Semantic interoperability in the Internet of Things; an overview from the INTER-IoT perspective[☆]

Maria Ganzha[a,c], Marcin Paprzycki[a,d], Wiesław Pawłowski[b], Paweł Szmeja[a], Katarzyna Wasielewska[a]

[a] *Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*
[b] *Institute of Informatics, University of Gdańsk, Poland*
[c] *Warsaw University of Technology, Warsaw, Poland*
[d] *Warsaw Management Academy, Warsaw, Poland*

Abstract

The Internet of Things (IoT) idea, explored across the globe, brings about an important issue: how to achieve interoperability among multiple existing (and constantly created) IoT platforms. In this context, in January 2016, the European Commission has funded seven projects that are to deal with various aspects of interoperability in the Internet of Things. Among them, the INTER-IoT project is aiming at the design and implementation of, and experimentation with, an open cross-layer framework and associated methodology to provide voluntary interoperability among heterogeneous IoT platforms. While the project considers interoperability across all layers of the software stack, we are particularly interested in answering the question: how ontologies and semantic data processing can be harnessed to facilitate interoperability across the IoT landscape. Henceforth, we have engaged in a "fact finding mission" to establish what is currently at our disposal when semantic interoperability is concerned. Since the INTER-IoT project is initially driven by two use cases originating from (i) *(e/m)Health* and (ii) *transportation and logistics*, these two application domains were used to provide context for our search. The paper summarizes our findings and provides foundation for developing methods and tools for supporting semantic interoperability in the INTER-IoT project (and beyond).

Keywords: Internet of Things, ontologies, semantic interoperability

[☆]This research has been partially supported by EU-H2020-ICT grant INTER-IoT 687283. Work presented here is an extension of results and continuation of the work reported in a conference paper [**?**].

Email addresses: `maria.ganzha@ibspan.waw.pl` (Maria Ganzha), `marcin.paprzycki@ibspan.waw.pl` (Marcin Paprzycki), `w.pawlowski@inf.ug.edu.pl` (Wiesław Pawłowski), `pawel.szmeja@ibspan.waw.pl` (Paweł Szmeja), `katarzyna.wasielewska@ibspan.waw.pl` (Katarzyna Wasielewska)

1. Introduction

The *Internet of Things* (IoT), conceptualized as an omnipresent network, consisting of physical or virtual objects/resources, equipped with sensing, computing, communication and actuating capabilities, can be seen as the most recent incarnation of, so called, *ubiquitous computing* [? ?]. With billions of sensors and actuators (*things*) already deployed, and combined into a number of domain-specific platforms, the vision of the *hyper-connected world* is closer than ever before.

Dealing with the vast amount of data produced by the *things*, their varying capabilities, and an exploding number of services, which they can offer (or require, to be "useful"), are among the biggest conceptual and technological challenges of our time. This challenge is further magnified by the typical ills of early-stage technology. With not much exaggeration it can be stated that "every IoT domain and every IoT vendor produces its own IoT platform." As a matter of fact, different "vendor groups" can be found in different domains, while not a single vendor can be seen as having an "upper hand" in being positioned across all IoT domains. Even the, EU "sponsored," FIWARE platform ([?]) has only a limited uptake. Furthermore, as typically happens in early stages of technology adoption, no real (accepted by most players) standards can be found and none can be expected to materialize in the near future.

It is possible to deal with the interoperability challenge on multiple levels of the software stack. However, it is our belief that the key to solving the problem is on the "highest" level. Specifically, common description and data representation frameworks, which will characterize the *things*, their capabilities and data they produce, in machine-readable and machine-interpretable form, are needed. Since the IoT can naturally be perceived as a "successor" of "the Web," it should not come as a surprise that approaches, which are believed to have a chance to be successful in the case of the latter, should be considered for the former. Henceforth, it is reasonable (today—June, 2016) to believe that semantic technologies, based on application of *ontologies* [?] have the best chance to facilitate interoperability among the *things*, as well as across the IoT platforms. Thus, ontologies should be used for semantic annotation, managing access, and resource discovery in the IoT. As a result, common interpretation of data and information, based on a shared ontology (or, more likely, multiple shared ontologies), is the best pathway to achieve semantic interoperability, which allows to exchange information such that the meaning of it will be automatically interpreted by the receiver in order to produce useful results.

We make these claims knowing very well that the original vision of the Semantic Web is still to be realized. For instance, as one can see from the, three-volume report, *Internet of Things Success Stories* [?], published by the Internet of Things European Research Cluster and Smart Action, actual semantic methods are still used almost exclusively within the research community. However, recent developments in the "world of information processing" (e.g. success of the Linked Data [?]) make us believe that widespread practical application of semantic technologies is just a matter of time. Note also that semantic tech-

nologies are often utilized within multi-agent systems (MAS). At the same time, MAS may provide the right set of mechanisms for implementing IoT interoperability (see, for instance, [? ? ? ? ?]. Therefore, the "push" to introduce semantic technologies into the IoT domain will come not only from within. It will also be facilitated by "outside technologies" that will be tried in the context of IoT interoperability.

This being the case, we have decided to take the bottom-up approach. To be able to apply semantic technologies, one has to have ontologies available. Therefore, we have delved into the state-of-the-art of ontologies in three areas: (1) general ontologies applicable to virtually any IoT platform, (2) ontologies in *(e/m)Health* and (3) ontologies in transportation and logistics. These ontologies, described in Sections 3-5, are considered in the context of use case scenarios introduced in Section 2. Finally, in Section 6, we outline a possible approach to use ontologies to achieve semantic interoperability among heterogeneous IoT platforms.

2. Application scenarios

We consider the issue of semantic interoperability in the IoT from the view point of two use cases related to different application domains. In general, the situation is as follows. We assume that two or more IoT platforms have been instantiated, likely by different vendors, using different technologies, to reach somewhat different goals. Due to the technical progress/change in the business model, stakeholders of these IoT platforms come to the conclusion that it would be beneficial if their platforms were able to "collaborate on voluntary basis" to achieve new set of goals. Obviously, since the platforms are already in place, it is practically impossible to replace one (or more) of them by another—to achieve a "joint system." Note that this is usually impossible (would be extremely rare) even if such platforms belong to the same stakeholder (e.g. after two companies merged). The needed amount of investment, to replace an existing and working deployment, would usually be prohibitive. Therefore, a different approach to establish interoperability has to be sought. Moving up the software stack and seeking "common understanding," through semantic interoperability is the solution considered here. Hence, let us describe in some more details generalizations of two specific use case scenarios that originate from the INTER-IoT project and provide the context for our work.

2.1. (e/m)Health

One of the fields where technology advances lead to collecting more and more data from heterogeneous resources (e.g. sensors) is the *(e/m)Health*. The term *mHealth* (mobile health, [? ?]) is relatively new, although Web-based healthcare (*eHealth*) was always seen as one of the main potential application areas for the IoT technologies. As a part of *eHealth*, *mHealth* focuses on the use of mobile, connected devices, to provide healthcare services, such as patient monitoring, medication management, gathering and storage of medical data, telemedicine services [?], etc.

The key technology for the mHealth are the wireless sensor networks (WSN) defined as distributed, interconnected sensor nodes that can cooperatively monitor human conditions such as, for instance, temperature, blood pressure, heart rate and motion. For example, body sensor networks (BSN) being a sub-type of WSN technology, allow collecting data streams concerning physical, physiological and behavioral human conditions, in order to process them in real time and/or store in dedicated data repositories (possibly, for further processing). Prototypes of such systems have demonstrated the potential in the areas of: round-the-clock patient monitoring both in hospital and out-patients, patients triage in case of a disaster, smart environments for elderly people, motion analysis for Parkinson's disease, etc. Note that in case of, for instance, disaster management and mass event monitoring cooperation of different BSNs becomes crucial.

Potential sources of information for the *(e/m)Health* can be divided into 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 information systems. Wearable devices can be further classified into: portable sensors for home use (e.g. glucose monitor), ambulatory sensors (e.g. Holter monitor), implantable sensors for continuous monitoring of physiological status, sensors embedded into e.g. assistive devices. At the same time, data collected about the patient can be classified into qualitative (e.g. health status analysis through questionnaires) and quantitative (collection and analysis of information concerning physiological parameters).

Note that data collected from different sources can describe the same facts about the patient, e.g. blood pressure can be measured in the ambulatory, or by portable medical sensors at home, heart rate can be measured by the Holter monitor, or by the wearable sensor being a part of the BSN. The context and accuracy of the measurements can be different in the case of each separate device, however, the IoT should provide means of interrelating such data (to understand that they represent the same parameter). Another example are health records generated by various health centers, with different vocabulary structures, that would require some form of mapping to make them "jointly presentable" to the doctor.

A sample high-level architecture of a cloud-based *(e/m)Health* solution consists of three main layers (analogically to any other application domain). At the lowest (first) layer, there are potential sources of medical data including wearable and non-wearable sensors, data files, medical systems, etc. Data originating from different systems/sources that are possibly described using different semantics are collected by an intermediate gateway (second layer), and can be further send to, processed and stored in the Cloud (third layer). In this context, applying the interoperability principle means that the heterogeneous data can be interrelated (and "combined") to be analyzed at the higher layer. Specifically, let us consider decentralized and mobile monitoring in an assisted livings' lifestyle use case that aims at developing an integrated IoT system for monitoring humans' lifestyle in order to prevent health issues result-

ing, mainly, from nutrition and physical activity disorders. It can be observed that most nutritional patients are treated at home, with only the most severe cases requiring hospital treatment. In many chronic diseases, risk factors can be monitored e.g. wrong lifestyles are expressed through raised blood pressure, raised glucose levels, abnormal blood lipids, and obesity. Let us consider two IoT platforms that allow monitoring risk factors. First one is a Cloud-based distance monitoring system collecting qualitative (questionnaires) and quantitative (physiological parameters) data. Measurements are performed with medical devices, e.g. weight scale, blood pressure monitor, that are wirelessly connected to the platform. Measurements are sent to the gateway that is installed on a smartphone/tablet/PC. The gateway sends measurements to the platform, to be stored in a Cloud. On the other hand, patient can wear lifestyle sensors (e.g. pedometer, accelerometer, gyroscope) that are part of the BSN and collect streams of data to send them to the (different) platform through another gateway application. The first approach is based on non-mobile remote monitoring based on non-wearable measurement devices, the second approach provides monitoring of mobile subjects through wearable devices organized as body sensor networks. Even though, these platforms are not interoperable from a technological point of view, their integration would produce a full-fledged mHealth platform, on top of which multitude services can be deployed. By exploiting an integrated IoT environment, the monitoring process can be decentralized from the healthcare center to the monitored patients' homes, and completed with data describing physical activity from on-body monitors. Examples of services that can be developed in integrated mHealth IoT environment include, among others, complex monitoring of patient's health, notification of doctor/patient in the case of any abnormality, comparative analysis of qualitative and quantitative data, e.g. concerning physical activity in the obesity treatment use case.

As mentioned above, medical data can be collected by numerous systems with various architectures. In the initial phase of the INTER-IoT project, data originating from two existing different platforms will be used. First, information collected from the Body Sensor Network (BodyCloud) and second, originating from medical devices placed in patients' home (the e-Care platform). Let us now briefly describe both data sources.

The BodyCloud [? ?] is one of the Cloud-enabled system architectures that integrates BANs (Body Area Networks) [? ?]; set of wireless wearable sensor nodes, usually coordinated by a static or mobile device that is used to monitor assisted livings, and a Cloud computing infrastructure. Cloud-based architectures for the *(e/m)Health* allow large-scale data sharing and processing via Cloud-available services. As stated in [?], in the BodyCloud, one can distinguish two main subsystems. First, *Body*—monitoring the state of the patient by means of wearable sensors and sending data to the Cloud. Second, *Cloud*—providing support for specific applications. Other platforms for assisted livings monitoring based on BANs/BSNs include: MobiHealth [?], CodeBlue [?], AlarmNet [?], LifeGuard [?], AID-N [?], SPINE [?], DexterNET [?], MITHril [?]. Note that, in principle, any subset of such platforms may need to be integrated into a common *(e/m)Health* system.

Example of a platform that processes data collected from non-wearable sensors is the e-Care, developed by Telecom Italia (including the commercial Nuvola It Home Doctor service [?]), were data can originate from medical devices at home (e.g. scales, ECG, glucometer, oximeter, spirometer), medical devices at hospitals (e.g. ambulatory), and lifestyle questionnaires. All data generated in the system is transmitted to a gateway device, e.g. tablet, smartphone, PC, from where it can be later transferred to the Cloud, and eventually analyzed by a doctor. Remote medical care that allows monitoring patients' health status, outside of a hospital, has recently gained popularity. This, in turn, resulted in other commercial solutions developed in that field, e.g. Comarch e-Care Platform, Remote Care Management Platform by Vivify Health. Furthermore, in [?] authors review various Cloud-based healthcare and biomedicine applications and discuss main issues and problems related to the use of such platforms for the storage and analysis of medical data. Discussed solutions use different service models and are dedicated to application domains ranging from genomics, molecular modeling to imaging and telemedicine. Here, again, it can be easily envision that multiple platforms of this type may need to cooperate (e.g. to provide ubiquitous medical services across the EU). Furthermore, while initially in the INTER-IoT project, only one BSN platform will be connected with one stationary platform, a more general situation of any combination of BSN-based and stationary systems can be put together to facilitate an *(e/m)Health* system.

Authors of [?] conducted a study in which they have analyzed 27 *(e/m)Health* projects, implemented across 20 regions in eight European countries with respect to the Suter's key principles for successful health systems integration. One of their conclusions was that interoperability and standardization of data is still the main technological barrier for full success of integrated personal health and care services projects. Here, recall that we believe that semantic technologies are the key to system interoperability. Therefore, in Section 4, we discuss existing ontologies that may be applied in this area, and initiatives that were undertaken in this field of research.

2.2. Transportation and logistics

The domain of transportation and logistics is very broad and includes many aspects of transportation, be it personal travel, city traffic management, transportation hub management, parcel or freight delivery, international shipments, and many others. Internet of Things in this field is realized by a multitude of types of systems, based on a myriad of sensors, both stationary and mobile. Those include, but are not limited to, smart city systems, vehicle and goods geolocation, road congestion analysis systems, recreational travel assist systems, and many others. Such systems serve many purposes, e.g. recreation, commerce, military/political operations and others. Since the INTER-IoT project focuses on the commercial side of transportation and logistics, with particular focus on port logistics, in what follows we also restrict our attention to these areas.

Transportation and handling of goods is a crucial part of nearly any business and there is a constant need for improvement in every facet of it. The

more popular include: reduction of fuel consumption, CO_2 emissions, driver turnover, waiting time, or storage space optimization. The Physical Internet Initiative [?] coined the term "Global Logistics Sustainability Grand Challenge," which describes challenges that the logistic community needs to face, in order to have a globally sustainable logistic network [?]. Those challenges span many economical, environmental and societal issues. However, most of them can be conceptualized in terms of *need for efficiency*, which actually is represented as decrease of cost, downtime and waste production, without a decrease in volume of transported entities (freight and/or humans).

In a constant rush to increase efficiency, every aspect of all logistic processes has seen increased automation and computerization. From the IoT perspective, a big increase in "sensorization" can be observed. The most relevant sensors in logistics gather real-time data about position and status of goods and/or transport vehicles, both en route and inside logistic centers or freight hubs. In order to optimize a supply chain (or a supply network) it is crucial to know the position of all the "moving elements" in the system. Further optimization comes from data acquired from outside of the system, i.e. status and ETA of vehicles and goods from collaborating entities. Many state-of-the-art warehouses include systems that track every item and dynamically optimize the warehousing process [? ?]. Parcel and freight delivery companies track their vehicles via GPS sensors [? ?], which might simply be enterprise versions of personal vehicle tracking systems. Some more advanced solutions include status tracking and map integration [?]. Transportation hubs such as airports [? ?] (eg. Heathrow's The Positive Boarding project [?]) and ports (eg. Hamburg's smartPort system [?]) have also seen an increase in use of "smart technologies."

Positional information can be relative or absolute. The latter is often presented in a standardized and globally understandable format, such as the GPS coordinates (latitude and longitude) and is used for tracking across large distances. Cargo tracking inside a warehouse or a hub often needs a more precise tracking system that includes elevation and relative position, e.g. the order, in which containers are stacked on top of each other [?], or what is their "area code" (e.g. terminal number etc). Status information is very heterogeneous and depends on the place of a system in the supply chain. For a warehouse, or a port, with multiple automated or manually operated devices, status might include state of a machine (busy, free, in need of maintenance, broken), freight space (free, occupied, reserved, free spaces left), while status of a vehicle might include fuel level, destination and source, ETA, driver and company information, and so on. At this point it is worth noting that, semantically speaking, there is a big overlap between the supply chain and either of its ends, i.e. production and delivery end-point (i.e. client). Theoretically, the same precision tracking systems would be useful in a production plant, warehouse, on the road, and on a store shelf. There is a great potential for interoperability when it comes to position and status tracking. Practically, the same functions are realized by different systems, often provided by competing companies. In any case a lot of this information needs to be available and processed in real-time, e.g. once a cargo truck arrives at a destination it receives additional information, which

terminal to approach, what is the waiting queue and so on. Such information is time-sensitive (i.e. it is true and relevant only for a limited time) as well as being available only on arrival and not before.

Ideally, in order to increase control over the supply chain, the status of every moving part in the system should be available on demand (as a real-time data). Increase in the number of sensors also increases the volume and detail of available data, and this enables deep analysis of processes. However, this process has also produced its own set of problems, one of which is the "data flood," [**? ?**] which occurs when a stream of data is too big to be properly analyzed. In this case, proper analysis of incoming data deluge takes too much time for the results to be relevant. Another notable problem is almost complete lack of interoperability. Multiple systems work well on their own, but cannot pass the relevant data to each other, which means that the potential optimization of each part of the supply chain cannot be "orchestrated." As a result, it is possible that, for instance, time saved in the warehouse (due to the IoT system) might be lost at an intermediate terminal (using a completely different IoT system).

Furthermore, observe that a single package originating from a smart warehouse usually stops being tracked (and trackable) as soon as it leaves that warehouse. Even if it is then smart-tagged, or otherwise introduced to another tracking system, it is (virtually) a completely different entity, with separate presences in multiple systems. In other words, along the route, the same physical object has many different descriptions in different systems (exists as separate entities "not related" to each other). This is because these systems do not transfer data from one to another (as they do not understand each other). In semantic data processing, this is a textbook case that benefits from instance alignment (entity resolution)—a process of identifying the same entity in different systems and linking it on the semantic level. A well executed semantic integration of individuals should enable transfer of information between different systems, e.g. a source warehouse might receive and understand positional information about a package tracked by a remote geolocation service. Another positive aspect of logistic systems, when it comes to integration, is that the most basic data, i.e. position (of vehicle or goods) is always relevant. This means that there is a very clear semantic connection between otherwise heterogeneous systems. A popular way of labeling goods—smart tagging—provides only identification services, while any status data is stored separately and accessed by means of the identifier. On the one hand, semantic integration by the means of identifiers is an easy way of entity resolution. On the other hand, the data formats and contents might result in a big interoperability challenge. For instance, a medical container might have an entirely different data schema than a freight container used to transport cars, or a container-less parcel.

The specific use case that drives our work concerns shipment management in a port, with multiple terminals. Here, different stakeholders, such as (among others): port authority, terminal owners, truck drivers, cargo owners, etc. may or may not have, more or less advanced, IoT systems. As in the general description, presented above, these systems do not talk with each other. In other words, truck arriving at the port gate is "invisible" to the terminal system. Similarly,

truck entering the terminal area does not have access to the terminal system, which has information where the container should be picked-from/delivered-to. As a result, truck driver has to seek further assistance to locate the specific place within the terminal area. However, while her own trucking company knows very well where the truck is located, this information is not available to the terminal system. Consequently, time of arrival of a specific truck to a specific location within the terminal remains a "mystery," until the truck materializes. It should be obvious how important establishing interoperability between pertinent IoT systems could be.

3. Ontologies in the Internet of Things

Keeping in mind the two use case areas, let us now focus our attention on more general issues concerning interoperability in IoT systems. Obviously, this problem has been (and still is) addressed by researchers on many levels/layers, including device [?], middleware [? ?], and service [?], while the semantic layer has received considerably less attention. Here, the integration of IoT data into the Web with semantic modeling and linked data approach was discussed in [?]. The *early stage of adoption* of semantic methods in the IoT is evident when one looks for available ontologies. Let us recall that practical use of semantic methods and tools requires formulation/existence of explicitly expressed ontologies, represented using one of ontology languages (currently RDF(S) or OWL). Therefore, let us discuss what is actually available for practitioners that would like to use semantic technologies in IoT environments.

The general observation is as follows. Most existing ontologies, capturing the IoT domain, were developed within individual research projects and, as a consequence, they typically are in a prototype stage, often incomplete and sometimes abandoned (upon project completion). A notable exception is the W3C SSN ontology, which was developed as a joint effort of several research organizations and became the standard ontology for the *semantic sensor networks*. For all practical purposes, this is the *only* ontology explicitly mentioned in [?]; if one doesn't count the *OpenIoT* ontology, which is a recent effort, based on the W3C SSN (see, below). However, while this ontology captures the domain of WSN, to be used in IoT applications it would require further elaboration of the details of the problem at hand. This is to be done in "more specific" sensor network ontologies that attempt at capturing further information about sensor capabilities, performance, usage conditions, and should enable contextual data discovery.

Among ontologies that have been developed in recent years, the following ones are worthy mentioning in the context of the INTER-IoT project. Additionally, [?] and [?] should be consulted for further references. Let us start from a short description of ontologies that, as far as we were able to establish, are no longer "under active development." Observe that some of them are more generic, while others focused on more domain-specific aspects of sensors and sensor networks.

- *CSIRO Sensor Ontology* [?]. 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 Commonwealth Scientific and Industrial Research Organization (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* [?]. The aim of the SWAMO project [?] 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* (OGC) standards, enabling geo-data consumption and exchange.

- *MMI Device Ontology* [?]. An extensible ontology of marine devices (hence, an ontology that is slightly more "domain-specific" than others) that integrates with models of sensor descriptions. Its main classes include component, system, process, platform, device, sensor, and sampler.

- *SEEK Extensible Observation Ontology (OBOE; [?])* is a suite of ontologies for modeling and representing scientific observations. It can express a wide range of measurement types, includes a mechanism for specifying measurement context, and has the 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 *SemSOS observation-centric ontology suite* [?], the *stimuli-centered ontology design pattern* [?], as well as the *OGC SensorML* standard [?] contributed to the development of the, mentioned above, *W3C Semantic Sensor Network ontology* (SSN) [?]. The W3C SSN [? ?] 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, SSOPlatform, MeasuringCapability, ConstraintBlock) which contains 41 concepts and 39 object properties. It directly inherits 11 concepts and 14 object properties from the top-level DOLCE-UltraLite ontology [?].

The W3C SSN ontology has been widely used, and both extended and specialized. Among the notable extensions are the *wireless sensor networks ontology; WSSN* [?], and *sensor cloud* ontology SCO [?]. The specializations include the *AEMET meteorological ontology* [?], *atmosphere observation ontology SWROAO* [?], *flood prediction ontology SemSorGrid4Env* [?], and (data) *stream annotation ontology SAO* [? ?]. A more recent *IoT lite* [?] ontology is a lightweight instantiation of the SSN that provides a general IoT knowledge model intended to limit processing time of ontologically demarcated resources.

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 [?] utilizes both cloud-computing and semantic methods and focuses on interoperable IoT deployments. At the sensor level, the *OpenIoT* utilizes the *XGSN* [?], an extension of the GSN middleware [?], which enables semantic annotation of (virtual) sensors. The *OpenIoT* ontology uses the *W3C SSN* ontology as the starting point. It has been combined with several well-known (at the time when it was being developed) vocabularies and ontologies (e.g. PROV-O provenance ontology, LinkedGeoData [?] and WGS84 geo-ontologies [?], LSM linked sensor middleware ontology [?], etc.). It was also augmented with Cloud-related concepts. By combining cloud-computing and sensing capabilities, the *OpenIoT* platform supported on-demand cloud-based access to the IoT resources, which was needed in the context of the *OpenIoT* project.

Another noteworthy IoT ontology is the Smart Appliances Reference Ontology [?] (SAREF). It describes a top-level perspective on IoT home appliances along with their functions and services. The model is generic enough to be used outside of the home environment, and includes concepts such as *device, sensor, actuator, service, state and function.* SAREF is most naturally applied at home, office or any limited public space.

Note that a popular IoT platform oneM2M [?] also has an ontology, but it has not been finalized (as of the time of writing of this paper). Latest version (0.6) can be found at [?]. The purpose of this oneM2M Base Ontology, as explicitly stated by authors, is to provide an ontology that other systems can align with (i.e. match it with their own ontologies) and, in this way, achieve interoperability. Interestingly, one of the side-goals of the Fiesta-IoT project ([?]) is to fix the interoperability problem between the oneM2M ontology and the FIWARE platform. However, the developed solution (see,[?]) has not been made public.

Henceforth, it can be claimed that (similarly to the way that the OpenIoT project proceeded) *any* project planning to *fuse Internet of Things and semantic technologies* should definitely start by taking full advantage of the W3C SSN ontology. Only then, it should extend it by adding concepts necessary to deal with intended application area(s). These concepts may be needed either on the "sensor level," or to represent concepts formalizing knowledge concerning application areas of interest themselves. Note that the need for adding concepts concerning sensors and sensing is not very likely, as the W3C SSN is quite comprehensive. Nevertheless, it may turn out that it does not capture some unique concepts related to use of sensors in a selected application area. This would be similar to the situation when we had to modify and expand the CoreGrid ontology to deal with resource management in the Grid (for more details, see, [? ?]).

As far as the domain ontologies are concerned, let us recall that the application areas that we are currently interested in are *(e/m)Health* and *transportation and logistics.* Therefore, let us proceed by summarizing what ontologies are available in these two areas.

4. (e/m)Health ontologies

As mentioned, the *(e/m)Health* use case aims at providing healthcare services for ambulatory and remote patient monitoring, where data is collected from different IoT platforms and integrated to provide homogeneous view on patient health record.

In this context, main sources of clinical information for both eHealth, in general, and mHealth, in particular are considered to be: body sensor networks based systems, non-wearable sensors instantiated in medical devices, and Electronic Health Records (EHR; defined as a complete electronic registry of all events and data related to the health status of a person). Note that, the EHR can be modeled in different ways, depending on the application context. Creating one reference model seems to be nontrivial or even impossible (e.g. taking into account different legal regulations around the world). On the other hand, one can hope that providing semantic interoperability between different EHR models can address at least the most important problems of standardization, data sharing and reuse. Furthermore, development of meta-level ontology/ontologies in the area of *(m/e)Health* would enable easier (and more generic) interoperability with other domains.

Therefore, let us summarize key existing standards for the healthcare-related data. However, let us immediately note that, despite existence of ontologies in the biomedical domain (see, for instance, [**?**]), which includes healthcare, and standards described below, use of semantic technologies has not seen widespread adoption in the *(m/e)Health* domain. The most notable existing developments include:

- *OBO*. Open Biomedical Ontologies [**?**] aim at providing a taxonomy shared across various biological and medical domains. The library of scientific ontologies was developed after specifying a set of best practices in ontology development, as an effort to foster interoperability. In the numerous groups of ontologies within the OBO, one can distinguish: *doid*—Human Disease Ontology, *cmo*—Clinical measurement ontology, *symp*—Symptom Ontology.

- *SNOMED Clinical Terms* (SNOMED CT) [**? ?**] 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, and 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 [**?**

?]. Even though the *ICDx* classifications and the *SNOMED CT* provide only terminologies (and not full-blown ontologies), what should be noted is the existence of mapping between different classifications. So far developed mapping mechanisms, between available classifications and terminologies, have potential for becoming foundations for providing interoperability within *e/m Health* IoT solutions.

- *ATC*. The Anatomic Therapeutic Chemical Classification of Drugs, also developed by the World Health Organization, is a pharmaceutical coding system including classification of therapeutic drugs and their therapeutic and clinical characteristics [?].

- *ICF*. International Classification of Functioning, Disability and Health is a taxonomy (also developed by the World Health Organization), structured around the following main components: body functions, body structures, activities and participation, environmental factors [?]. This classification describes human functioning and, as such, it can be useful when adding semantics to describe context in remote health monitoring.

- *LOINC*. Logical Observations Identifiers Names and Codes, is an ontology providing an universal code system for tests, measurements, and observations related to electronic health records [?].

- *CPT 4*—Current Procedural Terminology (4th Edition), is a taxonomy developed by the American Medical Association that describes medical procedures and services [?].

- *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*; [?]) standards that were later replaced by a more formal, and methodology-based, *version 3* (*HL7 v3* [?]) set of standards for the data exchange via point-to-point messaging. It should be noted that, with recent development in distributed computing and semantic technologies, *HL7 v3* is criticized for its poor interoperability and internal inconsistencies (see, [?]). As a result, in 2007, *HL7* developed *HL7 SAIF* [?] (Services-Aware Interoperability Framework) that provides consistency between all *HL7* artifacts—a foundation framework for further standardization 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 the semantic interoperability. It should be noted that while the *SAIF* proposes design paradigms for interoperability, it is not a full solution. Some criticism of the *SAIF* was presented in [?]. In 2014 the *HL7 FHIR* [? ?] (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*, a previous version of the standard, 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. is mobile applications [?] or in research projects [? ?]. The HL7 Version 3 Clinical Document Architecture (CDA) [?], specifies the structure and semantics of clinical documents for the purpose of exchange between healthcare providers and patients.

- *OpenEHR* [?] is a community working on interoperability and computability in the eHealth domain, with the main focus being 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 the *(e/m)Health*, one has to investigate medical data that are actually processed in the 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.

- *CEN/ISO EN13606*. Health informatics—Electronic health record communication, is a 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.

- *ISO/IEEE 11073*. Health informatics—Point-of-care medical device communication, is a European norm describing low level communication standards between medical, health care and wellness devices (e.g. weight scale, pulse oximeter) and with external computer systems [?]. Again, it is a very important norm that has to be considered. However, both these norms are not formal ontologies. Therefore, to be used for semantic processing, their underlying semantics has to be extracted and represented, preferably, in OWL.

- *Obesity management ontology* is a specific ontology proposed in [?], 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* [?] and *SNOMED CT* ontologies. It is of interest to the INTER-IoT project because it directly concerns the proposed test applications (see, Section 2.1).

Finally, while the standardized biomedical ontologies usually delve into details concerning particular biological issues, in [?] 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.

Let us note that, recently, several attention worthy projects have been, or are being, realized in the field of knowledge management and semantic description for the health-related data. The goal of the *Crystal—Critical System Engineering Acceleration* [?] project is to establish and push forward an interoperability specification and a reference technology platform, as a European standard for critical domains (one of them being health). One of the work packages within this project is *defining a common vocabulary for the healthcare use cases and apply it on the interoperability standard*. Unfortunately, the project website only contains the initial deliverable concerning health domain (from 2013), i.e. state-of-the-art of healthcare ontology [?], in which an exhaustive list of standards can be found. Another project is the *SemanticHealthNet—Semantic Interoperability for Health Network* [?], which ended in May 2015. It was focused on semantic interoperability of clinical and biomedical knowledge, in order to ensure efficiency of EHR systems, i.e. to develop an ontological framework, which is compliant with the SNOMED CT, HL7-CDA, EN/ISO 13606 and openEHR, and which allows to seamlessly exchange EHR data. The focus of that project was put on chronic heart failure as patient care exemplar and cardiovascular prevention as public health exemplar. The practical aspects of designing and implementing a system that gathers healthcare data from various sources was discussed in [?].

It should be noted that the aforementioned projects are focused on the eHealth (especially the EHR) rather than mHealth. Furthermore, the semantic interoperability is understood as the ability to exchange data related to the EHR between systems, automatically combining it with local data, and analyzing homogeneously. Moreover, existing projects are not realized in the context of IoT platforms, but tend to analyze existing standards and interrelate them. Research on semantic interoperability in the IoT ecosystem requires considering also issues related to data (measurements) gathered from heterogeneous sources. Note that data entered into the EHR systems is usually first verified by the doctor, whereas data generated in the IoT environment is mostly entered automatically, and can be eventually later evaluated by the doctor. Here, it should be also stressed that automatic data management in healthcare is very sensitive to privacy concerns and, although some existing solutions can be implemented [?], data security and privacy is a separate area. Moreover, it is largely independent of healthcare and, as such, it is out of scope of this paper.

At this moment it should be clear, that the *(e/m)Health* use case of IoT interoperability will find grounding in abundance of domain specific knowledge representing artifacts: ontologies, vocabularies, norms, etc. These are very likely to cover the complete domain of interest. However, the plenitude of standards, models and ontologies (that are in use and under development) as well as semantic and syntactic heterogeneity of health data may itself pose a considerable challenge when building interoperable solutions. Note also that majority of approaches listed above do not offer an RDF(S)/OWL represented ontology.

Therefore, ontology extraction will be necessary. Furthermore, it has been observed that organizations that develop competing approaches are not likely to seek consensus, but rather try to convince that their approach is better than the other ones. This "social aspect," which is rarely taken into account in engineering, may be an even greater challenge for interoperability of IoT platforms than the technical issues that are to be solved. The positive news is that semantic interoperability should provide a meta-level abstraction that may avoid directly dealing with conflicting conceptualizations of the area (see, also, Section 6).

5. Transportation/logistics ontologies

Let us now complete our survey by looking into existing ontologies in the transportation and logistics domain. As it turns out, found 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. As mentioned before (see, Section 2.2), due to the nature of our project, we are not interested in the generic or personal travel perspective, instead we are focusing on freight, cargo, and top-level transportation and logistics ontologies.

It was quite interesting to find out that, in logistics in particular, many ontologies cover specific (and narrow) areas [?] and very rarely describe a broad view of logistics and manufacturing. Furthermore, authors of [?] concluded that, in many cases, work on logistics ontologies ended at the design phase, without an actual full ontology delivered. Keeping this in mind, let us now describe in more detail the selected ontologies that are of particular interest to the INTER-IoT project.

- *OTN* 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 smart highway systems. The OTN is a realization and extension of the GDF [?]—Geographic Data Format—as a formal OWL ontology. The GDF itself is an ISO specification pimarily describing a way to store geographical information for an "intelligent transport systems." The OTN was used in [?], as part of an effort to improve ontology-driven interoperability between urban models. It was produced as part of the REWERSE [?] project and is publicly available at [?].

- *InterLogGrid* Logistic Grid Ontology [?] presents a service-oriented approach to logistics. It realizes the idea, presented in [?], 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 [?], aim of which was to "enhance the interoperability of logistics businesses." A cloud service [? ?] utilizing the Logistic Grid Ontology was one of the deliverables of this project. While

the results have been described in many publications and the cloud system is available online, it is hard to assess the outreach of the project. The URL of the ontology (interloggrid.org), as of the time of writing of this paper, is not associated with the project any more and, to the best of our knowledge, the ontology is not (possibly, no longer) publicly available. It is also not clear how the ontology was utilized within LOGICAL.

- *LogiCO* Logistics Core Ontology [?] is a model of a core (high-level) ontology for interoperable logistics operations. It was developed within the context of the iCargo [?] and the CASSANDRA [?] projects (both part of FP7 [?]), focused on international cargo transport and supply chains. Its creation was motivated by the need to provide interoperability for the enterprise (usually "non-semantic") logistics systems, especially when originally they were not designed to be interoperable. The project identified differences and similarities between terminologies (and taxonomies) used by publicly available commercial Internet logistic applications and pointed out the potential of semantic integration. CASSANDRA emphasized "visibility" of the supply chain and security of transport containers, in the context of long-distance (i.e. international, global) transport. The broader scope of the iCargo included production of ecosystem that would enable integration of services and information from different participating companies in a way that enables synchronization of logistic processes across participants. It also promised a dynamic system that reacts to constantly changing status of cargo, vehicles and infrastructure. The ontology itself [?] specializes the DOLCE+DnS Ultralite [?] standard. The current version of the ontology file is available at [?] and documented at [?]. It was created under the assumption that there can be no common ontology for all possible applications in logistics. Instead, the authors opted to model a network of ontologies based on common core (i.e. LogiCO), designed to be adaptable (extendable) for more specific logistic use cases.

- *LogiServ* Logistic Services Ontology and *LogiTrans* Logistics Transport ontology are two ontologies extending LogiCO. The LogiServ models business activities within the scope of logistics (i.e. logistic services). It was designed to support dynamic planning in a logistic environment with many stakeholders (i.e. transportation and terminal operators) and assist each in planning and optimizing their operations and cooperation. It is the underlying ontology for the iCargo Semantic Access Points (Semantic Gateways) available through the iCargo cloud. An example use case presenting a virtual company (a union of 150 truck service providers) is available at [?].

- *LogiTrans* extends both the LogiCO and the LogiServ and describes transport orders that are connected with transport services (from LogiServ). It details a kind of contract between provider and receiver of a transportation service by specifying requirements of a service user and a plan of realization of the service from the provider. It includes both physical

requirements, such as delivery and pickup locations, cargo type delivery times, as well as virtual properties, like delivery time constraints and legal requirements.

Both the LogiServ and the LogiTrans are available at [?], along with other ontologies.

- *Enterprise Ontology* [?] is an old and venerated ontology that describes very general concepts relevant to any business enterprise. The ontology itself is too high-level to be of direct use in a logistics domain. It has, however, inspired creation of many domain-specific ontologies. Among the relevant ontologies that specify the Enterprise Ontology are: an ontology for (i) mass customization for optimizing inter-organizational and distributed cooperation [?], (ii) production planning and control in a virtual enterprise environment [?], (iii) a supply chain ontology for semantic integration between heterogeneous supply chain information systems [?]. There are many more such ontologies that, although they intersect with the domain of interest, the intersection is not large enough for them to be applicable in our scenarios.

Let us also mention other, more low-profile, ontologies, which could become useful in our transportation and logistics use case. The *The Transport Disruption Ontology* [?], 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 [?] specification. Here, each *event* is defined in terms of its time of occurrence and place (location).

The *OTS* Ontology for Transportation Surveillance [?], is an ontology that models scenarios, in which vehicles are equipped with wireless sensors and organized into *WSN*s. The use case that led to the development of this ontology visualizes a scenario, in which a trucking company utilizes wireless sensors to monitor their fleet and goods, protect the cargo etc. While, at the first glance, the underlying assumptions of the *OTS* fit perfectly within the scope of the IoT transport scenario (i.e. the ontology is very promising), the authors say (see, [?]) that, as of yet the ontology is not a "complete enterprise ontology," but an "application ontology for the defined purpose" of the described use case.

The *VEACON*, is an ontology that describes transportation with the focus on vehicle accidents [?]. It presents a vision of smart vehicles interoperating in a Vehicular Network and using common semantics to communicate any information relevant to the road safety and, in particular, accidents. The communication is to take place within a neighborhood of smart cars, as well as with systems such as the GES (General Estimates System) [?] and with emergency services, management authorities, and police. The relevant work around the VEACON [?] contains a comprehensive description of research that includes usage of GES historical data, as well as crash tests performed in a physical environment. Despite our best efforts we could not find a publicly available VEACON ontology file.

Overall, it is clear that also in the area transportation and logistics there exists a number of ontologies that can adapted to be applied to the INTER-IoT use case application. On the other hand many ontologies (with notable exceptions) focus on a specific sub-area (e.g. production and manufacturing, transportation networks etc.) of transport and logistics. While they define some concepts that may be useful to us, they do not match our need well. Let us add that, to keep this paper focused, we did not include ontologies that we deemed too low-level (domain specific).

The decision, which ontology is to be used, as a starting point, 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 be a challenge rather than help for interoperability. Furthermore, it should be noted that we have not found an ontology that would be focused on port logistics. This means that, regardless of the choice of the starting ontology (ontologies), a lot of basic ontology engineering will be required to properly represent our domain of interest. In other words, one or more ontologies will have to be instantiated on the basis of data exchanged in the port.

6. Semantic interoperability

Thus far we have reviewed state-of-the-art in semantic representation of knowledge for: (i) the Internet of Things, (ii) medical applications, and (iii) transport and logistics. The main conclusions were that: (a) in each area a number of ontologies/vocabularies/standards exist, and (b) further work to achieve semantic interoperability will be required. This work may involve, among others, extraction of a "full-blown ontology" from messages exchanged within an IoT platform or database schema used (see, below). However, let us first assume that this work has been done for each IoT platform individually and reflect how the main goal of establishing semantic interoperability can be achieved.

Since the very beginning, IoT solutions were, and still are, mostly *use case centric*, resulting in the creation of various "IoT silos." As we have seen, from the discussion of both use cases, the "inter-silo" interoperability is of great importance, but achieving it *across* the IoT silos, or more generally, IoT platforms seems absolutely crucial for the future of the IoT. Interoperability can, of course, be "hard wired," enabling total control over the whole process, and giving many possibilities for optimization. Unfortunately, as an approach *hard wiring* does not scale, and obviously requires considerable amount of developer resources. Observe, for instance, amount of work to add another platform to already hard-wired ecosystem. There, most likely, a "bridge" will have to be instantiated between the new platform and each of the already connected ones. Hence, it is not a solution one should be looking for. In recent years, several much less ad-hoc interoperability architectures have been suggested.

Typically, such proposals make use of *agent*, *service*, or *middleware* based techniques. Service-oriented approaches have been surveyed in [**?**] and [**?**]. An interesting proposal of a service-oriented approach to logistics can be found in [**?**

]. The *Distributed Internet-like Architecture for Things (DIAT)*, introduced in [?], defines an interoperability architecture, which also addresses security and privacy issues. Most of these solutions still require to "manually" connect the services offered by the platforms that are being integrated. A slightly more flexible approach, presented in [? ?] assumes that the IoT platform architecture should be augmented with an extension (equipped with a RESTful API) to support integration into a network of heterogeneous *IoT hubs*.

Interestingly, the interoperability solutions mentioned above, either do not use semantic methods at all, or use them rather sparingly. Architectures using ontologies can be found at the bottom level of the IoT stack. The *A3ME* [?] proposes a generic middleware for heterogeneous sensor/actuator networks, in which devices are represented by *agents*. The *A3ME* architecture enables ad-hoc device discovery, semantic description exchange, as well as some basic interactions between devices. The approach presented in [?] is also an interoperability architecture at the sensor/perception layer. It is organized via *semantic brokers*, and conforms to the *Architectural Reference Model (ARM)* [?], developed within the European Lighthouse Integrated Project *IoT-A* [?].

IoT is also closely related to an interesting and important concept of *context-aware computing*. Solutions and tools based on the concept of context-awareness, which are currently available on the market, have been surveyed in [?]. Although interoperability has explicitly been discussed there, no use of semantic technologies have been mentioned. The surveyed platforms seem to approach the interoperability problem mostly utilizing *service-oriented* and *reactive* architectures. An interesting approach to context-awareness, using an ontology-based *context engine* has been proposed in [?]. Unfortunately, it does not tackle the interoperability issue.

When using semantic methods, one can obtain a much deeper understanding of the structure and services offered by the platforms being integrated. The problem of interoperability can then be approached using techniques of semantic integration [?], and to some extend, might be reduced to integration of their ontologies. There are many types of differences that need to be overcome to achieve semantic interoperability between ontologies. Those have been summarized in [? ?] and can be divided into 2 main groups: *language-level* and *ontology-level*.

A language-level difference means that ontologies are written in different formalisms, some of them possibly being more expressive than the others, or offering different sets of constructs. In such cases a *normalization* process needs to occur. Usually, this means a translation of all formalisms into the one used by the ontology that requires most expressiveness. Currently, with RDF(S) and OWL 2.0 [?] being the leading standards [? ?], there are many tools available for translating different formalisms from one to another.

A much harder problem occurs when integration of a formal ontology with an implicit one, perhaps reflected in a database schema or in a communication protocol specification and/or design documents is required. In such a case, there is little hope for an automated ontology integration at the language level. While there are tools available that can help this process by generating a simple

OWL ontology from documentation or by translating a database schema into an RDF(S) graph [? ?], such translations are often one-way only. This means that the obtained ontology can be only used, but it cannot be modified as there is no simple way of reflecting such modification in the original artifact.

We intend to explore the variety of language-level normalization tools in future works.

Assuming that the language normalization has been performed, one still needs to consider differences at the ontology level. In general terms, those arise when there are competing views on the same domain. The problem of *matching* ontologies (also known as *mapping* or *alignment*) has been extensively studied over the years (see, of instance, [? ?]) and many approaches to "solving" it have been proposed. Surveys of various available methods can be found in [? ? ? ?]. In [?] we present our research on existing ontology matching tools as well as an evaluation of their maturity and practicality of potential use.

Ontologies can also be seen as *presentations*, i.e., structures consisting of a signature/vocabulary and axioms/formulas in some formal language. As it turns out, the process of combining them can be neatly expressed using the language of the category theory [? ?]. In this setting, merging of ontologies corresponds to taking *colimits* of their *diagrams* in a suitable category [? ?]. An important feature of the categorical approach is that it can also be applied to *heterogeneous* ontologies, i.e., ontologies expressed in different languages [?]. This enables the treatment of both the language-level and ontology-level level differences within a single framework. The categorical approach also supports and advocates *modularity* in ontology design and usage.

Although numerous tools and methods for combining ontologies have been suggested, it should not come as a surprise that, in general, none of them work fully automatic. Nevertheless, using the semantic approach to interoperability still has many advantages. In particular, assuming one managed to combine ontologies of the IoT platforms, it is possible to employ *semantic reasoning* for the discovery and matching of data and various services offered by them.

From the reasoning point of view, ontologies are essentially *theory presentations* in a particular variant of the *description logic* [? ?], with *decidable subsumption* $\alpha \sqsubseteq \beta$ and *equivalence* $\alpha \equiv \beta$ of *concepts*. Many automated *reasoners*, including *RACER* [?], *Pelet* [?], and *Fact++* [?] have been developed and can be utilized in supporting the IoT platforms interoperability. An up-to-date list of available description logic reasoners can be found at [?]. Many of them have been reviewed and benchmarked in [? ? ?].

In summary, while semantic interoperability has considerable advantages over other approaches, achieving it is a multi-step process. First, formal representation in form of an OWL/RDF(S) ontology has to be made available. Second, ontology matching has to be applied, leading to a merged ontology, or development of some translation mechanism. This, in turn, may involve reasoners for discovery and matching of data and services. It should be stressed, that none of these steps can be done automatically in general, and it is unclear how good are the tools that can facilitate the work that has to be performed. However, it seems reasonable to postulate that the semantic concepts and methods,

including reasoning, should be considered to become a part of a future version of the *Architectural Reference Model* [**?**] for the IoT platforms.

7. Concluding remarks

The aim of this paper was to summarize results of our attempt at answering the question: what ontologies are available (and "ready to use") for the development of interoperable applications in the Internet of Things. We were particularly interested in (a) "general" IoT ontologies, and ontologies for our 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 ontologies dealing with various aspects of sensors and sensing (with different scope, granularity and generality). However, currently, the key role is being played by the W3C SSN ontology (the only W3C designated standard). It is by far the ontology that has seen the strongest uptake and inspired other projects, most notably the OpenIoT. The methodology used in engineering of the OpenIoT ontology bears promise for any other project aiming at development of interoperable IoT solutions (hence, the INTER-IoT project as well). In summary, this method starts with the W3C SSN ontology and extends it by a chain (or network) of ontologies. The purpose of the process is to capture/include domain specific concepts, either by linking and modifying existing domain ontologies or producing new low-level ontologies with the SSN at their core.

There exists a large number of ontologies/taxonomies/archetypes dealing with different aspects of *(e/m)Health*. Most of them are quite mature, and are under systematic development. The biggest repository of relevant ontologies is the *bioportal* that organizes the *OBO* ontologies. Regardless of the fact that ontologies stored there are used primarily in biological and medical research, it is natural that the semantic relationships with the area of healthcare are strong. However, there is a very strong competition between organizations developing and promoting individual approaches. Furthermore, majority of semantic representations of the area do not come as ontologies, but vocabularies, archetypes, norms, etc. This means that ontology extraction and formalization may be required as a part of reaching semantic interoperability.

When it comes to transportation and logistics, our findings agree with conclusions of [**?**] that, typically, organizations in transportation/logistics have their own "local" standards (often with poor, or outdated/not-interoperable formalization of semantics, or with no explicit semantics at all). Unfortunately, it is also not unusual for semantic models, in this area, 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 that were tried in practice. However, we have not been able to locate attempts at capturing semantics of port logistics.

In case of both INTER-IoT use-cases (mHealth, transport and logistics) abundance of low-level domain, or use case specific, ontologies may itself lead to challenges when building interoperable solutions within- and cross-domain.

Semantic interoperability, understood as building a common semantic representation of the world, depends heavily on whether ontology engineers can agree on a single model, or, at least, a set of compatible models representing the same domain. Ontology engineering practice, and slower than expected adoption of semantic technologies has, however, shown that no single domain has a global ontological standard (i.e. new solutions to the same problem are constantly being developed). On the other hand, the emergence and adoption of specific standards, when it comes to core ontologies (like the W3C SSN), is promising.

The unfortunate reality is that, even if semantic technologies are used in practice, they are often proprietary and a closely guarded intellectual property of individual organizations. Consequently, researchers and engineers do not have access to many ontologies that are very application specific on one hand, but also tried in practice and systematically verified, modified and extended. Fortunately, as is the case with publicly available ontologies, companies see the potential of interoperability when it comes to core ontologies and standards, which can be seen in publicly declared (and financially backed) support for technologies such as HyperCat or oneM2M.

Finally, when it comes to reaching semantic interoperability, there exists a number of approaches. Since the emergence of the OWL as an ontology formalization standard, most methods focus on the ontology-level (as opposed to the language level) integration. In addition to the more traditional and plentiful ontology matching/aligning/mapping techniques that exploit many aspects of semantic similarity, methods based on category theory, offering a unified, high-level view of the matching problem, should also be seriously taken into account. However, at least for the time being, the process of reaching semantic interoperability is very labor intensive, as the available tools have limited usability. Nevertheless, together with middleware, agent and service-oriented techniques, the semantic methods should be seriously considered as a potential solution to the problem of establishing interoperability between heterogeneous IoT platforms.