

Semantic technologies for the IoT – an Inter-IoT perspective

Maria Ganzha^{*†}, Marcin Paprzycki^{*}, Wiesław Pawłowski[†], Paweł Szmeja^{*} Katarzyna Wasilewska^{*}

^{*} Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warsaw, Poland

$\langle \text{firstname.lastname} \rangle @ \text{ibspan.waw.pl}$

[†]Institute of Informatics, University of Gdańsk, ul. Wita Stwosza 57, 80-952 Gdańsk, Poland

$w.pawlowski@inf.ug.edu.pl$

[‡] Warsaw University of Technology, ul. Koszykowa 75, Warsaw, Poland

Abstract—The Inter-IoT project is aiming at the design and implementation of, and experimentation with, an open cross-layer framework, to provide voluntary interoperability among heterogeneous Internet of Things (IoT) platforms. The project is driven by use cases from two domains: *(e/m)Health* and *transportation and logistics* in a port environment. While the Inter-IoT will provide interoperability across the software stack, here, we focus on the semantic interoperability. In this context, we present a concise overview of existing IoT-related semantic approaches, which might either be directly applicable to, or serve as a source of inspiration for, the Inter-IoT applications.

I. INTRODUCTION

Ubiquitous computing, proposed by M. Weiser in the late 80’s (**Weiser:1991**, **Friedewald.Raabe:2011**), is gradually becoming reality. The *Internet of Things* (IoT), conceptualized as an omnipresent network, consisting of physical and/or virtual objects/resources, equipped with sensing, computing and communication capabilities is its most recent incarnation. With billions of sensors and other *things* already deployed, and having Cloud and Big Data technologies at our disposal, we can clearly see that the vision of *hyper-connected* world is closer than ever before.

Dealing with volume of data produced by the *things*, their varying capabilities, and an exploding number of services, which they offer (or require, to be “useful”), are among the biggest conceptual and technological challenges of our time. But it’s not only the scale and the heterogeneity, which have to be considered. A dynamic and often unknown topology of the network, high possibility of conflicts between *things*, unknown availability of data-points, as well as various security issues are just a few examples of other problems to be tackled.

To deal with these challenges, common description and data representation frameworks, to characterize *things*, their capabilities and data they produce, in machine-readable and -interpretable form, are needed. Since the IoT can be naturally perceived as a “successor” of “the Web,” approaches, which are believed to have a chance to be successful in the case of the latter, are being considered for the former. Henceforth, it is reasonable to believe that semantic technologies, based on application of *ontologies* **Staab.Studer:2009** can facilitate

interoperability among the *things* and can be used for their semantic annotation, managing access, resource discovery and knowledge extraction. In this context, common interpretation of data and information, based on a global shared ontology, is crucial to achieve semantic interoperability. Here, we acknowledge that the vision of the Semantic Web is still to be realized. For instance, as one can see from **Cousin.all** semantic methods are still used almost exclusively within the research community. However, recent developments (e.g. success of Linked Data) make us believe that practical application of semantic technologies is just a matter of time. This being the case, in what follows we present a short overview of ontologies applicable to the IoT “in general,” as well as domain specific ontologies and standardization efforts related to the two use cases of the Inter-IoT project, i.e. *(e/m)Health* and *transportation and logistics*.

II. ONTOLOGIES IN THE INTERNET OF THINGS

The *early stage of adoption* of semantic methods in the IoT is evident already at the level of ontologies. Most of them were developed within individual research projects and, hence, they are prototypes, often incomplete, or abandoned (upon project completion). A notable exception is the W3C SSN ontology, developed as a joint effort of several organizations, which became the standard ontology for the *semantic sensor networks*. For all practical purposes, this is the *only* ontology explicitly mentioned in **Cousin.all** other than the *OpenIoT* ontology, which is based on the W3C SSN.

Sensors and sensor networks are the core component of the IoT. Sensor network ontologies attempt at capturing information about sensor capabilities, performance, usage conditions, and enabling contextual data discovery. Among recently developed ontologies, we start from a short description of ontologies that, as far as we were able to establish, are no longer “under development.” Observe that some of them were more generic, while others focused on more domain-specific aspects of sensors and sensor networks. Additionally, **Barnaghi.etal:2012** and **Perera.etal:2015** should be consulted for further references.

- **CSIRO Sensor Ontology** **CSIRO** It was an early attempt at development of a generic ontology for describing functional, physical and measurement aspects of sensors.

It was created at the CSIRO, Australia. Its main classes include sensors, features, operations, results, processes, inputs and outputs, accuracy, resolution, abstract and physical properties, and metadata links.

- **SWAMO Ontology SWAMO** The aim of the SWAMO project **Underbrink.etal:2008** was to use collaborative, distributed set of intelligent agents for supervising and conducting autonomous mission operations. SWAMO ontology enables automated decision making and responses to the sensor web environment. One of its advantages was compatibility with the Open Geospatial Consortium standards, enabling data consumption and exchange.
- **MMI Device Ontology MMI** An extensible ontology of marine devices (hence, a more “domain-specific” ontology) that integrates with models of sensor descriptions. Its main classes include component, system, process, platform, device, sensor, and sampler.
- **SEEK Extensible Observation Ontology (OBOE; OBOE)**. A suite of ontologies for modeling and representing scientific observations. It can express a range of measurement types, includes a mechanism for specifying measurement context, and has ability to specify the type of entity being measured. In this way it is focused more on the results produced by sensors than sensors themselves.

All these ontologies, as well as the observation-centric ontologies, e.g. the SemSOS ontology **Henson.etal:2009** and the Stimuli-Centered ontology **Janowicz.Compton:2010** contributed to the development of the *W3C Semantic Sensor Network ontology* (SSN); **SSN** The W3C SSN **Lefort.etal:2011**, **Compton.etal:2012** ontology is actually a suite of general purpose ontologies for describing sensors, their accuracy and capabilities, observations and methods used for sensing. Further information, concerning deployment and use of sensors is also captured. More specifically, the SSN consists of 10 conceptual modules (Deployment, System, OperatingRestriction, PlatformSite, Device, Process, Data, SSOPPlatform, MeasuringCapability, ConstraintBlock) which contain 41 concepts and 39 object properties. It directly inherits 11 concepts and 14 object properties from the DOLCE-Ultralite ontology **DolceDNSUltralite**

When considering semantic technologies applied to the IoT (in general), it is also crucial to mention the results of the, recently completed, EU-funded *OpenIoT* project. The OpenIoT open source platform **Soldatos.etal:2015** utilizes both cloud-computing and semantic methods and focuses on interoperable IoT deployments. At the sensor level, the *OpenIoT* utilizes the *XGSN Calbimonte.etal:2014* an extension of the GSN middleware **Aberer.etal:2006** for semantic annotation of (virtual) sensors. The *OpenIoT* ontology uses the *W3C SSN* ontology as its “base.” It has been combined with several well-known vocabularies and relations (e.g. PROV-O provenance ontology, LinkedGeoData and WGS84 geo-ontologies, LSM live sensor data management vocabulary, etc.) to model the necessary concepts. It was also augmented with cloud-computing related concepts. By combining cloud-computing and sensing capabil-

ities, the *OpenIoT* platform supported on-demand cloud-based access to the IoT resources.

Henceforth, it can be claimed that *any* project planning to *fuse IoT and semantic technologies* should definitely start by taking full advantage of the SSN ontology. It should then extend it, by adding concepts necessary to deal with intended application(s). These concepts may be needed either on the “sensor level” (not likely, as the SSN is quite comprehensive), or to deal with application areas. This is also precisely the way that we will proceed during the Inter-IoT project. Since our application areas are related to *(e/m)Health and transportation and logistics*, let us now look into ontologies available there.

III. (E/M)HEALTH ONTOLOGIES

The term *mHealth* (mobile health, **Nacinovich:2011**, **Bashshur.etal:2011**) is relatively new, although Web-based healthcare (*eHealth*) was always seen as one of the main application areas for the IoT technologies. As a part of *eHealth*, *mHealth* focuses on use of mobile, connected devices, to provide healthcare services, such as patient monitoring, medication management, management of medical data, and telemedicine services **Germanakos.etal:2005** Potential sources of information in *(e/m)Health* are, e.g. non-wearable sensors (e.g. weight scale, oximeter, ECG meter, static blood pressure monitor), wearable sensors (e.g. pedometer, accelerometer, gyroscope, wrist band, mobile health and fitness apps, in-ear devices), hospital information systems, radiology data, etc.

Note that data collected from different sources can describe the same fact(s) about the patient e.g. blood pressure can be measured in the hospital or at home (by a portable device connected to a Cloud); heart rate can be measured by Holter monitor, or by a wearable sensor being part of a Body Sensor Network (BSN **Fortino2014** a set of wireless wearable sensor nodes usually coordinated by a static or mobile device that is used to monitor assisted living). Here, BodyCloud **Fortino2014**, **Fortino:2013** is one of the cloud-enabled platforms that integrates BSNs with Cloud computing. Overall, cloud-based architectures for *(e/m)Health* facilitate data sharing and processing via Cloud services. Note that while the context and the accuracy of the measurements may differ, the IoT infrastructure should provide means to interrelate appropriate data.

On the other hand, Electronic Health Record (EHR), defined as a complete electronic registry of all events and data related to the health status of a person, is the main source of clinical information for both *eHealth*, in general, and *mHealth*, in particular. Note that, the EHR can be modeled in different ways, depending on the application context. Creating a single reference model seems to be nontrivial or even impossible (e.g. taking into account different legal regulations around the world). On the other hand, providing semantic interoperability between different EHR models can address the problems of standardization, data sharing and reuse. Furthermore, standardization in the area of *(m)eHealth* would enable easier and more generic interoperability with other domains.

In order to facilitate semantic interoperability in (*m/e*)Health, and to consider interoperability with different domains, e.g. *transportation and logistics* let us summarize key existing standards for the healthcare-related data. However, note that, despite existence of (i) ontologies in the biomedical domain (see, for instance, **bioontology**), which includes healthcare, and (ii) standards described below, use of semantic technologies has not seen widespread adoption in (*m/e*)Health systems. The most notable existing developments include:

- **OBO.** Open Biomedical Ontologies **OBO** aims at providing a taxonomy across various biological and medical domains. The library of ontologies was developed after specifying a set of best practices in ontology development, as an effort to foster interoperability. Among ontologies within the OBO, one can distinguish: *doid* – Human Disease Ontology, *cmo* – Clinical measurement ontology, *symp* – Symptom Ontology.
- **SNOMED CT.** SNOMED Clinical Terms **SNOMED.CT**, **SNOMEDCT** constitutes the taxonomy of medical terms used in clinical documentation and reporting. It provides the core general terminology for the Electronic Health Records (EHR). The *SNOMED CT* covers, among others, symptoms, diagnoses, procedures, body structures. It maps defined concepts into other international standards and classifications, e.g. ICD-9, ICD-10. Importantly, *SNOMED CT* is used by the World Health Organization, in an ongoing effort aimed at developing the ICD-11 (the new version of the international classification of diseases).
- **ICDx.** The International Classification of Disease coding standard is a classification of diagnosis developed by the World Health Organization **ICD9**, **ICD10** Even though the *ICDx* classifications and the *SNOMED CT* provide only terminologies (and not full-blown ontologies), for one of the Inter-IoT application domains, notable is the existence of mapping between different classifications. Mapping mechanisms, developed so far, between available classifications and terminologies, are potential foundations for providing interoperability within *e/m Health* IoT solutions.
- **HL7.** While the *OBO* standards have been focused on scientific ontologies, *HL7* is an international standards development organization, in the area of healthcare information technology. Initially, the *HL7* created *version 2* (*HL7 v2*; **HL7v2**) standards that were later replaced by a more formal and methodology-based *version 3* (*HL7 v3* **HL7v3**) of standards for the data exchange via point-to-point messaging. Unfortunately, with recent development in distributed computing and semantic technologies, *HL7 v3* is criticized for its poor interoperability and internal inconsistencies (see, **HL7.Landgrebe:2011**). As a result, in 2007, *HL7* developed *HL7 SAIF* **HL7.SAIF** (Services-Aware Interoperability Framework) that provides consistency between all *HL7* artifacts – a foundation framework for further standarization and a general upper ontology. Specifically, it provides a family of standards that explicitly describe the governance, behavioral, information, compliance and conformance semantics needed to achieve semantic interoperability. It should be noted that the *SAIF* proposes design paradigms for the interoperability, and not a full solution. Some criticism of the *SAIF* was presented in **HL7.Landgrebe:2011** In 2014 the **HL7 FHIR** **HL7.FHIR**, **HL7.Bender** (Fast Healthcare Interoperability Resources) was proposed, as a set of standards describing data formats and elements for exchange of medical data (resources). It supports exposing basic data elements, e.g. patients, admissions, medications, diagnostic information that can be referenced by their assigned URIs. Its focus is on providing a set of APIs to enable creation of interoperable *mHealth* applications. The *FHIR* was build on top of the *HL7 v2*; version of the standard that was already implemented in many systems, which provides a standardized ontology. This makes the *FHIR* a good candidate for the *IoT* applications. It is worthy noting that it is currently used in some *mHealth* projects, e.g. in mobile applications **Lamprinakos:2014** or in research projects **Beredimas:2015**, **Schwartz:2014**
- **OpenEHR** **OpenEHR** is a community working on interoperability and computability in the *eHealth* domain, with the main focus on the EHR. It has developed a set of specifications (archetypes) defining the reference model that can be used to implement specific clinical models. The *OpenEHR* enables usage of external healthcare terminologies e.g. *SNOMED CT*, *ICDx*. Note that, in order to apply semantic interoperability in (*e/m*)Health, one has to investigate medical data that are actually processed in healthcare systems. The *OpenEHR* is a notable source of information regarding EHR modeling that needs to be considered a crucial part of the (*e/m*)Health application domain of the Inter-IoT project. However, within the Inter-IoT, software and architectural issues will be also addressed.
- **CEN/ISO EN13606 ISO13606** is an European norm designed to achieve semantic interoperability in the Electronic Health Record communication. As an existing ISO norm, it has to be seriously taken into account in any project dealing with medical data.
- **Obesity management ontology** is a specific ontology proposed in **Hyun-Youngi.etal:2013** where authors looked at the problem of obesity management from the point of view of patient monitoring via mobile devices. The resulting ontology was linked to the existing *OBO relationships* **OBO.Smith.etal:2005** and *SNOMED CT* ontologies. It is of interest to the Inter-IoT project because it directly concerns one of the proposed test applications.

Finally, while the standardized biomedical ontologies usually go into detail about particular biological issues, in **Cameron.etal:2015** a general ontology for *mHealth* was

proposed. The goal of this work is to provide a consistent set of definitions and requirements and to properly define *mHealth* and its position among *eHealth*, biomedicine and healthcare. As such, it describes a view of *mHealth* as a set of services, providers, clients, devices, etc.

It should be clear, that the (*e/m*)*Health* use case of the Inter-IoT project will find grounding in abundance of domain specific ontologies, likely to cover most of needed functionality. However, the plenitude of standards, models and ontologies (that are in use and under development) may itself pose a considerable challenge when building interoperable solutions.

IV. TRANSPORTATION/LOGISTICS ONTOLOGIES

Let us now complete our survey by looking into existing ontologies in the transport and logistics domain. As it turns out, such ontologies span business perspectives of freight and production companies, transportation hubs (e.g. airports, train stations), transport infrastructure, mass transit, personal and business travel, and others. This broad range stems from multitude of issues and problems encountered in modern transport and logistics systems (see **Montreuil:2011**). In accordance with the Inter-IoT's use cases, we are *not* interested in the generic or personal travel perspective, instead focusing on freight, cargo, and top-level transportation and logistics ontologies.

It is quite interesting that, in logistics in particular, many ontologies cover specific (and narrow) areas **Strzelczak:2014** and very rarely describe a broad view of logistics and manufacturing. Furthermore, authors of **Strzelczak:2014** concluded that, in many cases, work on logistics ontologies ended at the design phase. Let us describe in more detail selected ontologies of particular interest to the Inter-IoT project.

- **OTN OTN, lorenz2005ontology** Ontology of Transportation Networks is a top-level ontology modeling general facets of transportation, traffic networks and locomotion. It describes many aspects of transportation relevant to, for instance, smart city transportation or a smart highway systems. The OTN is a realization and extension of the GDF **GDF** – Geographic Data Format – in a formal OWL ontology. The GDF itself is an ISO specification, describing mostly a way to store geographical information for an “intelligent transport systems.” The OTN was used in **metral2010ontology** as part of an effort to improve ontology-driven interoperability between urban models. The OTN was produced as part of the REWERSE REWERSE project.
- **Logistic Grid Ontology InterLogGrid** presents a service-oriented approach to logistics. It realizes the idea, presented in **scheuermann2012ofipols** of combining semantic technologies and cloud services, in order to enable semantically-driven description and application of logistic processes. It was developed within the LOGICAL project **LOGICAL** aim of which was to “enhance the interoperability of logistics businesses.” A cloud service **LOGICALCloud** utilizing the Logistic Grid Ontology was one of the results of this project.

- **Logistics Core Ontology** **daniele2013** is a model of a core (high-level) ontology for interoperable logistics operations. It was developed within the context of the iCargo **iCargo** and the CASSANDRA **CASSANDRA** projects, focused on international cargo transport and supply chains. Its creation was motivated by the need to provide interoperability for the enterprise (often non-ontological) logistics systems, especially when they were not designed to be interoperable. The ontology itself specializes the DOLCE+DnS Ultralite **DolceDNSUltralite** standard. According to our best knowledge, currently the ontology is not publicly available.

Let us also mention one more class of ontologies, which could become useful in our transport / logistics use case. The *The Transport Disruption Ontology* **TDO** is devoted to modeling events, which can have a disruptive impact on travel planning. It is based on the analysis of published disruption information and the road disruptions described in the DATEX II **DATEX** specification. Here, each *event* is defined in terms of its time of occurrence and place (location).

Overall, it is clear that also in the area transportation / logistics there exists a number of ontologies that can be applied to the Inter-IoT use case application. Here, the decision, which one is to be used in the port logistics scenario, will be based on the requirements analysis and detailed information concerning data that is already in use. However, as in the case of (*e/m*)*Health* ontologies, abundance of choices may turn out to be a challenge rather than help for interoperability.

V. CONCLUDING REMARKS

The aim of this paper was to report results of our attempt at answering the question: what ontologies are available (and “ready to use”) for development of interoperable applications in the Internet of Things. We were particularly interested in (a) “general” IoT ontologies, and ontologies for our two use case applications (b) (*e/m*)*Health* and (c) *port transportation / logistics*. The key results of our investigations are as follows:

- There exists a number of semantic sensor network ontologies (with different scope); however, currently, the key role is being played by the W3C SSN ontology.
- In the Inter-IoT we should proceed similarly to the OpenIoT project; i.e. take the W3C SSN ontology as a starting point and build the top level interoperability ontology around its concepts; by extending (and, unlikely, modifying) them.
- Applicability of the *OpenIoT* ontology will have to be evaluated in detail, in the context of the Inter-IoT project (both the use cases and the meta-level artifacts).
- There exists a large number of ontologies / taxonomies / archetypes dealing with different aspects of (*e/m*)*Health*. They are most developed / mature among those considered, and under systematic development.
- Our findings agree with conclusions of **daniele2013** that, typically, organizations in transport / logistics have their own “local” standards (often with poor formalization of semantics). It is also not unusual for semantic models

to be abandoned after few attempts at implementation (or even during the design phase). Despite this there are a few projects focused on interoperability that offer comprehensive working models.

- Abundance of domain / use case specific ontologies may itself lead to challenges when building interoperable solutions within- and cross-domain, as they will require “extra work” to make them interoperable. This may be particularly difficult when dealing with “production

systems” actually applying semantic data processing.

Knowledge gained during initial investigation, described above, will be further extended and applied to design ontologies for the two application domains and to provide generic interoperability mechanisms based on the development of a generic interoperability ontology of the IoT. This ontology will be used as the centerpiece for establishing interoperability allowing for, among others, data interoperability, message translation, etc.