New Facets of Semantic Interoperability: Adding JSON – JSON-LD Transformation Functionality to the BIG IoT API

Yulia Svetashova (0000-0003-1807-107X)[1], Stefan Schmid[1], and York Sure-Vetter[2]

[1]Robert Bosch GmbH, Corporate Research
Robert-Bosch-Campus 1, 71272 Renningen, Germany
{yulia.svetashova,stefan.schmid}@de.bosch.com
[2]Karlsruhe Institute of Technology, AIFB
Kaiserstr. 89, 76133 Karlsruhe, Germany
york.sure-vetter@kit.edu

Abstract. The BIG IoT project focuses on easy integration of the IoT data and services offered by existing IoT platforms and services based on semantic technologies. To enable IoT applications to consume data and services provided by heterogeneous systems and different stakeholders, a common set of ontologies and an RDF triple store with querying functionalities are used in the service discovery phase. The subsequent communication between such applications or services (as consumer) and the provider can be optimized as well. The process of automatically transforming JSON-serialized responses from an IoT data provider to linked data format is discussed and the benefits of this automation are explored in this poster contribution.

Keywords: Internet of Things · Semantic Interoperability · Linked data

1 Introduction

The Internet of Things (IoT) is rapidly growing both in terms of the number of entities (sensors, actuators, etc.) placed onto networks and the number of applications and services which operate them, collect sensor data and offer smart solutions. The growth of the IoT and the corresponding market demands lead to an increased complexity of interactions among the various IoT stakeholders, including *thing* manufacturers, platform providers, service and application developers. To unlock the full potential of the connected world, they need to consider the multi-domain nature of the IoT, heterogeneity of the protocols and data formats used, competition and/or co-operation with other vendors and new players entering the market. Thus, interoperability defined as "the ability of two or more systems or components to exchange data and use information" [1: 9] remains one of the key problems to solve in the field of IoT [2-4].

The idea of IoT ecosystems and ad-hoc collaboration among different stakeholders is gaining popularity quickly to overcome the shortcomings of existing proprietary vertical systems created for distinct product lines. Capable of reciprocal operation with others, elements of such cross-platform and cross-domain ecosystems must utilize some kind of shared semantics, i.e. either follow standardized semantic models or through use of "on-the-fly" concept alignment mechanisms. Sharing vocabularies and mapping varying concepts to selected reference ontologies have been proposed as semantic interoperability solutions long ago (see, for example, [7]). Semantic Web technologies however go far beyond simple platform and service integration and acquire new facets of usage in the context of IoT.

2 Semantic Technologies in the BIG IoT Project

The goal of the BIG IoT ("Bridging the Interoperability Gap of the Internet of Things", see more in [5]) project is to design and develop ecosystems for IoT stakeholders (1) who want to monetize their IoT resources (*providers*) and (2) need IoT resources for their applications (*consumers*). With its main focus on easy integration of existing platforms and run-time integration of resources, the project provides (1) a meeting point for providers and consumers – a *marketplace* which enables advertising and discovery of IoT offerings, gives uniform access to the distributed repositories and services, and billing & charging functionality, as well as (2) a common Web API and open source SDK.

One of the core concepts in the BIG IoT is the concept of *offering* representing a set of IoT resources: sensor data, information or services. Providers register their offerings on the marketplace using semantically annotated *offering descriptions* which are stored into an RDF triple store. Offering descriptions are grounded on semantic models common for all triples in a store: the semantics of each offering follows the BIG IoT Semantic Core Model (e.g. defining the offering type, license, price, endpoint URI, protocol, input data needed to use its functionalities and output data returned when an offering is consumed) enriched with domain independent and domain dependent semantic annotation.

Consumers discover the needed resources through the *(offering) query* specifying the properties of the offerings of interest. For each query a set of matching offerings is returned, and a consumer is able to subscribe and get access to them at run-time. The pair *offering description – offering query* provides a mechanism of data/service discovery within the plentitude of registered providers (see the flow in Figure 1). It is important to stress that offering description also contains an explicit mapping of the output fields to the corresponding semantic URIs (e.g. "longitude" : "http://schema.org/longitude").

Semantic Web technologies therefore form the core of the BIG IoT project: modeling providers, consumers and offerings according to a predefined RDF schema allows to conceptualize all activities needed for the collaboration among the various stakeholders of an IoT ecosystem. The meta-data coming from different sources is combined and stored in an RDF triple store. Sophisticated semantic queries as well as semantic matching (even in cases where participants use different semantic models or formats) are enabled.

3 BIG IoT API and SDK

To facilitate participation in the BIG IoT ecosystem, developers can download an SDK for the BIG IoT API. The SDK encompasses libraries for both providers and consumers, which translate function calls from the respective application or service code into interactions with the marketplace, or directly between a consumer and provider. The Provider Lib allows an IoT platform or service to register offerings. The Consumer Lib in turn allows an application or service to discover available offerings based on semantic queries, and to subscribe to offerings of interest.

In the presented work, we make use of semantic annotations already provided in the offering descriptions in order to empower consumers to use this knowledge when accessing data that are not yet semantically enriched by their providers. Specifically, we extend the Consumer Lib in order to automatically transform the obtained data and semantically annotate them using linked data frameworks.

4 JSON – JSON-LD Transformation in the BIG IoT API

Semantic matching during the IoT service discovery phase might result in a set of offerings which – when accessed – may differ, for instance, in terms of output data returned to the consumer. Many already deployed platforms or services use JSON as data format. However, different providers use typically different property names for the data offered, which, in conjunction with the lack of semantic information, makes it impossible for a consumer application to automatically integrate the data. As a consequence, the developer of a consumer application has to take care of the integration of data from different providers manually and at implementation time.

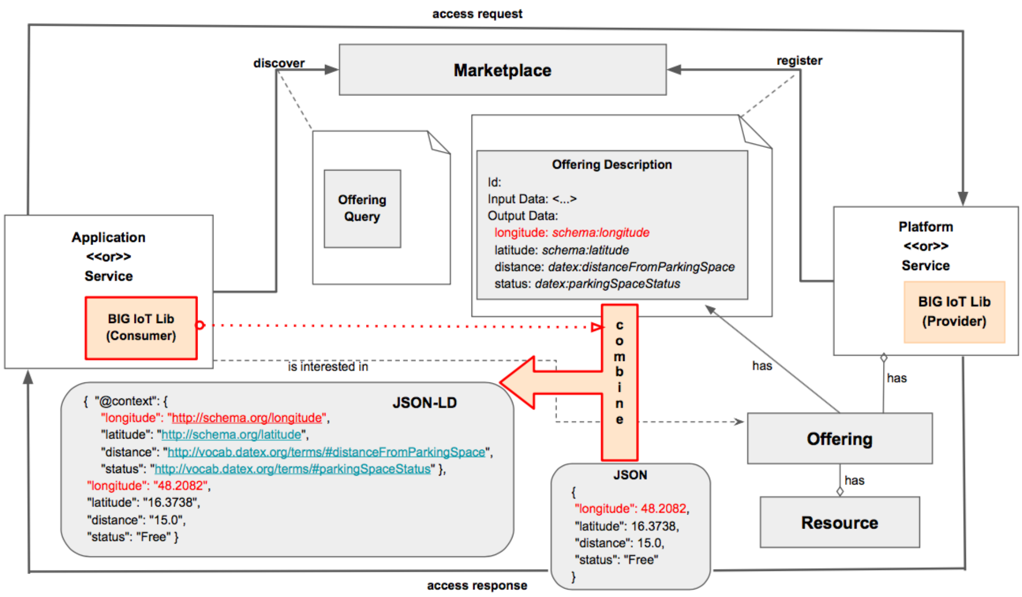

Fig. 1. JSON – JSON-LD Transformation as a Part of BIG IoT Consumer Lib

Figure 1 depicts our JSON – JSON-LD transformation approach in more detail. Instead of simply redirecting the providers' response data (JSON) to the subscribed consumer, the Consumer Lib first obtains the offering description of a particular data or service from the marketplace. The offering description contains the *JSON object keys to IRI/URI mapping* which is used as a basis for the fully automated JSON-LD `@context` extraction [8] during the JSON response processing. Then the Consumer Lib returns the response in the JSON-LD format to the consumer.

Transforming JSON-serialized data obtained from an external data source into linked data has many advantages from a consumer application point of view. Firstly, it enables it to compare (rank, filter) and fuse data coming from different providers, and to validate and check them for consistency. This is vital for future-proof IoT applications, which will be increasingly autonomous and thus rely on run-time selection of providers.

Secondly, the resulting data harmonization can help developers to minimize the efforts needed to integrate new data providers, and to access data from various sources in a uniform manner. Even if a consumer does not utilize the set of BIG IoT ontologies itself, it is easier to translate a single pre-defined vocabulary to the local schemata than to manage potentially conflicting keys from different sources. As JSON-LD is 100% JSON compatible [8], the existing JSON parsers and libraries can be re-used to process the harmonized data. Moreover, a consumer application can utilize generic visualization toolkits or APIs to inspect the incoming data visually.

Thirdly, as JSON-LD represents an instance of an RDF data model [8], it can be directly used with other linked data technologies. Thus, the JSON-LD representation of the aggregated output data can be either stored in an RDF triple store, which in turn will enable a consumer to leverage the power of SPARQL and inference, or passed to a real-time link-data reasoner (for instance, *ldfu* system [6]) to take advantage of inference based on live data obtained from multiple sources.

In conclusion, mapping varying JSON attributes to a single set of unambiguous identifiers gives each portion of data a well-defined structure, streamlines its processing, and makes data truly interoperable. For future-proof IoT consumer applications, which rely on dynamic (run-time) service discovery and integration of new data sources and services, we believe the proposed JSON – JSON-LD transformation is the best solution, as generic linked date frameworks can be used to integrate the data in a very flexible manner and with the extra benefit that JSON-LD capable data providers can also be considered without additional efforts.

References

1. IERC – European Research Cluster on the Internet of Things: IoT semantic interoperability: research challenges, best practices, recommendations and next steps. (2015).
2. James, M. et al.: The Internet of Things: Mapping the value beyond the hype. McKinsey Global Institute (2015): 3.
3. Jara, A.J. et al.: Semantic web of things: An analysis of the application semantics for the IoT moving towards the IoT convergence. International Journal of Web and Grid Services 10.2–3 (2014): 244–272.
4. Murdock, P. et al.: Semantic Interoperability for the Web of Things (2016): 1–19. https://doi.org/10.13140/RG.2.2.25758.13122 [Accessed 18 Jul 2017].
5. Schmid, S. et al.: An Architecture for Interoperable IoT Ecosystems. International Workshop on Interoperability and Open-Source Solutions. Springer, Cham (2016): 39–55.
6. Stadtmüller, S., Speiser, S., Harth, A. and Studer, R.: Data-Fu: a language and an interpreter for interaction with read/write linked data. In: Proceedings of the 22nd international conference on World Wide Web (2013). ACM: 1225–1236.
7. Vetere, G., and Lenzerini, M.: Models for semantic interoperability in service-oriented architectures. IBM Systems Journal 44.4 (2005): 887–903.
8. World Wide Web Consortium, 2014. JSON-LD 1.0: a JSON-based serialization for linked data. https://www.w3.org/TR/json-ld/ [Accessed 18 Jul 2017].