Evolution of DATEX II Standard towards Open Data

A Case Study on LOD-RoadTran18 project

Ana M. Delgado
IRTIC, Universitat de València, Valencia, Spain
ana.maria.delgado@uv.es

J. Javier Samper-Zapater
Departamento de Informática, Universitat de València, Valencia, Spain
Jose.J.Samper@uv.es

Silvia Del Campo
IRTIC, Universitat de València, Valencia, Spain
silvia.campo@uv.es

Juan J. Martínez Durá
Departamento de Informática, Universitat de València, Valencia, Spain
Juan.Martinez-Dura@uv.es

J. Macario Rocha
IRTIC, Universitat de València, Valencia, Spain
Jose.M.Siqueira@uv.es

José F. Garcia Calderaro
Departamento de Informática, Universitat de València, Valencia, Spain
j.francisco.garcia@uv.es

ABSTRACT

The DATEX II standard, originally conceived for traffic information exchange among control centers and road operators, has undergone a transformative journey, emerging as a fundamental European standard for open traffic data. Propelled by the ITS Directive, which positions DATEX II as the preferred language for traffic information publication at National Access Points, it stands as a cornerstone in the realm of open traffic data. This article explores the transformation of DATEX II, with a specific focus on the exemplary LOD-RoadTran18 project.

The analysis within this article intricately explores the significance of profiling DATEX II publications, emphasizing its pivotal role in augmenting comprehension and expanding usability for a broader audience. Furthermore, it sheds light on the strategic implementation of semantic enrichment, specifically through the LOD-DATEX format, as an indispensable tool in the open data ecosystem. This semantic enrichment not only enhances the reusability of DATEX II data but also acts as a bridge, fostering seamless integration with a diverse array of datasets beyond the domain of road traffic.

CCS CONCEPTS

• **Information systems** → World Wide Web; Web data description languages; Markup languages; Extensible Markup Language (XML); World Wide Web; Web data description languages; Semantic web description languages; Resource Description Framework (RDF); • **Applied computing** → Operations research; Transportation; Computers in other domains; Computing in government; E-government.

KEYWORDS

DATEX II, OPEN DATA, TRAFFIC INFORMATION, DATEX II PROFILES, LOD, ITS, WEB SEMANTIC

EATIS 2024, July 03–05, 2024, Praia, Cape Verde

ACM ISBN 979-8-4007-1733-8/24/07
https://doi.org/10.1145/3685243.3685292

ACM Reference Format:
Ana M. Delgado, J. Javier Samper-Zapater, Silvia Del Campo, Juan J. Martínez Durá, J. Macario Rocha, and José F. Garcia Calderaro. 2024. Evolution of DATEX II Standard towards Open Data: A Case Study on LOD-RoadTran18 project. In *12th Euro American Conference on Telematics and Information Systems (EATIS 2024), July 03–05, 2024, Praia, Cape Verde*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3685243.3685292

1 INTRODUCTION

The evolution of DATEX II standard [1] stands as a testament to the dynamic nature of traffic information management in Europe. Originating as a protocol primarily designed for traffic control centers to exchange information, DATEX II underwent a significant transformation with the advent of the ITS Directive 2010/40/EU [2]. Initially focused on ensuring interoperability and detailed data representation for control centers, it has now emerged as a de facto standard for open traffic data. This evolution, driven by directives such as ITS, reflects DATEX II's responsiveness to emerging needs, especially in the realm of open data.

The ITS Directive, endorsing DATEX II as the preferred language for traffic information publication at National Access Points (NAP), marked a pivotal moment. DATEX II transitioned from a niche protocol for control centers to a foundational standard for open traffic data. This shift not only expanded its technical role but also redefined its purpose, making DATEX II a crucial component in enhancing the accessibility, transparency, and usability of traffic-related information across diverse applications.

This article focuses on DATEX II's transformation, highlighting the LOD-RoadTran18 project [3] as a representative use case. As DATEX II expands beyond control centers, profiling DATEX II publications becomes critical for comprehension and increased utilization. The article underscores semantic enrichment through the LOD-DATEX format (LOD for Linked Open Data) as a crucial tool, enhancing reusability and enabling seamless integration with diverse datasets beyond road traffic.

In this Introduction, the main idea of the article is displayed, as well as the basic open data concepts and attributes. In Section 2, the LOD-RoadTran18 case study is analyzed, focusing in two main aspects: DATEX II profiling and semantic enrichment. Section 3 proposes future directions and finally the conclusions are presented.

1.1 Open Data Definition

According to the Open Data Handbook, the full Open Definition [4] summarizes the most important characteristics of such data: availability and access, re-use and redistribution and universal participation. In this article, we explore two aspects of the project LOD-RoadTran18 which contribute to these items, such as DATEXII profiling (enhances Availability and access and Universal Participation) an DATEXII LOD enrichment (facilitates Re-use and Redistribution and Universal Participation), and how they convert the narrow original purpose of DATEX II format into a broad scope format that encourages, above all, the Universal Participation principle.

1.2 Data Typology and Classification

Mahajan [5] introduces a typology to classify emerging transport data based on the concentric shell model by Backx [6], which was also used by Loenen and Grothe [7]. This model defines the attributes needed for a dataset to be considered as open data, which are the following: to be known, legally attainable, accessible, affordable, usable and distributable. The typology aids in clarifying the distinctions between public data, open data, and other categories, providing a nuanced understanding of the data landscape for transport modeling.

Building on said typology, in this article we want to propose the idea that DATEXII publications served on the National Access Point as ordered by the ITS Directive, can be considered as open data, as they accomplish mentioned attributes. Therefore, the proposed classification aligns with DATEX II's evolution towards an open data nature, according with the proposal of this article. In the following sections, the case study shows the two aspects in the project that reinforce this idea: profiling and semantic enrichment.

2 CASE STUDY: LOD-ROADTRAN18 PROJECT

The LOD-RoadTran18 project [3] was a CEF (Connecting Europe Facility) Action funded by the European Commission and coordinated by the authors of this paper. It stands as an exemplary case study, offering intricate insights into the real-world implementation of DATEX II within the realm of open traffic data.

The primary aim of the LOD-RoadTran18 project was to bridge the existing gap between DATEX II and the open data ecosystem. The undertaking involved the development of specialized mechanisms for DATEX II profiling, semantic enrichment in the LOD-DATEX format, and the creation of interoperable datasets. The case study prominently underscores the adaptability of DATEX II and its pivotal role in influencing and shaping open traffic data standards.

The "*Guidelines on recommended standard licences, datasets and charging for the reuse of documents*" (Commission notice 2014/C 240/01) [8] Category TRANSPORT & datasets "traffic information", 3.2 point recommends:

"*To facilitate the use of data in the public sector while significantly increasing the value of datasets for subsequent re-use, it is recommended that datasets be: [. . .] d) published in machine-readable and open formats (CSV, JSON, XML, RDF, etc.) to enhance accessibility; e) described in rich metadata formats and classified according to standard vocabularies (DCAT, EUROVOC, ADMS, etc.) to facilitate searching and interoperability; [. . .]*"

The **ITS Directive (2010/40/EU) [**2**]** and subsequent delegated regulations establish the necessary specifications to ensure compatibility, interoperability and continuity of the implementation and operational use of data and procedures to facilitate, as far as possible, the following types of information or priorities: EU-wide multimodal travel information services, EU-wide real-time traffic information services, free safety-related Minimum Traffic Information, an interoperable EU-wide eCall, information services and reservation services for safe and secure parking places for trucks and commercial vehicles. They establish the obligation to create a **National Access Point** with information that allows reuse of (public and private) data available in each country to develop potential markets. These regulations establish requirements to ensure the availability, dissemination, exchange and re-use of data.

Taking into account Commission notice 2014/C 240/01 guidelines and ITS Directive, activity 5 of the LOD-RoadTran18 project carried out an analysis of the implementation of the ITS Directive [2] in the Spanish NAP [9]. As a result of said analysis, some points of improvement for the Spanish NAP were showcased:

- An appropriate system with metadata management capability to allow discovery and access services through APIs.
- Disseminate data in LOD formats (RDF, OWL, etc.) because only XML formats were used currently on the Spanish NAP.
- Using not only metadata for describing datasets, but metadata for datasets catalogue.
- To extend the publication of other kinds of information through European Data Portal (EDP) [10] at least.

These results show the crucial role of the ITS Directive in the evolution of DATEX II to become an open data format. Because it is from these recommendations from where the development of the LOD-DATEXII comes.

In parallel, as a natural consequence of the ITS Directive, the European NAPCORE [11] project emerges to achieve the harmonization of the National Access Points, as they are defined by the mentioned directive. It must be noted the importance of reference profiles among the main goals of NAPCORE project.

2.1 DATEX II profiling for enhanced comprehension

According to the DATEX II website [12], the EU Delegated Regulations are supported by Recommended Reference Profiles (RRP's) which contain the minimum set of data-elements required to provide the information needed by each specific EU Delegated Regulation on ITS.

On the other hand, the EU-EIP [13] reference handbook has described harmonized uses cases for the provision of data from the road operator's point of view. These use-cases are supported by so called Recommended Service Profiles (RSP's). These RSP's are based on one or more RRP's and extended with data-elements described in the use case. It is notorious the effort put in defining this framework of Recommended Profiles in DATEXII standard, and it demonstrates the evolution of the standard to boost its usability. This evolution of DATEX II is particularly highlighted in the LOD-RoadTran18 [3] project, which underscores the increasing significance of profiling DATEX II publications.

None of the Spanish DATEX II NAP publications uses a profile to particularize their content. This has been identified as an improvement to make when updating the publications to DATEX II v3.0. There is an initiative by DGT to upgrade to DATEX II version 3.0, including the development of one or several custom DGT profiles. There was also a DATEX II v3.0 SRTI profile (Safety Related Traffic Information), even though this was not defined as a CEN standard. Therefore, LOD-RoadTran18 focused in SRTI following the ITS Directive and developed the ontology LOD-DATEX over the SRTI profile of DATEX II version 3.0 with more than a thousand elements.

Commission delegated regulation (EU) No 2013/886 [14] was the primary source of truth, even though it had to be elaborated into details of models actually used. Article 3 of said regulation "*List of road safety-related events or conditions*" lists categories of safety-related traffic information. Consequently, this categories of events considered as SRTI by the Commission delegated regulation (EU) No 2013/886 were taken in the project as the basis profile on which to develop the ontology.

As we can see from above, profiling DATEX II allows for the customization of data representations, tailoring them to meet the diverse needs of a broader audience and ensuring comprehension and usability for various stakeholders. This not only promotes broader adoption but also aligns DATEX II with the principles of open data accessibility.

2.2 Semantic enrichment in LOD-DATEX format

The LOD-DATEX format not only adheres to open data principles but also facilitates the seamless integration of DATEX II information with heterogeneous datasets. The resulting publication on the Spanish NAP [9] can be observed at Figure 1.

Considering DATEX II as a structured data format derived from XML, every DATEX II publication in the NAP can be initially regarded as having 3-star data quality [15], according to the Star rating system of data. As a logical consequence, pointed out in [16], enhancing the DATEX II format with semantic markup can elevate its quality level to 5 stars.

Semantic enrichment involves imparting explicit meaning to data, rendering it more interpretable and context-aware. Within the LOD-RoadTran18 project, the adoption of the LOD-DATEX format ensures that traffic data becomes an integral part of the broader Linked Open Data ecosystem. This interconnectedness allows DATEX II data to be seamlessly linked with other datasets, extending its utility beyond the domain of road traffic. Semantic enrichment not only enhances the value of DATEX II data but also aligns with the principles of transparency, traceability, and collaboration inherent in the open data paradigm.

This aspect is also highlighted in projects like CONECTA[17], where the semantic markup of published data can result in showing the link between apparently unrelated databases, expanding the scope of possible interpretations of said data and improving decision making at all levels.

The LOD-RoadTran18 project strategically employed semantic enrichment techniques to enhance the LOD-DATEX format (Figure 2). By incorporating standardized vocabularies and ontologies such as BTN100[18], DATEX II data gained additional context, making it more interpretable for both human users and machines. This approach not only fosters data integration but also lays the groundwork for more advanced analytics and cross-domain collaborations. Therefore, the semantic enrichment of DATEX II improves cross-dataset linkages and consequently decision-making based on the result of those linkages. This empowers the creation of new services, promotes innovation, improves efficiency and effectiveness of government services, enables impact measurement of policies and creates new knowledge from combined data sources: all of them

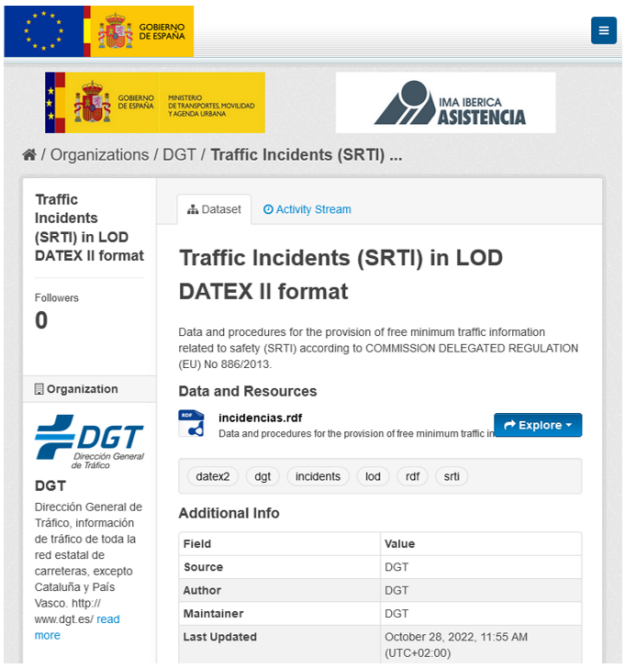

Figure 1: LOD-DATEX publication in Spanish NAP

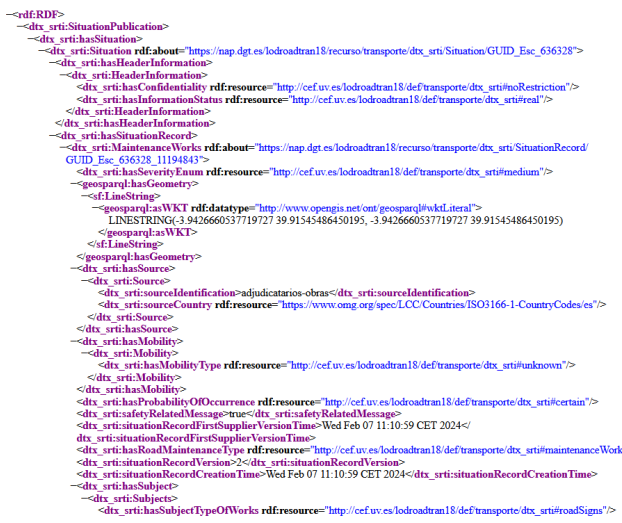

Figure 2: LOD-DATEX file fragment

areas of impact of open data according to the Open Data Handbook [4].

3 FUTURE DIRECTIONS AND CONCLUDING REMARKS

3.1 Anticipating the Future Trajectory of DATEX II both in Spain and in the rest of Europe

In the near future, the DGT will need to start a deep evolution of its DATEX systems, implying the adoption of the new version of DATEX II. This will involve the implementation of a profile and some extensions, according to the necessities identified in the LOD-RoadTran18 project.

Additionally, future extensions of the LOD-DATEX ontology are suggested. This can be developed over RTTI profile (Real Time Traffic Information), new Spanish location profile or other DATEX publications in the Spanish NAP or other European NAPs.

3.2 Collaborative Initiatives

A suggested collaboration could emerge between DGT and private service providers, in order to design a new DGT DATEXII profile which meets the necessities of the main traffic data consumers in Spain.

4 CONCLUSIONS

In conclusion, DATEX II has actively evolved, incorporating fundamental open data principles, profiling methodologies, and semantic enrichment. The tangible collaboration between DATEX II and groundbreaking projects like LOD-RoadTran18 stands as a concrete manifestation of these principles in practical, real-world applications.

ACKNOWLEDGMENTS

The content of this article is the result of the participation in the LOD-RoadTran18 project (Action No: 2018-EU-IA-0088), approved within the framework of the Call for Telecommunications CEF 2018-Open Public Data (CEF-TC-2018 -5). Our thanks to the General Directorate of Traffic (DGT) for their collaboration and help.

REFERENCES

[1] Datex II: 2024. https://datex2.eu/. Accessed: 2024-01-19.
[2] Consolidated text: Directive 2010/40/EU of the European Parliament and of the Council of 7 July 2010 on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport (Text with EEA relevance)Text with EEA relevance: 2023. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri$=$CELEX%3A02010L0040-20231220. Accessed: 2024-02-05.
[3] LOD-RoadTran18 Website: 2022. https://cef.uv.es/lodroadtran18/index.php/en/. Accessed: 2024-02-04.
[4] What is open data? 2024. http://opendatahandbook.org/guide/en/what-is-open-data/. Accessed: 2024-.
[5] Mahajan, V. *et al.* 2022. Data to the people: a review of public and proprietary data for transport models. 42, 4 (Jul. 2022), 415–440. DOI:https://doi.org/10.1080/01441647.2021.1977414.
[6] Backx, M. 2003. Gebouwen redden levens. Toegankelijkheidseisen van gebouwgegevens in het kader van de openbare orde en veiligheid. (2003).
[7] Van Loenen, B. and Grothe, M. 2014. INSPIRE Empowers Re-Use of Public Sector Information .
[8] Guidelines on recommended standard licences, datasets and charging for the reuse of documents 2014/C 240/01: 2014. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri$=$CELEX%3A52014XC0724(01). Accessed: 2024-02-05.
[9] Punto de Acceso Nacional de Tráfico y Movilidad: 2024. https://nap.dgt.es/en/. Accessed: 2024-02-07.
[10] The official portal for European data | data.europa.eu: 2024. https://data.europa.eu/en. Accessed: 2024-02-07.
[11] Mylonas, C. *et al.* 2023. Harmonization of National Access Points to Intelligent Transport Systems data: A data content and added value perspective. Elsevier BV.
[12] DATEX II Reference Profiles: 2024. https://docs.datex2.eu/reference_profiles/index.html. Accessed: 2024-02-05.
[13] European platform for road operators using Intelligent Transport Systems (ITS): 2021. https://www.its-platform.eu/. Accessed: 2024-02-05.
[14] COMMISSION DELEGATED REGULATION (EU) No 886/2013: 2013. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri$=$CELEX:32013R0886&from$=$BG. Accessed: 2024-02-05.
[15] 5-star Open Data: 2015. http://5stardata.info/en/. Accessed: 2024-02-04.
[16] Najdenov, B. *et al.* 2015. Automated Linked Data Generation from the Transport Administration Domain. (Nov. 2015).
[17] Soriano, F.R. *et al.* Smart Mobility Trends: Open Data and Other Tools | IEEE Journals & Magazine | IEEE Xplore. DOI:https://doi.org/https://doi.org/10.1109/MITS.2017.2743203.
[18] Gutierrez, J. *et al.* 2022. LOD-RoadTran18: Supporting the cross-border use of Road Traffic Data with Linked Open Data based on DATEX II. ACM.