levels in IoT



Web technologies. Towards this goal, a number of modelling approaches and ontologies used to annotate and describe the IoT data have been developed. Semantic descriptions and annotations are used to represent devices, real-world objects and events, and services and business process models. These semantic descriptions support the automated management and interaction of the different components of the IoT systems. In the following, we review some of the recent developments that use semantic technologies in the IoT domain.

4.1. Semantic Modeling and Ontology Development

Ontologies in IoT have been developed for a number of uses, including the description of sensor and sensor networks, IoT resources and services, smart "Things," etc. In this section we review some of the most important ontologies in the IoT domain and give a brief overview of the recent activities on the ontology developments in this field.

An early work on defining common interfaces and descriptions for IoT related data is provided by the Sensor Web Enablement (SWE) group at OGC. The main specifications defined by OGC are: Observations & Measurements (O&M), which defines a standard model and XML Schema for encoding real-time and archived observations and measurements of sensor data; Sensor Model Language (SensorML), which is a standard model to describe sensor systems and processes associated with sensor observations in an XML-based schema; Sensor Observations Service (SOS), which is a standard Web service interface for requesting, filtering, and retrieving observations and sensor system information; Sensor Planning Service (SPS), which is a standard Web service interface and acts as an intermediary between a client and a sensor collection management environment; PUCK Protocol, which defines how to retrieve a SensorML description and other information and can enable automatic installation, configuration and operation of sensor devices; SWE Common Data Model, which is used in nodes to exchange sensor related data; SWE service model, which defines data types used across SWE services. The PubSub Standards Working Group[3] is implementing the SWE standards to enable publish/subscribe functionality for OGC Web Services and define the methods to realise the core publish/subscribe functionality for a specific service binding (e.g., using SOAP, RESTful).

The models and interfaces provided by OGC define a standard framework for dealing with sensor data in heterogeneous environments. The primary representation models in SWE are encoded in XML, which has significant limitations in semantic interoperability and defining associations between different elements.

The W3C Semantic Sensor Networks Incubator Group has developed an ontology for describing sensors and sensor network resources, called the SSN ontology (Barnaghi et al., 2011; Compton et al., in press). The ontology provides a high-level schema to describe sensor devices, their operation and management, observation and measurement data, and process related attributes of sensors. It has received consensus of the community and has been adopted in several projects[4]. To model the observation and measurement data produced by the sensors, the SSN ontology can be used along with other ontologies such as the Quantity Kinds and Units ontology[5] and the SWEET ontology[6]. The SSN has also been used with domain ontologies to develop various smart Things ontologies, such as the smart product ontology (Nikolov et al., 2011).

However, the IoT domain is not only limited to sensors and sensor networks. The physical world objects (i.e., "*Things*"), also referred to as "*Entities of Interest,*" their features of interest, spatial and temporal attributes, resources that provide the data and their related service are other important features that need to be modelled. Autonomous integration of the IoT data and resources to the business process requires machine process-able descriptions of execution requirements. In De et al. (2011) a set of models for IoT entities, resources and services is described. An entity represents a 'Thing' in IoT and is the main focus of interactions by humans and/or software agents. This interaction is made possible through a hardware component, a 'device,' which allows the entity to be part of the digital world by mediating the interactions. The actual software component that provides information on the entity or enables controlling of the device is called a 'resource.' Finally, a 'service' has standardised interfaces and exposes the functionality of a device by accessing its hosted resources (De et al., 2011). Modeling of business processes by using semantically annotated resources that take dynamicity of the IoT environments into account is described in Meyer et al. (2011).

In general, to achieve autonomous and seamless integration of the IoT data in business applications and services, semantic description of different resources in the IoT domain is a key task. The aforementioned works are some examples of the recent efforts that have been made to address this issue. The semantic descriptions and annotations need to be provided at "Things" level, device and network level (e.g., W3C SSN ontology), service level (e.g., SemSOS) (Henson et al., 2009), and interaction and business process level (e.g., the IoT-aware business process modeling) (Guinard et al., 2010) to enable autonomous processing and interpretation of the data by different providers and users in the IoT domain.

4.2. Linked Sensor Data

Semantic annotations can describe IoT resources, services and related processes. However, often there is no direct association to the domain knowledge in the core models that describe the IoT data. Different resources, including observation and measurement data, also need to be associated with each other to add meaning to the IoT data. Effective reasoning and processing mechanisms for the IoT data, and making it interoperable through different domains, requires accessing domain knowledge and relating semantically enriched descriptions to other entities and/or existing data (on the Web). Linked Data is an approach to relate different resources and is currently adopted on the Web. The four principles, or best practices, of publishing data as linked data include (Berners-Lee, 2006):

1. Using URI's as names for things; everything is addressed using unique URI's.
2. Using HTTP URI's to enable people to look up those names; all the URI's are accessible via HTTP interfaces.

3. Providing useful RDF information related to URI's that are looked up by machine or people;
4. Linking the URI's to other URI's.

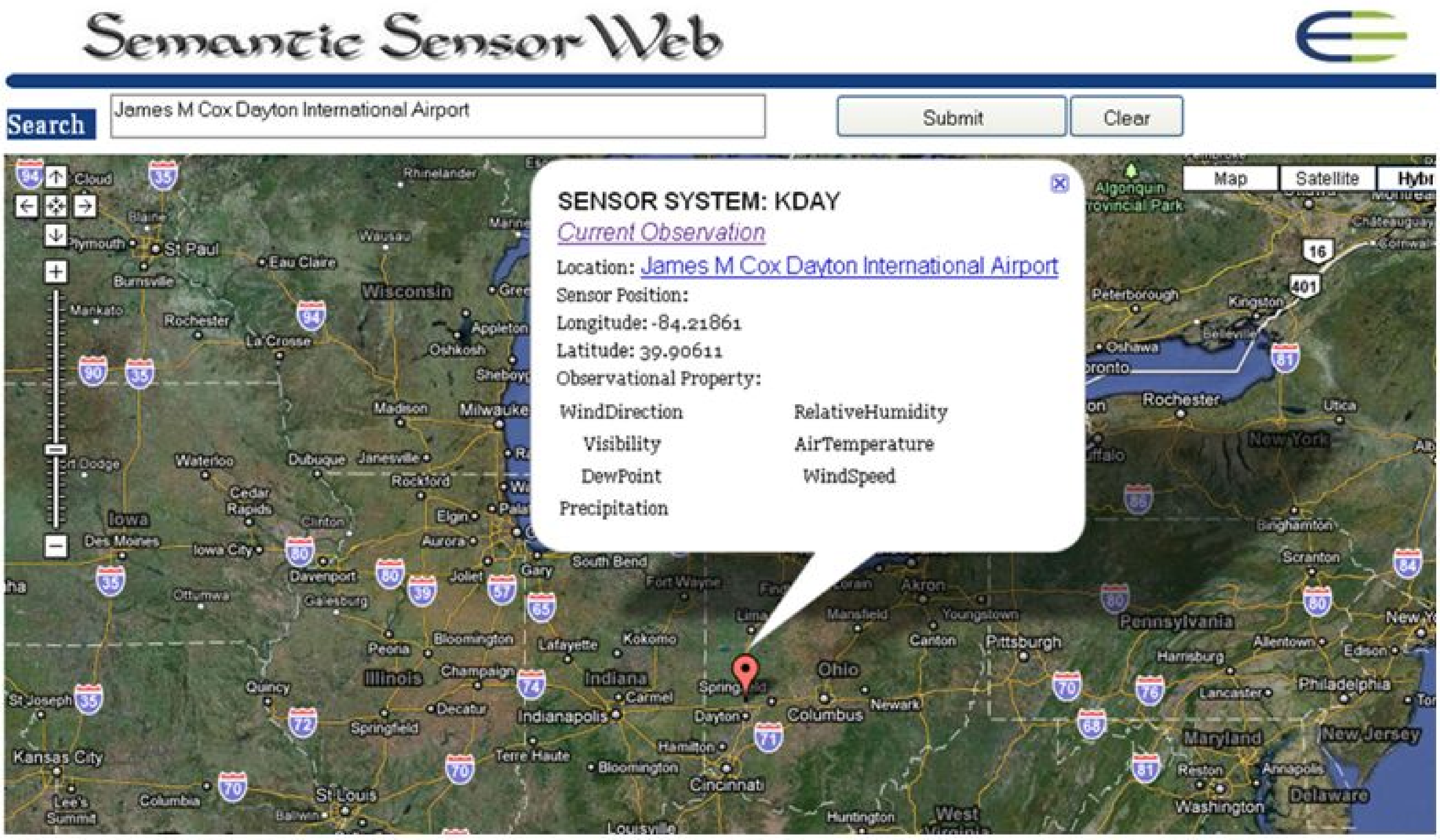
The current linked open data (LOD)[7] effort on the Web provides a large of number of inter-linked data represented in RDF accessible via common standard interfaces (Bizer et al., 2009). The linked data approach is also applied to the IoT domain by providing semantic data and linking it to other domain dependent resources such as location information and semantic tags; e.g., the work described in Patni, Henson, and Sheth (2010a) and Page et al. (2009). The linked data approach enables resources described via different models and ontologies to be interconnected. Linking the data to existing domain knowledge and resources also makes the descriptions more interoperable. Providing automated mechanisms for semantic tagging of the resources using the concepts available as linked data (e.g., such as those available on the LOD cloud[8]), and defining automated association mechanisms between different resources (e.g., based on location, theme, provider and other common properties) make the IoT data usable across different domains. The following are some sample use cases that use the linked data approach to describe the IoT data (e.g., sensor data).

Kno.e.sis linked sensor data: Linked Sensor Data is an approach to representing and publishing sensor descriptions and sensor observations on the Web using the Linked Data best practices. Publishing sensor data as Linked Data enables discovery, access, query, and interpretation of sensor data. Patni et al. (2010a) have developed an RDF dataset[9] containing expressive descriptions of ~20,000 weather stations in the United States and over 160 million sensor observations. In total, this results in over 1.7 billion RDF triples. The data originated at MesoWest[10], a project within the Department of Meteorology at the University of Utah, which has been aggregating weather data since 2002. On average, there are about five sensors per weather station measuring phenomena such as temperature, visibility, precipitation, pressure, wind speed, humidity, etc. In addition to location attributes such as latitude, longitude, and elevation, there are also links to locations in GeoNames[11] that are near each weather station. This dataset has been integrated with a semantically enabled Sensor Observation Service (SemSOS) (Henson et al., 2009) and has been used to enable sensor discovery queries based on named locations (e.g., *find sensors near Dayton International Airport*) rather than longitude and latitude coordinates (Pschorr, 2010). Figure 3 shows a screenshot of a demonstration application that is created using this data. The application allows browsing and accessing the individual data by selecting locations on a map or by searching for location concepts in GeoNames that are used to annotate the data (Patni et al., 2010b).

Sense2Web linked sensor data platform: Sense2Web provides graphical user interfaces to annotate the IoT data (i.e., resource description, real world entities and services) using concepts obtained from linked open data cloud (e.g., DBPedia[12] and GeoNames) and also other local domain ontologies. The annotated data is published as RDF triples and is available via a common SPARQL-end point (Barnaghi et al., 2010). Sense2Web has also implemented RESTful interfaces that enable direct publication, access and query of linked IoT data (De et al., 2012). This platform provides two different approaches to linked data; one using publically available linked data resources as domain knowledge to annotate the resources and second publishing the annotated data as linked data resources. Figure 4 shows a screenshot of the resource annotation and publication interface in Sense2Web.

Linked Sensor Data and RESTful serving of RDF and GML: Page et al. (2009) present an API to expose data from the Channel

Figure 3. Browsing Kno.e.sis linked sensor data

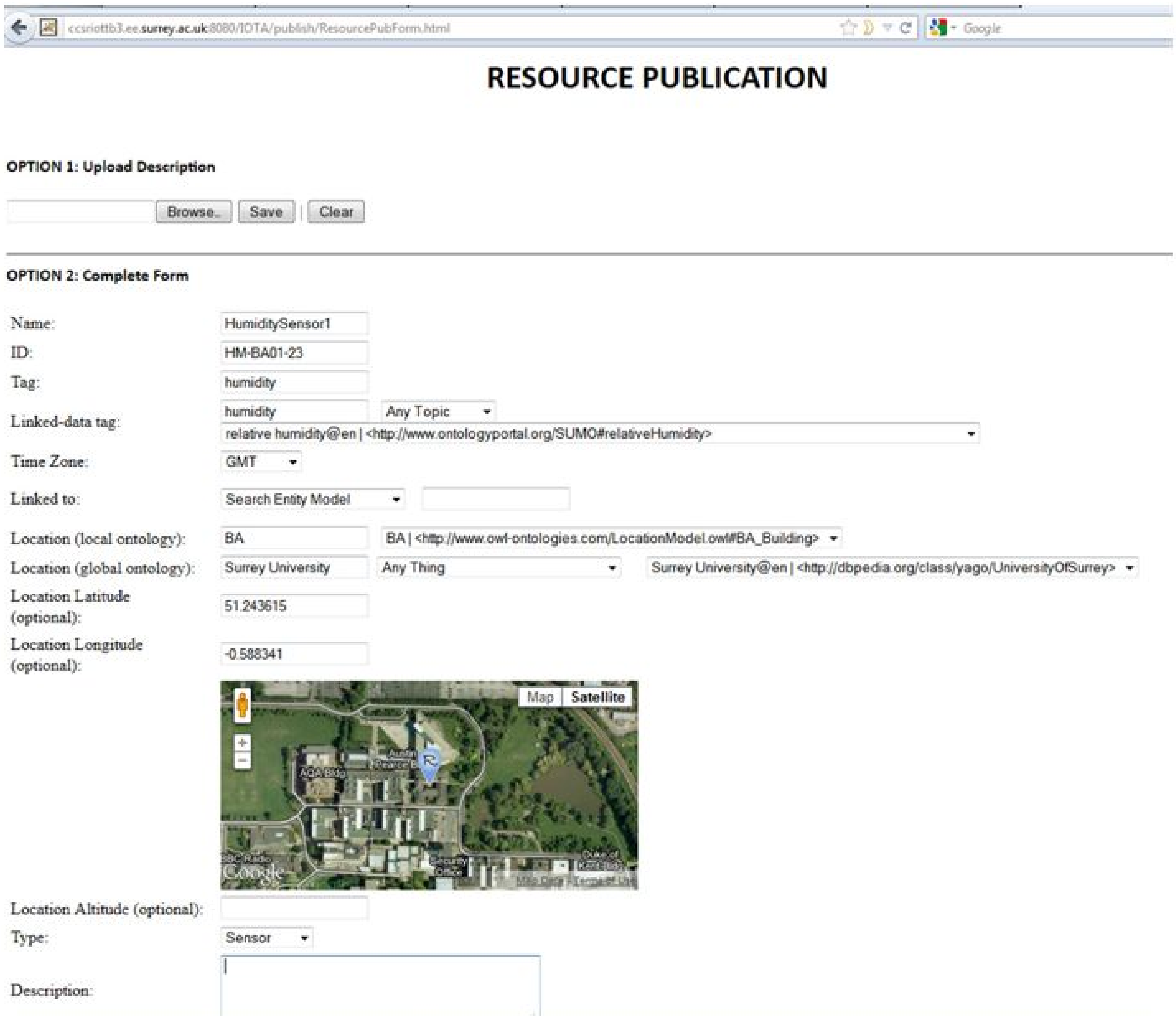


Coastal Observatory in the UK, using linked data principles. The presented API uses REST and linked data principles that allow supporting both web clients and the OGC GML[13] clients. The presented platform uses URIs and provides semantic annotations in the form of linked data to represent observation and measurement data. This enables supporting both legacy GML applications that refer to XML descriptions and semantic web clients that use enhanced semantic annotations to interpret and utilise the data.

SensorMasher: SensorMasher uses linked data principals to makes sensor data available on the Web (Le Phuoc, 2009). SensorMasher publishes sensor data as Web resources and enables users to describe the sensor data using semantic annotations. These semantic annotations are then used for discovery and automation support to construct mashups using data from different resources. Sensor data published in this platform can be accessed through SPARQL endpoints and RESTful services. Users can access the data in JSON, XML, and RDF formats. By exploring links between the data resources, users and mashup tools can traverse the sensor data. In addition, SensorMaher filters and identifies relations between different data sources, which enhances the process of integration of data and applications. Figure 5 shows a screenshot of the SensorMasher platform.

The described systems are a few samples of how the linked data principles can enhance access, querying, filtering, and integration of the IoT data. The linked open data can be also used as an abundant source of knowledge for annotating the IoT data. This not only promotes reuse of the existing knowledge, but also creates potential to design novel IoT resource and service discovery methods. Analyzing the links and semantic descriptions can also support integration of different data and construction of high-level abstractions from the data (for example, events or perceptions).

Figure 4. Resource annotation using linked open data concepts

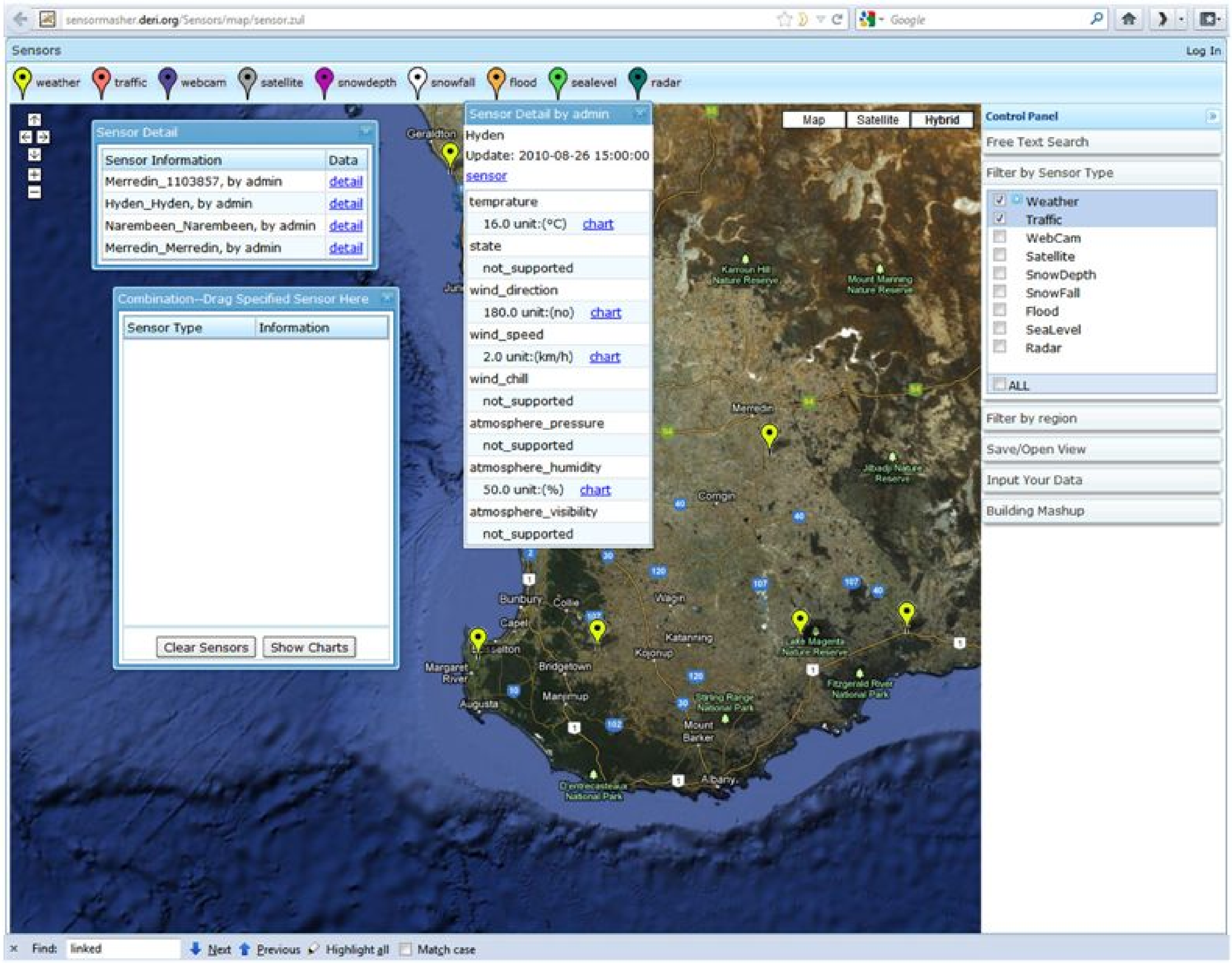


4.3. Data Abstraction and Knowledge Extraction

Processing and analyzing semantic descriptions for extracting knowledge and enabling enhanced interactions with the IoT resources depends on effective querying, analysis, and processing of the semantic data and links between the resources. The current query mechanisms for the Semantic Web are mainly based on SPARQL. The IoT data is often represented as streams and is distributed over different networks with diverse types of data. As the data is real-time and the attributes of data (i.e., quality attributes) can change over time, the query mechanisms for IoT need to address this dynamicity and agility. Querying and processing the semantic descriptions in large scale is also another important issue. There are already mature solutions to work with large-scale semantic descriptions (e.g., Oren et al., 2009; Hogan et al., 2010); however in the changing environment in IoT requires more efficient query and processing techniques. For example, in IoT, resources can appear and disappear over time, the data can be collected from different heterogeneous resources, and real time processing of data streams is required for event detection. In the past, the knowledge and data engineering efforts in IoT have mainly focused on developing infrastructures for the IoT data, such as publication, query and access. As a result, less attention is given to intelligent data processing that can exploit and process the IoT data, integrate it to the existing business processes and/or creates situation-awareness.

The observation and measurement is the low-level data that is captured by sensors, other devices or human users. This could be large volume of data related to an entity of

Figure 5. The SensorMasher platform



interest or an environment. However, the IoT data consumers (i.e., users and applications) are often interested in the high-level concepts that refer to machine-interpretable or human-understandable knowledge. A sample application of creating such high-level abstractions is discussed in Henson et al. (2012) where sensory observation data is used in a logical inference model to derive perceptions from the raw observations. The data abstraction and knowledge extraction processes to enable transforming low-level IoT data to high-level knowledge that refer to an event, a pattern, are comprehensible to the machines and human users, and play an important role in leveraging the full potential of IoT. The high-level abstractions, in relation to domain knowledge in different applications, can create a source of perception which will be the driving asset for developing intelligent applications and smart environments that use the IoT data.

4.4. Sensor Perception

The act of observation performed by heterogeneous sensors creates an avalanche of data that must be integrated and interpreted in order to provide knowledge of the situation. This process is commonly referred to as perception, and while people have evolved sophisticated mechanisms to efficiently perceive their environment – such as the use of a-priori background knowledge of the environment – machines continue to struggle with the task. More specifically, perception is the process of deriving abstractions from a set of sensor observations. Given some background knowledge – i.e., as a set of relations between entities (or "things") and their observable qualities – and a set of observations, the perception process identifies a set of entities that explain the set of observations (Henson et al., 2011).

The primary challenge of machine perception is to define efficient computational methods to derive high-level knowledge from low-level sensor observation data. Emerging solutions are using ontologies, such as the W3C SSN ontology, to provide expressive representation of concepts in the domain of sensing and perception, which enable advanced integration and interpretation of heterogeneous sensor data. For a model of perception to be useful for real-world situations, it should meet the following requirements (Henson et al., 2012):

Perception is an abductive process – An entity represented as an abstraction is not necessarily implied by the set of observations, but rather is a hypothetical explanation of the observations. Thus, perception is not a deductive process (in the first-order logic sense of the term), but rather an abductive process, meaning an *inference to the best explanation*.

Graceful degradation with incomplete information – Even with an incomplete set of observations, the perception process should still identify a set of explanatory entities. This property is referred to as graceful degradation with incomplete information, and it's often necessary since observing all possible qualities is usually impractical.

Abstractions should be generated efficiently – Sensors are constantly streaming observation data in real-time. Therefore, to be practically useful, the generation of abstractions should also be computed in *near-real-time*. In addition, in many applications, the perception process must compute abstractions of sensor observations within resource-constrained environments such as mobile devices or gateway nodes.

Integration with Semantic Web languages – The perception process must generate abstractions from observations and background knowledge encoded in Web languages. As discussed, much sensor data is now being annotated with a sensor ontology (i.e., SSN ontology), encoded in standard Web formats (i.e., RDF), and is increasingly being made available on the Web (i.e., as Linked Data).

4.5. Tools for IoT Resource Annotation and Data Query

Tools for the sensor and sensor data annotation and publication according to common ontology models are not only useful because of the functionalities they provide but also the roles they play to promote the reuse and wide adoption of the common models. However, to our knowledge, currently there are not many publicly available tools for these purposes (in the IoT domain). The annotation tool in Sense2Web allows users to annotate sensor data (i.e., resource, entity and service descriptions) according to the models presented in De et al. (2011). It also supports the linking of the sensor resources to concepts in linked open data cloud (Barnaghi et al., 2010). Designing more annotation and publication tools that can support widely accepted ontologies (e.g., W3C SSN) and making them publicly available are important for the IoT community to promote interoperability and encourage use of common annotation and description frameworks.

A number of tools that aim to address some of the specific requirements of the IoT data have also been developed, mostly to support the data queries and through extending the standard functionalities in the existing query languages such as SPARQL. The stSPARQL and stRDF extend the SPARQL query language and RDF representations with spatial and temporal dimensions to facilitate query on sensor data which is mostly time and location dependent (Kyzirakos, Koubarakis, & KaoudI, 2009). Continuous SPARQL (C-SPARQL) and streaming SPARQL are other extensions of the SPARQL query language to support continuous queries over streaming data (Barbieri et al., 2009; Bolles, Grawunder, & Jacobi, 2008). EP-SPARQL (Event Processing SPARQL) is an extension to SPARQL that enables processing complex events and stream reasoning

(Anicic et al., 2011). It is designed for timely detection of compound events within streams of simple events based on semantic reasoning with background knowledge.

Most of the tools developed in the semantic Web research can be used in IoT for resource and data query, browsing (for linked sensor data), reasoning, etc. However, much of the IoT data has its own characteristics and needs to be processed in specific ways. For example, IoT produces huge amounts of streaming data which requires continuous and timely processing methods that are able to handle large data throughput and at the same time to perform semantic reasoning. This can also able allow the IoT systems to continue to update their background knowledge by processing and interpreting the new observations related to continuously changing event which is referred to as continuous semantics in Sheth et al. (2010).

5. RESEARCH CHALLENGES

IoT describes a splendiferous future: a dynamic and universal network where billions of identifiable "things" (e.g., devices, people, applications, services, etc.) communicating with one another anytime and anywhere; things become context-aware, are able to configure themselves and exchange information, and show "intelligent/cognitive" behaviour when exposed to a new environment and unforeseen circumstances; intelligent decision-making algorithms will enable appropriate rapid responses, revolutionizing the ways business values are generated (Sundmaeker et al., 2010).

Back to the reality, the current research and developments are still too far from that vision. Diversity, heterogeneity and spatiotemporal dependency of IoT data and resources make physically interconnected things disconnected at semantic levels. Common frameworks are essential to describe and represent the data and to make it seamlessly accessible and processable across different domains. Still, we have not seen scalable methods which can derive actionable and reliable knowledge and create perceptions from the large amount of data generated by the physical devices and human sensors (Sheth, 2009b), especially when quality of data depends on many factors (e.g., sensing devices, environmental variables and data sources). Another distinctive characteristic of IoT compared to other research areas is the high dynamicity. IoT requires efficient mechanisms and methods that can handle large amount of data and respond to the identified phenomenon and events arising from the environment in a timely fashion. Furthermore, security and privacy issues, trust and reliability of the data are also important for IoT based applications and services, especially those in the business domain. In what follows, we present a detailed analysis on the major research challenges and opportunities related to applying semantic technologies into the IoT domain.

5.1. Dynamicity and Complexity

Real world data is more transient, subject to environment changes and it is mostly time and location dependent. While semantic technologies and semantic annotation help describe the meanings behind data and enable description of different attributes of the resources and networks that provide data, the pervasiveness and volatility of the underlying environments require continuous updates and monitoring of the descriptions. Although this dynamicity does not apply to all the real world resources, in many cases when the status of the resource (e.g., quality of measurement, energy profile, and network or power outage) changes the semantic descriptions need to be updated accordingly. Addressing this dynamicity and providing up-to-date descriptions that reflect the current state of the resources (and their data) become a challenging issue when scalability, diversity and network/resource constraints are taken into consideration. Another issue that hinders maintaining up-to-date semantic description of the IoT resources is mobility and ubiquity of the resources which imply continuous updates in real-time streaming data processing scenarios. The issues of dynamicity and complexity have a significant impact on many aspects of the IoT such as data and resource access services, semantic description publication and maintenance, data analysis, aggregation and mining. Further research on resource compensation and adaptation methods, semantic event processing and analysis, continuous semantic data processing mechanisms is needed to address these two issues.

5.2. Scalability

Creating semantic annotation frameworks and domain knowledge models for describing a large number of entities, devices and their related data is critical for knowledge and data engineering in IoT. The IoT data refers to different phenomena in the real world; so the semantic description and annotation of data need to be associated with domain knowledge of real world resources and entities. In some applications, initiatives such as Linked Open Data can be used as domain knowledge to describe thematic and spatial aspects of the IoT data; however, the community-driven knowledge sources such as Linked Open Data are prone to errors and inconsistency (due to lack of quality control). Many applications develop and maintain their own domain knowledge, but reuse and interoperability is an issue. Granularity of the descriptions (e.g., in describing the location data) is another important issue; the more precise terms and concepts used in describing the semantics, the more extensive will be the domain knowledge. Maintaining large-scale and distributed semantic data is never an easy task. In recent years there has been a number of works by the semantic Web community on introducing efficient approaches to store, process, reason and query large scale semantic data in distributed environments. However, what makes semantic data handling in IoT more challenging and fraught with technical difficulties is the scale of data generated by its corresponding resources, continuous changes in the state (and consequently description) of the resources and data and volatility of the IoT environments. The research in this area needs to address issues such as automated (or semi-automated) annotation of the resources, semantic association discovery and analysis (when resources appear or are deployed), efficient solutions to create linked IoT data and to explore and analyse the links between different resources. Creating tools and APIs for annotating the resources and observation and measurement data, constructing the semantic repositories and implementing light-weight services that allow accessing and querying the sensory data and resource descriptions are also essential in creating a scalable IoT.

5.3. Semantic Service Computing for IoT

The number of resources and the amount of data produced by the resources in IoT introduce scalability issues to all aspects of IoT. While semantic technologies are ideal for promoting interoperability, given that common ontology models are shared and widely reused, the adoption of service oriented computing enables increased scalability of IoT. The concept of IoT services that are able to expose capabilities of their corresponding resources defines the paradigm of service-oriented computing in IoT. This type of services is also referred to as "real-world services on physical devices" (Guinard et al., 2010), and there are a number of existing semantic service description models for this purpose (e.g., De et al., 2011; Henson et al., 2009; Bröring et al., 2009). The IoT services often operate in dynamic environments, and in some cases, the resources underlying these IoT services are mobile, unreliable, and capability-constrained. All these factors make the IoT services different from most existing legacy services on the Web.

The IoT services can be combined with other applications and services to compose complex, context-aware business services. In a service composition process involving the IoT services, adaptation and compensation are important design considerations to ensure continuous service access and reliable response to consumers' requirements. Automated service composition in a resource-constrained environment such as IoT is more challenging than in domains where reliable services are abundant. The research in IoT service computing needs to address automated and dynamic composition of services and adaptation/compensation mechanisms that can re-configure delivery and provisioning of services when context changes. Another key issue is creating lightweight service description and implementation solutions that encourage the use of the semantic Web services in resource constrained environments. As we have seen in the past, the use of the semantic Web services has not gained popularity as it was expected in the early days when different models and frameworks for annotating Web service with semantic data (i.e., models such as OWL-S[14] or WSMO[15] have not really been used) were introduced. The complexity involved in describing the services using the common semantic Web service frameworks has hindered wide adaptation of the semantic Web services. We argue that the concept of introducing the IoT data and resource capabilities as a service will soon change this paradigm. The need for introducing enhanced description frameworks for the IoT services (with diverse attributes, capabilities and qualities) and associated mechanisms that enable publishing, discovery, testing and provisioning of this type of services will become important issues for the research community (for example, some of the recent solutions such as SA-REST[16] or WADL[17] can be adapted for developing IoT service models).

5.4. Distributed Data Storage Query

With large volumes of data and semantic descriptions, efficiency of storage and data handling mechanisms become a key challenge; especially considering the scale and dynamicity involved. The streaming sensory data can be stored (together with their semantic descriptions and possibly domain knowledge) either temporarily or for longer spans of time. Designing and implementing repositories than enable publishing and accessing the semantic data in large distributed and dynamic environments, and providing efficient indexing and discovery mechanism are important issues in IoT. More efficient mechanism on information search and retrieval, indexing, query, and information access will be required to address the issues such as: data discovery and information analytics using semantic data distributed across many repositories; supporting real-time query and aggregation over multiple data streams; finding relevant data among many resources and providers; and subscribing to events and data that can be provided by different resources. Cloud computing is clearly a promising technical approach to address some of these challenges. However, the solutions for handling, maintaining, and processing the data still need to be scalable and efficient; simply putting a centralised and non-scalable solution in the cloud will not make it scalable or very efficient.

5.5. Quality, Trust, and Reliability Of Data

The IoT data is provided by different sensory devices or citizen sensors (Sheth, 2011). This data is prone to errors and quality changes. Different semantic description models such as the W3C SSN ontology offer a means to describe quality related aspects of data. However, the quality of observations and measurements can change over time, for example, changes in the environment, faults in devices, or errors in device settings. Inaccuracy and varying qualities in the IoT data are unavoidable. Detecting and filtering anomalies and false readings from the devices, along with reliable semantic descriptions of quality related attributes of the IoT data, can help detecting errors, and help retrieve and process the data according to different quality requirements. When data is provided by different resources, trust is another key issue. Trustworthiness of resources, identification of the source providing the data, and an understanding of accuracy and reliability of the data, can be supported by semantics describing quality and trust related attributes for the resources and providers. While semantics can play an important role for defining trust and reliability attributes, trust model development and its feedback and verification mechanisms are major issues that need to be addressed.

5.6. Security and Privacy

IoT data is often personal. It can describe our environment, the status of our homes and cities, or our personal health and activities. The mechanisms to provide and guarantee the security and privacy of data are curial issues in IoT. Semantics can help specify verification measures and requirements and provide machine-interpretable description of desired security and privacy requirements while sharing and communicating the IoT data (from the data publisher point of view). Provenance of data and analysis methods that can effectively utilise the provenance data are also important. As data is communicated over the Internet and can be shared with different parties and users, it is also important to define appropriate access control (authentication and authorization) mechanisms, e.g., who can use the data, what part of the data they can use, when and where they can use the data. Further development of IoT will also be highly dependent on developing reliable and efficient solutions that can support and maintain security and privacy requirements in the IoT domain. The research community needs to consider developing efficient privacy and security solutions that can be applied and used in resource-constrained environments with various types of devices and communication networks as a part of the IoT design.

5.7. Interpretation and Perception of Data

Creating high-level abstractions through machine perception from the IoT data is a key enabler for developing situation-aware applications that can intelligently respond to the changes in the real world. Perception is a primary basis of human intelligence and experience. Providing interpretation and analytics methods for machines to process and elucidate changes and events in the physical world will enable machines to perceive their surrounding environment. Semantic descriptions and background knowledge provided in machine-readable and interpretable formats, in cooperation with intelligent information analytics and data processing techniques, will support transforming enormous amount of raw observations created by machine and human sensors into higher-level abstractions that are meaningful for human or automated decision making processes. However, machine perception in IoT adds additional challenges to the problems that conventional AI methods have been trying to solve in the past. Examples of such additional challenges include: integration and fusion of data from different sources, describing the objects and events, data aggregation and fusion rules, defining thresholds, real-time processing of the data streams in large scale, and quality and dynamicity issues. The research is this field needs to develop solutions that can efficiently query and access the data from various sources considering their constraint environments, analyze the data and identify patterns and anomalies, associate the identified patterns with existing knowledge to create higher-level abstractions or new knowledge.

6. CONCLUSION

Adding semantics to different levels of IoT ensures that data originating from different sources is unambiguously accessible and process-able across different domains and users. Observation or measurement data collected from the real world can be semantically described to facilitate automated processing and integration in relation to domain knowledge and other existing resources in the cyber world; resources and components in the IoT framework (e.g., sensors, actuators, platform and network resources) can be described using semantic annotations to enable effective discovery and management of them; at a higher-level, the IoT services and their interfaces can also be semantically described to enable service discovery and composition as well as scalable access to the IoT data. Different knowledge engineering and machine learning techniques have also been used to process the IoT data and associated semantics to extract new knowledge and create perception from the physical world observations and measurements.

Initial work such as the W3C SSN ontology has shown success in describing common attributes of the IoT related resources (i.e., in this case sensor devices) by accommodating requirements from different stakeholders. However, the complexity of annotating and describing the resources and their data using detailed ontologies hinders the widespread adoption of comprehensive semantic models in IoT. IoT and using semantics in IoT are still in their early days. The IoT community requires coordinated efforts to define more vocabularies and description frameworks to represent resources, data and services in the IoT domain. Looking at the future prospect of using semantics in the IoT domain, lightweight and easy-to-use ontologies seem to have a better chance of being widely adopted and reused in order to create an interoperable platform across different domains and applications. Furthermore, providing automated or semi-automated methods and tools to annotate, publish and access the semantic descriptions also play essential roles in using semantic technologies to enhance processing and management of the data in IoT. Most of the existing semantic tools and techniques have been created mainly for Web resources and have not taken into consideration the dynamicity of the physical environments and the constraints of the IoT resources. Future work in this area should embrace dynamicity, volatility and scalability, and provide solutions that are easily adaptable to the resource constrained and distributed environments.

The linked data principles have been applied to the IoT domain to support creation of more interoperable and machine process-able data and resource descriptions (e.g., for sensors and sensor networks). Including domain knowledge and linking IoT resources to external data (e.g., the linked open data cloud or existing knowledge base) that describe different thematic, spatial and temporal concepts is also another key aspect in supporting effective interpretation and utilisation of the IoT data. The same principles should be applied in a broader range (not only for sensors) to create a truly interconnected network of Things.

In this paper, we have identified several research challenges in applying semantic technologies to IoT and outlined the challenges and future research. Most of these challenges and research issues are closely related to the dynamicity and pervasiveness of the IoT domain. While there are many other areas in IoT to which semantics can contribute, and the research community will continue exploring the novel use of semantic technologies in IoT, the multi-disciplinary nature of the IoT domain requires synergetic efforts from other fields such as service computing, data mining and social science to enhance the processing and utilisation of semantic data in the IoT domain.

ACKNOWLEDGMENTS

The authors would like to thank Professor Amit Sheth for his valuable comments and suggestions. Payam Barnaghi and Wei Wang would like to acknowledge support from the European Union Seventh Framework Programme IoT.est project; IoT.est: Internet of Things Environment for Service Creation and Testing (http://ict-iotest.eu/iotest/), contract number: 257521. Cory Henson would like to acknowledge that this research was supported in part by US National Science Foundation award no. 1143717 (III: EAGER — Expressive Scalable Querying over Integrated Linked Open Data).

REFERENCES

Anicic, D., Fador, P., Rudolph, S., & Stojanovic, N. (2011). EP-SPARQL: A unified language for event processing and stream reasoning. In *Proceedings of the 20th International Conference on World Wide Web* (pp. 635-644).

Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of things: A survey. *Computer Networks*, *54*(15), 2787–2805. doi:10.1016/j.comnet.2010.05.010

Barbieri, D. F., Braga, D., Ceri, S., Della Valle, E., & Grossniklaus, M. (2009). C-SPARQL: SPARQL for continuous querying. In *Proceedings of the 18th International World Wide Web Conference* (pp. 1061-1062).

Barnaghi, P., Compton, M., Corcho, O., Garcia Castro, R., Graybeal, J., & Herzog, A. …Page, K. (2011). *Semantic sensor network XG final report*. Retrieved from http://www.w3.org/2005/Incubator/ssn/XGR-ssn-20110628/

Barnaghi, P., Meissner, S., Presser, M., & Moessner, K. (2009). Sense and sens'ability: Semantic data modelling for sensor networks. In *Proceedings of the ICT Mobile Summit*.

Barnaghi, P., Mirko Presser, M., & Moessner, K. (2010, November). Publishing linked sensor data. In *Proceedings of the 3rd International Workshop on Semantic Sensor Networks (SSN), Organised in conjunction with the International Semantic Web Conference*.

Berners-Lee, T. (2006). *Linked data*. Retrieved from http://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*, 35–43. doi:10.1038/scientificamerican0501-34

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data - The story so far. *International Journal on Semantic Web and Information Systems*, *5*(3), 1–22. doi:10.4018/jswis.2009081901

Bolles, A., Grawunder, M., & Jacobi, J. (2008). Streaming SPARQL—Extending SPARQL to process data streams. In S. Bechhofer, M. Hauswirth, J. Hoffmann, & M. Koubarakis (Eds.), *Proceedings of the 5th European Semantic Web Conference: Research and Applications* (LNCS 5021, pp. 448-462).

Bormann, C., Castellani, A., & Shelby, Z. (2012). CoAP: An application protocol for billions of tiny internet nodes. *IEEE Internet Computing*, *16*(2), 62–67. doi:10.1109/MIC.2012.29

Bröring, A., Janowicz, K., Stasch, C., & Kuhn, W. (2009). Semantic challenges for sensor plug and play. In *Proceedings of the 9th International Symposium on Web and Wireless Geographical Information Systems* (pp. 72-86).

Compton, M., Barnaghi, P., Bermudez, L., García-Castro, R., Corcho, O., & Cox, S. (in press). The SSN ontology of the W3C semantic sensor network incubator group. *Journal of Web Semantics*.

Compton, M., Henson, C., Lefort, L., & Neuhaus, H. (2009). A survey of the semantic specification of sensors. In *Proceedings of the 8th International Semantic Web Conference in conjunction with the 2nd International Workshop on Semantic Sensor Networks*.

Corcho, O., & Castro, R. (2010). Five challenges for the semantic sensor Web. *Semantic Web Journal*, *1*(1-2), 121–125.

De, S., Barnaghi, P., Bauer, M., & Meissner, S. (2011). Service modeling for the Internet of Things. In *Proceedings of the 3rd Workshop on Software Services: Semantic-based Software Services* (pp. 949-955).

De, S., Elsaleh, T., Barnaghi, P., & Meissner, S. (2012). An Internet of Things platform for real-world and digital objects. *Journal of Scalable Computing: Practice and Experience*, *13*(1).

Evans, D. (2011). *The Internet of Things: How the next evolution of the internet is changing everything*. Retrieved from http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf

Fernández, J., Martínez-Prieto, M. A., Gutierrez, C., & Polleres, A. (2011). *Binary RDF representation for publication and exchange (HDT)*. Retrieved from http://www.w3.org/Submission/2011/SUBM-HDT-20110330/

Ganz, F., Barnaghi, P., Carrez, F., & Moessner, K. (2011). Context-aware management for sensor networks. In *Proceedings of the Fifth International Conference on Communication System Software and Middleware* (article 6).

Guinard, D., Trifa, V., Karnouskos, S., Spiess, P., & Savio, D. (2010). Interacting with the SOA-based Internet of Things: Discovery, query, selection, and on-demand provisioning of web services. *IEEE Transactions on Services Computing*, *3*(3), 223–235. doi:10.1109/TSC.2010.3

Henson, C., Pschorr, J. K., Sheth, A. P., & Thirunarayan, K. (2009, May 18-22). SemSOS: Semantic sensor observation service. In *Proceedings of the International Symposium on Collaborative Technologies and Systems*, Baltimore, MD (pp. 44-53).

Henson, C., Sheth, A., & Thirunarayan, K. (2012). Semantic perception: Converting sensory observations to abstractions. *IEEE Internet Computing*, *16*(2), 26–34. doi:10.1109/MIC.2012.20

Henson, C., Thirunarayan, K., & Sheth, A. (2011). An ontological approach to focusing attention and enhancing machine perception on the Web. *Applied Ontology*, *6*(4), 345–376.

Hogan, A., Pan, J. Z., Polleres, A., & Decker, S. (2010). SAOR: Template rule optimisations for distributed reasoning over 1 billion linked data triples. In *Proceedings of the 9th International Semantic Web Conference* (pp. 337-353).

IBM. (2012). *Bringing big data to the enterprise*. Retrieved July 10, 2012, from http://www-01.ibm.com/software/data/bigdata/

Kranenburg, R., Anzelmo, E., Bassi, A., Caprio, D., Dodson, S., & Ratto, M. (2011). The Internet of Things. In *Proceedings of the 1st Berlin Symposium on Internet and Society: Exploring the Digital Future*.

Kyzirakos, K., Koubarakis, M., & Kaoud, I. Z. (2009). *Data models and languages for registries in SemsorGrid4Env* (Deliverable D3.1 Version 1.0). Retrieved from http://www.semsorgrid4env.eu/files/deliverables/wp3/D3.1.pdf

McIlraith, A., Son, T. C., & Zeng, H. (2001). Semantic web services. *IEEE Intelligent Systems*, *16*, 46–53. doi:10.1109/5254.920599

Meyer, S., Sperner, K., Magerkurth, C., & Pasquier, J. (2011). Towards modeling real-world aware business processes. In *Proceedings of the Second International Workshop on Web of Things* (article 8).

Nikolov, A., d'Aquin, M., Giordanino, M., & Vildjiounaite, E. (2011). *Conceptual framework in SmartProducts, D.2.1.3: Final version of the conceptual framework*. Retrieved from http://projects.kmi.open.ac.uk/smartproducts/smartproducts-d-2-1-3-final.pdf

Oren, E., Kotoulas, S., Anadiotis, G., Siebes, R., ten Teije, A., & van Harmelen, F. (2009). Marvin: Distributed reasoning over large-scale Semantic Web data. *Journal of Web Semantics*, *7*(4), 305–316. doi:10.1016/j.websem.2009.09.002

Page, K., De Roure, D. C., Martinez, K., Sadler, J. D., & Kit, O. Y. (2009). Linked sensor data: RESTfully serving RDF and GML. In *Proceedings of the Second International Workshop on Semantic Sensor Networks*

Patni, H., Henson, C., & Sheth, A. (2010a). Linked sensor data. In *Proceedings of the International Symposium on Collaborative Technologies and Systems*.

Patni, H., Henson, C., & Sheth, A. (2010b, May 30-June 3). Provenance aware linked sensor data. In *Proceedings of the 2nd Workshop on Trust and Privacy on the Social and Semantic Web, co-located with ESWC*, Heraklion Greece.

Phuoc, D. L. (2009). SensorMasher: Publishing and building mashup of sensor data. In *Proceedings of the 5th International Conference on Semantic Systems*.

Preuveneers, D., & Berbers, Y. (2008). Encoding semantic awareness in resource-constrained devices. *IEEE Intelligent Systems*, *23*(2), 26–33. doi:10.1109/MIS.2008.25

Pschorr, J., Henson, C., Patni, H., & Sheth, A. (2010). *Sensor discovery on linked data*. Dayton, OH: Kno.e.sis Center, Wright University.

Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of Information Science*, *33*(2), 163–180. doi:10.1177/0165551506070706

Selvage, M., Wolfson, D., Zurek, B., & Kahan, E. (2006). *Achieve semantic interoperability in a SOA, Patterns and best practices*. Retrieved from http://www.ibm.com/developerworks/webservices/library/ws-soa-seminterop/index.html

Shelby, Z., & Bormann, C. (2009). *6LoWPAN: The wireless embedded Internet*. New York, NY: John Wiley & Sons.

Sheth, A. (2009a). *Semantic integration of citizen sensor data and multilevel sensing: A comprehensive path towards event monitoring and situational awareness*. Paper presented at From E-Gov to Connected Governance: The Role of Cloud Computing, Web 2.0 and Web 3.0 Semantic Technologies, Falls Church, VA.

Sheth, A. (2009b). Citizen sensing, social signals, and enriching human experience. *IEEE Internet Computing*, *13*(4), 87–92. doi:10.1109/MIC.2009.77

Sheth, A. (2011). Computing for human experience: Semantics empowered cyber-physical, social and ubiquitous computing beyond the Web. In *Proceedings of the OnTheMove Federated Conferences and Workshops*, Crete, Greece.

Sheth, A., Henson, C., & Sahoo, S. (2008). Semantic sensor Web. *IEEE Internet Computing*, *12*(4), 78–83. doi:10.1109/MIC.2008.87

Sheth, A., Thomas, C., & Mehra, P. (2010). Continuous semantics to analyze real-time data. *IEEE Internet Computing*, *14*(6), 84–89. doi:10.1109/MIC.2010.137

Spiess, P., Karnouskos, S., Guinard, D., Savio, D., Baecker, O., Sa de Souza, L. M., & Trifa, V. (2009). SOA-based integration of the internet of things in enterprise services. In *Proceedings of the IEEE International Conference on Web Services* (pp. 968-975).

Sundmaeker, H., Guillemin, P., Friess, P., & Woelfflé, S. (Eds.). (2010). *Vision and challenges for realizing the Internet of Things*. Brussels, Belgium: European Commission.

Zorzi, M., Gluhak, A., Lange, S., & Bassi, A. (2010). From today's intranet of things to a future internet of things: A wireless- and mobility-related view. *Wireless Communications*, *17*(6), 44–51. doi:10.1109/MWC.2010.5675777