

Article

Not peer-reviewed version

Semantic Modelling Approach for Safety-Related Traffic Information using DATEX II

[J. Javier Samper-Zapater](#)*, [Julian Gutiérrez-Moret](#), Jose Macario Rocha, Juan José Martínez-Durá, Vicente R. Tomás

Posted Date: 10 November 2023

doi: 10.20944/preprints202311.0671.v1

Keywords: Safety Road Traffic; Semantic web; LOD



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Semantic Modelling Approach for Safety-Related Traffic Information Using DATEX II

J. Javier Samper-Zapater ^{1,*} , Julián Gutiérrez-Moret ¹, Jose Macario Rocha ¹,
Juan José Martínez-Durá ¹ and Vicente R. Tomás ² 

¹ IRTIC - Research Institute on Robotics and Information and Communication Technologies. Universitat de València. Spain. jsamper@irtic.uv.es, jgutierrez@irtic.uv.es, Jose.M.Siqueira@uv.es, Juanjo@irtic.uv.es

² Department of Computer Science and Engineering, Universitat Jaume I. Castellón de la Plana, Castellón; Spain. vtomas@uji.es

* Correspondence: jsamper@irtic.uv.es

Abstract: The significance of Linked Open Data datasets for traffic information goes beyond just including open traffic data. It extends to incorporating links to other relevant thematic datasets available on the web. This enables federated queries across different data platforms from various countries and sectors, such as transport, geospatial, environmental, weather, and more. Businesses, researchers, national operators, administrators, and citizens at large can benefit from having dynamic traffic open data connected to heterogeneous datasets across different Member States. This paper focuses on the development of a semantic model that enhances the basic service of accessing open traffic data through a LOD-enhanced Traffic Information System in alignment with the ITS Directive (2010/40/EU). The objective is not limited to just viewing or downloading data but also to improve the extraction of meaningful information and enable other types of services that are only achievable through LOD. By structuring the information using the RDF format meant for machines and employing SPARQL for querying, LOD allows for comprehensive and unified access to all datasets. Considering that the European standard DATEX II is widely used in many priority areas and services mentioned in the ITS Directive, LOD DATEX II was developed as a complementary approach to DATEX II XML. This facilitates the accessibility and comprehensibility of European traffic data and services. As part of this development, an ontological model called *dtx_srti*, based on the DATEX II Ontology, was created to support these efforts.

Keywords: Safety Road Traffic; Semantic web; LOD

1. Introduction

Domain ontologies express conceptualizations that are specific to a particular domain [1]. In transportation, many solutions have emerged. For example, in [2] a vehicle accident ontology has been defined, in [3] an ontology for traffic management and control has been defined and tested in a multiagent system. In [4], an implementation of a semantic web service discovery system for road traffic information has been developed. Details of existing traffic ontologies are analyzed in [5].

In distributed environments, such as the transportation, data is the key element for developing Intelligent Transport Systems (ITS). The specification of DATEX II XML/UML [6] provides a description of concepts and data structures pertaining to traffic. However, it is important to note that this description is primarily focused on the syntactic aspects and does not incorporate semantic meaning. Therefore, there is a necessity for a semantic model that describes the contents of DATEX II elements to facilitate linking this data with other vocabularies and ontologies available on the internet. Linked Open Data (LOD) emerges as a methodology for publishing and interlinking structured data on the web, following the principles of the Semantic Web. LOD offers a standardized framework for sharing and connecting data from diverse domains and sources, enabling data integration, enrichment, and reuse. The utility of LOD can be described through various aspects, including data integration and enrichment, data discovery, data reuse, interoperability, and data quality.

Its usefulness has been demonstrated in a variety of domains, such as healthcare [7,8] or education [9,10].

In the transportation domain, the use of non-LOD formats produces some interoperability problems. For example, in the European project *CROCODILE corridor* [11], to make possible the exchange information between operators, road administrators and end-users, a middleware was created to provide translation of differently *situation_publication*¹ types between the different national implementations of DATEX II. However, the middleware requires to be upgraded when new data concepts are added to any DATEX implementation, making this solution not efficient nor viable. So, we consider that the use of LOD would be a much more efficient solution, given that sharing a same semantic model (Ontology for DATEX II) could avoid the interoperability problems since it avoids semantic ambiguity of specific terms.

In the CEF Action LOD-RoadTran18 with number 2018-EU-IA-0088 [12], our primary aim was to facilitate the effective reuse of real-time road traffic data in the Czech Republic and Spain. To achieve this objective, we developed the DATEX II Safety Road Traffic Information (SRTI) ontology called *dtx_srti*. This ontology, built on the foundation of DATEX II, underwent rigorous testing on diverse datasets using SPARQL queries at a small scale [13]. In addition, we invested considerable effort in creating tools for model conversion, adaptation, and validation, with specific attention given to the geospatial characteristics of the traffic data.

In order to accomplish our goals, we thoroughly analyzed existing DATEX II models, LOD general concepts, as well as relevant vocabularies and models related to traffic information. This paper provides comprehensive details on the development of the semantic model, which consists of a main ontology and two secondary ontologies. We delve into the most pertinent aspects, providing elaborate explanations of the decisions made regarding the current and potential future usage of vocabularies and datasets. The paper covers the definition of concepts, their interrelationships, and the utilization of individual instances instead of the generic and specific data types found in the UML/XML schema of DATEX II version 3.2 [6].

The purpose of this paper is to provide a detailed exposition of the *dtx_srti* ontology, offering a semantic description of its contents based on the UML data model developed as part of the CEF Action.

The main contributions of this research are summarized as follows:

- A groundbreaking semantic modeling approach for traffic information was introduced through the development of a novel ontology known as *dtx_srti* [14]. This ontology, accompanied by a set of secondary ontologies, serves as a comprehensive semantic vocabulary specifically designed to represent the SRTI DATEX II profile [15] in accordance with the Commission Delegated Regulation (EU). The primary purpose of this vocabulary is to facilitate seamless mapping between DATEX II version 3.2 and Linked Open Data (LOD) formats, enabling efficient interoperability and data exchange. By leveraging these semantic resources, a significant step forward has been taken in advancing the representation and integration of traffic information within the context of dynamic data utilization.
- Regarding the implementation aspect, the ontologies put forward in this study were constructed using RDF/OWL (Resource Description Framework/Web Ontology Language) [16]. In order to enhance the depth and breadth of knowledge representation, these ontologies have been interconnected with relevant external ontologies. To consolidate data in the RDF standard [17], a set of mapping functions has been devised, enabling automatic storage in a shared RDF repository and Endpoint service. Building upon this foundation, an array of sophisticated SPARQL queries has been formulated, designed as an API service. Furthermore, to foster widespread adoption

¹ A *situation_publication* is a identifiable instance of a traffic/travel situation comprising one or more traffic/travel circumstances which are linked by one or more causal relationships.

within the research community, a user interface has been developed, streamlining the utilization of these resources. This user-friendly interface aims to facilitate seamless exploration and interaction with the ontologies, empowering researchers to harness their full potential.

- To ensure the robustness and accuracy of the developed ontologies, a comprehensive validation process was undertaken. Multiple tests were conducted, involving various queries on traffic incident data, administrative units, and road information. Furthermore, federated queries were designed to establish connections between the data generated by our LOD Converter and other SPARQL Endpoints, including the esteemed Spanish National Geographic Institute (IGN) [18]. These tests not only explored the potential of leveraging linked data from diverse datasets but also assessed the intricacies and challenges associated with executing complex SPARQL queries. In doing so, the tests aimed to identify and rectify common errors and shortcomings, contributing to the refinement and optimization of the overall system [19].

This paper is organized as follows: Section 2 highlights the required legal and technical aspects related with the research. Section 3 describes the semantic approach, focusing on the OWL Ontology. The definition and implementation of the Linked Open Traffic Data Model is described on Section 4. The procedure to validate this approach is described in Section 5. Finally, Section 6 concludes with key remarks and future works. All this content not only provide a picture of semantic modelling within Safety Traffic Data but also an example of research and implementation that can be partially replicated for the creation and analysis of other ontologies.

2. Exploration of legal and technical framework for enabling actionable linked open traffic data.

The initial phase of the research involved a thorough examination of the regulatory landscape encompassing the Public Sector Information (PSI) and Intelligent Transport Systems (ITS) Directives, both at the European and national levels. The primary objective was to identify the key legal and procedural considerations that should guide the development of a semantic model and subsequent architecture. To accomplish this, the analysis was divided into two distinct tasks: 1) Scrutinizing legal and procedural aspects associated with the PSI Directive, which governs the access and reuse of public sector information and 2) Evaluating legal and procedural aspects relevant to the ITS Directive, which outlines the framework for intelligent transport systems.

By delving into these areas, the research sought to establish a comprehensive understanding of the regulatory landscape, providing essential insights to inform the subsequent development of a robust semantic model and architecture.

2.1. Analysis of the Public Sector Information Directive

The primary objective of this task was to conduct a comprehensive review of the existing legislation pertaining to Public Sector Information (PSI) at both the European and national levels. This analysis encompassed key documents and related materials, including:

- The Open Data Directive (Directive (EU) 2019/1024), as well as the previous PSI Directives 2003/98/EC and its later revision 2013/37/EU.
- Commission Notice 2014/C 240/01, titled "Guidelines on recommended standard licenses, datasets, and charging for the reuse of documents."
- Directive 2007/2/EC, which pertains to the establishment of an Infrastructure for Spatial Information in Europe (INSPIRE).

Additionally, the analysis involved an examination of open data initiatives, such as the APORTA Initiative of the Spanish Government [20], action plans, national strategies in each country, and the initial assessment of compliance levels.

2.2. Analysis of the Intelligent Transport Systems Directive

This task involved a comprehensive analysis of the norms outlined in the Intelligent Transport Systems (ITS) Directive, specifically focusing on the publication of traffic data. The analysis

encompassed the types of data to be published, the manner in which such data should be made available, and the implementation of the National Access Point (NAP) for road traffic information to meet the specified requirements.

A review of the state of the art was conducted, taking into consideration the ITS Directive 2010/40/EU [21], subsequent delegated regulations, associated legislation, and relevant standards. Notably, insightful information from analytical documents of the EU EIP project [22], which focused on European harmonization within the ITS Directive, was also taken into account.

The analysis concluded that the ITS Directive and subsequent delegated regulations provide the necessary specifications to ensure compatibility, interoperability, and continuity in the implementation and operational use of data and procedures. These specifications aim to facilitate the following types of information or priorities:

1. Provision of EU-wide multimodal travel information services.
2. Provision of EU-wide Real-Time Traffic Information (RTTI) services.
3. Data and procedures for the provision of free Safety-Related minimum Traffic Information (SRTI).
4. Harmonized provision for an Interoperable EU-wide eCall.
5. Provision of information services for safe and secure parking places for trucks and commercial vehicles.
6. Provision of reservation services for safe and secure parking places for trucks and commercial vehicles.

3. Semantic issues

A comprehensive analysis was conducted to explore semantic aspects, focusing on metadata approaches within the contexts of Intelligent Transport Systems (ITS) and Open Data. This research aimed to investigate the rules and standards implemented in different countries. For instance, the Interoperability Technical Standard (NTI-RISP) [23] was thoroughly examined in Spain. This standard aims to establish common conditions for the selection, identification, description, format, use, and provision of public sector documents and information resources. It forms a part of the Spanish National Interoperability Scheme [24].

NTI-RISP utilizes its own controlled vocabulary and defines specific metadata elements, some of which are mandatory. It also establishes various taxonomies for specific metadata. For example, it adopts the W3C Time Ontology [25] to specify time-related property values and incorporates standards such as ISO-8601 (date-time) and RFC4646 (tags for identifying languages). Furthermore, NTI-RISP outlines the acceptable values for the "Geographic coverage" property, which involve the use of identifiers related to geographic resources within the Spanish territory at the level of autonomous communities, cities, and provinces [26].

Additionally, in the case of DCAT-AP [27], the URIs are defined based on the Languages Name Authority List [28] provided by the Publications Office of the EU. Furthermore, the Catalogue of Single Point Access Coordinated Metadata [29], recommended by the EIP project [22] for the implementation or enhancement of any National Access Point (NAP), suggests the use of NUTS 0-3 [30] levels (Nomenclature of Territorial Units for Statistics) for the "area covered by publication."

During the analysis, it was also highlighted that NTI-RISP includes 22 primary sectors (listed in Annex IV of the regulation) [31], with one of them being the "Transporte" (Transport) concept that encompasses communication and traffic-related issues.

4. Development and implementation of a comprehensive Linked Open Traffic Data Model

Upon establishing the project's initial scope, a thorough analysis of the DATEX II model was conducted, with a primary focus on the Safety Related Traffic Information (SRTI) profile [15]. This profile represents a subset of the complete DATEX II model (version 3.2). However, after extensive discussions, it was decided to extend the model beyond the SRTI profile by introducing new terms and relationships. This expansion allows the model to accommodate various types of information beyond

However, in our model, we have chosen to retain this attribute, as it is utilized in the full DATEX model to specify the end of a period if necessary (with a multiplicity of 0..1). Furthermore, we have adopted the W3C Time ontology [25] for temporal properties, as recommended by NTI-RISP. This ontology allows us to define properties with values related to temporal aspects.

Table 2. Features of the “Validity” package. Source “4b Classes and enumerations SRTI profile” document.

Class name	OverallPeriod
Attribute End	overallStartTime
Designation	Validity time specification
Definition	Represents the specific date and time that marks the beginning of the bounding period of validity.
Multiplicity	1..1
Type	DateTime

Figure 2 illustrates a portion of the UML schema related to validity in the DATEX II 3.2 model.

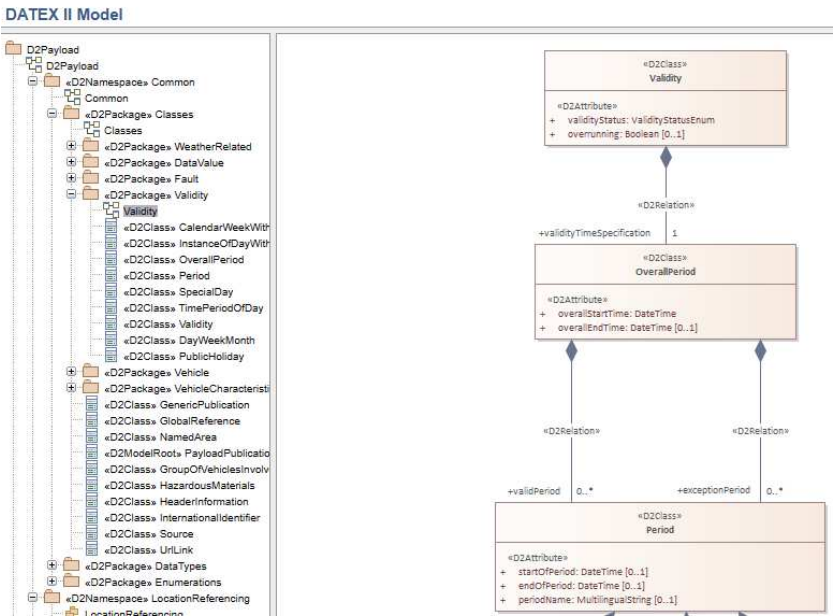


Figure 2. Validity Class. Source DATEX II Model (datex2.eu).

Regarding the DATEX Model version 3.2, the properties within the Overall Period had specific data type restrictions as follows:

- *overallEndTime* max 1 xsd:dateTime
- *overallStartTime* exactly 1 xsd:dateTime

To specify these restrictions within the model, several steps were taken. Firstly, the TIME W3C ontology [25] was imported, taking into consideration the indications provided by [35] regarding intervals: "Proper intervals are intervals whose extremes are different. Among other things, this allows using standard interval calculus and defining relations between intervals." By importing TIME, the *dtx_srti:OverallPeriod* was defined as a Temporal Entity, specifically a *time:ProperInterval*.

By designating the *dtx_srti:OverallPeriod* as a *time:ProperInterval*, instances of it were created with the properties *time:hasBeginning* and *time:hasEnd*, which have a range of *time:Instant*. These properties accurately specify the beginning and end of the validity period for a Situation Record (*dtx_srti:SituationRecord*). The related definitions in our model are presented in Listing 1.

Listing 1: Definitions and relationships for dtx_srti:Validity concept

```
##Validity
dtx_srti:Validity dtx_srti:validityTimeSpecification exactly 1 dtx_srti:OverallPeriod
dtx_srti:Validity dtx_srti:hasValidityStatus exactly 1 dtx_srti:ValidityStatusEnum

##OverallPeriod
dtx_srti:OverallPeriod rdfs:subClassOf time:ProperInterval
dtx_srti:OverallPeriod rdfs:subClassOf time:hasBeginning exactly 1 time:Instant
dtx_srti:OverallPeriod rdfs:subClassOf time:hasEnd max 1 time:Instant

## ValidityStatusEnum
dtx_srti:ValidityStatusEnum rdfs:subClassOf dtx_srti:PayLoadEnumerations
dtx_srti: ValidityStatusEnum owl:equivalentTo {active , definedByValidityTimeSpec , planned , suspended}
```

Furthermore, in accordance with the W3C Time Ontology, we can apply various interval relations between time periods within our instances if needed. Figure 3 depicts the dtx_srti:Validity class and its relationships using the Protégé software tool.

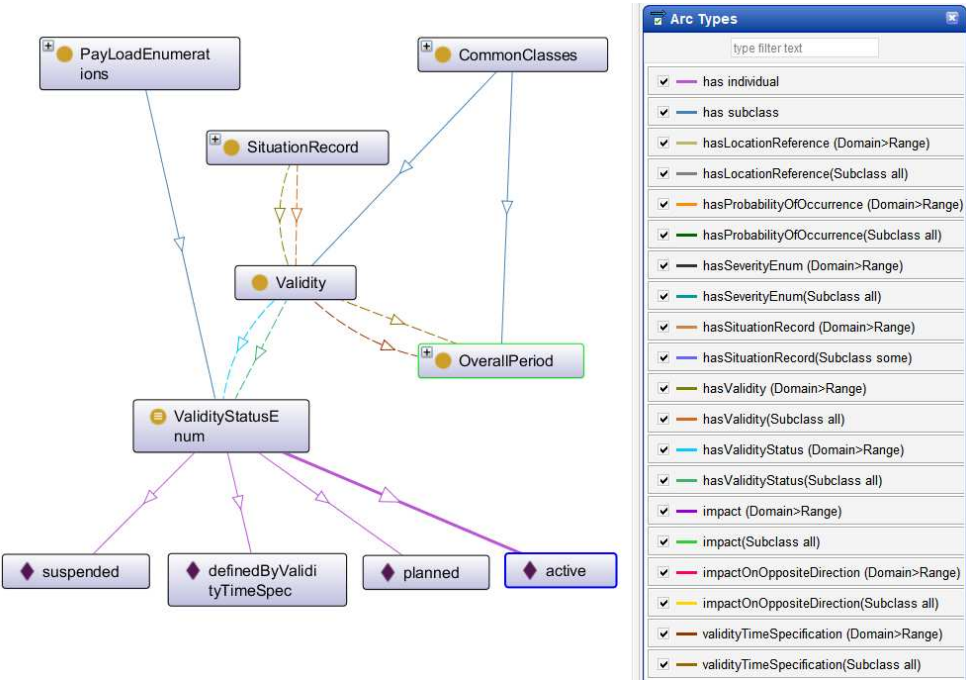


Figure 3. Representation of the Validity Class and its relationships in the dtx_srti main ontology (Caption in Protégé-Ontograf).

Figure 4 illustrates how the property *hasValidity* allows for the establishment of a validity interval for a *SituationRecord*. In addition, Listing 2 presents an example of a partial representation of an instance, considering a *SituationRecord* with a specified validity interval. For this, the use of terms from the W3C TIME ontology is crucial. For example, the term "Instant" will allow us to specify the beginning and end of any interval, as specified for the *dtx_srti:Validity* concept (Listing 1).

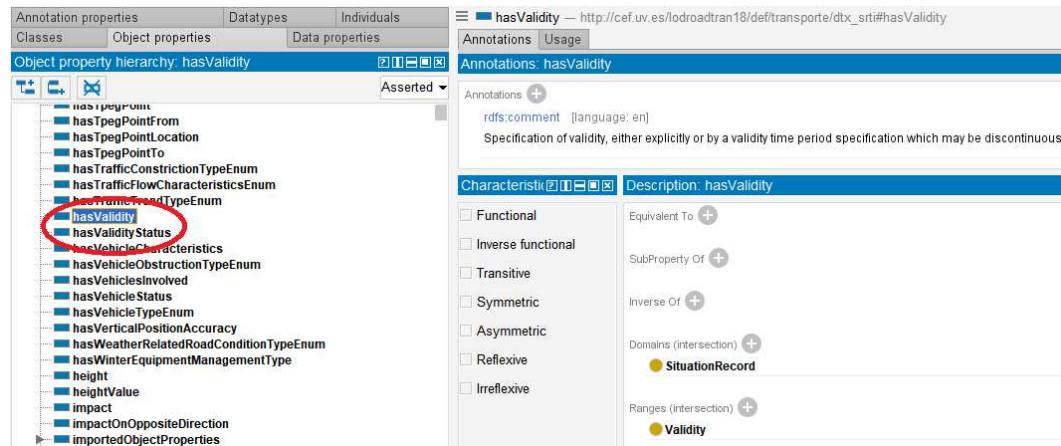


Figure 4. Representation of the hasValidity Object Property and Its Relationships in the dtx_srti Main Ontology

Listing 2: SituationRecord Individual with validity intervals

```
<dtx_srti:hasValidity>
<dtx_srti:Validity>
<dtx_srti:validityTimeSpecification>
  <dtx_srti:OverallPeriod>
    <time:hasEnd>
      <time:Instant>
        <time:inXSDDateTime
          rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">
            2021-11-30T17:00:00.019Z
        </time:inXSDDateTime>
      </time:Instant>
    </time:hasEnd>
    <time:hasBeginning>
      <time:Instant>
        <time:inXSDDateTime
          rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">
            2021-11-30T07:00:00.019Z
        </time:inXSDDateTime>
      </time:Instant>
    </time:hasBeginning>
  </dtx_srti:OverallPeriod>
</dtx_srti:validityTimeSpecification>
</dtx_srti:Validity>
</dtx_srti:hasValidity>
```

The *SituationRecord* specification offers a robust framework and provides the capability to query the model using SPARQL, allowing us to retrieve information based on the following criteria:

- Retrieve information about a *SituationRecord* with an *OverallPeriod* that starts at the time "2022-01-04 T 08:27:04.963 + 01:00" (refer to Listing 3).
- Retrieve information about a *SituationRecords* that started after a specific date, indicated by an *OverallStartPeriod* greater than "2021-01-04 T 08:26:04.963 + 01:00". The results should be sorted in ascending order by the *OverallStartRecord* (refer to Listing 4).

Listing 3: TIME SPARQL query 1

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX dtx_srti: <http://cef.uv.es/lodtran18/def/transporte/dtx_srti#>
PREFIX time: <http://www.w3.org/2006/time/>
SELECT DISTINCT ?situationRecord ?PO ?Validity ?SRM ?CT ?FSVT ?V ?VT ?OverallPeriod ?instant
WHERE {
  ?situationRecord a dtx_srti:SituationRecord;
    dtx_srti:hasProbabilityOfOccurrence ?PO;
    dtx_srti:hasValidity ?Validity;
    dtx_srti:safetyRelatedMessage ?SRM;
    dtx_srti:situationRecordCreationTime ?CT;
    dtx_srti:situationRecordFirstSupplierVersionTime ?FSVT;
    dtx_srti:situationRecordVersion ?V;
    dtx_srti:situationRecordVersionTime ?VT.
  ?Validity dtx_srti:validityTimeSpecification ?OverallPeriod .

  ?OverallPeriod time:hasBeginning ?instant .
  ?instant time:inXSDDateTime "2022-01-04T08:27:04.963+01:00"^^xsd:dateTime }

```

Listing 4: TIME SPARQL query 2

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX dtx_srti: <http://cef.uv.es/lodtran18/def/transporte/dtx_srti#>
PREFIX time: <http://www.w3.org/2006/time/>
SELECT distinct ?situationRecord ?ProbabilityOfOccurrence ?safetyRelatedMessage ?CreationTime
?FirstSupplierVersionTime ?Version ?VersionTime ?OverallStartPeriod
WHERE {
  ?situationRecord a dtx_srti:SituationRecord;
    dtx_srti:hasProbabilityOfOccurrence ?ProbabilityOfOccurrence;
    dtx_srti:hasValidity ?Validity;
    dtx_srti:safetyRelatedMessage ?safetyRelatedMessage;
    dtx_srti:situationRecordCreationTime ?CreationTime;
    dtx_srti:situationRecordFirstSupplierVersionTime ?FirstSupplierVersionTime;
    dtx_srti:situationRecordVersion ?Version;
    dtx_srti:situationRecordVersionTime ?VersionTime.
  ?Validity dtx_srti:validityTimeSpecification ?OverallPeriod .
  ?OverallPeriod time:hasBeginning ?instant .
  ?instant time:inXSDDateTime ?OverallStartPeriod

  FILTER (?OverallStartPeriod > "2021-01-04T08:26:04.963+01:00"^^xsd:dateTime). }
ORDER BY ?OverallStartPeriod

```

4.2. Units of Measure

In order to express units of measure, we conducted a review of the main vocabularies pertaining to this topic, including the Ontology of units of Measure (OM) [36], Quantities, Units, Dimensions and Data Types (QUDT) [37], Custom Data Type (CDT) [38], and others. After careful consideration, we selected the QUDT vocabulary. QUDT consists of a collection of vocabularies that represent various quantity and unit standards, and it offers solutions to problems in this domain. We utilized the QUDT ontology, specifically focusing on the class *qudt:QuantityValue*, which is defined as a value representing a quantity numerically in relation to a chosen unit of measure.

Additionally, the *qudt:unit* is an *ObjectProperty* with a range of *qudt:Unit*. It serves as a reference to the unit of measure associated with a quantity of interest, whether it is a variable or a constant. The concept of *qudt:QuantityValue* imposes a restriction on the "unit" property, allowing it to have exactly one unique value from the general concept *owl:Thing*. Moreover, the vocabulary *qudt-unit* [39] contains numerous units of measure, each of which is a subtype of *qudt:Unit*. For instance, Listing 5 illustrates the specification of the *unit:KiloM* (kilometre) unit.

Listing 5: qudt-unit:KiloM specification

```
unit:KiloM
  a qudt:DerivedUnit ;
  a qudt:Unit ;
  dct:terms:description "A common metric unit of length or distance. One kilometer equals exactly 1000 meters, about
0.621 371 19 mile, 1093.6133 yards, or 3280.8399 feet. Oddly, higher multiples of the meter are rarely used; even the
distances to the farthest galaxies are usually measured in kilometers. "^^rdf:HTML ;
  qudt:allowedUnitOfSystem sou:CGS-EMU ;
  qudt:allowedUnitOfSystem sou:CGS-GAUSS ;
  qudt:allowedUnitOfSystem sou:SI ;
  qudt:conversionMultiplier 1000.0 ;
  qudt:dbpediaMatch "http://dbpedia.org/resource/Kilometre"^^xsd:anyURI ;
  qudt:hasDimensionVector qkdv:A0E0L1I0M0H0T0D0 ;
  qudt:hasQuantityKind quantitykind:Length ;
  qudt:informativeReference "http://en.wikipedia.org/wiki/Kilometre?oldid=494821851"^^xsd:anyURI ;
  qudt:isScalingOf unit:M ;
  qudt:prefix prefix:Kilo ;
  qudt:symbol "km" ;
  qudt:ucumCode "km"^^qudt:UCUMcs ;
  qudt:uneceCommonCode "KMT" ;
  qudt:unitOfSystem sou:SI ;
  rdfs:isDefinedBy <http://qudt.org/2.1/vocab/unit> ;
  rdfs:label "Kilometer"@en-us ;
  rdfs:label "Kilometre"@en ;
```

Thus, in our model, the new concepts related to units of measure are subclasses of *qudt:QuantityValue*. They also impose similar restrictions on the *qudt:unit* property, but with specific ranges that correspond to individual values (units of measure from *qudt:unit*) instead of the general class *owl:Thing*. Additionally, these concepts have a *numericValue* data type property with a multiplicity restriction and a range of XSD (XML Schema Definition) types, in accordance with the DATEX II model, as depicted in Figure 5 and Table 3.

The subsequent step involves importing the individuals from the *qudt* vocabulary using the prefix *qudt:unit*. Specifically, we import the units "Meter", "Degree", "Kilometer per Hour", and "Kilometer."

Taking all the aforementioned factors into consideration, we have developed various concepts pertaining to units of measure in our model.

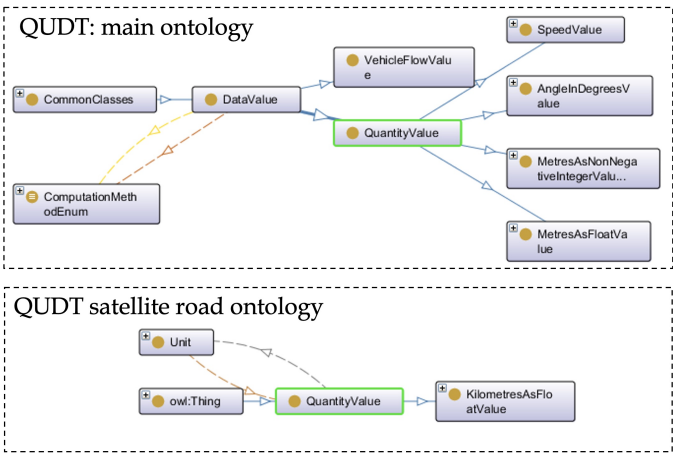


Figure 5. New classes based on QUDT in the main ontology dtx_srti. (top box) and in the satellite roads ontology (bottom box).

Table 3. List of imported individuals from QUDT (qudt-unit).

UoM	URI with prefix	Ontology	Used In:
http://qudt.org/vocab/unit			Classes
Meter	qudt-unit:M	dtx-srti	MetresAsFloatValue , MetresAsNonNegativeIntegerValue
Degree	qudt-unit:DEG	dtx-srti	AngleInDegrees
Kilometer per Hour	qudt-unit:KiloM-PER-HR	dtx-srti	SpeedValue
Kilometer	qudt-unit:KiloM	roads	KilometresAsFloatValue

Figure 6 depicts the description of the newly introduced classes. Each class is associated with two properties. The first property, *numericValue*, has a range that corresponds to a specific XSD type, while the second property, *unit*, refers to an individual from the *qudt* vocabulary.

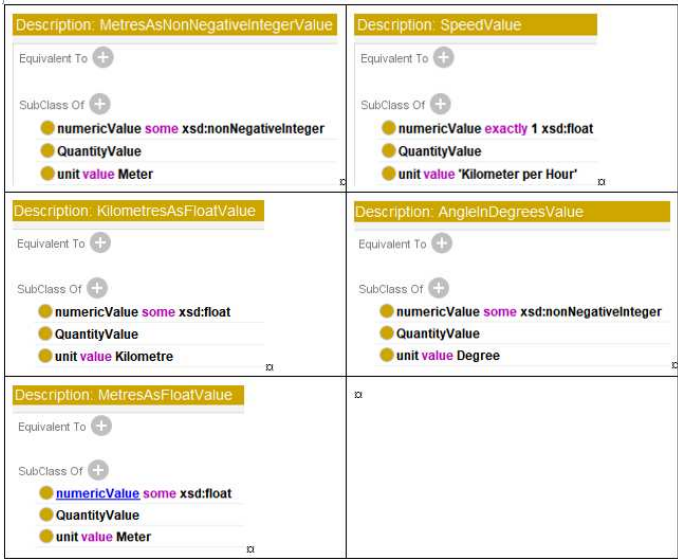


Figure 6. Description of the new classes in the semantic model, which are based on the QUDT vocabulary

In Listing 6, we have a partial instance of a *SituationRecord*, demonstrating the usage of the *qudt:numericValue* and *qudt:unit* properties to specify a distance of 90.0 meters.

Listing 6: SituationRecord Individual with Quantity Value

```
<dtx_srti:DistanceFromLinearElementStart>
  <dtx_srti:distanceAlong>
    <qudt:QuantityValue>
      <qudt:numericValue rdf:datatype=http://www.w3.org/2001/XMLSchema#float> 90.0
    </qudt:numericValue>
    <qudt:unit rdf:resource="http://qudt.org/vocab/unit#M"/>
  </qudt:QuantityValue>
</dtx_srti:distanceAlong>
</dtx_srti:DistanceFromLinearElementStart>
```

4.3. Units of GeoSPARQL

In [40], the OGC GeoSPARQL is introduced as a vocabulary that enables the representation of geospatial data in RDF. It extends the SPARQL query language to handle geospatial data and is designed to accommodate both qualitative spatial reasoning and quantitative spatial computations.

To express geographical features using this vocabulary, we imported the GeoSPARQL ontology. In our model, the *dtx_srti:SituationRecord* class was defined as a subclass of *subClassOf geosparql:Feature*. This inheritance relationship allows us to utilize the *hasGeometry* property for our concepts. Figures 7 and 8 depict the usage of the *geosparql: Feature* in our model and the subclass relationship with the *SituationRecord* concept.

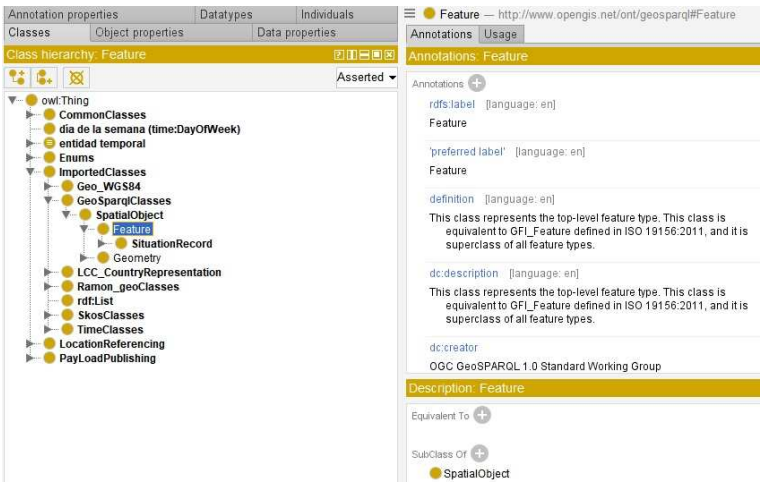


Figure 7. Use of geosparql:Feature.

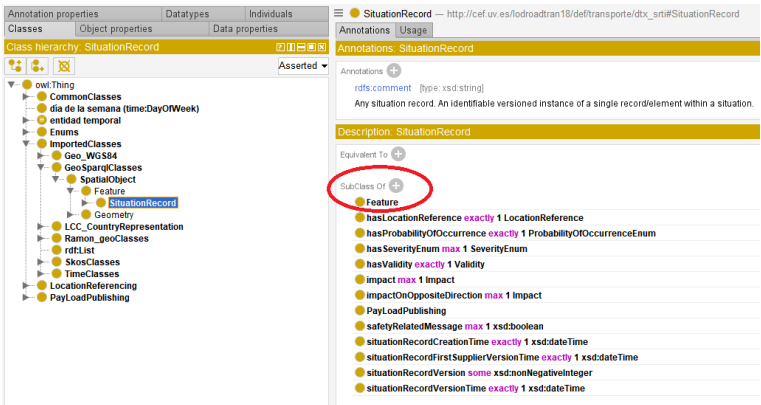


Figure 8. Definition of dtx_srti:SituationRecord using a subClassOf relationship with geosparql:Feature.

Listing 7 provides an example of an instance of a *SituationRecord*, demonstrating how a geometry feature can be added to it.

Listing 7: SituationRecord individual with geometry property

```
<owl:NamedIndividual
  rdf:about="http://cef.uv.es/lodtran18/def/transporte/dtx_srti#SituationRecord/GUID_Suc_612387_612387">
  <rdf:type rdf:resource="http://cef.uv.es/lodtran18/def/transporte/dtx_srti#SituationRecord"/>

  <geosparql:hasGeometry
    rdf:resource="http://cef.uv.es/lodtran18/def/transporte/dtx_srti#geoGUID_Suc_612387_612387"/>
</owl:NamedIndividual>

<owl:NamedIndividual rdf:about="#geoGUID_Suc_612387_612387">
  <rdf:type rdf:resource="http://www.opengis.net/ont/sf#LineString"/>
  <geosparql:asWKT rdf:datatype="http://www.opengis.net/ont/geosparql#wktLiteral">LINESTRING(-3.02388881
    37.101178705838,-3.02316586 37.100498995838)</geosparql:asWKT>
</owl:NamedIndividual>
```

4.4. Other sources used.

Other vocabularies were also used to specify several DATEX II concepts in the model. Table 4 shows these vocabularies.

Table 4. Ontologies used for several DATEX II concepts.

Concept defined in DATEX II	Ontology used for
AdministrativeAreaName	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio
NutsNamedArea	https://ec.europa.eu/eurostat/ramon/ontologies/geographic.rdf , http://data.europa.eu/nuts/code, http://data.europa.eu/euodp/repository/ec/estat/nuts/nuts.rdf
CountryCode and SubdivisionCode	http://www.omg.org/spec/LCC/Countries/CountryRepresentation/
Language Representation	https://www.omg.org/spec/LCC/Languages/LanguageRepresentation/
PointCoordinates	http://www.w3.org/2003/01/geo/WGS84, https://datos.ign.es/def/geo_core

For the CountryCode or SubdivisionCode, LCC Country Representation from OMG [41] has been used in the model. Figure 9 shows the concepts Alpha2Code, Country and CountrySubdivision imported from LCC for our model and Figure 10 the definition of the first one.

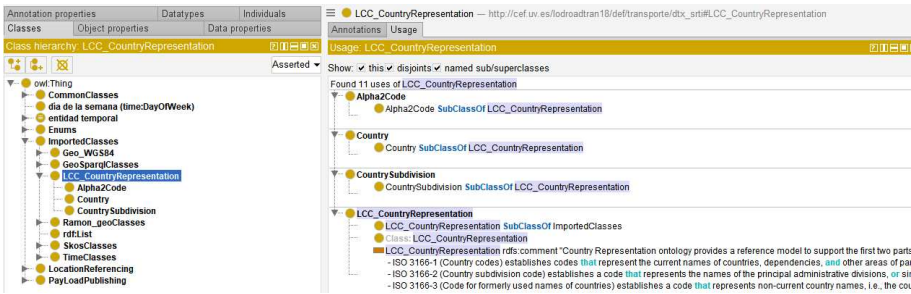


Figure 9. Imported “Alpha2Code” , “Country” and “CountrySubdivision” classes from LLC Country Representation.

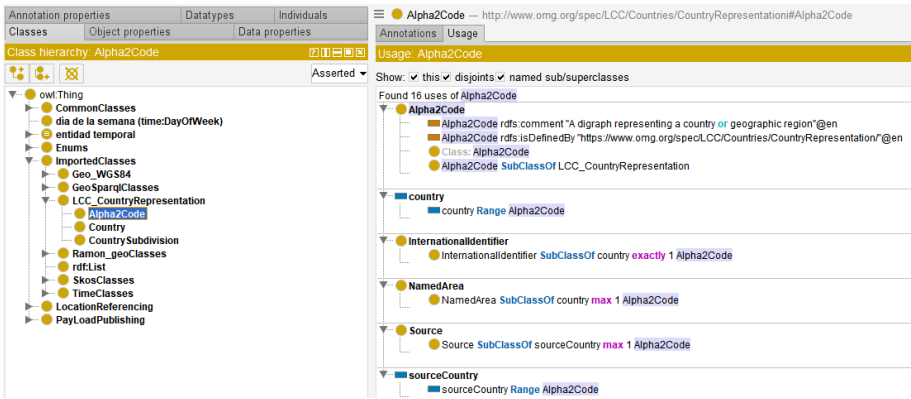


Figure 10. “Alpha2Code” definition.

5. Validation process

The technical system for getting traffic data from different sources, tagging, and processing it as LOD, together with subscription and querying capabilities, was designed and tested in order to evaluate the concept. The developed traffic data ontology was used as support. An Interface to query the linked open data (SPARQL) was proposed and manually tested.

5.1. Ontology editor

The Protégé editor [42] is available both as a desktop application and a web-based version. The development of the ontology has primarily been carried out using the Protégé desktop application. This choice was driven by its user-friendly interface, expandability through third-party plugins, and the development team’s experience with it. Protégé, developed by Stanford University, is a powerful, free, and open-source ontology editor. It enjoys widespread adoption by academic, government, and corporate users who employ it to build knowledge-based solutions in various domains such as biomedicine, e-commerce, and organizational modeling.

Protégé provides several features, including the ability to gather statistics on metrics related to classes and properties, as demonstrated in Figure 11 for the DATEX II SRTI ontology. Additionally, Protégé includes reasoners that facilitate the detection of inconsistencies within the model, making it easy to identify and resolve such issues.

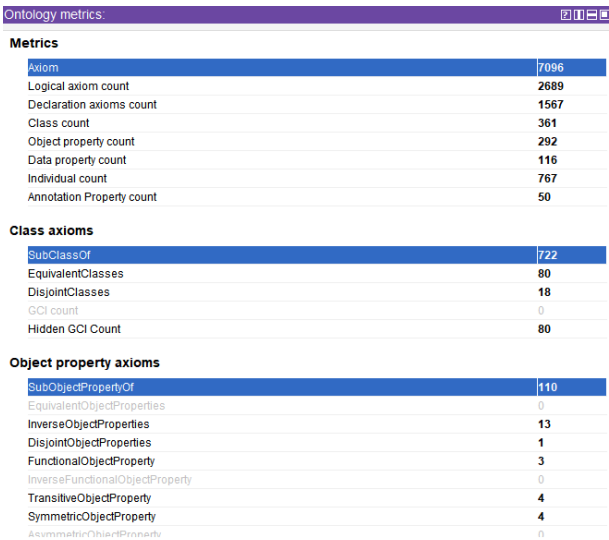


Figure 11. Statistics and metrics in Protégé Desktop Environment

the publication of RDF data in compliance with LOD standards. It is a Java web application based on Spring and Jena, providing W3C standard-compliant IRI dereferencing. LodView is easily configurable and deployable for developers, significantly enhancing the end user's experience when accessing HTML representations of RDF resources. Figure 14 depicts a SituationRecord resource using LodView.

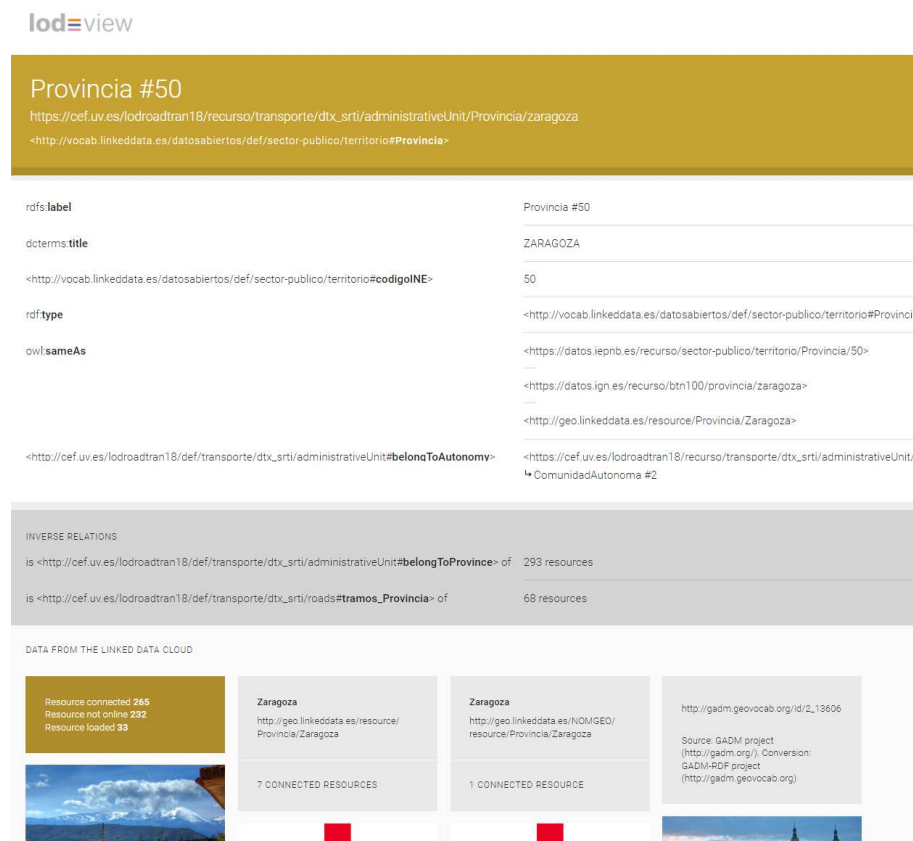


Figure 14. Visualisation of a Situation Record resource through LodView.

Finally, the finalized models, consisting of a main ontology and two satellite ontologies, were shared and made available through our server. The ontologies and their respective URLs are as follows:

1. Ontology of DATEX II-based traffic events SRTI
https://cef.uv.es/lodroadtran18/def/transporte/dtx_srti/
2. Ontology of administrative units of the Spanish and Czech territory.
https://cef.uv.es/lodroadtran18/def/transporte/dtx_srti/administrativeUnit/
3. Ontology of roads of the Spanish and Czech national road network.
https://cef.uv.es/lodroadtran18/def/transporte/dtx_srti/roads/

Moreover, the main ontology is also available via one of the best repositories about ontologies:
https://lov.linkeddata.es/dataset/lov/vocabs/dtx_srti

6. Conclusions

This paper introduces the *dtx_srti* ontology for semantic modeling of road traffic data and metadata within the context of SRTI. As of the time of writing, *dtx_srti* stands as the only semantic model based on DATEX II V3.2 and SRTI profile in accordance with the ITS Directive. This model enhances the basic service of accessing open traffic data provided by the traffic National Access

Point (nap.dgt.es), the Spanish Open Data Portal (datos.gob.es), and the European Data Portal (data.europa.eu). It not only facilitates data viewing and downloading but also improves the extraction of data meaning and enables other services that are only achievable using LOD.

The significance of having a LOD dataset for traffic information lies not only in the inclusion of data but also in the incorporation of links to other thematic datasets available on the web. This allows for federated queries that span across datasets from different platforms and sectors, not limited to just traffic data. Incorporating external data from diverse sources can add value to the tools and provide potential benefits for various applications.

The entire research, in addition to the semantic modelling partially shown in this article has allowed the deployment of traffic information data with metadata collected and it has been made available to the public through Open Data systems and made harvestable by respective national data catalogues and the European Data Portal. Both technical and legal issues in national and European context were addressed. The Action considered the general and technical requirements/constraints established by the European Data Portal. Once the requirements for being harvested by the European Data Portal were identified, the Action started the federation process. It serves as an integration architecture that allows interoperability between platforms.

Storing historical data (situation records) in SPARQL endpoint will help researchers to perform time and location-based searches, which significantly will improve usability of the LOD-SRTI resource. This, however, will considerably increase the size and will need to be managed in future.

On the other hand, the possibilities of implementation of measuring impact can be summarised in two approaches: measuring the impact of each dataset by making them visible (likes, downloads...) or measuring usage through connections to a given system. The availability of tools that make possible to know the behaviour of users and the level of use of the data is necessary: to demonstrate the success of the Action; to compare the results with other systems; and potentially to establish policies in the area of transport. Understanding user demands beyond compliance with the 2019 PSI Directive enables policies to be established to create high-value data and thematic data ecosystems. Also, "data sharing" between the public and private sector is recommended, respecting intellectual property and privacy.

Acknowledgments: This work was supported by the CEF Action Number 2018-EU-IA-0088, "Supporting the cross-border use of Road Traffic Data with Linked Open Data based on DATEX II".

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Guarino, N. *Formal Ontology in Information Systems*; IOS Press, 1998.
2. Barrachina, J.; Garrido, P.; Fogue, M.; Martinez, F.; Cano, J.C.; Calafate, C.; Manzoni, P. VEACON: a Vehicular Accident Ontology Designed to Improve Safety on the Roads. *Journal of Network and Computer Applications* **2012**, *35*. doi:10.1016/j.jnca.2012.07.013.
3. Tomás, V.; García, L. A cooperative multiagent system for traffic management and control. 2005, pp. 52–59. doi:10.1145/1082473.1082804.
4. Samper Zapater, J.J.; Llidó Escrivá, D.M.; Soriano García, F.R.; Martínez Durá, J.J. Semantic web service discovery system for road traffic information services. *Expert Systems with Applications* **2015**, *42*, 3833–3842. doi:https://doi.org/10.1016/j.eswa.2015.01.005.
5. Katsumi, M.; Fox, M. Ontologies for Transportation Research: A Survey. *Transportation Research Part C: Emerging Technologies* **2018**, *89*, 53–82. doi:10.1016/j.trc.2018.01.023.
6. DATEX Technical Board. DATEX II. Data Exchange II. <https://www.datex2.eu/>.
7. Perera, G.; others. Linked Data and the Future of Health Information Exchange. *Journal of AHIMA* **2012**, *83*, 38–42.
8. Lamy, J.B.; others. Toward a Formalization of the Process to Select IMIA Yearbook Best Papers. *Methods of Information in Medicine* **2014**, *53*, 1–10.
9. d'Aquin, M., On the Use of Linked Open Data in Education: Current and Future Practices; 2016; pp. 3–15. doi:10.1007/978-3-319-30493-9_1.

10. Vrandečić, D.; Krötzsch, M.; Völkel, M.; et al.. Wikidata: A Free Collaborative Knowledge base. *Communications of the ACM* **2013**, *57*, 78–85.
11. Cocodril partners. Cooperation as a solution to european cross-border traffic. CROCODILE 3 corridor. European Project. CEF programme. 2018-2022. <https://www.its-platform.eu/its-corridors/crocodile/>.
12. LOD-RoadTran18. Supporting the cross-border use of Road Traffic Data with Linked Open Data based on DATEX II (LOD-RoadTran18). Action number: 2018-EU-IA-0088. <https://cef.uv.es/lodroadtran18>.
13. Harris, S.; Seaborne, A. SPARQL 1.1 query language: Tech. rep., 2013.
14. J. Gutierrez and J. Javier Samper. An RDF/OWL vocabulary to represent SRTI DATEX II. https://cef.uv.es/lodroadtran18/def/transporte/dtx_srti.
15. DATEX II. SRTI. Safety Related Traffic Information messages. <https://docs.datex2.eu/profiles/srti.html>.
16. W3C OWL Working Group. OWL 2 Web Ontology Language Document Overview (Second Edition). <https://www.w3.org/TR/owl2-overview/>.
17. Schreiber, G.; Raimond, Y. RDF 1.1 primer: Tech. rep. <https://www.w3.org/TR/rdf11-primer/>.
18. Espinoza-Arias, P.; García-Delgado, M.; Corcho, O.; Vivas-White, P.; Potti-Manjavacas, H. A sustainable process and toolbox for geographical linked data generation and publication: a case study with BTN100. *Open Geospatial Data, Software and Standards* **2019**, *4*. doi:10.1186/s40965-019-0060-4.
19. Almendros-Jiménez, J.M.; Becerra-Terón, A. Discovery and diagnosis of wrong SPARQL queries with ontology and constraint reasoning. *Expert Systems with Applications* **2021**, *165*, 113772. doi:https://doi.org/10.1016/j.eswa.2020.113772.
20. Spanish Government. APORTA Initiative. <https://datos.gob.es/en/about-aporta-initiative>.
21. Comision, E. ITS directive. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:EN:PDF>.
22. Lubrich, P. Coordinated Metadata Catalogue – An Introduction. *5th Follow-Up Member States Expert Meeting Del. Reg. (EU) 2017/1926* **2019**.
23. Official National Bulletin (BOE). Reuse of information resources in Spain (in Spanish). <https://www.boe.es/eli/es/res/2013/02/19/4>.
24. Official National Bulletin (BOE). National interoperability scheme (in Spanish). <https://www.boe.es/eli/es/rd/2010/01/08/4/con>.
25. Cox, S.; Little, C. Time ontology in OWL: Tech. rep. <https://www.w3.org/TR/owl-time/>.
26. Spanish Government. Identification of geographical coverage in Spain. <https://datos.gob.es/es/recurso/sector-publico/territorio>.
27. European commission, JoinUp. How to use the MDR data themes vocabulary? DCAT application profile implementation guidelines Web site. <https://joinup.ec.europa.eu/release/dcat-ap-how-use-mdr-data-themes-vocabulary>.
28. Union, E. EU Vocabularies. <https://op.europa.eu/en/web/eu-vocabularies/controlled-vocabularies>.
29. Peter Lubrich et al.. Monitoring and Harmonisation of National Access Points in Europe. https://www.its-platform.eu/wp-content/uploads/ITS-Platform/AchievementsDocuments/NAP/EUEIP_Coord.MetadataCatalogue_v2.0_191115.pdf.
30. Eurstat. NUTS. Nomenclature of Territorial Units for Statistics, by regional level.
31. Spanish Government. Taxonomy of primary sectors. <https://datos.gob.es/en/kos/sector-publico/sector>.
32. Gutierrez, J.; Samper, J.J.; Delgado, A.; Rocha, J.; Petr, B.; Zuzana, P. LOD-RoadTran18: Supporting the Cross-Border Use of Road Traffic Data with Linked Open Data Based on DATEX II. Proceedings of the 11th Euro American Conference on Telematics and Information Systems; Association for Computing Machinery: New York, NY, USA, 2022; EATIS '22. doi:10.1145/3544538.3544651.
33. Samper, J.J.; Adell, F.J.; García, J.F.; Martínez, J.J.; Vidal, J. Semantic Traffic Applications Based on DATEX II. Proceedings of the 2009 Euro American Conference on Telematics and Information Systems: New Opportunities to Increase Digital Citizenship; Association for Computing Machinery: New York, NY, USA, 2009; EATIS '09. doi:10.1145/1551722.1551729.
34. Pieter Colpaert. Linked DATEX II. <https://lov.linkeddata.es/dataset/lov/vocabs/datex>.
35. Knowledge Craver. Guidelines for using the W3C time ontology. <http://knowledgecraver.blogspot.com/2014/07/guidelines-for-using-w3c-time-ontology.html>.
36. Rijgersberg, H. Ontology of units of Measure (OM) 2.0 . <https://github.com/HajoRijgersberg/OM>.

37. QUDT organization. QUDT catalog. Quantities, Units, Dimensions and Data Types Ontologies. <https://www.qudt.org/2.1/catalog/qudt-catalog.html>.
38. M. Lefrançois and A. Zimmermann.. Custom Datatypes. Towards a web of Linked Datatypes, 2020.
39. QUDT organization. QUDT Information. qudt.org/vocab/unit.
40. Perry, M.; Herring, J. OGC GeoSPARQL. A geographic query language for RDF data, 2012.
41. Object Management Group. Country Representation from OMG. <http://www.omg.org/spec/LCC/Countries/CountryRepresentation/>.
42. Musen, M.A. The protégé project: a look back and a look forward. *AI Matters* **2015**, *1*, 4–12. doi:10.1145/2757001.2757003.
43. Spanish traffic Directorate (DGT). Spanish national access point. <https://nap.dgt.es/sparql>.
44. Spanish National Geographic Institute. Metadata catalogue. <https://www.ign.es/csw-inspire/srv/eng/main.home>.
45. Spanish Government. Incidencias de Tráfico (SRTI) en formato LOD DATEX II. <https://datos.gob.es/es/catalogo/e00130502-incidencias-de-traffic-srti-en-formato-lod-datex-ii>.
46. European Union. The official portal for European data. Traffic Incidents (SRTI) in LOD DATEX II format. <https://data.europa.eu/data/datasets/https-datos-gob-es-catalogo-e00130502-incidencias-de-traffic-srti-en-formato-lod-datex-ii?locale=en>.
47. LODView Team. LODView. Giving data a new shape. <https://lodview.it>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.