LOV4IoT: A second life for ontology-based domain knowledge to build Semantic Web of Things applications

Amelie Gyrard*, Christian Bonnet†, Karima Boudaoud‡, Martin Serrano*,
*Insight Center for Data Analytics, NUI Galway, Ireland, Email: {amelie.gyrard, martin.serrano}@insight-centre.org
†Eurecom, Sophia Antipolis, France, Email: {bonnet}@eurecom.fr
‡Laboratoire I3S-CNRS/UNSA, Sophia Antipolis, France, Email: karima@polytech.unice.fr

***Abstract*—Semantic Web of Things is a new field combining Semantic Web and Internet of Things technologies to be surrounded by smart objects and applications connected to the Web. On one hand, one of the Linked Open Data applications, called DataHub aims at referencing datasets, on the other hand, the Linked Open Vocabularies (LOV) references more than 400 ontologies. However, we discovered that more than 200 ontology-based projects relevant for IoT are not referenced on such tools since domain experts are not aware of them nor of the semantic web best practices. We propose the Machine-to-Machine Measurement (M3) framework, available online, to rapidly design and develop semantic-based cross-domain IoT applications by reusing as much as possible the domain knowledge (ontologies, datasets and rules). To achieve this goal, there are challenging steps: (1) referencing and classifying semantic-based projects relevant for IoT, (2) re-engineering a dataset of interoperable domain rules to deduce high-level abstractions from sensor data, (3) re-engineering an interoperable cross-domain knowledge to combine domains, and (4) assisting developers in designing IoT applications by designing pre-defined templates. In this article, we are focused on referencing and classifying semantic-based projects relevant for IoT by designing the Linked Open Vocabularies for Internet of Things (LOV4IoT) dataset. We also design a dataset of interoperable domain rules to deduce high-level abstractions from sensor data by designing the Sensor-based Linked Open Rules (S-LOR). This work has been applied to two uses cases: (1) redesigning a security and cross-domain knowledge base to assist users in suggesting security mechanisms to secure their applications, and (2) designing semantic-based IoT applications embedded in Android-powered devices.**

***Index Terms*—Semantic Web of Things (SWoT), Internet of Things (IoT), Web of Things (WoT), Semantic Sensor Networks (SSN), Semantic Web technologies, Linked Open Data (LOD), Linked Open Vocabularies (LOV), Linked Open Rules (LOR), Ontologies, Reasoning, Security.**

I. Introduction

To build the Semantic Web of Things [1] there is a need to reuse as much as possible knowledge already designed by existing domain-specific and sensor-based applications. To easily combine all sensor data together to provide promising cross-domain applications there is a need to: (1) unify domains to ease cross-domain interoperability, (2) interconnect data in a uniform way by adding an explicit information using semantic web technologies, and (3) reuse as much as possible domain knowledge (ontologies, datasets and rules). In our previous work [2], we proposed an approach to automatically enrich cross-domain data to later reason about them by reusing domain knowledge. We introduced the need of the concept called 'Linked Open Rules' to reason about sensor data and infer new knowledge [3]. We intuitively want to reuse as much as possible domain knowledge that are already described in domain-specific applications. For example, we could interlink the weather and transportation system domains (e.g., if foggy then switch on the fog lamp).

The domain ontologies that we are interested in are those related to sensor or actuator data in transportation systems, meteorology, building automation, agriculture, healthcare, affective science, etc. These ontologies are implemented by domain experts and not by semantic web experts. Therefore, they are not as generic as those referenced in Linked Open Vocabularies (LOV)[1] such as the Dublin Core[2] ontology, defined for describing documents that can be used in other ontologies (e.g., describe authors, title). However, the domain ontologies provide a domain-specific knowledge that can be reused and interlinked with each other to build promising Internet of Things (IoT) applications.

According to the literature, reusing the domain knowledge already defined by domain experts is not popular, even with the new trend "Linked Data" to share data on the web. To begin with, there is a need to build a dataset referencing domain knowledge relevant for IoT. Due to several major interoperability issues, it is essential to re-design the domain knowledge in a interoperable way. Once the domain knowledge is interoperable, inferring high level abstractions from sensor measurements becomes easy and reusable. Usually, reasoning on sensor data is based on machine learning and hidden in a black box, so it cannot be easily reused by other applications. A second major need is to combine domain expertise to build cross-domain IoT applications.

The main contributions of this paper are referencing and classifying semantic-based projects relevant for IoT by designing the Linked Open Vocabularies for Internet of Things (LOV4IoT) dataset. We also design a dataset of interoperable domain rules to deduce high-level abstractions from sensor data by designing the Sensor-based Linked Open Rules (S-

[1] http://lov.okfn.org/dataset/lov/
[2] http://dublincore.org/documents/dcmi-terms/

LOR). The rules have been manually extracted from the LOV4IoT dataset. Those 2 components are integrated within the M3 framework [4].

Section II presents related works and their limitations. The M3 framework explains solutions to the limitations presented above in section III, particularly, the Linked Open Vocabularies for Internet of Things (LOV4IoT) dataset, its knowledge extraction and its interaction with other components. The implementation of the LOV4IoT and the M3 interoperable domain knowledge are explained in section IV, the use cases in section V and the evaluation in section VI. Finally, we conclude the paper in section VII.

II. Related Works and their Limitations

In this section, we demonstrate that domain knowledge relevant for IoT is constantly redefined. In addition, we clearly explain and compare the major limitations of Semantic Web of Things (SWoT) frameworks.

A. Domain Knowledge Constantly Redefined

We take the meteorology example to demonstrate it is redefined in various domains such as building automation, intelligent transportation systems or healthcare, as depicted in Figure 1. Paul Staroch [5] designs weather ontologies and rules related to `Atmospheric Pressure`, `Cloud Cover`, `Humidity`, `Precipitation`, `Solar Radiation`, `Sun Position`, `Temperature` and `Wind`. Most of these concepts and rules are redefined by other domain-specific applications. In intelligent transportation systems, Morignot et al. [6], Kannan et al. [7], Fuchs et al. [8] design ontologies and describe that `Weather` conditions, `Emotion Driver` and `Road` type have to be taken into account when driving. Weather concepts were already defined in the weather ontologies but are not reused. Ruta et al. [9] design rules that could be reused in other intelligent transportation systems, (e.g, switch on the fog lamp when it is foggy). In the healthcare domain, Hazdic et al. [10] describe relationships between the `Environmental Factors (Climate, Noise and Pollution)`, `Emotion` and explain that the `Weather` affects the emotional state. In the agricultural domain, the weather domain knowledge is still redefined. Kim et al. [11], Bansal et al. [12], Walisadeera et al. [13] and Miao et al. [14] design ontologies related to smart farms by considering the climate. Ye et al. [15] [16], Bae [17], Kim et al. [18] and Zografistou [19] redefine the same activity rules such as detecting if the person is `locatedIn` the room to automatically switch on the light, `watching TV`, `sleeping`, `washing`, etc. Kofler et al. [20], Vasileios et al. [21] and Preuveneers et al. [22] describe concepts and rules to build a smart home to turn on/off the lights according to the weather.

B. Indexing Domain Knowledge

We found various semantic web tools to index ontologies and datasets: the Linked Open Vocabularies (LOV) [23] [24]

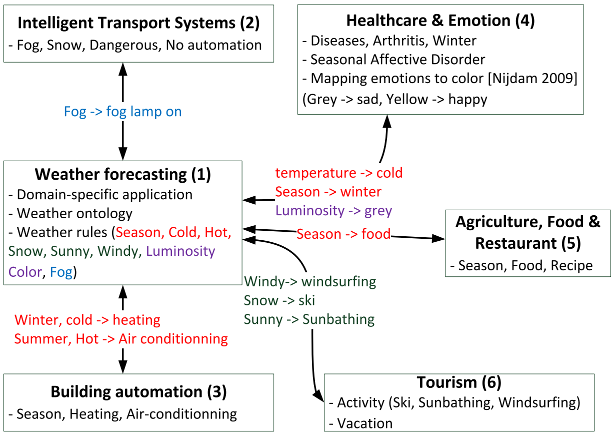

Fig. 1. Domain knowledge (ontologies, datasets, rules) can be reused in other applications.

catalogue for ontologies, the DataHub[3] project for datasets and semantic search engines such as Sindice[4], Watson[5] or Swoogle[6]. Unfortunately, since domain experts do not publish their ontologies online and do not follow semantic web guidelines, their domain ontologies or datasets are not referenced on these tools. Noy et al. [25] introduce a methodology to build ontologies and recommend to reuse the domain knowledge. This step is almost never respected by domain experts.

C. Semantic Web of Things Frameworks

Sheth et al. [26] are the first authors introducing the need to semantically annotate sensor networks, they coined the term "Semantic Sensor Web". Sheth et al. [26] and Wang et al. [27] are the first works to propose the idea to reason on semantic sensor data, but do not provide any concrete solutions that we can reuse. Henson [28] explains the idea of 'semantic perception', to interpret and reason on sensor data. His work his based on abductive logic framework and Parsimonious Covering Theory (PCT). He clearly explains that the development of background knowledge to interpret sensor data is a difficult task and out of the scope of this work. The Spitfire [1] project proposes the new concept called "Semantic Web of Things". They use ontologies for the discovery mechanisms for sensors. They are not focused on the interpretation of sensor data to generate IoT applications. Barnaghi et al. [29] introduce the need to share and integrate information across different domains to infer new knowledge. Ruta et al. [30] propose the concept of SWoT framework and introduce the need of reasoning, but do not propose the idea to reuse and interlink the domain ontologies and rules. Ganz et al. [31] propose the Knowledge Acquisition Toolkit (KAT), a machine learning-based approach to infer high-level

[3]http://datahub.io/
[4]http://datahub.io/
[5]http://watson.kmi.open.ac.uk/WatsonWUI/
[6]http://swoogle.umbc.edu/

	Reusing domain knowledge	Cross-domain (Interlinking domain knowledge)	Reasoning	Constrained Devices	Security
Sheth, Henson et al. 2008	No	No	Yes.	No (Sensor Web Enablement)	No
Wei Wang, Barnaghi et al. 2009-2013	No	No	Yes.	No (Sensor Web Enablement)	No
Henson 2008-2013	No	No	Yes ("Semantic Perception")	Yes (Parsimonious Covering Theory)	No
Ganz 2011-2014	No	No	Yes ("High level abstraction")	Yes (Sensor SAX in gateways)	No
Ruta 2010-2012	No	No	Yes	Yes (Android)	No
Spitfire 2011	No	No	No	Yes (CoAP)	No
Gyrard 2012-2014	Yes (LOV4IoT)	Yes (M3 interoperable domain knowledge)	Yes (S-LOR: Sensor-based Linked Open Rules)	Yes (works both on cloud and android-based devices)	Yes (STAC)

Fig. 2. Limitations of SWoT frameworks

abstractions of sensor data in gateways to reduce the traffic in network communications. They do not reuse the domain knowledge already designed to deduce high-level abstractions. The W3C Semantic sensor Networks (SSN) ontology [32] does not provide neither interoperability for domain concepts nor a basis for reasoning.

D. Heterogeneous Semantic Languages and Tools

Domain expert use heterogenous semantic web rules languages such as SPARQL Inferencing Notation (SPIN)[7] or Semantic Web Rule Language (SWRL). Frequently, they use the Semantic Web Rule Language (SWRL) language. However, it is really hard to combine SWRL rules with each other since they are designed with heterogeneous softwares (e.g., ontology editors and inference engines). Most importantly, the domain knowledge is thought differently. Rules are frequently defined as `owl:Restriction` in the ontology in different ways.

E. Summary

In this section, we identify five important criteria (see 2) where existing Semantic Web of Things Frameworks have shortcomings:

1) Reusing Domain Knowledge: Frequently, the domain knowledge is not reused since ontologies, datasets and rules are not shared on the web. It is essential to: (1) build a dataset referencing domain knowledge relevant for IoT, (2) encourage domain experts to share their domain knowledge online to latter reuse it, and (3) design solutions to assist developers in reusing the domain knowledge.

2) Interlinking Domain Knowledge: Existing works do not propose to reuse and interlink the domain ontologies already implemented by domain experts. The ontology, the dataset and the related rules could be reused in a similar domain (e.g., building automation) or in disparate domains (e.g., weather, eHealth, affective science, tourism) which is possible only if the cross-domain knowledge is interoperable with each other.

[7]http://spinrdf.org/

3) Reasoning: Existing reasoning methods do not provide a way to share and reuse interoperable rules to deduce high-level abstractions from sensor data. Frequently, the reasoning part is hidden in the black-box and cannot be reused. We need an innovative approach to have interoperable domain rules to easily interpret sensor data. These rules should be easy to reuse and combine.

III. The Machine-to-Machine Measurement (M3) Framework

We present in this section, the Machine-to-Machine Measurement (M3) framework as depicted in the Figure 3. It assists users in reusing the domain knowledge and interpreting sensor measurements to build IoT applications.

A. M3 architectural overview

The M3 framework (Figure 3) is composed of several layers as following. The **perception layer** handles hardware devices such as sensors, actuators and RFID tags. The **data acquisition layer** gets SenML [33] sensor data from M2M devices. Sensor data are provided by different projects which do not use the same terms to describe sensors (e.g., they are synonyms such as precipitation and rainfall sensors). We encounter the same issue with sensor measurements (e.g., 't', 'temp' and 'temperature' represent the temperature measurement). This layer converts sensor data in a unified way using standardized semantic web languages such as (RDF/XML)[8], RDFS and OWL. We designed the Machine-to-Machine Measurement (M3) ontology[9], an extension of the W3C SSN ontology [32], more precisely of the `ssn:ObservationValue` concept. The M3 ontology is the key component for interoperability. Our ontology is innovative since it unifies sensor data to provide a basis for reasoning. Further, the M3 ontology enables to interlink different domain knowledge relevant for IoT to build cross-domain applications. The **persistence layer** stores the M3 ontology, the M3 domain knowledge (e.g., ontologies, datasets and rules), semantic sensor data and inferred sensor data in a triple store (e.g., Jena TDB). A triple store is a database to store our semantic sensor data. We also store M3 SPARQL queries and M3 rules in files. The first essential dataset called Linked Open Vocabularies for IoT (LOV4IoT)[10] references and classifies more than 200 domain knowledge relevant for IoT. We design a second valuable dataset composed of IoT application templates to easily find the interoperable M3 ontologies, M3 datasets and M3 rules to generate IoT applications. Finally, the third valued Sensor-based Linked Open Rules (S-LOR) dataset [3] is composed of M3 rules compliant with M3 domain ontologies and datasets enabling domain interlinking. The **knowledge management layer** is responsible for finding, indexing, reusing and combining domain-specific knowledge (e.g., smart home, intelligent

[8]http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/Overview.html

[9]http://www.sensormeasurement.appspot.com/m3

[10]http://www.sensormeasurement.appspot.com/?p=ontologies

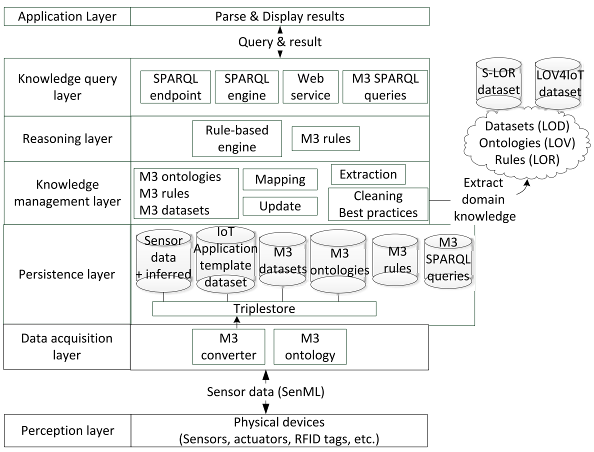

Fig. 3. The M3 conceptual architecture

transportation systems, etc.) to update the LOV4IoT and S-LOR dataset. Then, the domain knowledge is redesigned to upgrade M3 domain ontologies, M3 datasets and M3 rules. The M3 domain knowledge is interoperable and structured in the same way which eases future tasks such as reasoning on sensor data and generating IoT application templates. The **reasoning layer** infers high-level knowledge from sensor data. It uses reasoning engines (e.g., Jena rule-based inference engine) and M3 rules referenced in the S-LOR dataset. For instance, when the cloud cover is equal to 0 okta, M3 rules can deduce that it is sunny. Sunny is the result of the M3 rule and is interlinked to the M3 tourism dataset describing activities when it is sunny. The **knowledge query layer** executes SPARQL (a SQL-like language) queries on inferred sensor data. We propose web services embedding or generating M3 SPARQL queries relevant for the IoT applications to save time and make life easier to the developers. The **application layer** employs an application (e.g., running on the cloud or on smart devices) which parses and displays the results to end users. For instance, the M3 framework suggests activities according to the weather forecasting (e.g., catamaran when it is windy). STAC[11] is an other cross-domain application in the security domain. STAC suggests the most suitable security mechanisms to secure sensor data, communications (between sensors, gateways, mobile phones, cloud) or IoT applications.

B. Linked Open Vocabularies for IoT (LOV4IoT)

We pursued a deeper analysis of domain knowledge relevant for IoT, the research questions are as following. What domains do sensors use? Which ontologies exist that cover each domain? What reasoning exit that cover each domain to interpret sensor data? Is the ontology publicly accessible e.g., downloadable from a website? Which technologies or tools are used to implement the ontology or rules? Does the ontology follow the semantic web best practices?

[11]http://www.sensormeasurement.appspot.com/?p=stac

To exploit the domain knowledge expertise and facilitate IoT application development, we design the Linked Open Vocabularies for Internet of Things (LOV4IoT) catalogue. It references more than 200 ontology-based works in various domains such as health care, building automation, food, agriculture, tourism, security, transportation and smart city. We discover, identify, study and reference these works as depicted in Figure 4 since: (1) sensors and their measurements are described, (2) they can be used to design new cross-domain use cases (e.g., the naturopathy application to combine health, weather and smart kitchen), (3) the projects are based on ontologies, (4) the projects designed rule-based systems, (5) domain experts published their works in conferences, (6) they explained why they integrate semantics, (7) they describe how they evaluate ontologies, and (8) the ontology or dataset code could be used to implement our scenarios.

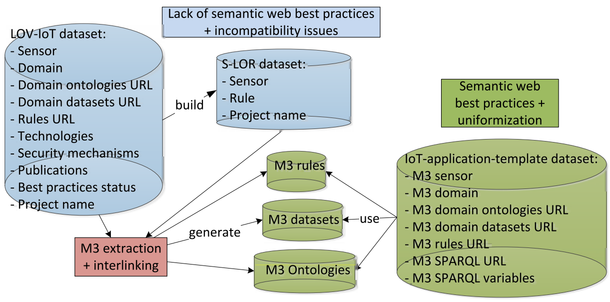

Fig. 4. The LOV4IoT dataset used to build an interoperable domain knowledge

The LOV4IoT catalogue is an essential step, to later build an interoperable domain knowledge necessary to generate IoT application templates. This catalogue is both available for machines as an Resource Description Framework (RDF) dataset[12] and for human as an HTML table accessible online[13] with the described information.

C. The M3 Interoperable Domain Knowledge

We exploit the LOV4IoT dataset by reusing domain ontologies, datasets and rules and by making them interoperable as depicted in Figure 4 and 5. Figure 4 explains the relationships between the LOV4IoT catalogue, the interoperable knowledge and the dataset composed of IoT application templates. We re-engineer the domain knowledge found in LOV4IoT to make it interoperable; an important step to combine domains with each others to build cross-domain applications.

Figure 5 explains the interoperable domain knowledge extracted from the LOV4IoT. To begin with, rules are extracted from domain ontologies because they are frequently designed as `owl:Restriction`. We re-design them to be compliant with our M3 framework and M3 ontology. They are implemented as ”IF THEN ELSE” rules.

[12]http://www.sensormeasurement.appspot.com/dataset/lov4iot-dataset
[13]http://www.sensormeasurement.appspot.com/?p=ontologies

In addition, we rewrite the domain ontologies and datasets to be compliant with our M3 framework; an essential step for domain interlinking. We distinguish ontologies, datasets and rules. The domain knowledge is upgraded according to the Linked Open Vocabularies (LOV)[14] [23] recommendations. We explain where the domain knowledge comes from (e.g., title, rights, authors, licenses). Further, we improve the M3 domain knowledge with semantic web tools such as the OOPS project[15] to detect common ontology pitfalls. If encountering errors when submitting on LOV, we check ontologies on validation tools such as Vapour[16] and TripleChecker[17] and the syntax with RDF validator. We followed the Linked Data principles[18] to create a well-designed RDF dataset.

Sharing the new domain knowledge online and reference it on semantic web tools. Suggesting M3 rules to the Linked Open Rules[19] which is still a work in progress. The domain ontologies could be suggested to the LOV catalogue and the semantic search engines such as Watson and Swoogle. The M3 domain datasets could be suggested to the Linked Open Data, the DataHub project[20] and on semantic search engines such as Sindice[21]. The name of the ontology (namespace) and the location of the ontology are the same (URI deferencable).

The last step consists in integrating ontology matching tools to automatically align the domain knowledge to infer additional knowledge and combine domains. To ease the task of ontology matching tools, we add labels and comments and rewrite the domain knowledge in a universal language (English).

Updating the M3 domain knowledge with new sensors, measurements, units or domains is simple and can seamlessly interoperate with the existing environment. M3 is an interoperable cross-domain knowledge. For instance, it enables to build a cross-domain knowledge that we called naturopathy to combine healthcare, weather, agriculture and affective sciences domains. M3 also provides STAC, a security cross-domain knowledge to combine security for sensor networks, web, network management, wireless or cellular networks. New domains such as smart city, smart home, environment and intelligent transportation systems have been easily integrated in M3.

D. *Sensor-Based Linked Open Rules (S-LOR)*

Sensor-Based Linked Open Rules (S-LOR) [3] is a dataset referencing rules to deduce high-level abstractions from sensor data. Rules are extracted from this dataset to make them interoperable and reusable. Ideally, M3 rules are defined once and reused by other systems. Figure 6 shows the deduction of high-level abstraction from sensor data (e.g., heavy rain).

[14]http://lov.okfn.org/dataset/lov/
[15]http://oeg-lia3.dia.fi.upm.es/webOOPS/index-content.jsp
[16]http://validator.linkeddata.org/vapour
[17]http://graphite.ecs.soton.ac.uk/checker/
[18]http://linkeddata.org/
[19]http://www.sensormeasurement.appspot.com/?p=rule
[20]http://datahub.io/fr/
[21]http://sindice.com/

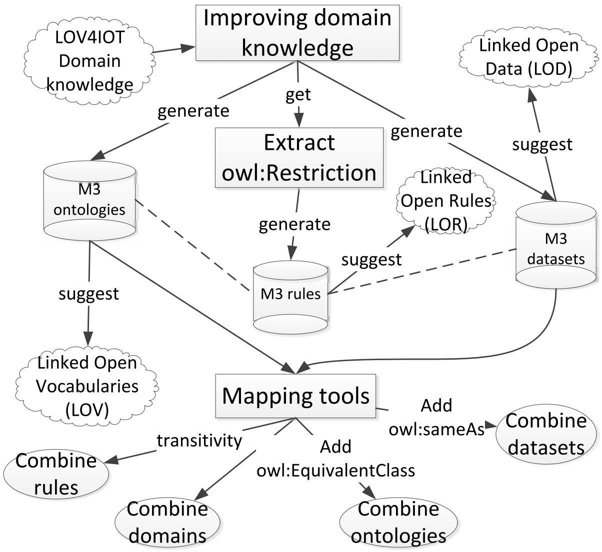

Fig. 5. Extracting and combining M3 domain knowledge

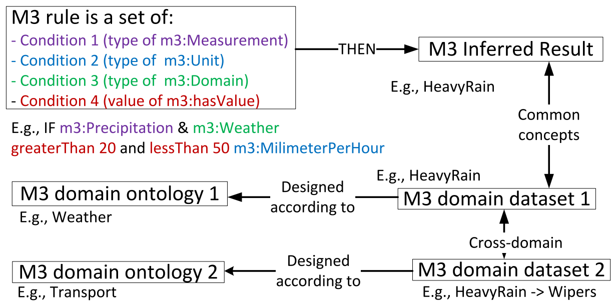

Fig. 6. M3 rules combining domain knowledge

Heavy rain is the result of the reasoning engine by taking into consideration the M3 measurements, M3 units, M3 domains and M3 values. The deduction is also defined in M3 domain datasets (e.g., weather and transportation) which enables to enrich original sensor data. M3 connects two domains in this example, tourism and transportation thanks to the common terms described in an interoperable manner in M3 rules and M3 datasets.

E. *Generating IoT Application Templates*

The M3 framework generates IoT application templates according to the sensors and domains employed by the users. This is possible thanks to the IoT application dataset (see Figure 4) reusing the M3 interoperable domain knowledge and M3 rules. For instance, the user chooses a sensor and the domain (e.g., LightSensor and Weather) and the M3 approach finds IoT application templates using the sensor and combined it with other domains. The sequence diagram is depicted in Figure 7. For the given example, the M3 framework proposes

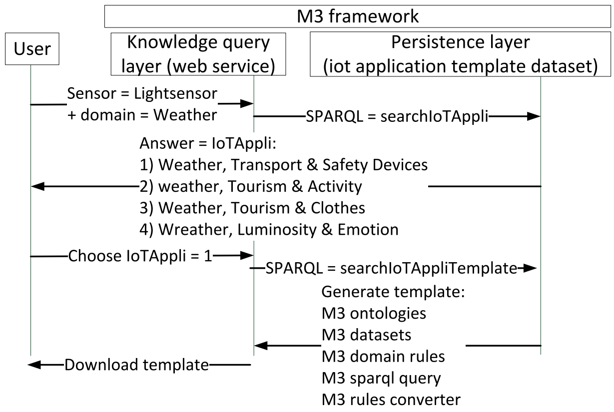

Fig. 7. Generating IoT applications

one cross-domain template "Weather, Transport & Safety Devices" to suggest safety equipments in the car according to the weather (e.g., wiper) deduced by the M3 framework. Once, the user chooses a template, the M3 framework will automatically generates the M3 domain ontologies, M3 datasets, M3 rules and M3 SPARQL queries needed to build the IoT application as depicted in Figure 7. The templates are defined in our application IoT template dataset. For each template we indicate sensor used, domains, M3 domain ontologies, datasets and rules relevant to build the IoT application template.

F. Summary: Added value of the M3 framework

We design an innovative Machine-to-Machine Measurement (M3) framework to build cross-domain IoT applications. It is composed of the following components filling the gap explained in section II-E.

1) Reusing Domain Knowledge: The Linked Open Vocabularies for Internet of Things (LOV4IoT) dataset references and classifies more than 200 ontology-based projects relevant for IoT in various domains such as building automation, eHealth, intelligent transportation system, tourism, agriculture, affective science, food, security and weather forecasting.

2) Interlinking Domain Knowledge: The M3 interoperable domain knowledge extracted from the LOV4IoT dataset enables to build cross-domain applications and suggestions.

3) Reasoning: The Sensor-based Linked Open Rules (S-LOR) is a dataset referencing rules to build interoperable M3 domain rules to deduce high-level abstractions from sensor data.

IV. IMPLEMENTATION

The M3 framework code has been released on Github[22] under GNU GPLv3 license. Numerous documentations are provided and available[23] to encourage the reusability of the M3 framework and its components. LOV4IoT, S-LOR and IoT application template datasets and the M3 interoperable domain knowledge have been designed with standardized semantic web languages such as RDF, RDFS and OWL. The S-LOR catalogue contains more than 100 M3 rules to deduce measurements provided by sensors in various domains. For instance, high-level abstraction are deduced from luminosity, temperature, wind and precipitation measurements. In healthcare, symptoms (e.g., fever) are deduced from physiological sensor values such as heart rate, blood pressure, temperature or blood glucose level. The IoT application template dataset contains more than 30 IoT templates: half of them are cross-domain applications, and the others are just to deduce high-level abstractions from sensor data without combining heterogeneous domains.

These datasets have been employed to easily build several scenarios: (1) the naturopathy application that combines meteorology, healthcare, affective science and smart home domains to suggest foods, home remedies and recipes according to the emotional state, the user's diseases, diets, allergies or the weather, (2) the tourism application to suggest clothes and activities according to the weather, and (3) the transportation application to suggest safety equipments in the car according to the weather. These cross-domain scenarios are accessible on our web site web and has been developed with the following technologies: the Jena framework to design semantic web applications, Google App Engine to publish the prototype online, Ajax to query M3 web services, HTML5, CSS3, Javascript and Bootstrap for the user interface.

V. USE CASES

The M3 framework has been employed by two uses cases: Android developers and the security domain.

A. Android-based cross-domain IoT applications

M3 is highly modular and flexible for constrained devices. M3 has been used by Android-based mobile developers to build cross-domain Semantic Web of Things applications [34] [35]. A first connection to Internet is required to download the M3 interoperable domain knowledge. Then, the entire M3 process is done locally without internet access. The mobile phone gets sensor data from a gateway (e.g., Raspberry Pi) and converts them with the M3 converter. Then, it executes the reasoning engine with M3 rules to deduce high-level abstractions from sensor data. Finally, it loads M3 cross-domain knowledge and provides M3 suggestions by executing the M3 SPARQL query. AndroJena[24], a lightweight version of Jena has been employed for the implementation.

B. STAC (Security Toolbox: Attack & Countermeasure)

Our proposed methodology for building an interoperable domain knowledge has been applied in a different context, the security domain. A cross-domain security application, called STAC (Security Toolbox: Attack & Countermeasure) has been built to assist developers in suggesting security

[22]https://github.com/gyrard/M3Framework
[23]http://sensormeasurement.appspot.com/?p=documentation
[24]http://code.google.com/p/androjena/

mechanisms to secure their applications.[36].STAC is a cross-domain application because it combines security in various domains such as cellular (2G, 3G, 4G), wireless (Bluetooth, Wi-Fi, Wimax), sensor networks, web applications or network management.

The STAC application is based on an interoperable security cross-domain security knowledge base which has been re-engineering following the methodology explained in this paper. The knowledge base comprises an ontology and a dataset describing relationships between security mechanisms, attacks, security properties (e.g., authentication), OSI model, advantages and drawbacks of technologies.

For instance, the developer chooses a technology (e.g., WiFi). Then all related attacks (e.g., Steal NIC) and security mechanisms specific to the WiFi technology are displayed. The developer can choose a security mechanism (e.g., WPA1) to obtain additional information: the security property satisfied (e.g., authentication) and the advantages/disadvantages (e.g., deprecated). The STAC application is available on the web[25].

VI. Evaluation

Ontologies have been classified in the LOV4IoT catalogue according to their domain and status. Figure 8 shows that we have encouraged domain experts to share their domain knowledge online (ontology, dataset and rule), and some of them have improved their ontology according to the semantic web best practices. Although getting domain knowledge is still ongoing, about 80 ontologies are online including 13 already referenced by LOV. Sixty percent of the domain knowledge: (1) is shared online, (2) is shared online and the semantic web best practices are followed, so the LOV catalogue referenced the ontology, (3) is a work in progress, the ontology will be published online soon, or (4) cannot be shared online since the ontology is lost, confidential or the project is finished. Concerning the forty percent remaining, we do not know yet if we can obtain the domain knowledge.

We have published our M3 and STAC domain knowledge online according to the semantic web best practices to facilitate knowledge sharing and reuse. The STAC ontology and dataset has been extracted from security ontologies designed by security experts. The STAC knowledge base has been validated by the semantic web community, since it is now referenced by LOV. Further, we evaluated the M3 and STAC interoperable domain knowledge with tools such as Triple-Checker[26], RDF validator[27], Oops[28] and used the reasoner Hermit and Fact++. We fixed as far as possible errors and took into considerations recommendations provided by these tools.

By evaluating this work, we learnt and summarized a set of semantic web best practices by referencing and classifying a set of tools and guidance to use them to fix errors [37]. Further, we have shared lessons learned on our web site, more than 91

M3 or SenML Domain name	#Total number ontologies	# No answer	# ontology online	# ontology lost	# ontology in development	# referenced by LOV
BuildingAutomation	33	10	8	4	7	3
Health	35	11	12	7	5	0
Weather	9	1	6	0	0	2
Agriculture	8	5	2	1	0	0
Environment	7	4	3	0	0	0
Emotion	6	2	2	0	0	2
Transport	28	12	5	5	5	1
Tourism	27	11	9	4	2	1
City	2	1	2	0	0	0
Security	24	7	10	1	2	4
Food, Beverage, Restaurant	23	10	9	0	3	1
Total	202	78	69	20	25	13

Fig. 8. Ontology Status

countries and 8,267 persons have visited our web site since December 2013.

A software performance evaluation has been performed with a dataset of 8 KB of sensor data. The reasoning time takes less 30 milliseconds even with 50 rules to load. Due to the modularity and the flexibility of M3, performance of the semantic-based reasoning is good. Thanks to the IoT application template dataset, we do not load the entire M3 domain knowledge, but a subset related to sensors required in the application. Further evaluations are considered as future works.

VII. Conclusion and Future Work

Domain knowledge (ontologies, datasets and rules) should be shared and reused rather than constantly redefined in numerous domains. This will enable to build the Semantic Web of Things to easily create new cross-domain IoT applications. We presented in this paper the M3 framework composed of: (1) the LOV4IoT dataset referencing more than 200 ontology-based projects relevant for IoT, (3) the M3 interoperable domain knowledge to provide cross-domain applications, (4) IoT application templates to assist developers in designing semantic-based IoT applications, (5) S-LOR to share and reuse interoperable M3 domain rules to deduce high-level abstractions from sensor data.

We synthesize lessons learned and propose the guidelines to publish, share, interlink and reuse the domain knowledge (ontology, dataset, rule) according to the semantic web best practices. We disseminate our work to ETSI M2M, oneM2M, W3C Web of Things, W3C SSN ontology and domain experts.

As future work, we envision to automatically extract the domain knowledge and redesign it to enrich the M3 interoperable domain knowledge. To achieve this task, there is a need to build a tool to clean and improve existing ontologies. Another tool is necessary to extract rules to enrich S-LOR. We also need to upgrade current ontology matching tools to automatically generate the cross-domain knowledge. Last but not least, a thorough evaluation would be required to improve the quality of the M3 framework.

Acknowledgment

The authors would like to thank the semantic web experts (Bernard Vatant, Ghislain Atemezing, Raphael Troncy, Fabien

[25]http://www.sensormeasurement.appspot.com/?p=stac
[26]http://graphite.ecs.soton.ac.uk/checker/
[27]http://www.w3.org/RDF/Validator/
[28]http://oeg-lia3.dia.fi.upm.es/oops/index-content.jsp

Gandon and Payam Barnaghi) for their valuable feedback and expertise and colleagues, friends and students for assisting in the implementation. This work is supported by the Com4Innov Platform of Pole SCS[29] and FIESTA-IoT[30] from the European Horizon 2020 Programme with the Grant Agreement No. CNECT-ICT-643943..

REFERENCES

[1] D. Pfisterer, K. Romer, D. Bimschas, O. Kleine, R. Mietz, C. Truong, H. Hasemann, A. Kroller, M. Pagel, M. Hauswirth *et al.*, "Spitfire: toward a semantic web of things," *Communications Magazine, IEEE*, vol. 49, no. 11, pp. 40–48, 2011.

[2] A. Gyrard, C. Bonnet, and K. Boudaoud, "Enrich machine-to-machine data with semantic web technologies for cross-domain applications," in *WF-IOT 2014, World Forum on Internet of Things, 6-8 March 2014, Seoul, Korea*, Seoul, KOREA, REPUBLIC OF, 03 2014.

[3] A. Gyrard, C. Bonnet, and K. Boudaoud, "Helping IoT application developers with sensor-based linked open rules," in *SSN 2014, 7th International Workshop on Semantic Sensor Networks in conjunction with the 13th International Semantic Web Conference (ISWC 2014), 19-23 October 2014, Riva Del Garda, Italy*, 10 2014.

[4] A. Gyrard, S. K. Datta, C. Bonnet, and K. Boudaoud, "Cross-domain Internet of Things application development: M3 framework and evaluation," in *FICLOUD 2015, 3rd International Conference on Future Internet of Things and Cloud, August 24-26, 2015, Rome, Italy*, Rome, ITALY, 08 2015.

[5] P. Staroch, "A weather ontology for predictive control in smart homes," Master's thesis, 2013.

[6] E. Pollard, P. Morignot, and F. Nashashibi, "An Ontology-based Model to Determine the Automation Level of an Automated Vehicle for Co-Driving," in *16th International Conference on Information Fusion*, Istanbul, Turquie, Jul. 2013. [Online]. Available: http://hal.inria.fr/hal-00838680

[7] S. Kannan, A. Thangavelu, and R. Kalivaradhan, "An intelligent driver assistance system (i-das) for vehicle safety modelling using ontology approach," *International Journal of UbiComp*, vol. 1, no. 3, pp. 15–29, 2010.

[8] S. Fuchs, S. Rass, and K. Kyamakya, "Integration of ontological scene representation and logic-based reasoning for context-aware driver assistance systems," *Electronic Communications of the EASST*, vol. 11, 2008.

[9] M. Ruta, F. Scioscia, F. Gramegna, and E. Di Sciascio, "A mobile knowledge-based system for on-board diagnostics and car driving assistance," in *UBICOMM 2010, The Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2010, pp. 91–96.

[10] M. Hadzic, M. Chen, and T. S. Dillon, "Towards the mental health ontology," in *Bioinformatics and Biomedicine, 2008. BIBM'08. IEEE International Conference on*. IEEE, 2008, pp. 284–288.

[11] T. Kim, N. Bae, M. Lee, C. Shin, J. Park, and Y. Cho, "A study of an agricultural ontology model for an intelligent service in a vertical farm," *International Journal of Smart Homes*, vol. 7, no. 4, 2013.

[12] N. Bansal and S. K. Malik, "A framework for agriculture ontology development in semantic web," in *Communication Systems and Network Technologies (CSNT), 2011 International Conference on*. IEEE, 2011, pp. 283–286.

[13] A. I. Walisadeera, G. N. Wikramanayake, and A. Ginige, "An ontological approach to meet information needs of farmers in sri lanka," in *Computational Science and Its Applications–ICCSA 2013*. Springer, 2013, pp. 228–240.

[14] L. Miao, J. Chun, and J. C. Yoshiyuki HIGUCHI, "Ontology-based user preferences bayesian model for personalized recommendation," *Journal of Computational Information Systems*, vol. 9, no. 16, pp. 6579–6586, 2013.

[15] J. Ye, G. Stevenson, and S. Dobson, "A top-level ontology for smart environments," *Pervasive and mobile computing*, vol. 7, no. 3, pp. 359–378, 2011.

[16] L. Coyle, S. Neely, G. Stevenson, M. Sullivan, S. Dobson, P. Nixon, and G. Rey, "Sensor fusion-based middleware for smart homes," *International Journal of Assistive Robotics and Mechatronics*, vol. 8, no. 2, pp. 53–60, 2007.

[17] I.-H. Bae, "An ontology-based approach to adl recognition in smart homes," *Future Generation Computer Systems*, 2013.

[18] E. Kim and J. Choi, "An ontology-based context model in a smart home," in *Computational Science and Its Applications-ICCSA 2006*. Springer, 2006, pp. 11–20.

[19] D. Zografistou, "Support for context-aware healthcare in ambient assisted living," Master's thesis, 2012.

[20] C. Reinisch, M. Kofler, F. Iglesias, and W. Kastner, "Thinkhome energy efficiency in future smart homes," *EURASIP Journal on Embedded Systems*, vol. 2011, p. 1, 2011.

[21] E. Vasileios and G. Antoniou, "A real-time semantics-aware activity recognition system," 2012.

[22] D. Preuveneers, J. Van den Bergh, D. Wagelaar, A. Georges, P. Rigole, T. Clerckx, Y. Berbers, K. Coninx, V. Jonckers, and K. De Bosschere, "Towards an extensible context ontology for ambient intelligence," in *Ambient intelligence*. Springer, 2004, pp. 148–159.

[23] P.-Y. Vandenbussche and B. Vatant, "Metadata recommendations for linked open data vocabularies," *Version*, vol. 1, pp. 2011–12, 2011.

[24] F. Scharffe, G. Atemezing, R. Troncy, F. Gandon, S. Villata, B. Bucher, F. Hamdi, L. Bihanic, G. Képéklian, F. Cotton *et al.*, "Enabling linked-data publication with the datalift platform," in *Proc. AAAI workshop on semantic cities*, 2012.

[25] N. F. Noy, D. L. McGuinness *et al.*, "Ontology development 101: A guide to creating your first ontology," 2001.

[26] A. Sheth, C. Henson, and S. Sahoo, "Semantic sensor web," *Internet Computing, IEEE*, vol. 12, no. 4, pp. 78–83, 2008.

[27] W. Wei and P. Barnaghi, "Semantic annotation and reasoning for sensor data," *Smart Sensing and Context*, pp. 66–76, 2009.

[28] C. A. Henson, "A semantics-based approach to machine perception," Ph.D. dissertation, Wright State University, 2013.

[29] P. Barnaghi, W. Wang, C. Henson, and K. Taylor, "Semantics for the internet of things: early progress and back to the future," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 8, no. 1, pp. 1–21, 2012.

[30] M. Ruta, F. Scioscia, and E. Di Sciascio, "Enabling the semantic web of things: Framework and architecture." in *ICSC*, 2012, pp. 345–347.

[31] F. Ganz, P. Barnaghi, and F. Carrez, "Automated semantic knowledge acquisition from sensor data," 2014.

[32] M. Compton, P. Barnaghi, L. Bermudez, R. Garcia-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog *et al.*, "The ssn ontology of the w3c semantic sensor network incubator group," *Web Semantics: Science, Services and Agents on the World Wide Web*, 2012, http://www.w3.org/2005/Incubator/ssn/ssnx/ssn.

[33] C. Jennings, J. Arkko, and Z. Shelby, "Media types for sensor markup language (senml)," 2012.

[34] S. K. Datta, A. Gyrard, C. Bonnet, and K. Boudaoud, "One M2M architecture based user centric IoT application development," in *FICLOUD 2015, 3rd International Conference on Future Internet of Things and Cloud, August 24-26, 2015, Rome, Italy*, Rome, ITALY, 08 2015. [Online]. Available: http://www.eurecom.fr/publication/4600

[35] S. K. Datta, C. Bonnet, A. Gyrard, R. P. Ferreira da Costa, and K. Boudaoud, "Applying Internet of Things for personalized healthcare in smart homes," in *WOCC, 24th Wireless and Optical Communication Conference, October 23-24, 2015, Taipei, Taiwan*, Taipei, TAIWAN, PROVINCE OF CHINA, 10 2015. [Online]. Available: http://www.eurecom.fr/publication/4627

[36] A. Gyrard, C. Bonnet, K. Boudaoud *et al.*, "An ontology-based approach for helping to secure the etsi machine-to-machine architecture," *IEEE International Conference on Internet of Things 2014 (iThings)*, 2014.

[37] A. Gyrard, M. Serrano, and G. Atemezing, "Semantic web methodologies, best practices and ontology engineering applied to internet of things," in *WF-IOT 2015, World Forum on Internet of Things, 14-16 December 2015, Milan, Italy*, 2015.

[29] http://www.pole-scs.org/

[30] http://www.fiesta-iot.eu/