Semantic-enhanced National Access Points to Multimodal Transportation Data

Alessio Carenini, Marco Comerio, and Irene Celino

Cefriel – Politecnico di Milano

viale Sarca, 226 – Milano, Italy

E-mail: {alessio.carenini,marco.comerio,irene.celino}@cefriel.com

Abstract. The establishment of National Access Points (NAPs) to transportation data represents a first step towards the realization of semantic interoperable travel information services. This paper presents the Semantic Assets Manager, i.e., an organized collection of asset types (i.e., datasets, APIs, ontologies, and schemas) enhanced by tools for their publication, governance and discovery that represents a semantic-enhanced solution for the establishment of a NAP to multimodal transportation data.

1   Introduction

Semantic interoperability in the transportation sector is one of the European Commission challenges: establishing an interoperability framework enables European travel and transport industry players to make their business applications 'interoperate' and provides the travelers with a new seamless travel experience, accessing a complete multimodal travel offer which connects the first and last mile to long distance journeys.

The EU-report [5] on the provision of EU-wide multimodal travel information services highlighted several barriers for the realization of comprehensive travel information services in the EU, such as insufficient accessibility of travel and traffic data and the lack of travel and traffic data interoperability. Key enablers to address these barriers are (i) ensuring that users have *access* to the right scope of data and information with the appropriate data sharing mechanism, and (ii) making travel and traffic data *interoperable* with a common set of data exchange standards.

A first step towards the realization of multimodal travel information services is the establishment of National Access Points (NAPs) to multimodal transportation data as defined in the recent Commission Delegated Regulation (EU) 2017/19261 [6], which establishes the specifications necessary to ensure the accessibility, exchange and update of static and dynamic transportation data. According to the regulation, each Member State shall set up a NAP; transport authorities, transport operators and transport infrastructure managers of each Member State shall provide travel and traffic data to the NAP defined according to specific standards (e.g., NeTEx[1]) and described using national application profiles (e.g., DCAT-AP[2]). As emerged from the Annual NAP Report

---

[1] http://netex-cen.eu/

[2] https://joinup.ec.europa.eu/release/dcat-ap-v11

2017 [6], only a few countries have planned their NAPs, and very few operators are ready to be compliant with the requested standards.

This paper presents the *Semantic Assets Manager*, i.e., an organized collection of different asset types (i.e., datasets, APIs, ontologies, and schemas) enhanced by tools for their publication, governance, and discovery. The Semantic Assets Manager represents a semantic-enhanced solution for the establishment of NAPs. Moreover, it promotes the *sharing of semantic converters*, i.e., software components supporting transport authorities, operators and infrastructure managers in transforming their data into the correct formats/standards.

The paper is organized as follows. Section 2 describes the requirements for a Semantic-enhanced National Access Point. Section 3 proposes our solution. Section 4 outlines future and ongoing work.

2 Requirements for a National Access Point

According to [6], EU Member States shall set up the first version of their NAPs within December 2019. For this reason, the first requirement to be considered is that the proposed solution must be market-ready.

Additional requirements come from the following categories of stakeholders involved in the management/usage of the NAP:

- **NAP Manager**: is in charge of realizing the NAP governance by managing the technical and procedural aspects. The NAP Manager needs to provide a trusted environment with clearly defined authentication and authorization procedures. Moreover, the NAP Manager needs support in quality assurance and lifecycle management of the published assets.
- **Transport authorities**, **transport operators** and **transport infrastructure managers**: need support in publishing their data in the NAP and making them discoverable. Moreover, they potentially need support in turning their data in the required formats/standards.
- **Intelligent Transport System (ITS) providers**: need support in discovering and using datasets and API from the NAP for ITS app development.
- **Travelers**: need support in discovering and using NAP information for their journey planning.

In summary, a solution for the establishment of a NAP should address the following requirements:

R1. Authorization and authentication mechanisms;

R2. Tools for the publication of the datasets, their metadata descriptions and storage;

R3. A dataset lifecycle management that includes an "IN REVIEW" status where a specific NAP management committee verifies the quality of the dataset and its compliance with the EU regulation;

R4. Trustworthiness in the datasets management;

R5. Tools for the discovery of the datasets;

R6. Semantic converters of datasets between different format/standards.

3   The Semantic Assets Manager

The Semantic Assets Manager is our semantic-enhanced solution for the establishment of a NAP to multimodal transportation data. To be market-ready and to reuse basic existing features, we decided to start from the popular and open source WSO2 Governance Manager[3] that, out of the box, satisfies requirements R1-R4. We configured the Semantic Assets Manager by extending WSO2 in order to manage the following asset types: (i) *ontologies,* shared reference domain ontologies; (ii) other *schemas,* e.g., XSD definitions of the standards, such as NeTEx and GTFS[4]; (iii) *datasets,* travel and traffic data of different transport modes; (iv) *service descriptions,* made available by transport stakeholders to dynamically access data; (v) *exploration APIs,* supporting the discovery of available assets; (vi) *converter services*, transforming data between different standards. Each asset type is published according to a specific semantically-enhanced metadata profile and associated with a specific lifecycle management. The Semantic Assets Manager[5] is composed of two Web application, namely the *publisher[6]* and the *store[7]*, supporting the publishing and the discovery of the assets, respectively.

To further strengthen the satisfaction of R4, we have included blockchain technology [9]: a transaction is added to a distributed ledger each time an asset changes its lifecycle state. Such transaction contains all accounting information, enabling all the ecosystem participants to check the complete history of an asset on a trusted platform.

To cover also R5, the WSO2 Governance Manager is extended with modules supporting the transformation, storage, and discovery of the published assets descriptions in RDF, thus enabling more expressive queries on such data. Following the approach in [3,8] of SPARQL query templates, we introduce a new asset type, the exploration API, for an intelligent asset discovery (e.g., finding all transport datasets/services proving information about bike sharing in a specific Italian region). When a new exploration API method is created, the Semantic Asset Manager automatically generates its OpenAPI/Swagger documentation[8]. The benefits introduced by this asset type are: (i) a better control on RDF data access; (ii) improved usability for users involved in dataset discovery; (iii) easy integration with external systems.

Finally, to cover also R6, the semantic converters proposed in [2] and based on the approach in [4], are also added as assets. The semantic converters are software components supporting the conversion between transportation standards. As an example, a specific semantic converter enables the translation of transportation schedule, geographic and fare information expressed in GTFS to a NeTEx specification preserving the original meaning.

---

[3] http://wso2.org

[4] http://gtfs.org

[5] A detailed description of the Semantic Assets Manager is available at: https://bit.ly/2JG8zL8

[6] https://ns3056488.ip-213-32-26.eu:9443/publisher (usr:guest, psw: $ISWC2018_guest#)

[7] https://ns3056488.ip-213-32-26.eu:9443/store (usr:guest, psw: $ISWC2018_guest#)

[8] https://www.openapis.org/

The Semantic Assets Manager satisfies the FAIR principles[9] since the assets, accessible through public APIs, are associated with metadata profiles and are defined according to specific standards that make them interoperable and reusable. Moreover, our solution overcomes most of the limitations of data management platforms mentioned in [1] since we provide (i) access to the source code to permit customization and extensions, (ii) APIs to support easy integration, and (iii) unique identifiers for the assets to improve findability.

4   Ongoing and Future Work

Together with a network of transport stakeholders, we are defining a governance framework for the Semantic Assets Manager, to enable its adoption and sustainability as a semantic-enhanced NAP for multimodal transportation data. Moreover, we are currently exploring the use of SHACL Advance Features[10] to improve the automation of ontology-based annotations and mappings and to enhance further the approach used for the realization of semantic converters.

References

1. Amorim, R. C., Castro, J. A., Da Silva, J. R., & Ribeiro, C. (2017). A comparison of research data management platforms: architecture, flexible metadata, and interoperability. *Universal Access in the Information Society*, *16*(4), 851-862.
2. Carenini, A., Dell'Arciprete, U., Gogos, S., Kallehbasti, M., Rossi, M., Santoro, R. ST4RT – Semantic Transformations for Rail Transportation. In Proc. TRA 2018.
3. Daga, E., Panziera, L., Pedrinaci, C.: A BASILar Approach for Building Web APIs on top of SPARQL Endpoints. In Proc. of SALAD@ESWC 2015: 22-32.
4. Della Valle, E., Cerizza, D., Celino, I., ... Monteleone, G. (2007). SEEMP: a semantic interoperability infrastructure for e-government services in the employment sector. *The Semantic Web: Research and Applications*, 220-234.
5. EU Commission (2016). Study on ITS Directive, Priority Action A: the Provision of EU-wide Multimodal Travel Information Services. Available at: http://ec.europa.eu/transport/themes/its/consultations/doc/2015-its-mmtips/consultation-report.pdf
6. EU Commission (2017). Commission Delegated Regulation (EU) 2017/1926. Available at: https://bit.ly/2zSoOFv
7. Hendriks L., Jorna R., et Al. (2017). Monitoring and Harmonisation of National Access Points in Europe. Available at: https://www.its-platform.eu/filedepot_download/1971/6320
8. Meroño-Peñuela, A., Hoekstra, R. (2017). Automatic Query-Centric API for Routine Access to Linked Data. International Semantic Web Conference (2) 2017: 334-349
9. Swan, Melanie. Blockchain: Blueprint for a new economy. O'Reilly Media, Inc., 2015.

---

[9] https://www.force11.org/group/fairgroup/fairprinciples
[10] https://www.w3.org/TR/shacl-af/