

The InfraRisk Ontology: Enabling Semantic Interoperability for Critical Infrastructures at Risk from Natural Hazards

Dumitru Roman^(✉), Dina Sukhobok, Nikolay Nikolov,
Brian Elvesæter, and Antoine Pultier

SINTEF, Pb. 124 Blindern, 0314 Oslo, Norway
{dumitru.roman,dina.sukhobok,nikolay.nikolov,
brian.elvesater,antoine.pultier}@sindef.no

Abstract. Earthquakes, landslides, and other natural hazard events have severe negative socio-economic impacts. Among other consequences, those events can cause damage to infrastructure networks such as roads and railways. Novel methodologies and tools are needed to analyse the potential impacts of extreme natural hazard events and aid in the decision-making process regarding the protection of existing critical road and rail infrastructure as well as the development of new infrastructure. Enabling uniform, integrated, and reliable access to data on historical failures of critical transport infrastructure can help infrastructure managers and scientist from various related areas to better understand, prevent, and mitigate the impact of natural hazards on critical infrastructures. This paper describes the construction of the InfraRisk ontology for representing relevant information about natural hazard events and their impact on infrastructure components. Furthermore, we present a software prototype that visualizes data published using the proposed ontology.

Keywords: Ontology · Infrastructure components · Natural hazards · Events

1 Introduction

A natural hazard can be defined as a natural process that poses a threat to human life or property [1]. Extreme natural hazard events have the potential to cause devastating impacts to infrastructure networks, resulting in significant economic losses. In Europe, the number of disasters due to natural hazards increased in recent decades due to a combination of climate change effects, and changes in physical and social systems. For the period between 1998 and 2009, natural hazards and technological accidents caused nearly 100,000 fatalities and affected more than 11 million people, at the same time resulting in overall economical impact of about 200 billion euro [2]. These natural hazards have included hydrometeorological hazards (e.g. storms, floods) and geophysical hazards (e.g. landslides, earthquakes). Thus, floods, along with storms, are natural hazards that cause the highest economic losses in Europe. The flood-related losses in the EEA member countries over the period from 1998 to 2009 amounted to more than

60 billion euro [3]. In [4], the authors estimated that the expected annual damage (EAD) from flooding events in Europe may increase to 23.5 billion euro by 2050.

Given the potential economic losses caused by natural hazards, it is necessary to analyse the effects of natural hazards on the infrastructure, in particular critical infrastructure. Reliable transport infrastructure is of a great value to society as it facilitates the effective transportation of people and goods. The EU transport network has over 4.5 million km of paved roads and 212,500 km of rail lines¹. Transport infrastructure plays a fundamental role in the EU and the ability to transport goods safely, quickly and cost-efficiently is highly important for international trade and economic development [5]. The complex interdependency of European infrastructure networks results in spreading the interruptions in infrastructure networks to many parts of Europe.

In this context, novel methodologies and tools are needed to analyse the potential impacts of extreme natural hazard events and aid in the decision-making process regarding the protection of existing critical road and rail infrastructure as well as the development of new infrastructure. One example of initiative addressing such aspects was the InfraRisk project.² The project aimed to develop reliable stress tests on European critical infrastructure using integrated modelling tools for decision-support. An important aspect of the project was to set the foundations for the development of a Geographical Information System (GIS) knowledge base of major global infrastructure failures, enabling users uniform, integrated, and reliable access to data on historical failures of critical transport infrastructure. The potential users of such a knowledge base are infrastructure managers, but also researchers (risk management, transportation, civil engineering, natural sciences, etc.). The knowledge base can serve as a case study for the events an infrastructure manager might consider important, and provide them with data of good/bad practices of managing solutions during and after the event.

In order to share common understanding of the data structure among the knowledge base users and enable semantic interoperability of infrastructure failure related data, the InfraRisk ontology was developed. Although the ontology was primarily developed to support data sharing and data usage within the scope of the InfraRisk project, it is generic and can be used for publishing and integrating various kinds of infrastructure components and natural hazards data. In this paper we present the design and implementation of the InfraRisk ontology for describing infrastructure failures due to the natural hazard events. Furthermore, we present a software prototype developed to consume data using the proposed ontology and interactively visualize information about various infrastructure components and natural hazards. The contributions of this paper are thereby two-fold:

1. First, we describe the InfraRisk ontology (design, implementation) for enabling semantic interoperability for critical infrastructures at risk from natural hazards;
2. Second, we propose a software prototype to visualize infrastructure components and natural hazards data made available using the developed ontology.

¹ http://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/site/en/facts_and_figures.html.

² <https://www.infrarisk-fp7.eu/>.

The rest of the paper is organized as follows. Section 2 discusses related work. Section 3 describes the development process of the InfraRisk ontology. Section 4 illustrates an example of using the ontology for data publishing and integration, and presents the software prototype developed to visualize the data made available using the developed ontology. Section 5 summarizes the paper and outlines directions for future work.

2 Related Work

Defining and modeling natural hazards and their consequences is inconsistent across various natural hazard studies, databases and vocabularies. We analyzed the most common natural-hazard and infrastructure related terminologies and vocabularies in order to use available ontological knowledge in the ontology development process.

The terminology used by UNISDR (United Nations office for Disaster risk reduction)³ defines *natural hazard* as a process, phenomenon or human activity associated with natural processes and phenomena that may cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation. Natural hazards, as well as any other hazardous event can cause a *disaster* – a serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental *losses* and impacts, which exceeds the ability of the affected community or society to cope using its own resources. It is further commented that disasters are often described as a result of the combination of: the exposure to a hazard; the conditions of vulnerability that are present; and insufficient capacity or measures to reduce or cope with the potential negative consequences. Disaster impacts may include loss of life, injury, disease and other negative effects on human physical, mental and social well-being, together with damage to property, destruction of assets, loss of services, social and economic disruption and environmental degradation.

The International Federation of Red Cross and Red Crescent Centres (IFRC)⁴ defines *disaster* as a sudden, calamitous event that seriously disrupts the functioning of a community or society and causes human, material, and economic or environmental *losses* that exceed the community's or society's ability to cope using its own resources. Though often caused by nature, disasters can have human origins. *Natural hazards* are considered as types of disasters and are defined as naturally occurring physical phenomena caused either by rapid or slow onset events.

The concepts *Natural Hazard*, *Loss* and *Event* as a generalized concept representing occurrence of a particular set of circumstances, are key concepts used in various terminologies and vocabularies in this field. The above mentioned concepts were also taken as the most general concepts in the domain and were used as a basic classes in the developed ontology.

³ <https://www.unisdr.org/>.

⁴ <http://www.ifrc.org/en/>.

A few linked open vocabularies have emerged to capture natural hazard consequence data. For example, the Management of a Crisis Vocabulary (MOAC)⁵ provides a minimum set of classes and properties for describing crisis management activities. The vocabulary is specifically designed to aid the disaster information managers to carry out activities in response to a disaster, but doesn't describe losses related to natural hazard events.

None of existing terminologies and vocabularies cover aspects related to how natural hazard events affect infrastructure components and therefore they were found unsuitable for a direct use in the InfraRisk project. Nevertheless, the analysis of existing terminologies and vocabularies has provided us a good baseline for ontological knowledge to be built upon in the development of the InfraRisk ontology.

3 InfraRisk Ontology Development Process

The InfraRisk ontology was developed in accordance with existing guidelines and methodologies for ontology development process, in particular the one proposed by [6]. As a first step in the ontology development process we defined the ontology domain, scope and purpose, and requirements. After that we analyzed existing ontologies in order to find a way to refine and extend them for our particular domain and task. The next step was to collect domain knowledge to determine important terms in the ontology and build and refine a conceptual model using Object-Role Modelling (ORM). Finally, the conceptual model was realized in a concrete language (RDFS/OWL). Defining the domain and scope of the ontology is described in Sect. 3.1. The resulting conceptual model is discussed in Sect. 3.2. The process of the ontology implementation is discussed in the Sect. 3.3.

3.1 Defining the Scope of the Ontology

The main purpose of the ontology creation was to relate global major infrastructure failures with natural hazard events. Although the ontology was primarily developed to follow the InfraRisk project's focus on critical transport infrastructure (more specifically, European Ten-T core network⁶) and high-impact natural events, one can expect the developed ontology to be applicable in a wider critical infrastructure context.

In order to define the scope of the ontology, a set of competency questions were developed with infrastructure components and natural hazards experts in the project. The followings are samples of competency questions used in the process:

- Which tunnels/bridges are located in country X?
- Which road bridges have collapsed between 1990 and 2014 in region Y and were triggered by floods?
- Which infrastructure failures were triggered by the 2003 flood of Danube river?
- What were the consequences (monetary loss) of the collapse of bridge Y in 2010?

⁵ <http://www.observedchange.com/moac/ns/>.

⁶ https://ec.europa.eu/transport/themes/infrastructure/ten-t-guidelines/maps_en.

- How many tunnels collapsed in region Y due to floods during 2001-2011 and incurred monetary losses more than X amount euro?
- What were the events that triggered the collapse of bridge Y in 2002?
- What were the infrastructure failure events that were triggered by no distinguishable natural hazards or by low/moderate/high/"black swan" natural hazards?
- Which original natural hazard events caused cascading hazard events (and subsequently caused infrastructure failure)?
- Which type of infrastructure failure causes the biggest losses relatively to other types?
- Due to the flood event in country X, which bridges had to be closed to traffic?
- What were the casualties due to the collapse of bridge Y in 2010?
- Which highway segment on infrastructure X could not be accessed due the earthquake in 2010?

The scope of transport infrastructure therefore covers road and rail transport infrastructures and their elements. The scope of natural hazards covers disasters affecting road and rail transport infrastructures and their elements. This includes components such as bridges (single or viaducts), tunnels, off ramps, embankments and slopes, and road and rail surface segments. In terms of natural hazards we consider events such as floods, earthquakes, landslides and any cascading hazards.

3.2 Building the Conceptual Model

Conceptual modeling methodologies have proven to be very effective for building information systems in a graphical interface at the high level of abstraction. Conceptual data schemes and ontologies have a lot of similarities, as both model concept relations and rules (constraints) [7]. The idea of reusing conceptual modeling techniques for ontology development is proposed by several authors (e.g. [7–9]) and provides a lot of advantages such as ability to use numerous existing conceptual modeling tools and methods.

A conceptual model for the InfraRisk ontology was developed with the help of the ORM (Object Role Modeling) data modeling approach. ORM models consist of objects (mapped to classes in the ontology) playing roles (relations) [10]. One advantage of using this technique is that ORM diagrams can be translated into pseudo natural language statements. This enables non-computer scientists (e.g. infrastructure experts) to evaluate the developed model. The ORM model for the InfraRisk ontology defines a conceptual model that relates major global infrastructure failures with natural hazard events. In the following we provide a brief description of key aspects covered by the conceptual model: infrastructure components and events (consequences and natural hazards).

Infrastructure components

An *Infrastructure* represents a transport mode, e.g., *Road* or *Rail* in our context. It has a *name*, *description* and a *geographical feature*. An infrastructure consists of one or more *Infrastructure Components*, e.g. *Bridge*, *Tunnel*, etc. (see Fig. 1). Each component has a *name*, *description*, a number of *lanes*, and a *geographical feature*. An Infrastructure Component can be connected to other Infrastructure Components. As mentioned above,

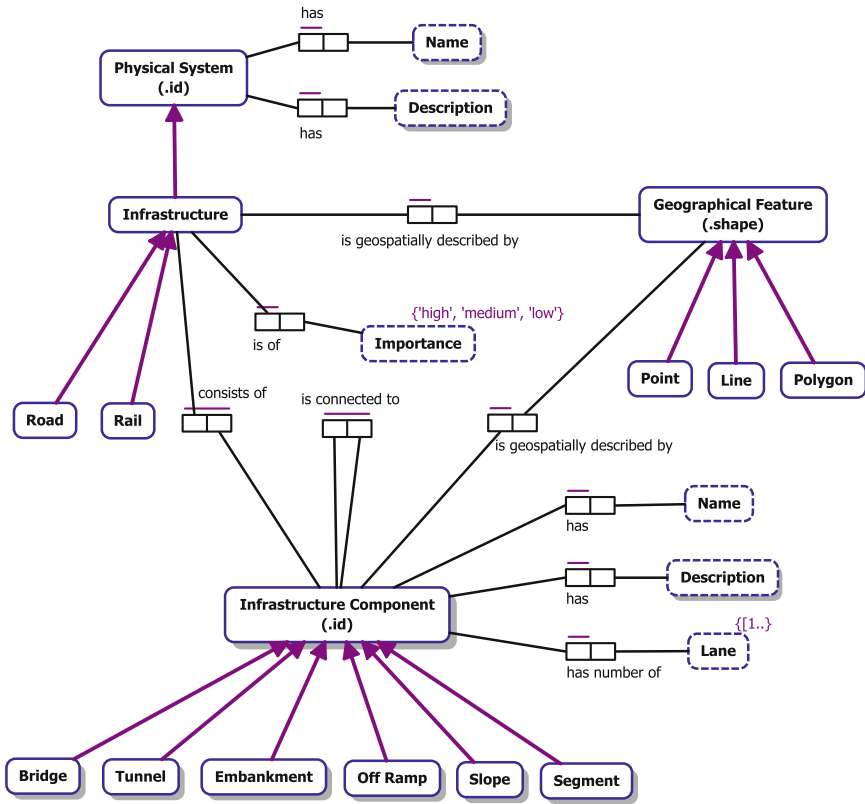


Fig. 1. ORM model for Infrastructure Component

the scope of this ontology is limited to components such as *bridges*, *tunnels*, *embankments*, *off ramps*, *slopes* and *segments* (e.g. of a road or rail line). Each of these infrastructure component types has its own set of properties as shown in the ORM models in Appendix A.

Events

An *Event* represents an incident where a *Natural Hazard* or *Infrastructure Component Failure* has occurred. It has a *name*, *description*, *location*, *date* and *consequence* (see Fig. 2). An *Infrastructure Component Failure* concerns the *full* or *partial collapse* of an *Infrastructure Component*.

A *Consequence* represents the expected losses in a specific location as a result of a given event. The *Consequence* can be a *Monetary Loss*, *Societal Loss* or *Usability Problem* concerning closure of or reduced traffic on an *Infrastructure Component* (see Fig. 3). The conceptual model distinguishes between three types of *Natural Hazard* events, namely *Earthquakes*, *Floods* and *Landslides*. The conceptual models for natural hazards are detailed in Appendix B.

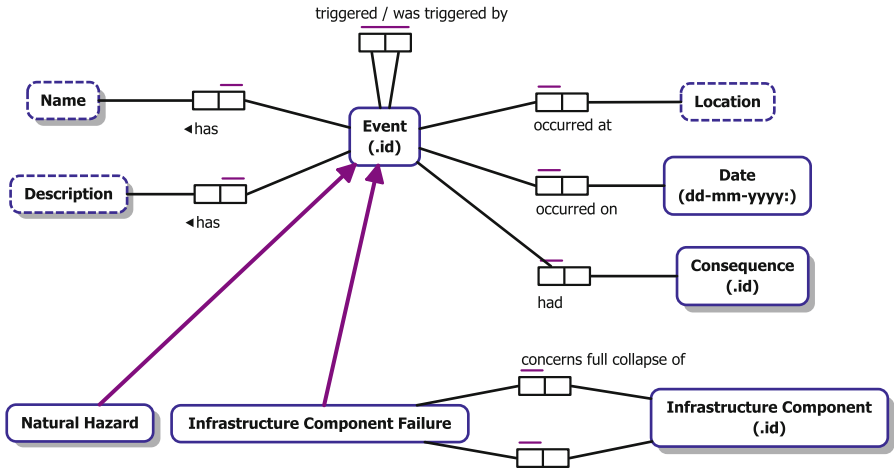


Fig. 2. ORM model for Event and Infrastructure Component Failure

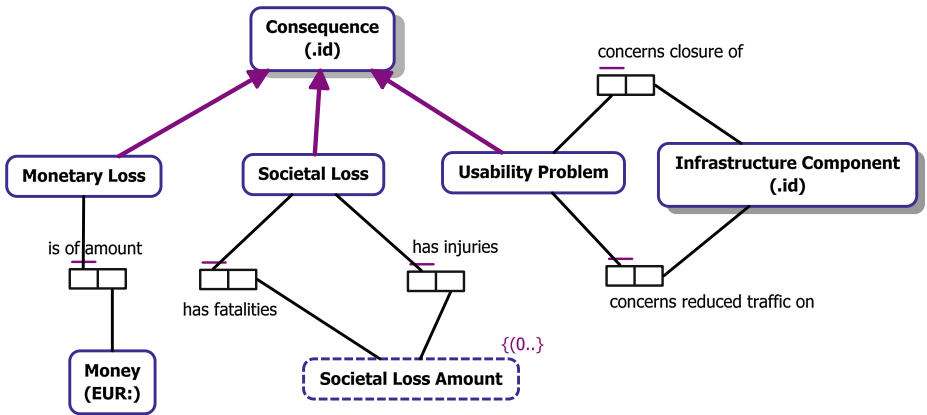


Fig. 3. ORM model for Consequence

3.3 Realizing the Conceptual Model in RDFS/OWL

The next step in the development of the InfraRisk ontology was its realization in a concrete language that can be used for publishing data. The InfraRisk conceptual model was specified using class hierarchy in RDFS⁷/OWL⁸ and implemented in the Neologism vocabulary publishing platform⁹ (see Fig. 4). RDFS is the most basic schema language commonly used in the semantic Web to model concepts, properties and their

⁷ <https://www.w3.org/TR/rdf-schema/>.

⁸ <https://www.w3.org/TR/webont-req/>.

⁹ <http://neologism.deri.ie/>.

All terms at a glance

Classes: Bridge | Consequence | Earthquake | Embankment | Event | Flood | GeographicalFeature | Infrastructure | InfrastructureComponent | InfrastructureComponentFailure | Landslide | Line | Mass | MonetaryLoss | NaturalHazard | OffRamp | Point | Polygon | Rail | Road | RockLandslide | Segment | Slope | SocietalLoss | Soil | SoilLandslide | Tunnel | UsabilityProblem

Properties: concernsFullCollapseOfInfrariskCom | concernsPartialCollapseOfInfrariskCom | hadConsequence | hasBridge | hasDepth | hasDescription | hasDuration | hasEvent | hasExtent | hasFatalitiesSocietalLoss | hasFaultMechanism | hasFlowVelocity | hasHazardMap | hasHeight | hasInfrariskStructure | hasInjuriesSocietalLoss | hasIntensity | hasMagnitude | hasName | hasNumberOfLanes | hasObject | hasRiskMap | hasRunOut | hasScourVulneGrade | hasSeismicDesign | hasShakeMap | hasShape | hasShapeMap | hasSlopeGrade | hasTypology | hasVolume | hasWaterLevel | hasInfrariskCom | isFormedBySoil | isGeospatiallyDescribedBy | isMadeOfConstMat | isOfAmountMoney | isOfImportance | isOfRock | isOfShock | isOfSoil | isOfSoilCatg | isOnRock | isOnSoil | isSusceptibleToWaterInfr | occurredOnDate | occurredOnDate | resultedInDisplacementOf

Fig. 4. Screenshot of the InfraRisk ontology in Neologism

relationships and characteristics (such as domains and ranges of properties). In its turn OWL is built upon RDFS and provides a larger vocabulary for web ontology modelling and can be used to model more advanced relationships.

In order to cover geospatial attributes of natural hazard and infrastructure component data, the InfraRisk ontology makes use of OGC GeoSPARQL standard [11]. The OGC GeoSPARQL standard supports representing and querying geospatial data on the semantic Web. It defines a vocabulary for representing geospatial data in RDF, and also provides an extension to the SPARQL query language for processing geospatial data.

The InfraRisk RDFS/OWL vocabulary is openly available¹⁰ and contains 28 classes and 48 properties. The vocabulary is licensed under Creative Commons Attribution (CC BY)¹¹ and is available as an endpoint¹² via the DataGraft platform¹³.

4 Software Prototype

Five datasets have been published on the DataGraft [12, 13] platform during the InfraRisk project using the ontology as a central reference model. The source datasets included data about natural hazard events (floods and landslides) in Europe that

¹⁰ <http://vocabs.datagraft.net/infrarisk>.

¹¹ <https://creativecommons.org/licenses/by/3.0/>.

¹² <https://rdf.datagraft.net/4037543173/db/repositories/infrarisk-vocabulary-1>.

¹³ <https://datagraft.io/>.

resulted in failures of critical transport infrastructure during the period 1972-2016. Data was obtained from InfraRisk project partners.

The DataGraft platform's warehouse for the RDF data is realized through the Semantic Graph Database-as-a-Service (DBaaS) component – a fully managed, cloud-based version of GraphDB¹⁴ semantic graph database (triplestore). To meet the requirements of working with geospatial data and linked data, the DBaaS component introduces support for GeoSPARQL. The implementation of the GeoSPARQL specification in DBaaS is delivered as an additional plug-in for the GraphDB engine.

After the data has been published, it was possible to perform queries on data related to historical failures of critical transport infrastructure using the SPARQL query language. In order to query geographic information, GeoSPARQL extension functions for spatial computations are used. The following shows a SPARQL query retrieving infrastructure failures that occurred in 2015.

```
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX geof: <http://www.opengis.net/def/function/geosparql/>
PREFIX : <https://www.infrarisk-fp7.eu/vocabs/#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xs: http://www.w3.org/2001/XMLSchema#

SELECT distinct ?eventCause ?eventDescription ?coordinates
?hasFatalities ?hasInjuries

WHERE {
  [] :hasinfrariskCom ?infrastructureComponent;
    :isGeospatiallyDescribedBy ?point.
  ?infrastructureComponent rdf:type :InfrastructureComponent .
  ?point geo:asWKT ?coordinates .
  ?infrastructureComponentFailure a :InfrastructureComponentFailure .
  ?infrastructureComponentFailure :icAssociation ?infrastructureComponent .
  ?event rdf:type ?eventClass .
  ?eventClass rdfs:subClassOf :Event .
  ?event ?eventAssociation ?infrastructureComponent .
    ?causedBy :hasEvent ?event .
    ?causedBy :hasName ?eventCause .
  ?event :hadConsequence ?societalLoss .
  ?societalLoss rdf:type :SocietalLoss .
```

Table 1 shows a result sample of the above SPARQL query.

Thus, using the InfraRisk ontology one can represent and query integrated data from originally heterogeneous data sources. The results of the queries can further be

¹⁴ <http://graphdb.ontotext.com/>.

Table 1. Result excerpt of the SPARQL query

Event cause	Event description	Coordinates	Has fatalities	Has injuries
Storm Frank	Bridge subsided and partially collapsed	POINT (53.688716 - 1.840771)	no	no
Rockfall	Rockfall blocked & damaged tracks	POINT (46.749723 8.642357)	no	no
Mudflow due to heavy rain	Mudflow caused by thunderstorm covered road and trapped several cars	POINT (34.886522 - 118.904150)	no	yes

visualized in various tools. A graphical user interface (GUI) application prototype was developed to visualize the data published using the ontology on a map using various interaction mechanisms. The prototype’s GUI is based on the open source MASTER application¹⁵. It is an HTML5 application which can be used on smartphones, tablets and desktop computers. In addition to the map view, the application was integrated with the Google Street View technology. It allows the user to navigate along the roads photographed by Google. This mode provides an interesting alternative for viewing of hazard events (see Fig. 5).

Data about infrastructure components and events published using the InfraRisk ontology are retrieved via SPARQL queries and the results are presented in the GUI application (see Fig. 6). The output data is formatted using JSON.

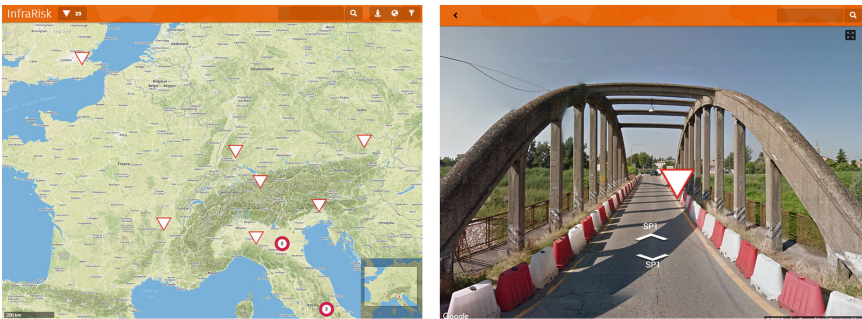


Fig. 5. Visualizing infrastructure events in the software prototype

¹⁵ <https://github.com/SINTEF-9012/mobileMaster>.



Fig. 6. Example of event details visualized in the software prototype

5 Summary and Outlook

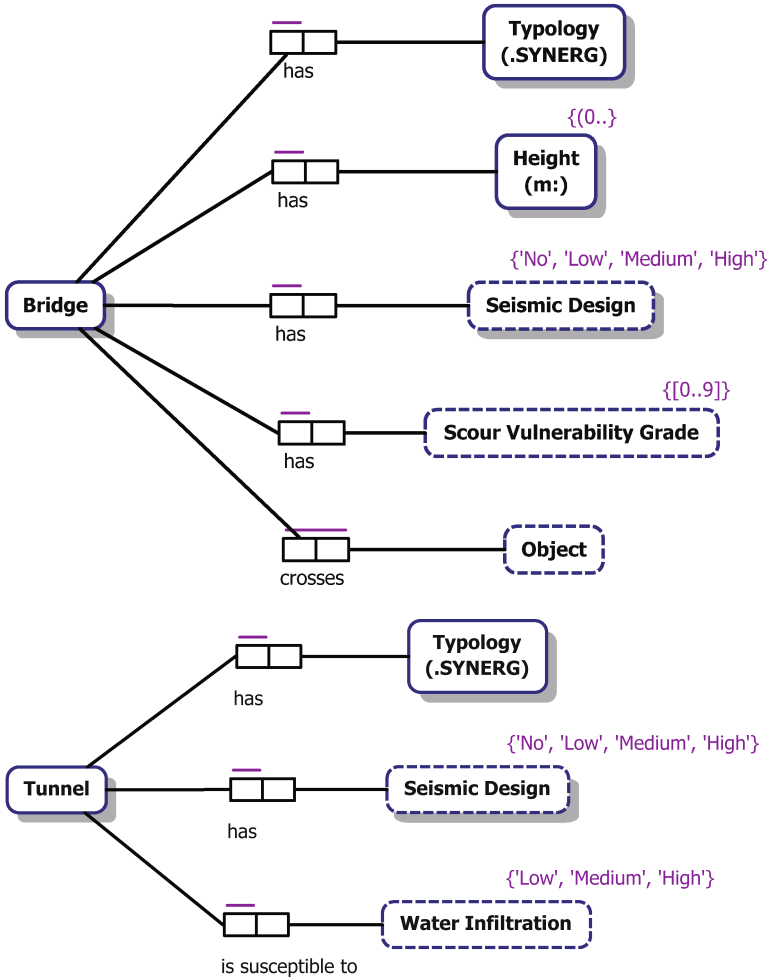
This paper provided an overview of the InfraRisk ontology developed in order to assist publishing and integration of data about infrastructure failures due to natural hazard events. The ontology was developed in accordance with well-known ontology development guidelines. The ontology aimed to relate critical transport infrastructure with high-impact natural hazard events. The ontology was used to integrate and publish datasets about natural hazard events resulted in failures of critical transport infrastructure as Linked Open Data.

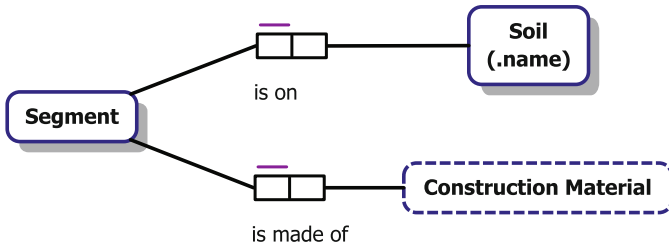
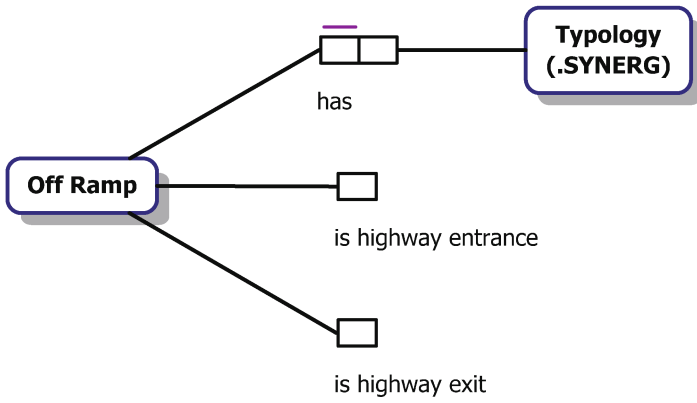
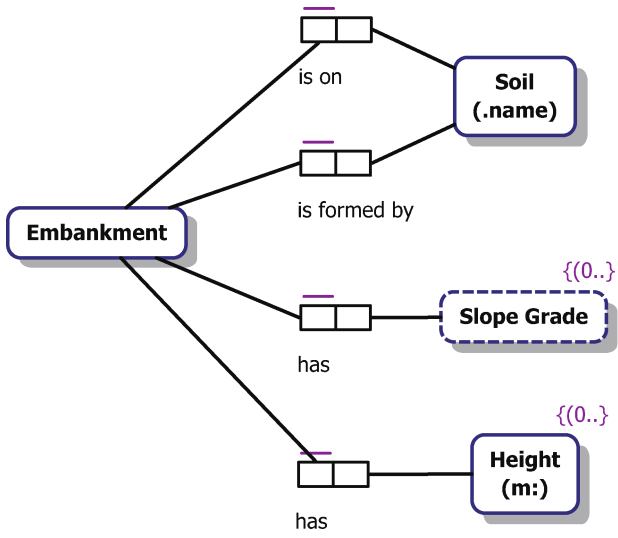
Furthermore, we developed a software prototype that visualizes data about infrastructure components and natural hazards published using the proposed ontology.

The ontology and the software prototype were developed based on the InfraRisk project's focus on European critical transport infrastructure, however they are suitable for use with infrastructure components failure data from other sources than those defined in the project. Future activities will be related to publication of data from various sources using the ontology, standardization of the ontology, and further improvements to the developed prototype.

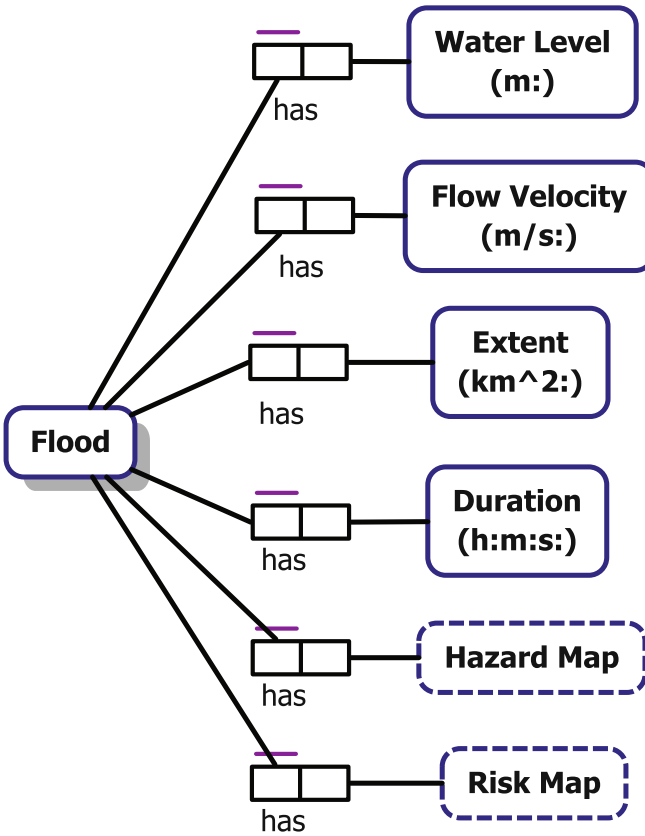
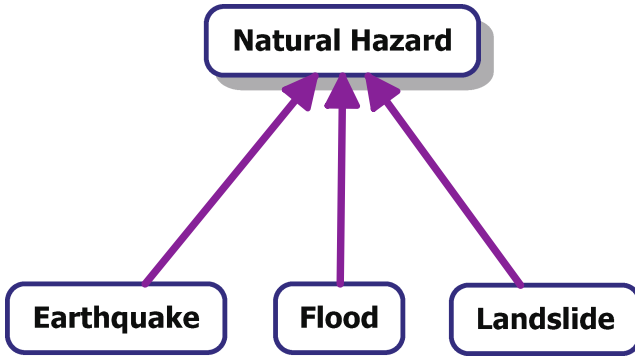
Acknowledgements. The work in this paper is partly supported by the EC funded projects InfraRisk (Grant number: 603960) and proDataMarket (Grant number: 644497). The authors would like to thank Pierre Gehl (UCL), Khaled Taalab (UCL), Pieter van Gelder (PSCT), Yuliya Avdeeva (PSCT), Maria Jose Jimenez (CSIC), Mariano Garcia-Fernandez (CSIC), Bryan T. Adey (ETH), Miguel José Segarra Martínez (Dragados) and Mark Tucker (ROD), as well as other InfraRisk project partners for their involvement and contributions to the development of the InfraRisk ontology.

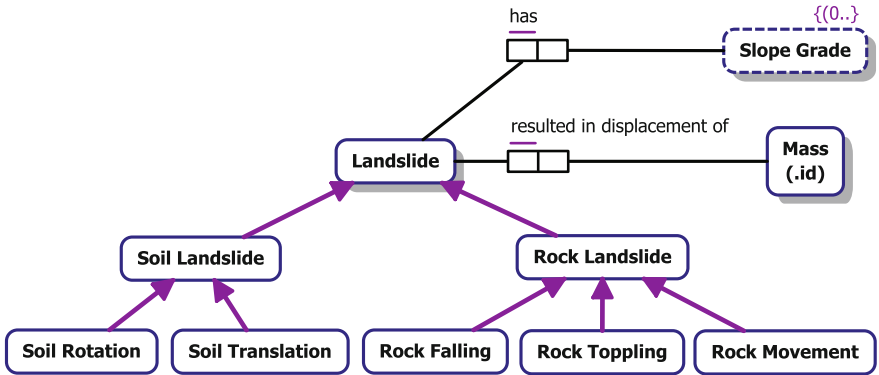
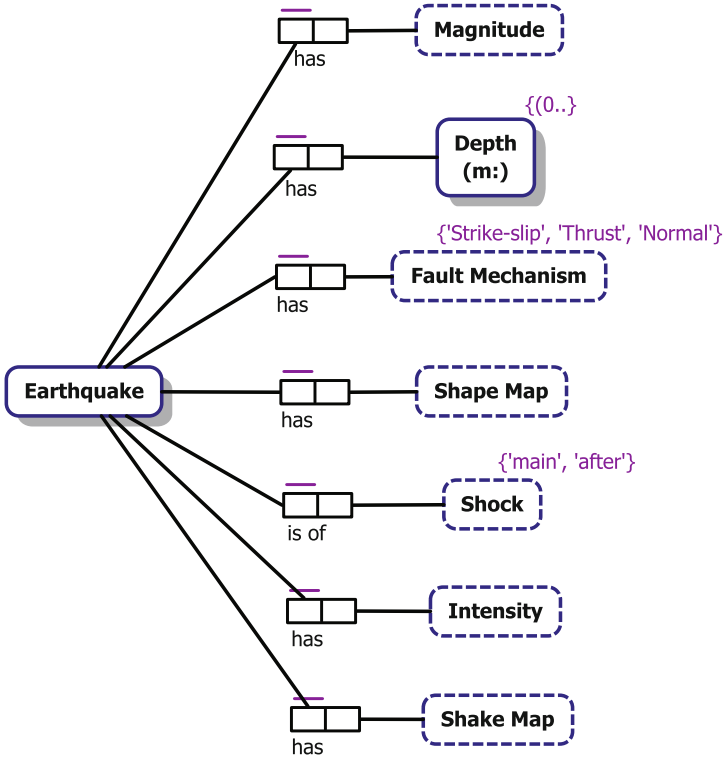
Appendix A. ORM Models for Various Infrastructure Components

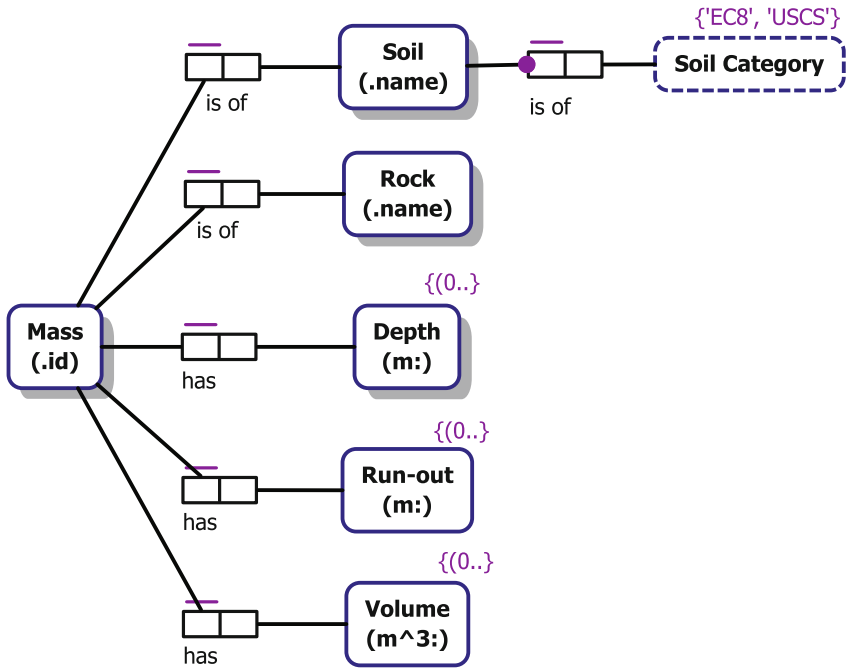




Appendix B. ORM Models for Various Natural Hazards







References

- Hyndman, D., Hyndman, D.: Natural hazards and disasters. Cengage Learn. (2016)
- Mapping the impacts of natural hazards and technological accidents in Europe. An overview of the last decade (2010)
- Exploring nature-based solutions. The role of green infrastructure in mitigating the impacts of weather- and climate change-related natural hazards (2015)
- Brenden, J., et al.: Increasing stress on disaster-risk finance due to large floods. *Nat. Clim. Change* **4**(4), 264–268 (2014)
- Eurostat regional yearbook, Luxembourg: Publications office of the European Union (2016)
- Noy, N.F., McGuinness, D.L.: Ontology development 101: A guide to creating your first ontology (2001)
- Jarrar, M., Demey, J., Meersman, R.: On using conceptual data modeling for ontology engineering. In: Spaccapietra, S., March, S., Aberer, K. (eds.) *Journal on Data Semantics I*. LNCS, vol. 2800, pp. 185–207. Springer, Heidelberg (2003). doi:[10.1007/978-3-540-39733-5_8](https://doi.org/10.1007/978-3-540-39733-5_8)
- Cranefield, S., Haustein, S., Purvis, M.: UML-based ontology modelling for software agents. In: *Proceedings of the Workshop on Ontologies in Agent Systems, 5th International Conference on Autonomous Agents*, Montreal, Canada (2001)
- Baclawski, K., Kokar, Mieczyslaw K., Kogut, Paul A., Hart, L., Smith, J., Holmes, William S., Letkowski, J., Aronson, Michael L.: Extending UML to Support ontology engineering for the semantic web. In: Gogolla, M., Kobryn, C. (eds.) *UML 2001*. LNCS, vol. 2185, pp. 342–360. Springer, Heidelberg (2001). doi:[10.1007/3-540-45441-1_26](https://doi.org/10.1007/3-540-45441-1_26)

10. Halpin, T.: Object-role modeling: Principles and benefits. *Int. J. Inf. Syst. Model. Des. (IJISMD)* **1**(1), 33–57 (2010)
11. OGC GeoSPARQL - A Geographic Query Language for RDF Data (2012)
12. Roman, D., Nikolov, N., Putlier, A., Sukhobok, D., Elvesæter, B., Berre, A., Ye, X., Dimitrov, M., Simov, A., Zarev, M., Moynihan, R.: DataGraft: one-stop-shop for open data management. To appear in the *Semantic Web Journal (SWJ) – Interoperability, Usability, Applicability* (2017). IOS Press, ISSN: 1570–0844, doi:[10.3233/SW-170263](https://doi.org/10.3233/SW-170263)
13. Roman, D., Dimitrov, M., Nikolov, N., Putlier, A., Sukhobok, D., Elvesæter, B., Berre, A., Ye, X., Simov, A., Petkov, Y.: DataGraft: simplifying open data publishing. In: Sack, H., Rizzo, G., Steinmetz, N., Mladeníc, D., Auer, S., Lange, C. (eds.) *ESWC 2016. LNCS*, vol. 9989, pp. 101–106. Springer, Cham (2016). doi:[10.1007/978-3-319-47602-5_21](https://doi.org/10.1007/978-3-319-47602-5_21)