Semantic Web 14 (2023) 613–616
DOI 10.3233/SW-223278
IOS Press

Editorial of transport data on the web

David Chaves-Fraga [a,b,*], Pieter Colpaert [c], Mersedeh Sadeghi [d] and Marco Comerio [e]
[a] *Universidad Politécnica de Madrid, Spain*
E-mail: david.chaves@upm.es
[b] *KU Leuven, Belgium*
[c] *Ghent University – imec, Belgium*
E-mail: pieter.colpaert@ugent.be
[d] *University of Cologne, Germany*
E-mail: sadeghi@cs.uni-koeln.de
[e] *CEFRIEL, Italy*
E-mail: marco.comerio@cefriel.com

1. Preface

Whether you are planning your next trip abroad or want a package delivered to your doorstep, chances are high that you will need a chain of services provided by multiple companies. Transport is inherently a geographically and administratively decentralized domain composed of a diverse set of actors, – from public transport authorities to vehicle sharing companies, infrastructure managers in different sectors (road, rail, etc.), transport operators, retailers, and distributors. As a result, it suffers vast data heterogeneity, which, in turn, brings severe challenges to data interoperability. However, such challenges have also been posed in other domains such as the Internet of Things [18], agriculture [11], building data management [17], biology [7] or open data [2], which have found their solutions using semantic web technologies. However, despite several research contributions [6,14,19,23,25], public-funded projects[1,2] or academic-industry events,[3,4] we have not yet seen a wide adoption of semantic technologies in the transport domain.

We may only guess the inhibitors for adopting Linked Data in this domain: i) the SPARQL query language is not built for optimal path planning, and ii) RDF is perceived as highly conceptual by industry experts. We argue that SPARQL does not fit well with the concerns that typically matter to route planners (e.g., calculating the optimal Pareto path [4]). While calculating a path with SPARQL is feasible through property paths, controlling the path planning algorithm, which can hardly be done in SPARQL, is the core concern of route planners. On the other hand, the transport domain is dominated by different standards (e.g., NeTEx,[5] or DATEX II[6]) and vocabularies, which are based on legacy data exchange technologies (e.g., XML or RDB). However, to construct a distributed and scalable architecture that addresses the current needs of this domain, the Web and its associated technologies (i.e., the Semantic Web) are the key resource.

---

[*] Corresponding author. E-mail: david.chaves@upm.es.
[1] https://cordis.europa.eu/project/id/826172
[2] https://cordis.europa.eu/project/id/777522
[3] GIRO Workshop at KG Conference: https://giro.cefriel.it/.
[4] Semantics for Transport Workshop at SEMANTiCS conference: https://sem4tra.linkeddata.es/.
[5] https://netex-cen.eu/
[6] https://www.datex2.eu/

Although these technologies are still not able to perform route planning itself, they can be beneficial for a route planning algorithm to make the source selection and get the right data. Furthermore, data interoperability can be achieved using standard ontology-based access technologies such as R2RML [9] or RML [10], which ensures transparent and sustainable data integration approaches to generate RDF graphs that facilitate their adoption in the transport domain. In this direction, semantic technologies can also leverage mapping to transform legacy data models into semantic-based standards and models [3,5,13]. Semantic technologies can also play an important role beyond data interoperability in the transportation domain and leverage service interoperability (as a key enabler of conception such as smart mobility [27]) and many other relevant aspects such as security, privacy and access control [12,22], business agreements [26], booking and ticketing [1], traffic management [28], and travel service recommendation [8,15,16].

We believe that researchers, industry partners, and stakeholders should work together to build a data ecosystem in the transport domain using the Web as its central support platform. It will improve multimodal experiences and end-to-end services, increasing the use of more innovative, sustainable, and efficient mobility solutions.

2. Topics

The special issue of transport data on the web included the following relevant topics:

a. Decentralized data management

- Query languages and methods for a Web of Transport Data
- Creating automated alignments between datasets using the various specifications
- Recommendations for mobility specification builders to raise interoperability between specifications
- Alignments with general-purpose Web API specifications such as W3C Web payments or SOLID
- Interfaces between enterprise data and the Semantic Web

b. Exploitation of transport data on the web

- Challenges and opportunities for route planning and ticketing
- Personalized route planning taking into account data stored on your client or personal data space
- Anonymization in the mobility space for data sharing

c. Benchmarking web infrastructure

- Comparison of ticketing API architectures
- Comparing route planning API architectures
- Publishing and querying planned transport data and their live updates on the Web

d. Knowledge Graphs and Ontologies in the transportation

- Aligning regional vocabularies with international domain models (e.g., Transmodel[7]).
- Reusing existing linked datasets (e.g., Geonames, OpenStreetMap) for the Mobility-as-a-Service use case.
- Bridging the gap with non-RDF specifications (e.g., GTFS,[8] GBFS[9]).

3. Content

The special issue attracted four submissions, covering our key research topics: semantic web technologies and transport data. At least three researchers reviewed each paper, and three papers were accepted after two or three rounds of review. In the following, we provide a broad overview of the accepted papers.

[7] http://www.transmodel-cen.eu/
[8] https://developers.google.com/transit/gtfs
[9] https://github.com/NABSA/gbfs

- The first paper "Applying the LOT methodology to a Public Bus Transport Ontology aligned with Transmodel: Challenges and Results" [21] presents the translation from an excerpt of the Transmodel reference model to an ontology in the specific context of Public Bus transport. The authors follow the well-known LOT [29] methodology to develop the ontology. The resource also provides a set of competency questions to validate the vocabulary, and a set of mapping rules in RML [10] to allow the declarative generation of RDF graphs from (semi)structured data sources.
- The second paper "Urban IoT Ontologies for Sharing and Electric Mobility" [24], describes a set of vocabularies to enable interoperability between municipalities and public organizations with IoT service providers. The authors also follow the LOT methodology [29] to develop the vocabularies, together with a set of competency questions and some RDF examples in JSON-LD serialization of the bikes service from the city of Milan.
- The last accepted paper "Publishing planned, live and historical public transport data on the Web with the Linked Connections framework" [20] presents a system architecture to support cost-efficient publishing of dynamic public transport schedules and historical data. The authors also present an extensive empirical evaluation including route planning query performance based on data fragmentation size, dataset publishing costs, and a comparison with traditional route planning engines.

References

[1] K. Angele, D. Fensel, E. Huaman, E. Kärle, O. Panasiuk, U. Şimşek, I. Toma and A. Wahler, Semantic web empowered E-tourism, in: *Handbook of e-Tourism*, 2020, pp. 1–46. doi:10.1007/978-3-030-05324-6_22-1.
[2] F. Bauer and M. Kaltenböck, Linked open data: The essentials, *Edition mono/monochrom, Vienna* **710** (2011).
[3] A. Carenini, U. Dell'Arciprete, S. Gogos, M.M.P. Kallehbasti, M. Rossi, R. Santoro et al., ST4RT – Semantic transformations for rail transportation, in: *Proceedings of 7th Transport Research Arena TRA 2018*, 2018, pp. 1–10. https://re.public.polimi.it/bitstream/11311/1068073/1/Contribution_10301_fullpaper.pdf.
[4] Y. Censor, Pareto optimality in multiobjective problems, *Applied Mathematics and Optimization* **4**(1) (1977), 41–59. doi:10.1007/BF01442131.
[5] D. Chaves-Fraga, A. Antón, J. Toledo and O. Corcho, Onett: Systematic knowledge graph generation for national access points, in: *SEM4TRA-AMAR@ SEMANTiCS*, 2019.
[6] D. Chaves-Fraga, F. Priyatna, A. Cimmino, J. Toledo, E. Ruckhaus and O. Corcho, GTFS-Madrid-Bench: A benchmark for virtual knowledge graph access in the transport domain, *Journal of Web Semantics* **65** (2020), 100596. doi:10.1016/j.websem.2020.100596.
[7] H. Chen, T. Yu and J.Y. Chen, Semantic web meets integrative biology: A survey, *Briefings in Bioinformatics* **14**(1) (2013), 109–125. doi:10.1093/bib/bbs014.
[8] D. Corsar, P. Edwards, J. Nelson, C. Baillie, K. Papangelis and N. Velaga, Linking open data and the crowd for real-time passenger information, *Journal of Web Semantics* **43** (2017), 18–24. doi:10.1016/j.websem.2017.02.002.
[9] S. Das, S. Sundara and R. Cyganiak, *R2RML: RDB to RDF Mapping Language, W3C Recommendation, World Wide Web Consortium (W3C)*, 2012. http://www.w3.org/TR/r2rml/.
[10] A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens and R. Van de Walle, RML: A generic language for integrated RDF mappings of heterogeneous data, in: *Ldow*, 2014.
[11] B. Drury et al., A survey of semantic web technology for agriculture, *Information Processing in Agriculture* **6**(4) (2019), 487–501. doi:10.1016/j.inpa.2019.02.001.
[12] A. Fiaschetti, F. Lavorato, V. Suraci, A. Palo, A. Taglialatela, A. Morgagni, R. Baldelli and F. Flammini, On the use of semantic technologies to model and control security, privacy and dependability in complex systems, in: *International Conference on Computer Safety, Reliability, and Security*, Springer, 2011, pp. 467–479. doi:10.1007/978-3-642-24270-0_34.
[13] S. Kalwar, M. Sadeghi, A.J. Sabet, A. Nemirovskiy and M. Rossi, Smart: Towards automated mapping between data specifications, in: *33rd International Conference on Software Engineering and Knowledge Engineering, SEKE 2021*, Vol. 2021, Knowledge Systems Institute Graduate School, 2021, pp. 429–436. doi:10.18293/SEKE2021-161.
[14] M. Katsumi and M. Fox, Ontologies for transportation research: A survey, *Transportation Research Part C: Emerging Technologies* **89** (2018), 53–82. doi:10.1016/j.trc.2018.01.023.
[15] Y. Li, M. Zhu, H. Chen and K.-M. Chao, A service-oriented travel portal with semantic package recommender, in: *The 2010 14th International Conference on Computer Supported Cooperative Work in Design*, IEEE, 2010, pp. 384–389. doi:10.1109/CSCWD.2010.5471944.
[16] A. Moreno, A. Perallos, D. López-de-Ipiña, E. Onieva, I. Salaberria and A.D. Masegosa, A novel software architecture for the provision of context-aware semantic transport information, *Sensors* **15**(6) (2015), 12299–12322. doi:10.3390/s150612299.
[17] P. Pauwels et al., Semantic web technologies in AEC industry: A literature overview, *Automation in Construction* **73** (2017), 145–165. doi:10.1016/j.autcon.2016.10.003.
[18] A. Rhayem et al., Semantic web technologies for the Internet of things: Systematic literature review, *Internet of Things* **11** (2020). doi:10.1016/j.iot.2020.100206.

[19] J.A. Rojas, M. Aguado, P. Vasilopoulou, I. Velitchkov, D.V. Assche, P. Colpaert and R. Verborgh, Leveraging semantic technologies for digital interoperability in the European railway domain, in: *International Semantic Web Conference*, Springer, 2021, pp. 648–664. doi:10.1007/978-3-030-88361-4_38.
[20] J.A. Rojas, H. Delva, P. Colpaert and R. Verborgh, Publishing planned, live and historical public transport data on the web with the linked connections framework, *Semantic Web* (2022).
[21] E. Ruckhaus, A. Anton-Bravo, M. Scrocca and O. Corcho, Applying the LOT methodology to a public bus transport ontology aligned with transmodel: Challenges and results, *Semantic Web* (2021). doi:10.3233/SW-210451.
[22] M. Sadeghi et al., A semantic-based access control approach for systems of systems, *ACM Applied Computing Review* **21**(4) (2022). doi:10.1145/3512753.3512754.
[23] M. Sadeghi, P. Buchníček, A. Carenini, O. Corcho, S. Gogos, M. Rossi and R. Santoro, SPRINT: Semantics for PerfoRmant and scalable INteroperability of multimodal transport (2020), 1–10. doi:10.3030/826172.
[24] M. Scrocca, I. Baroni and I. Celino, Urban IoT ontologies for sharing and electric mobility, *Semantic Web* (2021). doi:10.3233/SW-210445.
[25] M. Scrocca, M. Comerio, A. Carenini and I. Celino, Turning transport data to comply with EU standards while enabling a multimodal transport knowledge graph, in: *International Semantic Web Conference*, Springer, 2020, pp. 411–429. doi:10.1007/978-3-030-62466-8_26.
[26] M. Scrocca, M. Comerio, A. Carenini and I. Celino, Modelling business agreements in the multimodal transportation domain through ontological smart contracts, *CoRR* (2022). arXiv:2209.05463. doi:10.48550/arXiv.2209.05463.
[27] F.R. Soriano, J.J. Samper-Zapater, J.J. Martinez-Dura, R.V. Cirilo-Gimeno and J.M. Plume, Smart mobility trends: Open data and other tools, *IEEE Intelligent Transportation Systems Magazine* **10**(2) (2018), 6–16. doi:10.1109/MITS.2017.2743203.
[28] B. Van de Vyvere, P. Colpaert, E. Mannens and R. Verborgh, Open traffic lights: A strategy for publishing and preserving traffic lights data, in: *Companion Proceedings of the 2019 World Wide Web Conference*, 2019, pp. 966–971. doi:10.1145/3308560.3316520.
[29] P.-Y. Vandenbussche, G.A. Atemezing, M. Poveda-Villalón and B. Vatant, Linked open vocabularies (LOV): A gateway to reusable semantic vocabularies on the web, *Semantic Web* **8**(3) (2017), 437–452. doi:10.3233/SW-160213.