Smart City Artifacts Web Portal

Noorani Bakerally[(⊠)], Olivier Boissier, and Antoine Zimmermann

Univ Lyon, MINES Saint-Etienne, CNRS, Laboratoire Hubert Curien UMR 5516,
42023 Saint Etienne, France
{noorani.bakerally,olivier.boissier,antoine.zimmermann}@emse.fr

Abstract. In the smart city domain, many projects and works are generating essential information. Open and efficient sharing of this information can be beneficial for all parties ranging from researchers, engineers or even governments. To our knowledge, there is currently no full-fledged semantic platform which properly models this domain, publishes such information and allows data extraction using a standard query language. To complement this, we developed and deployed the Smart City Artifacts web portal. In this paper, we present our approach used within this platform and summaries some of its technical features and applications.

Keywords: Smart cities · Linked data · Semantic technologies · Semantic web

1 Introduction

Currently, as far as we know there is no semantic platform or central repository for keeping track of artifacts generated from projects or works in the field of smart cities. By artifact, we refer to any piece of work, for example a vocabulary, application or deliverable. Most well-known smart city projects, like km4City[1], could improve the way they publish their data. For instance, the entry points to their web portal are mostly human readable and can be enriched with a SPARQL endpoint to be machine queryable. Moreover the platform is not self-describing and one has to manually go through much details before discovering important things such as underlying model used to structure data, public datasets, RDF dumps or SPARQL endpoints. An exception is Ready4SmartCities [3]. They do provide a web platform which lists ontologies and datasets for smart cities[2] which both can be downloaded as RDF. However, the platform itself does not follow some of the Linked Data principles and best practices and does not make provenance information explicit for vocabularies. As a result, it may not be possible for a person or machine to explore their data, find specific resources and relate other data to it.

To complement existing smart city web portals, we propose the Smart City Artifacts (SCA) web portal[3] which gathers information about smart city projects

---

[1] http://www.disit.org/drupal/?q=node/6056 on 29/02/2016.
[2] http://www.ready4smartcities.eu/ on 29/02/2016.
[3] http://opensensingcity.emse.fr/scans/.

and their artifacts while conforming to Linked Data principles and best practices. In the rest of this paper, we summarize some of the technical features and applications of the SCA web portal to demonstrate how we achieved the following tasks: (1) summarize the development of SCA ontology to provide a metamodel for smart city projects and artifacts in Sect. 2; (2) explain how we set up the SCA web portal to publish and answer queries related to smart city information in Sect. 3; (3) show how we conform to semantic standards, Linked Data principles and best practices in Sect. 4; (4) identify some use cases where the portal can be beneficial in Sect. 5.

2   SCA Ontology

The SCA web portal provides information about smart city projects and artifacts. We need a metamodel in the form of an ontology to structure all this information. No such ontology was found on ontology repositories but many such ontologies, like those shown in Table 1, were found which could be leveraged to describe some aspects of the metamodel.

To develop the SCA ontology, we chose to maximize vocabulary reuse. Most of the metadata we provide come from the domains of the ontologies listed in Table 1. When a relevant term could not be found in an existing vocabulary, we created the term within our ontology and ensured that the dereferenceable content provides the appropriate semantics. Part of the SCA ontology is shown in Fig. 1. All entities in the figure are linked to `sca:Domain` and `muto:Tag` via the `dc:subject` and `muto:hasTag` respectively. An instance of `sca:Domain` and/or `muto:Tag` is linked to a DBpedia resource via `rdfs:seeAlso`. More details about the whole ontology can be found on Github.[4]
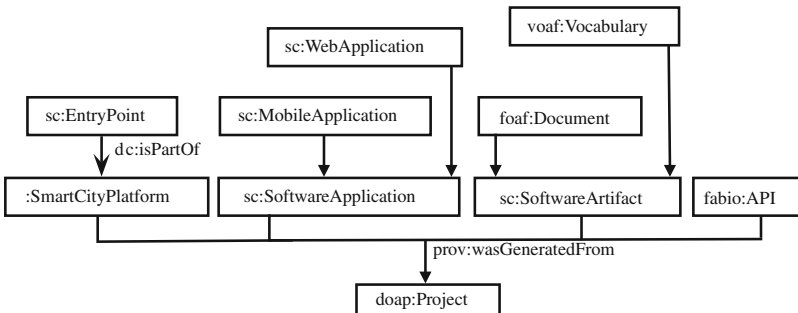

Fig. 1. Part of SCA ontology

Table 1. Reused vocabularies

Name	Prefix	Namespace IRI
Dublin core terms	dc	http://purl.org/dc/terms/
Description of a project	doap	http://usefulinc.com/ns/doap
The PROV ontology	prov	http://www.w3.org/ns/prov
Friend of a friend	foaf	http://xmlns.com/foaf/0.1/
Vocabulary of a friend	voaf	http://purl.org/vocommons/voaf
Schema.org	sc	http://schema.org/
Modular unified tagging Ontology	muto	http://purl.org/muto/core
FRBR-aligned bibliographic ontology	fabio	http://purl.org/spar/fabio/
DBpedia ontology	dbowl	http://dbpedia.org/ontology/
Ontology metadata vocabulary	omv	http://omv.ontoware.org/ontology#

3   SCA Web Portal

In this section, we describe the features of the SCA web portal and summarize its architecture.

3.1   User Features

All information on the SCA web portal comes from an RDF dataset structured according to the SCA ontology. The SCA web portal provides numerous entry points to visualize important resources from its dataset. For example, it provides an entry point to visualize a list of projects and then from this list, the user can choose a project to get more details about it. Viewing this project is like visualizing a specific resource in the dataset. The portal shows important links from that resource to other resources and literals. In the case of viewing a project, project's details like title and links to artifacts or documents are shown. From there, the user can choose another resource, like an artifact, and continue navigating in the dataset through the portal.

The portal also provides search facilities. Search can be performed using keywords, domain, tags and types. For example, a user can search for all resources of a particular type, e.g. vocabulary, or search resources related to domain or tags. External tools are used to augment services provided to the user. When viewing details about a particular vocabulary, the user can visualize it graphically (using Web VOWL[5]), detect ontology pitalls (using Ooops[6]) or validate namespaces (RDF Triple Checker[7]).

---

[5] http://vowl.visualdataweb.org/webvowl.html.
[6] http://oops.linkeddata.es.
[7] http://graphite.ecs.soton.ac.uk/checker/.

3.2   Architecture

Like most web application, the portal has 3 layers: a data, application and web Layer. At the data layer lies the dataset. It is structured as per the SCA ontology and enriched with links to resources from the DBpedia dataset. The portal communicates with the dataset through a SPARQL endpoint provided by Apache Fuseki[8], a SPARQL Server. At the Application layer lies the portal itself. It was developed using a Python microframework[9] and follows the Model-View-Controller pattern. The controller handles all requests, when required, it connects with the SPARQL server using a SPARQL endpoint interface[10] to fetch content and finally serves it in a particular format through content negotiation. The SPARQL endpoint is exposed on the web.[11]

At the Web layer resides Apache HTTP Server which expose the portal on the web. For all SPARQL queries over HTTP, Apache HTTP Server acts like a proxy server between the client and the SPARQL server. The portal also provides a web-based form[12] where users can directly make request on the SPARQL endpoint, get results on the same interface and download the result in different format.

4   Conformance to Semantic Standards and Best Practices

The conformance of a web platform to semantic standards and Linked Data principles and best practices can be evaluated at numerous levels. Lóscio *et al.* [4] list many principles and best practices which are currently in the standardization pipeline. Many principles and best practices we apply come from this list. At the vocabulary level, we have maximized vocabulary reuse to enhance semantic interoperability. At the dataset level, resources are linked to other resources from the LOD dataset, like DBpedia, for data enrichment and discoverability. It would have been ideal to link resources from our dataset to Ready4SmartCities's dataset.[13] However, this is not possible because none of the resources defined in their dataset come from their own namespace.

Moreover, all resources in our dataset which have an IRI within our namespace are dereferenceable with an almost equivalent representation of that resource both in human form (`HTML`) and machine readable form (`rdf/xml,n3,nt,turtle`) through content negotiation. Vocabulary of Interlinked Dataset (VoID) [1] is used to provide a self-description of the portal's dataset, facilitating automated data discovery. The VoID file can be obtained by requesting RDF data on the root IRI[14] of the portal.

---

[8] https://jena.apache.org/documentation/serving_data/.
[9] http://flask.pocoo.org/.
[10] https://rdflib.github.io/sparqlwrapper/.
[11] http://opensensingcity.emse.fr/scans/ds/query.
[12] http://opensensingcity.emse.fr/scans/query.
[13] http://smartcity.linkeddata.es/rdf/ontologyRDF.ttl on 29/02/2016.
[14] http://opensensingcity.emse.fr/scans.

5   Use Cases

In this section, we outline some scenarios in which the SCA web portal can be used to demonstrate the added value it can bring in.

Scenario 1: The usual scenario is navigating through the SCA web portal using a browser to obtain information. Users can go on the Projects Page[15] to view all projects. There, they can search using full-text search or through particular domain or tags. They can also, view and browse through all the details of artifacts, visualize vocabularies and apply external tools as mentioned above on vocabularies. Through this, anyone engaged in smart city projects can obtain an overview of the state of the arts in this field and find information or artifacts appropriate to their use.

Scenario 2: Consider a case where a user wants to create an HTML table of all smart city projects and the vocabularies they generated. Such content can be obtained using automated script which requests for RDF data on the *Projects Page*'s IRI. From there, after getting a list of projects' IRIs, the script exploits the provenance relationship to obtain all artifacts of type Vocabulary generated for each project.

Scenario 3: The European Union open data portal[16] provides data about many datasets which is queryable through a SPARQL Endpoint.[17] A user may want to search for all datasets from this portal and see if any of their dataset relates to a string literal containing a particular domain or tag found in SCA data. This is a possible SPARQL query[18] for such an information request.

Scenario 4: Consider an information request where someone from France wants to locate all smart city platforms in France having entry points which are related to bicycle and assume that the person only know the french word "Bicyclette" of bicycle. At first, it may seem trivial to incorporate multilingual and geographic operation. But, such an information request can be formalized as a SPARQL query[19] by exploiting the RDF links to resources from DBpedia.

6   Conclusion and Future Improvements

Smart city is becoming an important and hot topic of research in many countries. Much information about smart city projects and artifacts is scattered on the web. Through this work, we have shown that applying semantic standards, Linked Data principles and best practices has enabled us to efficiently centralize, enrich, publish and serve complex information requests using SPARQL queries and LOD dataset links. To ensure that the portal continues to benefit the international

---

[15] http://opensensingcity.emse.fr/scans/projects.

[16] https://open-data.europa.eu/fr/data.

[17] http://open-data.europa.eu/sparqlep.

[18] http://tinyurl.com/ztny6fh.

[19] http://tinyurl.com/gslgcfy.

community for smart cities, we intend to incorporate collaboration features such as providing a web form to submit new details about projects and artifacts. Also, to further enrich the dataset, we aim to densify the link set from SCA data to other LOD datasets. It is important to realize that when setting up a semantic platform, the Linked Data principles [2] are not the only 4 principles that developers have to follow. Instead, in a given context, these 4 principles generate a number of other principles, patterns and best practices which have to be considered to ensure the platform provides linked data and contributes to further realize the vision of the Semantic Web.

References

1. Alexander, K., Cyganiak, R., Hausenblas, M., Zhao, Y.: Describing linked datasets with the VoID vocabulary, W3C interest group note 03. W3C IG note, World Wide Web Consortium (W3C), 3 March 2011
2. T Berners-Lee. http://www.w3.org/designissues/linkeddata.html. Accessed 11 June 2012
3. Garcıa-Castro, R., Gómez-Pérez, A., Corcho, O.: Ready4smartcities: ICT roadmap and data interoperability for energy systems in smart cities. In: 11th Extended Semantic Web Conference (ESWC 2014) (2014)
4. Lóscio, B.F., Burle, C., Calegari, N.: Data on the web best practices. Working draft, W3C, January 2016. https://www.w3.org/TR/dwbp/