Smart Cities Semantics and Data Models

Antonio J. Jara[1(✉)], Martin Serrano[2], Andrea Gómez[4], David Fernández[4], Germán Molina[3], Yann Bocchi[1], and Ramon Alcarria[4]

[1] Mobility Lab, University of Applied Sciences Western Switzerland (HES-SO), Technopole. 3, 3960 Sierre, Switzerland
{antonio.jara,yann.bocchi}@hevs.ch
[2] Insight Research Centre, National University of Ireland Galway (NUIG), Galway, Ireland
martin.serrano@insight-centre.com
[3] HOP Ubiquitous S.L. (HOPU), Luis Buñuel. 6, Ceutí, 30562 Murcia, Spain
german@hopu.eu
[4] Escuela Técnica Superior de Ingeniería en Telefomunicaciones, Universidad Politécnica de Madrid (UPM), Madrid, Spain
andrea@hopu.eu, ralcaria@upm.es

Abstract. Data models and semantics are a key aspect for the valorization of data in cross-domain applications and to obtain knowledge/insights beyond the original applications (vertical use cases). An important role of Big Data and a key fundament of its success is this capacity to discover and extract new knowledge beyond the original use of data, in order to learn, optimize processes and understand the hidden rules of our world. This works presents the different data models from standardization bodies such as IEEE PAR2530, ITU-T FG DPM, ETSI ISG CIM and oneM2M, W3C SSN, OMA LwM2M etc. An analysis and comparative among all of them and also the opportunities to link them in order to guarantee that we can obtain the major value through co-operation among cities and different departments. This work is contextualized in the principles from the Open and Agile Smart Cites (OASC) and linked initiatives focused on data management cross-cities and large scale pilots.

Keywords: Smart cities · Data models · Internet of things
Semantics · ETSI ISG CIM · ITU-T · oneM2M · FIWARE
Open and agile smart cities · OASC

1 Introduction

Smart Cities present an opportunity for valorizing digitalized services and Internet of Things (IoT) infrastructure with the creation of disruptive and innovative services and solutions. These innovations through exploitation of data correlation, open data and data analytics will deal with cities challenges such as environmental sustainability, citizens' engagement, economic growth and citizens' mobility. This work carries out an analysis and discussion around the standardization and data modelling actions that are working towards defining a common semantic descriptions for smart city data to enable cross-domain and advanced services development in smart cities.

Á. Rocha and T. Guarda (eds.), *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)*, Advances in Intelligent Systems and Computing 721,
https://doi.org/10.1007/978-3-319-73450-7_8

For this purpose, we have identified the main activities from IEEE, ETSI, ITU-T, W3C and OMA present examples of data that can be collected from cities, discuss issues around this data and put forward some preliminary thoughts for creating a semantic description model [1] to describe and help discover, index and query smart city data.

The Table 1 presents a summary of the key datasets that we can find in a Smart City based on the datasets from CityPulse [2] and FIESTA-IoT [3].

Table 1. Different types of data models in a Smart City.

Category	Description	Publisher (Data owner)	Data access
Mobility	City Maps (Roads, street names, Points of Interests (POIs), stations, etc.)	Government, Private Companies and Organizations	OpenStreetMap (Open Data), APIs accessible (Google Maps) and proprietary/closed GIS such as ESRI
	Public transport schedules	Government and Private Companies	Open Data or Web-accessible content
	Transport authority updates and events (road status, road works…)	Government and Private Companies	Waze (APIs accessible), Web-accessible data and Open Data
	Car Parks/Parking meters	Government and Private Companies	APIs accessible and Open Data
	Traffic (Number of vehicles passing between two points and speed)	Government and Private Companies	Private data, APIs accessible and Open Data
	Opportunistic and Crowd monitoring	Government, Private Companies and Organizations	Nomadic Sensing (NOSE)
Environmental data	Air Quality (Air Quality Index - AQI, Particles concentration and gases concentration)	Government	Private data, APIs accessible and Open Data
	Temperature/Humidity	Government	Private data, APIs accessible and Open Data
	Noise level/Noise maps	Government	Private data, APIs accessible and Open Data
Tourism	Points of Interest, Time schedule, events, museums, bar/discos etc.	Cultural Groups	APIs and Web available data
Demographic	Population, Crowd Monitoring, Ages, Cultural Level, Economical Level	Government and organizations	Open Data
Co-creation (Citizens engagement)	Opinions, decisions, surveys, etc.	Government and organizations	Siidi (Organicity)

2 Smart Cities Semantics and Data Models

This section introduces different available data models used in the majority of smart cities deployments and European actions such as FIESTA-IoT, Synchronicity, Organicity, CityPulse etc.

2.1 M3-Lite

The M3-lite is a taxonomy that enables testbeds to semantically annotate the IoT data produced by heterogeneous devices and store them in a federated datastore such as FIESTA-IoT. In this taxonomy, we classify devices, the domain of interests (health, smart home, smart kitchen, environmental monitoring, etc.), phenomena and unit of measurements.

2.2 W3C SSN

W3C SSN describes not only sensor device capabilities, but also organises the sensors into systems and describes processes that model sensor operations and can work across multiple domains. The goal is to correlate measurement data with capabilities of sensors (and sensor systems), however the descriptions about observation and measurement data are generic and cannot be used to annotate the data with domain knowledge - specific to applications. Therefore, SSN by itself cannot be used to describe smart city services (scenarios) in detail, as each service has its own quality requirements, relies on its own set of sensors, has different demands on data ownership (security, privacy concerns) etc.

Previous research has suggested building a linked-data approach for stream annotation [4]. According to this approach, external domain knowledge about the data can be provided on request - and can be specific per service rendered (e.g. quality description, sensor capabilities, etc.). The model proposed in [5] describes some basic, common attributes on the data stream but delegates details about the specific streams to other models (linked-data models).

2.3 oneM2M Ontology SAREF Ontology

The oneM2M base ontology and the SAREF ontology are described in the oneM2M TS00122 and ETSI TS1032643. OneM2M base ontology aims to provide a high level ontology for the IoT market in order to provide a minimal set of common knowledge that enables the cross-domain syntactic and semantic interoperability. oneM2M ontology is very abstract and general, thereby oneM2M expects external ontologies that describe a specific domain of interest in a more detailed way to be mapped to the oneM2M base ontology. Additionally, oneM2M defines how to internetwork between devices and things from different domains is enabled.

For the specific domains, it is where ETSI ISG Context Information Management, ETSI ISG Smart Cities data models, and other widely accepted ontologies such as SAREF plays a crucial role. These ontologies are described in the coming two sections. As an example of this base ontology and the specific domains one, we can explore the following Figure, where oneM2M plays a key role as an interconnection with telecommunications infrastructure, and other information systems. However, ETSI ISG CIM offers all the extensions for the Context management required to develop Smart Apps and provide contextualized Open Data (Fig. 1).

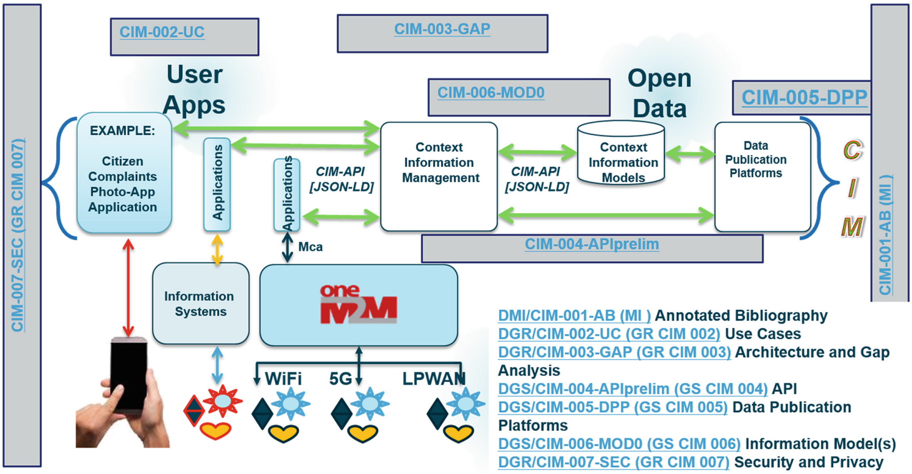

Fig. 1. ETSI ISG CIM as an example of extension and specialization for oneM2M base ontology towards the User Applications and re-use of data, which extends to oneM2M which is mainly focused on the interconnection among platforms and devices.

2.4 SAREF Ontology

SAREF aims to provide a common knowledge for the domain of Smart Appliance, especially on the energy consuming aspect. Compared to oneM2M base ontology, it is less high level and more applicable to describe devices. The mapping between oneM2M base ontology and SAREF is performed by oneM2M.

2.5 FIWARE and ETSI ISG CIM

ETSI ISG Context Information Management (ETSI ISG CIM). FIWARE initiative and platform through the Orion Context Broker component identified a key market need for IoT and Smart Cities; it is the management of context in a scalable and standardized way. For this purpose, FIWARE defined on the one hand, OMA NGSI interfaces to offer a homogeneous access to data, and on the other hand, a set of data models being standardized by ETSI ISG CIM.

Context Information provides the meta-data structure for sensors measurement and also other data feeds from video, social media etc. Even when context is very simple to understand by human being, in order to provide artificial intelligence capabilities to smart systems, it is crucial to formalize and provide much more details about the context and make it available in conjunction with the data. A Context Information

Management (CIM) system acts as a clearing-house for publishing, discovering, monitoring and maintaining data according to relevant contexts for smart applications.

"ETSI ISG CIM will specify protocols running on top of IoT platforms and allowing exchange of data together with its context, this includes what is described by the data, what was measured, when, where, by what, the time of validity, ownership, and others. That will dramatically extend the interoperability of applications, helping smart cities to integrate their existing services and enable new third-party services", as stated by the ETSI ISG CIM convenor, Lindsay Frost. ETSI ISG CIM has been focused on developing specifications for a common context information management API, data publication platforms and standard data models. A practical example is presented in the coming Fig. 2.

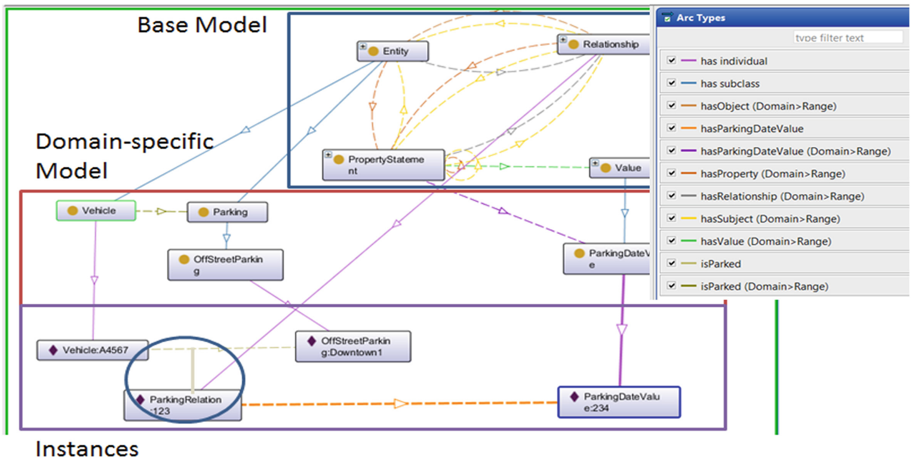

Fig. 2. ETSI ISG CIM domain-specific domain mapping of the base model [6, 7].

2.6 OMA LwM2M and IPSO Smart Objects

IPSO Application Framework from the IPSO Alliance defines in collaboration with OMA LwM2M, a set of RESTFul interfaces for the definition and management of resource lists, batch, sensors, parameter, actuators and binding tables of resources. For example, the semantic IPSO Application Framework has chosen SenML over JSON with the usage of the Unified Code for Units of Measure (UCUM). These initial semantic capabilities allow avoiding the initial mistakes from CoAP such as the use of inappropriate unit codes such as 23 °C for temperature, when it is according to the UCUM standard means velocity of light, and consequently this should be 23 CEL.

The current semantic capabilities from IPSO Application Framework are very basic in order to offer a very simple and lightweight solution. In details, the protocol over which IPSO Smart Object are defined is OMA LwM2M.

OMA LwM2M Device Management is a protocol for device management, the use this protocol in M2M requires efficient message formats and transport replacement such as CoAP, and Core Link Format. For that reason, Lightweight OMA DM has chosen CoAP to provide the core functionalities of HTTP (GET, PUT, POST, DELETE commands) in a reduced footprint.

OMA LwM2M offers key functionalities for the device management such as remote firmware upgrade, remote diagnostics, information reporting (read data, write data and subscribe to events) and it focuses on providing mechanisms for asynchronous and synchronous communication, store, forward and caching mechanism for optimizing the communication, and security with mechanisms to provide two way authentication and secure communication channels.

A detailed example of OMA LwM2M objects are presented in the references [8, 9] for the Smart Spot device, where several sensors focused on air quality, noise, temperature, humidity, crowd monitoring and connectivity management are properly modelled.

3 Analysis and Discussion

The main challenge that arise for the IoT is to make a proper usage and exploitation of the IoT potential to build more powerful applications and services.

The support for heterogeneous and legacy devices integration is being integrated thanks to the IoT context brokers and middlewares that enable the interfacing of heterogeneous protocols through a homogenized and harmonized interface. Additionally, these entities such as context brokers are enabling the capacity to integrate more details about context that facilitates the exploitation of the data and content provided by the sensors with knowledge engineering technologies.

For that reason, the current steps for the IoT are focused on the importance of metadata to build intelligent solutions; and there is where emerging context-aware systems such as the proposed by ETSI ISG CIM will play a key role, and where the platforms such as FIWARE via the Orion Context Broker and the several implementations of oneM2M such as OpenMTC and OM2M are supporting this data integration and data brokering as the core element.

A very relevant feature is the alignment that is happening between ETSI ISG CIM with ETSI oneM2M, at the same time that ETSI oneM2M with OMA LwM2M, and the IPSO Smart Objects and OMA LwM2M.

Thereby, we are finding a proper ecosystem definition where oneM2M, FIWARE, ETSI ISG CIM and OMA LwM2M are co-existing and playing a clear complementary role.

In details, OMA LwM2M is supported by the oneM2M, which provides an international initiative that will play a very relevant role to propose the standards for the syntactic and semantic information.

oneM2M defines the abstraction layers, using the same format. This will ease the creation of the higher-layers for the IoT and M2M that enables a high-level modeling of real world entities, development of applications, and finally huge quantities of data collection. oneM2M will also offer support and solutions to facilitate the development of vertical industries and new markets.

However, oneM2M has play a very clever and smart approach relying on ETSI ISG CIM and FIWARE for the development and interfacing of the Smart Applications and data exploitation, i.e., the interconnection with mobile and Web Apps developers.

At the same way, oneM2M is also relying on OMA LwM2M for the devices management and a more native interconnection with the devices addressing the physical features, sensors configuration, fine tuning, and calibration etc. as required several times for the proper use of the sensors, its configuration, and maintenance.

oneM2M is extending and coordinating with other institutions such as HGI, Broadband Forum, OSGi, Continua Alliance for healthcare devices, ZigBee Alliance for smart metering devices etc. Therefore, they are presenting a very inclusive approach.

The following Figure presents our vision of oneM2M integration for the physical devices with a support for OMA LwM2M for the devices management, in conjunction with FIWARE with a support for ETSI ISG CIM for the Apps interfacing, and also support for Open Data (CKAN) and external applications (Fig. 3).

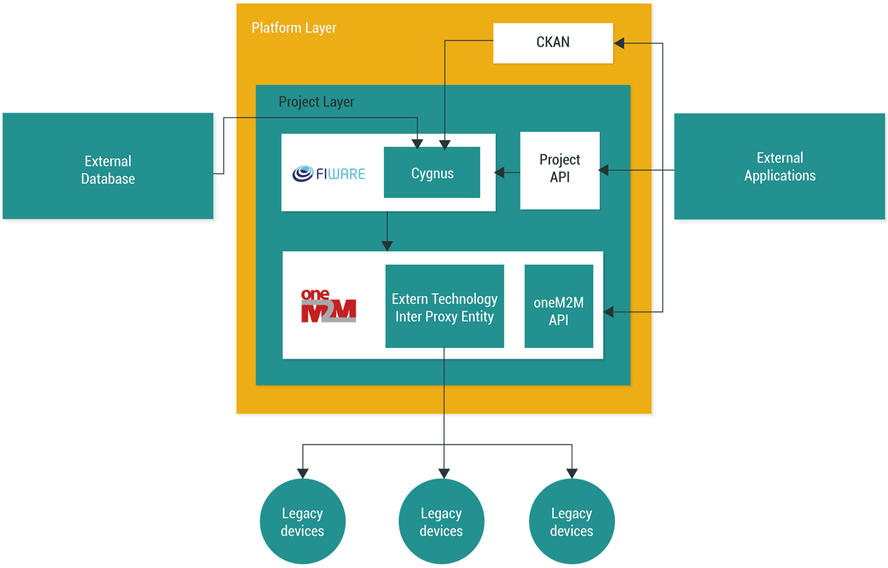

Fig. 3. FIWARE and oneM2M cooperation in conjunction with ETSI ISG CIM and OMA LwM2M.

4 Conclusions

The market is moving from vertical solutions where the sensors are stove-piped (one device per application) to specific platforms for its application in pre-defined use cases towards a more open market, where the sensors will be re-used, shared and accessed by a wide range of different applications.

The pending challenges cover the development of tools and protocols for dynamic interoperability, semantic discovery reasoners, mechanisms to re-adapt devices in case of change of context, ontologies repository, and in general toolkits that allows the semantic integration and exploitation from the IoT.

Thereby, the semantics will be managed through the different phases inside of a use case, their heterogeneous devices integrated, and they will be defined interfaces among the different components involved. In addition, these interfaces are based on very agile and flexible technologies such as RESTFul architecture, which supports a resource-oriented solution, simplifying and optimizing resource manipulations for a broad range of devices and solutions that enabled a quick and efficient application development.

The market is also moving from proprietary and complex protocols to open approaches such as HTTP, OMA LwM2M, CoAP (IPSO), oneM2M (OpenMTC and OM2M), and MQTT (Eclipse Foundation).

The current status and evolution of the IoT in the Smart Cities market is driven by the semantic and context-aware data models. The data exploitation of the data is the next step after of the provisioning of architectures and solutions with context-awareness. Interoperability and semantic-annotated models definitely increase the re-usability of the IoT resources outside the use cases and scope in which they were originally deployed and designed. This is a key need for Smart Cities use-cases and emerging IoT markets focused on large scale pilots. An example of data opportunities is Synchronicity http://synchronicity-iot.eu, a data marketplace for Smart Cities.

Acknowledgments. This work is sponsored by HES-SO University. This work has been carried out in collaboration with the National University of Ireland Galway (NUIG) in the framework of the European Project FIESTA-IoT (Grant Agreement CNECT-ICT-643943) and the experiment FINETUNE; a the same time this work has included results carried out in collaboration with SmartSDK (Grant Agreement 723174) and SynchroniCity (Grant Agreement 732240). Finally, Andrea Gómez collaboration is supported as part of the scholarship (Industrial Doctorate) by UCAM.

References

1. Bischof, S., Karapantelakis, A., Nechifor, C., Sheth, A.P., Mileo, A., Barnaghi, P.: Semantic Modelling of Smart City Data. Kno.e.sis Publications. The Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis) (2014). http://corescholar.libraries.wright.edu/knoesis/572
2. CityPulse: EU FP7 CityPulse Project, Grant No. 603095. http://www.ict-citypulse.eu
3. FIESTA-IoT: EU H2020 FIESTA-IoT project “Federated Interoperable Semantic IoT/cloud Testbeds and Applications”, Grant No. 643943 (2017)
4. Barnaghi, P., Wang, W., Dong, L., Wang, C.: A linked-data model for semantic sensor streams. In: 2013 IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on Green Computing and Communications (GreenCom), and IEEE Cyber, Physical and Social Computing, pp. 468–475 (2013)
5. Compton, M., Barnaghi, P., Bermudez, L., Garcıa-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K., Kelsey, W.D., Phuoc, D. L., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K.: The SSN ontology of the W3C semantic sensor network incubator group. Web Semant. Sci. Serv. Agents World Wide Web **17**, 25–32 (2012)

6. ETSI ISG CIM: Context Information Management. https://portal.etsi.org/tb.aspx?tbid=854&SubTB=854
7. Jara, A.J., et al.: An analysis of context-aware data models for smart cities: towards fiware and etsi cim emerging data model. Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **42** (2017)
8. HOP Ubiquitous: Homard overview and OMA LwM2M (2017). https://storage.googleapis.com/softfire/homard%20specifications.pdf
9. HOP Ubiquitous: Smart Spot Firmware (2017). https://storage.googleapis.com/softfire/smart-spot-specifications%20firmware-063d.pdf