IDISFER, an Ontology to Model Extreme Floods-Related Processes

Cécile Saint-Marc, Cicely Pams Capoccioni, Gilles Saussine, Dorine Chenier, Paule-Annick Davoine, Marlène Villanova-Oliver

To cite this version:

Cécile Saint-Marc, Cicely Pams Capoccioni, Gilles Saussine, Dorine Chenier, Paule-Annick Davoine, et al.. IDISFER, an Ontology to Model Extreme Floods-Related Processes. World Congress on Railway Research (WCRR 2016), May 2016, Milan, Italy. hal-01383641

HAL Id: hal-01383641
https://hal.inria.fr/hal-01383641

Submitted on 19 Oct 2016

IDISFER, an Ontology to Model Extreme Floods-Related Processes

Cécile SAINT-MARC[1, 2], Cicely PAMS CAPOCCIONI[1], Gilles SAUSSINE[1], Dorine CHENIER[1], Paule-Annick DAVOINE[2], Marlène VILLANOVA-OLIVER[2].

[1] SNCF Ingénierie et Projets. La Plaine Saint-Denis. France

[2] Laboratoire d'Informatique de Grenoble, Université Grenoble-Alpes. Saint-Martin-d'Hères. France

Contact: cicely.pams@reseau.sncf.fr

Abstract

Extreme floods occur in large territories and may cause significant damages to the railway system. As they are relatively rare, historical information about extreme floods is often not well known and used, whereas it could provide insights about: areas where the infrastructure is the most vulnerable, the resilience of the components of the railway system, and the financial and social impacts of extreme events. This paper deals with a framework to study historical extreme floods that can be used to model the flood processes impacting the railway system: IDISFER ontology. Ontologies are a flexible and interoperable data storage format, that is computer-readable and which can effectively process data. IDISFER includes all the relationships identified in the literature between natural flood phenomena and incidents on the railway system, causality relationships between several successive incidents (known as the 'domino effect') and chronological relationships between actions on the infrastructure and operations that are necessary to return the system to normal.

1. Introduction

Intense floods, sometimes called extreme floods, marked railway History from its early beginnings. These events occur in large territories and may impact every component of the railway system: infrastructure, operations, maintenance and human systems. Operational feedback on extreme floods provides information about areas where the infrastructure is the most vulnerable, about the resilience of railway system components and the socio-economic impacts of extreme events, such as the cost of infrastructure work, interruption of operations and the social impacts in areas where the train service is disrupted. But these events are rare (they occur every 20 or 100 years) and so the information about them is limited.

To investigate these issues, in 2014, SNCF RESEAU I&P initiated a research project to study the use of historical data about extreme past floods, in cooperation with the computer science Lab of Grenoble (LIG). It included a PhD thesis, based on four case studies of remarkable floods in France. The aim was to collect and analyze information about past flood phenomena, about their impacts on the railway system and about the management and restoration of the system after the events, up to the point where rail traffic conditions are returned to normal. Historical data was collected from the railway archives of SNCF, which holds back to the time of the first national railway lines (1840's). Complex flood-related processes were modelled with an ontology, called IDISFER, a format that allows information about past events to be stored and retrieved.

This article first deals with the choice of the ontology format and its relevance for the railway and the natural-risk domains. Secondly, related works about domain-related ontologies, on which the IDISFER concepts are aligned, are reviewed. Thirdly, the IDISFER ontology and the method used to build it are presented. Finally, the potential use of IDISFER to analyze past flood events is presented.

2. A computer science ontology: what for?

2.1. Definitions

In computer science, an ontology is a system for knowledge representation, which includes all concepts, properties and interrelationships between them that exist for a particular domain. As a dictionary of concepts, ontologies may have different forms and complexities: from simple glossaries or more structured taxonomies up to formal ontologies and logical inferences (Ushold and Gruninger 2004, in Kruchten et al. 2007).

An ontology consists of a set of 'triples' of concepts, that include an axiom, such as 'hasStation', which connects two objects together, for example a railway line and a railway station. General concepts are called 'classes' or types of object. An 'individual' is an instance of a class, that corresponds to a real world object. 'Relations' link classes and/or individuals to one another. Classes and individuals may also be related to 'attributes' via 'property' relations (Figure 1). In an ontology, every concept and relationship is documented to explain their meaning and to avoid ambiguity of terms.

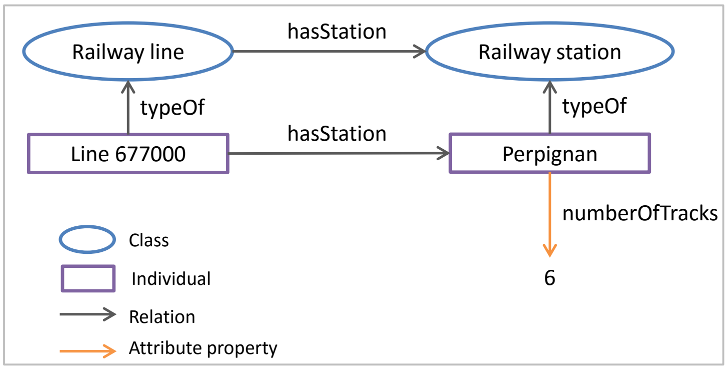

Figure 1: Example of ontology triples, classes, individuals and properties

2.2. Interests of ontologies

Ontologies are expressed in a computer readable format, such as RDF/OWL language. Formal ontologies can be used to organize information and store data, such as databases do. They can then be queried to retrieve data and to undertake analysis. One of their strengths is that they define semantics independently of the data representation (Lorenz et al. 2005). As a consequence, they are a more flexible and sharable format than databases. Flexible because any stakeholder may take, re-use or modify a part of the ontology schema and append its own extending concepts, without the need to modify the whole existing structure (Morris et al. 2015). Sharable because, by nature, they are designed independently of the format, the size and the structure of the data to be stored (Lorenz et al. 2005). Finally, their construction process may enable stakeholders to agree on a shared vocabulary. They are appropriate to be used as standard exchange formats between partners (Köpf 2010) and are an effective tool to guarantee data interoperability, which is needed in an open competitive and trans-national railway landscape.

2.3. Benefits of an ontology describing floods for the railway management

In the field of natural risks, ontologies have been developped either to assess the level of risk in local territories (e.g. Scheuer et al. 2013), to monitor at-risk events in real-time (e.g. Kollarits and Wergles 2009) or to set a common vocabulary and interoperable software for the actors involved in crisis management (e.g. Bénaben et al. 2008; Babitski et al. 2009). Regarding railway risk management, some previous work have planned the practical use of ontologies. Ontologies may allow inferences to be made to show which lines would be unavailable in the case of a failure at a given point of the network (Morris et al. 2015). They can hasten communications between railway actors involved in the management of an unexpected fault (operation, traffic management, maintenance), thanks to a shared information structure, and they can feed data into expert decision-support tools, thus improving efficiency of the event management (Umiliacchi and Henning 2008). They also can be an asset for predictive maintenance: by analyzing historical data, they can identify patterns of events which contribute to possible faults and evaluate the probability that they occur in a certain time period (Umiliacchi and Henning 2008).

Interviews with experts have identified that the knowledge about historical floods and their impacts may have three other uses. First, it may be used to improve knowledge about the infrastructure managers' network, by identifying which components (objects) of its system were damaged and which were robust during past events. The knowledge about past floods may produce statistics about flood types and the associated kinds of damage to the railway. Historical data about the costs of rebuilding and the length of traffic disruption during and after the flood may be used to conduct cost-benefit analyses about the building of protective structures and about the economical and social impacts of lengthy traffic disruption in the case of major events. Finally, the crisis management phase could base on establishing which events and issues could happen in each area and re-use of past procedures which were effective.

3. Related work about domain-related ontologies

An ontology that describes floods and their impacts on the railway system should deal with hydro-climatic phenomena involved in floods, the railway system and the kinds of impact these natural phenomena may have on the railway system. This section describes some existing related ontologies, from which IDISFER may re-use concepts (this step is called 'ontology alignment').

3.1. Ontologies of the railway system

Railway ontologies focus on different levels of detail, from an overall level, such as the Railway Domain Ontology (RDO) (Köpf 2010) to a very detailed level for specific problem solving, for example vehicle/track integration (Lu et al. 2006). Here, we focus on an overall level, firstly because floods may affect every component (object) of the railway system and secondly because macro-level information about generic components impacted by past floods is often the only level available.

A first existing formalization of railway data is the RailML[1] standard for data exchange. It has described Infrastructure, Rolling Stock and Timetabling, in an XML format. However, it is not a formal ontology, so it lacks some functionality. The first ontology that described the entire railway system in its complexity was the InteGRail Railway Domain Ontology (Umiliacchi and Henning 2008). This ontology has described four parts of the railway system: rolling stock, infrastructure, train operation and traffic management. The RDO has been used as a basis for the ongoing RaCoOn (Railway Core Ontologies) project (Morris et al. 2015), which aimed to model the rail domain in more details than the RDO, by a set of railway domain ontologies.

Other ontologies have dealt with specific aspects of the railway domain and may be a source for inspiration: rail traffic (Inkster 2009), train timetabling (Mohan and Arumugam 2005) or validation of new railway line plans (Lodemann and Luttenberger 2010).

3.2. Ontologies of natural risks and floods

A risk is defined as the co-occurrence in time and space of a hazard (occurence of a potentially dangerous phenomenon) and of vulnerable exposed elements (human, economical, environmental).

---

[1] https://www.railml.org/en/

As a consequence, ontologies of natural risks deal with five different aspects to describe at-risk situations (Hajji 2005; Kruchten et al. 2007; Bénaben et al. 2008; Scheuer et al. 2013):

- the physical territory subject to a risk (topography, hydrology),
- exposed elements (people, infrastructures, businesses…),
- natural phenomena and their properties (intensity, probability, extent),
- the impacts of natural events on the exposed elements (or damages) and
- the actors, procedures, resources and territories of crisis management.

Natural risks include *phenomena* (or *dangers*), which are the kinds of hazard that might happen (e.g. a flood). The realization of a phenomenon is called an *event* (e.g. the flood of October 3rd 1988 in Nîmes, France) (Bénaben et al. 2008).

Properties may be associated with a flood phenomenon. Its nature can be an overflow by a watercourse or a lake, a submersion by the sea, a flash flood, an increase of the water table level or a surface run-off. Its hydro-meteorological causes may be a storm, a tropical storm, high long-term precipitation records, fast snow melt or an engineering failure (EM-DAT DB - CRED 2009; Lang and Coeur 2014). To describe the characteristics of hydrogeological elements, it is possible to use and extend high-level environmental ontologies such as the SWEET[2] ontology.

3.3. Ontologies of events and impacts of natural events on infrastructures

According to Scheuer et al. (2013), a flood event may cause damage to infrastructure, populations and the environment. They have distinguished two classes of infrastructure: material infrastructure (such as buildings) and institutional infrastructure (the administrative environment). For example, impacts on the railway trigger the setting up of a crisis unit. Moreover, some damaging events may cause other events, which is known as the ‘domino effect’ in natural risk terminology. A conceptual ontology, which includes domino effects and their associated damages, has been developped to study disaster events (Dubos-Paillard and Provitolo 2012). The ‘event’ class has two mereological relationships: a causal relationship (an event may cause an event) and a composition relationship (a sub-event may compose an upper-event). In this ontology, events and damage ‘bear on’ (i.e. ‘impact’) territorial elements.

3.4. Limits of previous work

Existing ontologies about the railway system, natural risks and their impacts have introduced a lot of useful concepts. The next step is to merge them in a single ontology scheme in order to describe the complex interactions between these systems. Indeed, among the ontologies we explored, it did not seem that an ontology about the impacts of natural risks on infrastructure networks exists. Neither did we find any ontology of natural risks that dealt with the post-crisis period, for example the rebuilding steps for infrastructure, operations management during traffic disruption or the length of the period before the system is returned to its normal state. It could be useful to post-analyze the events and to evaluate their costs. Finally, to describe the impacts of a flood on the railway system, it seems appropriate to consider this system at a meso-detailed level. However, ontologies about the railway system have remained at a macro-level, such as the RDO. The RaCoOn ontology is promising but the project seems to be still ongoing and the ontology is not fully published yet.

The IDISFER ontology aims at filling the gaps in the ontology concepts and at merging existing ones to enable data analysis about flood events that have impacted the railways. It includes causality relations between natural phenomena and cascading incidents on the railway system, which are of importance when studying major risks. It also includes a temporal ontology to describe chronological relations between the events, until the restoration of normality in railway operations.

---

[2] https://sweet.jpl.nasa.gov/

4. IDISFER Ontology: modelling the impacts of floods on the railway system

4.1. A participative method to build the ontology

To gather the concepts to be included in the ontology, we organized a cooperative workshop with railway system experts of the French railway company (SNCF) working in different domains of expertise: maintenance, operations and infrastructure components. This exercise aimed to collect every object of the railway system, their level of vulnerability against floods and how much they contribute to the functioning of the system in the case of malfunction. It was organized in a *World Café* format: five tables were distributed in a room with a preliminary components tree diagram, for a different domain of the railway system, provided on each table. Participants were divided into groups of three to five persons with different expertise. They then had to complete the branches and leaves of the graph. After 10 minutes, the participants changed tables and continued to complete the diagram on another railway system domain. This method enabled information on a complex topic to be gathered in an informal atmosphere, according to a well-defined plan, that is favorable for co-operative working. The collected concepts could then be utilized in a formal ontology.

4.2. Presentation of IDISFER ontology

4.2.1. Classes

IDISFER ontology comprises four sub-sections. The current version of IDISFER includes one part describing the railway system, one about natural phenomena involved in floods, one about the impacts of natural events on the railway system (called ‘incidents’), and one about disturbances of the rail operations and the reactions necessary to restore the damaged system. Figure 2 represents a simplified schema of the overall structure of IDISFER.

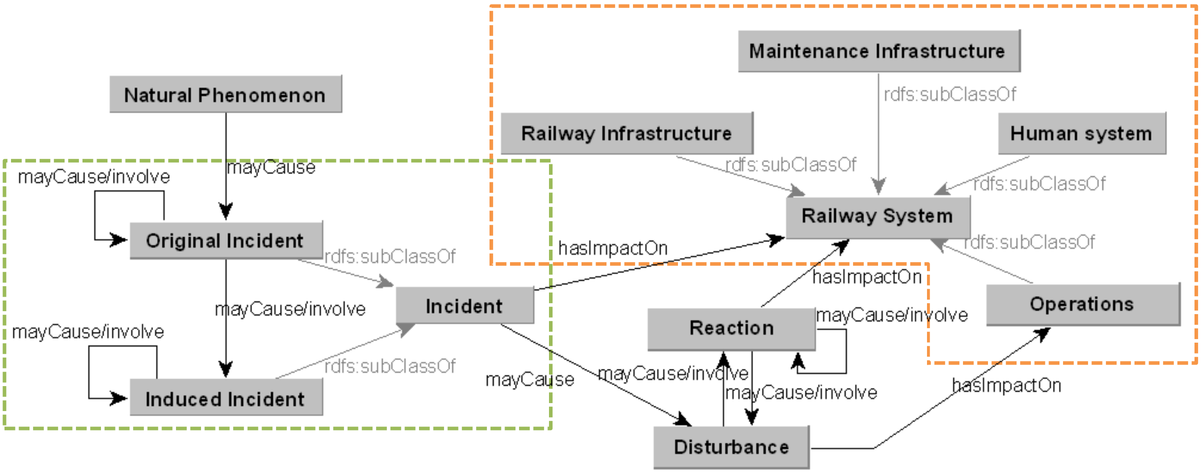

Figure 2: Simplified RDF Schema of IDISFER Ontology. The two main parts (in number of classes) are Incidents (left dashed frame) and Railway System (right dashed frame).

The two main parts of IDISFER, regarding the number of classes, are those for the railway system and for flood-related railway incidents. The railway system includes four sub-parts: railway infrastructure (68 classes), infrastructure for maintenance (8 classes), operations (41 classes, with 3 about stations that are shared with railway infrastructure) and the human system (1 class, which is ‘commuting travel of agents’ that may be impacted by floods). The incidents are divided in two parts: original incidents (19 classes), which are classes of incidents that may be a direct consequence of a natural phenomenon, and induced incidents (16 classes), which cannot be a direct consequence of a natural phenomenon but may be involved due to previous incidents. For example, a platform heave or a ground movement (original incident) may cause a surface deviation of the rail (induced incident).

4.2.2. Properties

Natural phenomena, incidents, distrubances and reactions are types of ‘events’, as described by the event ontology LODE (Troncy et al. 2010). Each event is characterized by a place and a date.

For spatial referencing, we used two approaches: linear referencing for events that occured along railway lines, following the concepts of the ISO norm 19148:2012 (ISO 2012), and latitude/longitude for events which occured elsewhere in the environment, connecting as much as possible spatial features to placenames expressed in the Geonames open database (GeoNames 2014).

To describe temporal information, we have used the Timeline Ontology (Raimond and Abdallah 2007) that is based on the ISO norm 19108 about Temporal Schema and on the W3C’s ontology OWL-Time (W3C 2006). The Timeline Ontology uses two main concepts: *instants* (a punctual position in time) and *intervals* (an extent in time). Instants and intervals are located on a *timeline*, relative to one or several *temporal coordinates systems*. An instant has a date, either defined precisely (atDate, atDateTime) or less precisely (atYear, atMonth, atDuration, meaning that the date is located during a fuzzy interval). An interval ‘starts’ and ‘ends’ at particular dates and has a *duration*.

As IDISFER is intended to deal with historical information, which is sometimes imprecise or uncertain, every attribute of an event (date, location, description) may be associated with an *imperfection* property. Imperfection descriptors consist of qualitative terms (e.g. ‘inaccurate’, ‘inconsistent’, ‘unreliable’), that refer to types of imperfection described in the imperfection taxonomy in (Saint-Marc et al. 2016, under review).

IDISFER ontology is shared under a Creative Commons CC-BY-SA 4.0 licence at this address: *http://lig-membres.imag.fr/saintmar/ontology/IDISFER.owl*.

5. Conclusion and outlooks

Ontologies are a good solution to model extreme event processes. The innovation of IDISFER is to merge several domains of interest: the railway system and natural risks, with in-chain events characterizing extreme risks. IDISFER includes 170 classes, 12 objects relations and 34 data properties. It was implemented with the software Protégé 4.3 and used to model three historical floods up to today that occurred in 1930, 1940 and 2001. Although it tries to be as complete as possible, IDISFER do not pretend to be exhaustive and might still evolve in the future. Indeed, the building of an ontology is an iterative and a dynamic process and expresses one way to model a domain among many alternatives, while the best way depends on its application (Mohan and Arumugam 2005). The main interests of ontologies are to formalize concepts and to be highly flexible and re-usable: ontologies are “achieving interoperability between multiple representations of reality [...] and between such representations and reality, namely human users and their perception of reality” (Hepp 2007, in Lodemann et al. 2010).

To continue the DHI-Avenir project, data stored in IDISFER will be used to statistically analyze the possible impacts of floods that will occur in the future at a given location and of a given hydro-climatic type. Knowing the past costs of rebuilding and the length of traffic disruption, preventive maintenance and works may be planned. Historical narratives of flood events will be shown to experts in a visual form, in an interactive and animated cartographical display (currently in development) in order to efficiently analyze information and make decisions.

6. References

Babitski G, Probst F, Hoffmann J, Oberle D (2009) Ontology Design for Information Integration in Disaster Management. Inform 2009 3120–3134.

Bénaben F, Hanachi C, Lauras M, et al (2008) A metamodel and its ontology to guide crisis characterization and its collaborative management. In: Fiedrich F, Van de Walle B (eds) Proceedings of the 5th International ISCRAM Conference. Washington DC (USA), pp 189–196

CRED (2009) EM-DAT General Classification. http://www.emdat.be/classification. Accessed 27 Feb 2015

Dubos-Paillard E, Provitolo D (2012) Analysis of the 2011 Japan's triple disaster through an ontology of the domain of risk and disaster. In: Complex System and Design Management. Paris (France), p 14

GeoNames (2014) About GeoNames. http://www.geonames.org/about.html. Accessed 27 Feb 2015

Hajji H (2005) Gestion des risques naturels : Une approche fondée sur l'intégration des données. Institut National des Sciences Appliquées de Lyon

Hepp M (2007) Ontologies: State of the Art, Business Potential, and Grand Challenges. In: Hepp M, De Leenheer P, Sure Y (eds) Ontology Management. Springer US, Boston (MA, USA), pp 3–22

Inkster T (2009) Linked Railway Data Project. http://ontologi.es/rail/#project. Accessed 9 Apr 2015

ISO/Technical Committee 211 (2012) ISO 19148:2012, Geographic information - Linear referencing.

Kollarits S, Wergles N (2009) MONITOR - an ontological basis for risk management.

Köpf H (2010) InteGRail - Publishable Final Activity Report.

Kruchten P, Woo C, Monu K, Sotoodeh M (2007) A Human-Centered Conceptual Model of Disasters Affecting Critical Infrastructures. In: Van de Walle B, Burghardt P, Nieuwenhuis C (eds) Proceedings of the 4th International ISCRAM Conference. Delft (Netherlands), p 19

Lang M, Coeur D (2014) Les inondations remarquables en France. Quae

Lodemann M, Luttenberger N (2010) Ontology-Based Railway Infrastructure Verification, Planning Benefits. In: International Conference on Knowledge Management and Information Sharing - KMIS 2010. Valencia (Spain),

Lorenz B, Ohlbach HJ, Yang L (2005) Ontology of Transportation Networks. Munich (Germany)

Lu J, Roberts C, Lang K, et al (2006) The Application of Semantic Web Technologies for Railway Decision Support. In: Gupta JND, Forgionne GA, Mora T. M (eds) Intelligent Decision-making Support Systems: Foundations, Applications and Challenges. Springer-Verlag, London, pp 321–337

Mohan R, Arumugam G (2005) Constructing Railway Ontology using Web Ontology Language and Semantic Web Rule Language. Int J Comput Technol Appl 2:314–321.

Morris MC, Easton J, Roberts C (2015) Ontology in the Rail Domain: The Railway Core Ontologies. In: 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - IC3K 2015. Lisbon (Portugal), pp 285–290

Raimond Y, Abdallah S (2007) The Timeline Ontology. http://motools.sourceforge.net/timeline/timeline.html. Accessed 4 Mar 2015

Saint-Marc C, Villanova-Oliver M, Pams Capoccioni C, et al (2016) Representation and Visualization of Imperfect Geohistorical Data about Natural Risks: A Qualitative Classification and its Experimental Assessment. In: Agile 2016. Helsinki (Finland), p 21

Scheuer S, Haase D, Meyer V (2013) Towards a flood risk assessment ontology - Knowledge integration into a multi-criteria risk assessment approach. Comput Environ Urban Syst 37:82–94. doi: 10.1016/j.compenvurbsys.2012.07.007

Troncy R, Shaw R, Hardman L (2010) LODE: une ontologie pour représenter des événements dans le web des données. In: 21ème Journées Francophones d'Ingénierie des Connaissances (IC'2010). Nîmes (France), p

19

Umiliacchi P, Henning U (2008) Standardized Data Interchange Between Railway Systems: an Integrated Railway Information System. In: WCRR 2008. Seoul (Korea),

Ushold M, Gruninger M (2004) Ontologies and semantics for seamless connectivity. J SIGMOD Rec 33:58–64.

W3C (2006) Time Ontology in OWL. http://www.w3.org/TR/2006/WD-owl-time-20060927. Accessed 4 Mar 2015