Ontology Summit 2015 Communique - Internet of Things: Toward Smart Networked Systems and Societies

Lead Editors: Mark Underwood and Michael Gruninger
Co-Editors: Ken Baclawski, Mike Bennett, Gary Berg-Cross, Torsten Hahmann, Leo Obrst, Ram Sriram

Introduction

We are witnessing another phase in the evolution in computing and communication. The Internet, which spans networks in a wide variety of domains, is having a significant impact on every aspect of our lives. The next generation of networks will extend beyond physically linked computers to include multimodal information from biological, cognitive, semantic, social, and sensor networks. This paradigm shift will involve symbiotic networks of people, intelligent devices, and mobile personal computing and communication devices (mPCDs), which will form net-centric societies or smart networked systems and societies (SNSS). mPCDs are already equipped with a myriad of sensors, with regular updates of additional sensing capabilities. Additionally, we are witnessing the emergence of “intelligent devices,” such as smart meters, smart cars, etc., with considerable sensing and networking capabilities. Hence, these devices – and the network -- will be constantly sensing, monitoring, and interpreting the environment – this is sometimes referred to as the Internet of Things (IoT). And as local and wide area networks became almost secondary to the WWW (World-Wide Web), users and their usage patterns will become increasingly visible. This will have significant implications for both the market for advanced computing and communication infrastructure and the future markets – for nearly 4.5 billion people -- that net-centric societies will create.

Smart networked systems and societies will result in better quality of life, reduced threat from external sources, and improved commerce. For example, assume a scenario where people at various locations suffer from flu-like symptoms. In a net-centric society, mPCDs will send vital signs and other associated information to appropriate laboratories and medical centers. These centers will analyze the information, including searching the Internet for potential solutions, and will aid in determining possible causes for this phenomenon. Based on the diagnosis, people will be directed to the nearest clinic for treatment. Here we have several types of information flowing through the net: data from mPCDs; location information; images; video; and audio.

The development of a trusted, secure, reliable, and interoperable net-centric computing environment will need technologies that can assure a flexible and scalable system allowing the application of diverse and robust privacy requirements, thus enabling the trusted and meaningful growth of net-centric infrastructures for the benefit of all societies. One such technical challenge is that the network consists of things (both devices and humans) which are heterogeneous, yet need to have seamless interoperability. Devices need to interoperate and data needs to to compatible to be integrated. This requires the development of standard terminologies which capture the meaning and relations of objects and events. Creating and testing such terminologies will aid in effective recognition and reaction in a network-centric situation awareness environment. The primary goal of this summit to discuss the role of ontologies in the development of smart networked systems and societies.

Several key issues were addressed within the Ontology Summit, especially:

1. Making the case for IoT ontologies
2. How ontologies are used in IoT
3. The challenge of scalability
4. Ontology-based standards for IoT

The Case for IoT Ontologies

Ontologies play a significant role in the realization of SNSS. For example, a considerable amount of data passes through the network and should be converted into higher abstractions that can be used in appropriate reasoning. This requires the development of standard terminologies which capture objects and events. Moreover, such terminologies must align with the intended semantics of generic and domain-specific concepts. Creating and testing such terminologies will aid in effective recognition and reaction in a network-centric situation awareness environment. This involves identifying a methodology for development of terminologies for multimodal data (or ontologies), developing appropriate ontologies, both foundational (such as time, situation, events) and domain specific, developing testing methods for these ontologies, demonstrating interoperability for selected domains (e.g., healthcare, situational awareness), and using these ontologies in decision making.

Sensors are most closely in touch with the outside world and are thus are a big part of IoT since they provide an observational basis for data about things of interest. Since sensors are a big embedded part of the sensing and processing infrastructure of IoT, this results in many Big Data challenges related to semantic heterogeneity. Data can be hard to use because it is in different formats, uses inconsistent naming conventions, and is often provided at a low level of abstraction that makes it difficult to integrate it with other knowledge bases and software systems. To address these challenges, the Semantic Sensor Network Ontology (SSNO) was developed by W3C SSN-XG (2011) to help process and understand sensor information, and to allow the discovery, understanding, and querying of sensor data. SSNO is an ontology for describing networked sensors and its output by introducing a minimal set of classes and relations centered around the notions of *stimuli*, *sensor*, and *observations*. It includes different operational, device related and quality of information attributes that are related to sensing devices, and it describes the operational range, battery and power and environmental ranges that are specified for sensor devices.

Upper Ontologies such as DOLCE can also play a role in extending other IoT ontologies. There are broader Device Ontologies which can leverage some of the Physics Domain Ontology available in DOLCE with its well organized, concept-based vocabulary. DOLCE also has a pattern for situation ontologies.

Of course, sensors are only one small part of the picture. Ontologies for time, duration, and dates are needed in order to capture the distinction between snapshots of measurements and the dynamic behaviour of an embedded system. Ontologies for location are required for scenarios in which the smart objects on the network are widely distributed geographically.

Events are a key concept that play a critical role in many IoT applications. In some scenarios, events create context by connecting people, things, places, and time; approaches such as the Simple Event Ontology (SEM) can be used to annotate events in these contexts and support retrieval of information. However, there are many scenarios in which there is a need to compose events into larger activities and to link events together to recognize patterns of behaviour.

Finally, IoT systems are not all passive -- in many scenarios, smart objects are enabled to make decisions and *act* autonomously in particular contexts. Many existing event ontologies need to be extended to represent this notion of agency.

How Ontologies are Used in IoT

There are several IoT applications that have utilized ontologies to various degrees. These applications include manufacturing, healthcare, and disaster management. Scenarios that include complex event processing require ontologies that have extensive axiomatizations in expressive logics such as first-order logic. In particular, manufacturing processes have complex causal and temporal structures, and complex event processing requires reasoning over situations and events. Typical ontology use scenarios in ontology mapping and decision support are described below.

Ontology Mapping

The wide array of sensors within an IoT application and the variety of data that they provide leads inexorably to the problem of integrating the ontologies that are associated with these sensors. A typical application requires the interconnection of algorithms and hardware for multiple existing networks (such as a medical network and a transportation network that provides traffic data). One approach is to select an existing ontology to bridge such networks, or to combine existing ontologies in various domains and use these ontologies to integrate systems [e.g., Quantities, Units, Dimensions; Semantic Sensor Networks; Foundation Model of Anatomy; Symptom Ontology; Human Disease Ontology] . Other approaches explicitly address the problem of mapping between ontologies. The simplest approaches manually map JSON entities to target ontologies. In the Hyper/CAT approach (see http://www.hypercat.io/standard.html), servers provide catalogues – an array of URIs -- of resources, annotated with metadata -- to clients. In the most sophisticated approaches we find Inference-based Mapping, in which the mappings between ontologies can be achieved using an inference engine (or AI theorem provers).

In many IoT applications, there are two fundamentally different approaches to interoperability. In the first approach, we find centralized processing of spatially distributed and heterogeneous sensor data (Semantics in the Cloud). Data is collected in different settings by various kinds of sensors/things/persons, and all sensor observations are sent to the cloud for semantic annotation and processing. The challenge is to describe the various sources correctly to allow semantic integration. In the second approach, there is local processing (Semantics at the Edge), in which local intelligent sensor networks perform in-place computing. The challenge here is in using ontologies to smartly aggregate, filter, process, access, and respond to sensor data.

Decision Support for IoT

Many IoT applications, ranging from complex event processing and situation awareness to manufacturing, use automated inference from ontologies to assist in the decision making and to

implement smart objects that can automatically act and react to changing situations. The critical issues in the deployment of IoT focus on three questions:

1. What kinds of axiomatizations are required for IoT ontologies?
2. How are the axioms of an ontology used in IoT applications?
3. How can ontology-based solutions scale up to realistic IoT scenarios?

A commonplace maxim invoked by many Semantic Web practitioners is “A little semantics goes a long way.” The critical issue is to identify, for a given IoT application, exactly what ontological approach is adequate. If ontologies are being used to annotate IoT data, then lightweight taxonomies can have a major impact by enabling the interpretation of data by other software applications. Nevertheless, SPARQL and RDF models are not adequate for all tasks; while SPARQL is great for querying a knowledge base, it is less ideal for fetching objects, and it is cumbersome when working with dynamic data. Applications based on complex event processing require more expressive axiomatizations of events, states, and causality.

Scalability

The number, volume and variety of sensor data, whether delivered in real time as data streams or processed as stored batches, results in Big Data challenges (e.g. heterogeneity challenging integration, interpolation and summarization, filtering, compression). Many Big Data issues are common to sensor networks, such as the explosion of standards and reliance on metadata vocabularies such as the idea of things within IoT like services, users, networks, concentrators/aggregators and devices called “resources.” In the face of these challenges we can ask whether light-weight sensor ontologies scale, and what are the realistic ontological commitments for big heterogeneous data.

One aspect that distinguishes IoT scenarios from other applications of ontologies is the role of physical constraints. A sensing/actuating task that requires the co-operation and coordination of thousands of devices (within an Internet of billions), might be impractical due to memory, processing, and energy constraints. The interplay between these constraints and the semantic content of the ontology remains to a large extent unexplored.

The challenge of scalability also arises in the design of ontologies. With the size and increasing complexity of IoT, extensible and modular approaches are useful, if not essential. Approaches for developing small, focused ontologies customized to the available sensors and sensor data might be necessary, but it is an open research question as to whether the combination and integration of a large number of such ontologies is feasible.

Scalability is influenced by the different application case studies that drive the need for more semantics in sensor networks, and these approaches can be contrasted in the following table:

Sensor data discovery and integration	In-network data stream processing
``Offline": happens after the fact	``On-line": happens when and where the data is collected
Somewhat centralized: only need to integrate data from different data collection servers	Completely decentralized: Each device is both sensor and data processor, with sensors making individual or collaborative decisions
Full datasets (with broad spatial and temporal scope) are available	Only small spatial and temporal window of data accessible
Can utilize full available computational power	Limited in processing power (sensor device limitations, including bandwidth and energy consumption)
Can employ complex ontologies	Limited to small tailored ontologies
Typical semantic problems: ● Integration problems arising from variety ● Context of data and sensors ● Provenance	Typical semantic problems: ● Ontologies can be deployed on sensors ● Integrating and maintaining ontologies across sensors.

Standards Integration

Ontology Summit 2009 explored ontology-based standards, and one of the key insights that arose from that work is that specifying an ontology for a standard enables more effective deployment of the standard and easier integration with other overlapping standards. There is also a symbiotic relationship between standards and ontologies -- the terminology within any standard provides the initial set of concepts which are axiomatized within an ontology, and the

specification of the ontology provides rigorous, unambiguous semantics for the terminology of the standard.

What are the relevant or de facto standards involved in the adoption of ontologies for the Internet of Things? There have been several IOT Ontology success stories. The W3C Semantic Sensor Network Ontology (OWL 2) and the OGC Sensor Web Enablement project (including SensorML, a Transducer Model Language, a Sensor Observations Service, Sensor Planning Service) efforts were cited by speaker Henson (Bosch). The GraphOfThings project incorporates SPARQL and the Continuous Query Evaluation over Linked Stream (CQELS) tool. Intellego leverages OWL, RDF and the SSN Ontology.

A decade-old example that predated IoT's entry into common parlance was Project Drishti (Ran, Helal, & Moore, 2004). The investigators sought to integrate data streams from RFID tags, GPS and wireless networks to aid the visually impaired in common navigation tasks. There were numerous other integrations in the wearable and ubiquitous computing literature, even in science fiction.

Fast forward to the present and the number of data sources has multiplied. Big Data is competing with IoT for attention – and legitimately so, as noted in the 2014 Ontology Summit. This has created terrific momentum, especially for Big Data and the Apache stack which owns most of the developer mindshare about this paradigm shift. A convergence of open source projects, cloud computing and a steady march toward web-enabled applications has facilitated big data, but has the same occurred for IoT? There does not seem to be an IoT equivalent for the shift represented by the Apache stack with Hadoop at its center.

It seems clear that there are many efforts underway, and that full coordination with standards or Standards Developing Organizations is not a prerequisite for building a workable system. Benefits from using ontology-based standards in IoT may be more evident as systems mature than at this early stage of IoT work simply because more things will be interconnected. A complex system requiring many different human and organizational roles, processing speed and volume might need an ontology as its associated sensor grid shifts beneath it.

Challenges

Software Support We lack tools for a wide range of tasks, including for semantic annotation and ontology validation. Furthermore, most applications still rely on manual methods for integration. There is also demand to create tools for ontology visualization and interoperability testing.

What ontologies are needed for supporting today's envisioned IoT applications? Much existing work for modeling IoT resources focuses primarily on sensors and sensor networks and is

modeled by SSNO. Most of the existing IoT or sensor-related ontologies represent IoT devices only partially (e.g. as sensing devices), so extensions will be required to include other entities and their relationship to actuator devices. A broader view of IoT resources including other important resources and devices such as actuators, IoT gateways, data aggregators and servers is needed. Work to develop ontologies for these is underway.

Beyond Semantic Sensor Network Ontologies How do we handle going beyond SSN with an Open Source Cloud solution for the Internet of Things (OpenIoT)? Challenges include sensor annotation, sensor mobility & efficient data harvesting and data quality.

What Kinds of Axioms are Needed? Is the priority work and opportunity for ontologies to be used to annotate IoT data, or to more fully represent and model sensors and data in order to analyse/understand it?

Semantic Annotation How can we provide an ontological base for generating semantic annotations of open source internet-connected objects? The challenge would be to obtain open sensor information in a standard encoding that is understandable by users and their software

Semantic Registry for IoT Entities, built on top of DUL and SSNO[1]. Besides the registration of IoT things, abstractions of technological heterogeneity are also required. Such abstract semantic heterogeneity leads to the need to use heterogeneous domain ontologies to semantically annotate data of IoT entities.
Ontology Evolution How can we characterize how ontologies change in order to address future IoT applications?

Forecasts

Ontology Development There will be a number of efforts to enhance and extend IoT ontologies such as SSNO. More ambitious extensions of SSNO will support the extraction of knowledge from the raw sensor data, enabling the understanding of the ``big picture" of what is happening by explicitly representing the interactions between complex processes and events that cannot be captured by a single signal alone.

Ontology Embedding The increased use of smart devices, store-and-forward, embedded intelligence automated data fusion (perhaps especially for geospatial aspects) suggests that ontology embedding could become a design pattern. The pattern could be used in building intelligent IoT, but ontology embedding within sensor systems themselves is possible. Metadata for discovery and provenance from devices are possible starting points.

[1] Some initial work along these lines can be found at
http://purl.org/IoT/iot-ontology.owl
http://ai-group.ds.unipi.gr/kotis/ontologies/IoT-ontology

Automated Deployment of IoT Apps in Unknown Environments Approaches such as the Semantic Smart Gateway Framework will be extended to support full automation in terms of uncovering the semantics of IoT entities as well as aligning their semantics in cases of disagreement.

Exploitation of (Lazy) Developer Pain Points Known problem areas in IoT exist across many different types of sensors. These include security, privacy, signal noise, reliability, configuration management, infrastructure dependency and other known architectural nuisances. A standard solution in any of these areas could catch on because it would solve a well-defined problem that is tangential to an architect or sponsor's main system objectives.

Specialized Engines Reusable, high-complexity solutions might take hold to implement mathematical solutions in certain spaces, such as Gruninger's work with PSL in ERP or Spencer Breiner's category theory.

Cloud Impact Because cloud engines such as Watson will provide complex building blocks for architects, the challenge may be taken up by small groups or even sole developers working in green field problem spaces.

Fun Hardware Syndrome Sometimes collateral innovations co-occur with fun hardware developments. The smart car, or low cost commercial unmanned vehicles could spur ontology-rich solutions. The reasons for such developments are connected both to standards and to the attitudes (plus and minus) about existing standards.

Integrated Development Environment Innovation Will IoT need its own integrated development environment? Test and development beds for IoT will likely require new combinations of devices, simulations, test data, standards, scalability exercises and more.

Recommendations

1. IoT ontologies need to deal with dynamic time varying data vs. the often static Semantic Web. In particular, more work is needed on the development of event ontologies for targeted domains, building from core ontologies.
2. Use design patterns toward ontology virtualization: Given a set of ontology design patterns and their combination into micro-ontologies, one can abstract the underlying axiomatization by: dynamically reconfiguring patterns in a plug and play style; bridging between different patters as micro-theories; providing ontological views and semantic shortcuts that suit particular provide, user, and use case needs by highlighting or hiding certain aspects of the underlying ontological model; and mapping between major modeling styles

3. Integrating SSNO with other Web standards and ontologies is a near-term focus for work. In particular, there is a need to support applications that combine SSNO with PROV-O (for data provenance), CoAP (Constrained Application Protocol), and RDF Data Cube vocabulary. There are also many applications based on biomedical ontologies dealing with sensors in medical devices.
4. Ontology reuse is key. Of course, ontology reuse issues are not unique to IoT but there are some good ontologies such as SSN and PROV that provide some starting points for representing sensors, sensor also networks, observations, etc.
5. Link your data and descriptions to other existing resources
6. Semantics are only one part of the solution and often not the end-product so the focus of the design should be on creating effective methods, tools and APIs to handle and process the semantics. Query methods, machine learning, reasoning and data analysis techniques and methods should be able to effectively use these semantics.
7. A critical obstacle in the widespread adoption/application of ontologies to earth science and sensor systems is the lack of tools that address concrete use cases. Developers will need to focus on those tools and techniques that support the deployment of ontologies in IoT applications.
8. Create an IoT equivalent to Google Search to identify the scope of available end points for different application domains.
9. A more coordinated effort is required to compile IoT case studies which can serve as the basis for ontology reuse and the design of new ontologies. Key areas included Sensor integration, Smart Grid, and Smart Healthcare.

Terminology

- *Internet of Things.* The Internet of Things (IoT) is a term that is being used to denote a network – typically the Internet -- of devices that constantly monitor the environment and can result in "intelligent actions." These devices can range from simple sensors to complex systems such as automobiles and buildings. There are several views of IoT in vogue. For example, ITU (International Telecommunication Union) and IERC (IoT-European Research Cluster) define IoT as "a global network infrastructure with self-configuring capabilities based on standard and interoperable communication protocols where physical and virtual things have identities, physical attributes and virtual personalities, use intelligent interfaces and are seamlessly integrated into the information network." (See Internet of Things – From Research and Innovation to Market Deployment, Vermesan, O. and Friess, P. (editors), 2014, River Publishers, Aalborg, Denmark).
- *Cyber-Physical Systems.* Cyber-physical systems (CPSs) extend IoT by adding a control and decision making layer. Again, several views of CPSs exist. One commonly used definition is provided in http://varma.ece.cmu.edu/summit/index.html, which places an emphasis on embedded systems and the tight coupling between hardware and software. CPSs will play an increasingly important role in the next generation industrial systems.
- *Cyber-Physical Human Systems.* When humans take an active role in CPSs we have Cyber-physical Human Systems (CPHSs). These systems can be viewed as socio-technical systems, with a symbiotic relationship between the human and the physical device.
- *Cyber-physical Social Systems or Smart Networked Systems and Societies.* Social networks, such as Facebook and Twitter, primarily connect people to one another. These networks are playing very important roles in people's lives today, from how some of them behave and interact with one another, to change in human resources processes, how companies market and sell products and services, developments in healthcare and smart (electrical) grid systems, and even roles in politics and democratic uprisings. Social networks have been used both to curtail and to propagate freedom of speech. When these networks are combined with CPSs, we have Smart Networked Systems and Societies (SNSS), which are also known as Cyber-physical Social Systems (CPSS) or Internet of Everything (IoE).