Ontology-based Disruption Scenario Generation for Critical Infrastructure

Paolo Trucco[a*], Boris Petrenj[a] and Massimiliano De Ambroggi[a]
[a] Politecnico di Milano, Milan, Italy

Abstract: Critical Infrastructures (CIs) are exposed to a wide spectrum of threats which vary in nature and can be either internal or external. THREVI2 project aims at assisting Authorities and Operators to get comprehensive information of all potential disruption scenarios relevant for CIP through creating a comprehensive and multi-dimensional all-hazards catalogue for CI. It consists of two ontologies (*CI systems* and *Hazards & Threats* affecting CI) connected through vulnerability and (inter)dependency models. Its final implementation in a software tool (PATHFINDER) is aimed to support analysts in specifying the overall system of systems and generating a set of relevant disruption scenarios.
The paper presents the adopted ontology development process and describes the main features of the final integrated set of ontologies. CI ontology covers *Energy*, *Transport*, *Water* and *Telecommunications* sectors, comprising 11 subsectors in total – each being described through two sub-ontologies (physical and functional) interconnected within the service delivery topology. Hazard & Threat ontology systematically characterises different typologies of events, their attributes, types and possible effects to CI systems. The validation process of the final ontologies is also described. Finally, the progress and challenges in modeling interdependencies is discussed, as well as further developments that will take place.

Keywords: Critical Infrastructure, Scenario Generation, Ontology, Disruption

1. INTRODUCTION

Critical Infrastructures (CIs) are exposed to a wide spectrum of hazards and threats which vary in nature (natural, technological, human-intentional or non-intentional) and, that can be external to the infrastructure (e.g. flood, chemical explosion, terrorist attack) or internal (e.g. technical failure, sabotage, human error). As such, hazard and threat assessment is a key element within CIP strategies and CI risk assessment, in particular:

- It is part of the CI identification process (e.g. the EC Directive 2008/114/EC requires to develop "worst case scenarios", to simulate the failure of a potential ECI, in order to assess the transboundary impacts on other Member States);
- At the infrastructure level, risk assessment is applied to define accurate protection measures (e.g. European CI operators have to include in the Operator Security Plan "a risk analysis based on major threat scenarios, vulnerability of each asset, and potential impacts").

Though, it is very difficult to get an overview of all existing hazards and related vulnerability of CI, since the existing information most often focuses only on one type of hazard or on the vulnerability of one type of target (a single infrastructure or asset). This is mainly due to the fact that hazard, threat and vulnerability assessments can be very specific in nature and related to different disciplinary fields. It is therefore difficult for Authorities or operators to get comprehensive information of all potential disruption scenarios relevant for CIP.

THREVI2 project aims at contributing to this major issue in CIP activities thanks to the achievement of two main objectives:

- to create a comprehensive and multi-dimensional all-hazards catalogue for critical infrastructures;

[*] paolo.trucco@polimi.it

- to develop a software tool (PATHFINDER) to support the analyst in specifying the overall system of systems and generating a set of relevant disruption scenarios.

To this end, an ontology-based approach has been implemented and two interconnected ontologies have been developed:

- an ontology of Hazards and Threats affecting CI;
- an ontology of physical and functional topologies of Critical Infrastructure systems;

merged through specific vulnerability and dependency entities.

The paper presents the ontology development process adopted in the project and describes the main features of the final integrated set of ontologies. The paper is organized as follows. The following section explains the ontology and this kind of the project approach. Section 3 goes through phases of ontology development, including data collection and analysis, and presents the main results. In Section 4 the validation process has been shortly described. Section 5 summarizes the remaining issues and the path forward regarding the final tool (PATHFINDER) development.

2. THE ONTOLOGY DEVELOPMENT PROCESS

An ontology can be defined as '*a formal description of entities and their properties, relationships, constraints, behaviours*' [1]. Or simpler, an ontology is a *specification of a conceptualization* [2]. Ontologies are used to capture and share knowledge about some domain of interest. Ontology deals with questions concerning what entities exist or can be said to exist, and how such entities can be grouped, related within a hierarchy, and subdivided according to similarities and differences. It is used to describe the concepts and relationships that are important in a particular domain, providing a vocabulary for that domain as well as a computerized specification of the meaning of terms used in the vocabulary. Ontologies range from taxonomies and classifications, database schemas, to fully axiomatised theories. In recent years, ontologies have been adopted in many business and scientific communities as a way to share, reuse and process domain knowledge. Ontologies are now central to many applications such as scientific knowledge portals, information management and integration systems, electronic commerce, and semantic web services. [3]

Different ontology languages provide different facilities. The most recent development in standard ontology languages is OWL from the World Wide Web Consortium (W3C). OWL makes it possible to describe concepts but it also provides new facilities. It has a richer set of operators and it is based on a different logical model which makes it possible for concepts to be defined as well as described. Complex concepts can therefore be built up in definitions out of simpler concepts [4]. Furthermore, the logical model allows the use of a reasoner which can check whether or not all of the statements and definitions in the ontology are mutually consistent and can also recognise which concepts fit under which definitions. The reasoner can therefore help to maintain the hierarchy correctly [4].

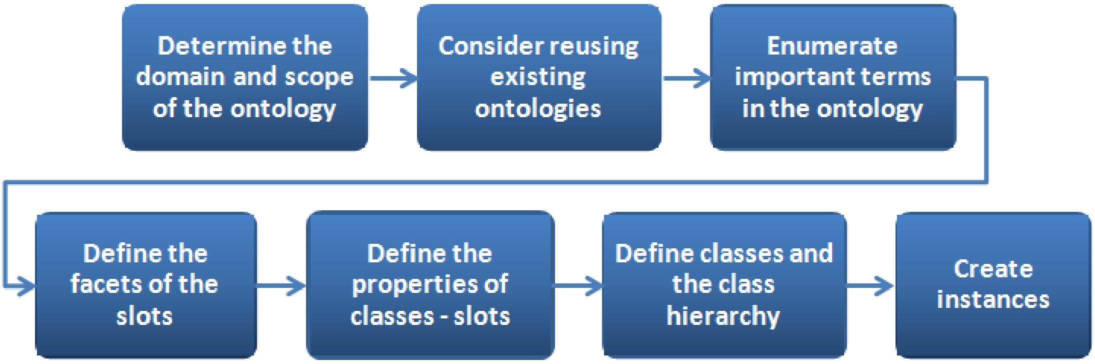
In the first place we have to organise knowledge in specific domain, using scattered data and various sources. For the two main domains (*Critical Infrastructure* and *Hazards & Threats*) it is necessary to determine what concepts exist, and to describe and classify them within the domain in a systematic way. Ontology is a structure that allows creating a conceptual map organizing elements within a domain, using *classes*, *properties* and *instances*. Any class can contain many subclasses organized on different levels. An instance is an "object" within the ontology domain which is described using the relevant classes and properties. A property is a directed binary relation that specifies class characteristics; generally, they are attributes of instances and sometimes act as data values or as link to other instances. Properties may possess logical capabilities such as being transitive, symmetric, inverse and functional. Properties may also have domains and ranges. An ontology together with a set of individual *instances* of classes constitutes a *knowledge base*. [5]

We use taxonomies to describe how different classes are related by organising them into groups and/or hierarchies (according to level of detail). Adopting a standardised description is also important for systematic connection between the taxonomies to the other parts of the ontology.

To make ontology development description clearer we describe typical steps in an ontology development process, as presented in Figure 1; it includes [5]:

- defining classes in the ontology;
- arranging the classes in a taxonomic (subclass–superclass) hierarchy;
- defining slots and describing allowed values for these slots;
- filling in the values for slots for instances.

Figure 1: Process of Ontology Development (adapted from [5])


Determine the domain and scope of the ontology
Consider reusing existing ontologies
Enumerate important terms in the ontology
Define the facets of the slots
Define the properties of classes - slots
Define classes and the class hierarchy
Create instances

The ontology development process consisted of the following activities:

- Defining the purpose and context of the ontology;
- Collecting the data needed for the ontology development;
- Data analysis and development the ontology based on the data;
- Evaluation, verification and update.

Ontology development is an iterative process. It is also a model of reality of the world. In order to make sure that the concepts in the ontology reflect reality, that the ontology fits intended use, and that we have an adequate level of detail, we are facilitating ontology development through:

- continuous alignment and evaluation between project WPs;
- validation carried out through interviews with experts in the domain.

3. ONTOLOGY-BASED SPECIFICATION OF CI SYSTEMS AND RELATED THREATS

3.1. Critical Infrastructure System Ontology

For the Critical Infrastructure System Ontology, the covered sectors include *Energy*, *Transport*, *Water* and *Telecommunications* for a total of 11 different infrastructure subsectors (Figure 2). Each CI is described by means of two interconnected sub-ontologies: one for the physical specification of CI topology, the other for the functional specification. The overall infrastructure ontology framework is organized in three parts – assets, functions and services (Figure 3) which were subsequently linked with the service delivery process (Functional ontology). The general concept of physical specification is to arrive at a complete and systematic physical description of the infrastructure thanks to a standardized nomenclature and definition of its most relevant elements. The goal for this effort was to deliver this capability by a mixed and harmonized used of international standards. More specifically, the Critical Infrastructure System ontology has been developed using different data sources:

- *Regulatory* – standards and codes adopted or required by government and public bodies
- *Technical/Professional* – standards and codes developed by industrial or professional associations, standardisation bodies, etc.
- *Scientific* – modelling and descriptive methods adopted in studies and tools reported in scientific literature

Figure 2: Covered infrastructure sectors and sub-sectors

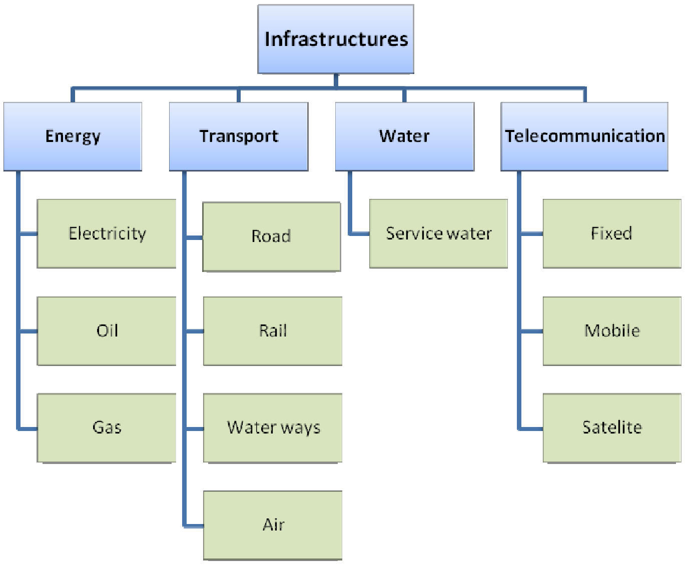


Figure 3: General CI ontology framework

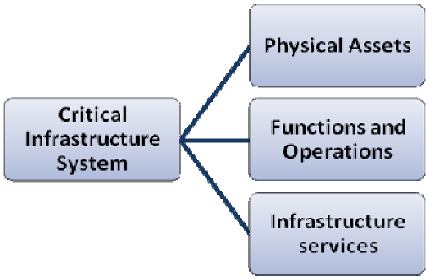
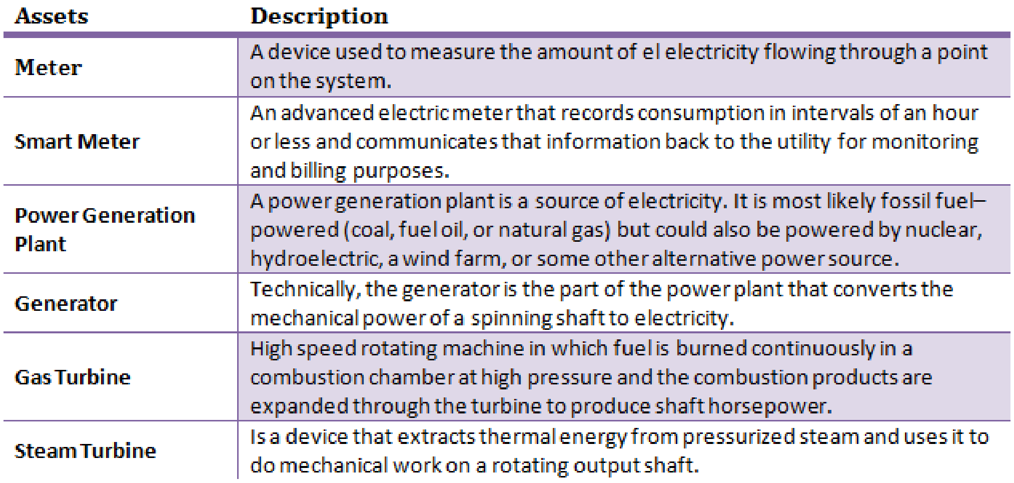


Table 1: Summary of used references

Document type	Number of sources
Regulatory	**24**
National	(17)
International	(7)
Technical	**25**
National	(11)
International	(14)
Scientific	**13**
TOTAL	**62**

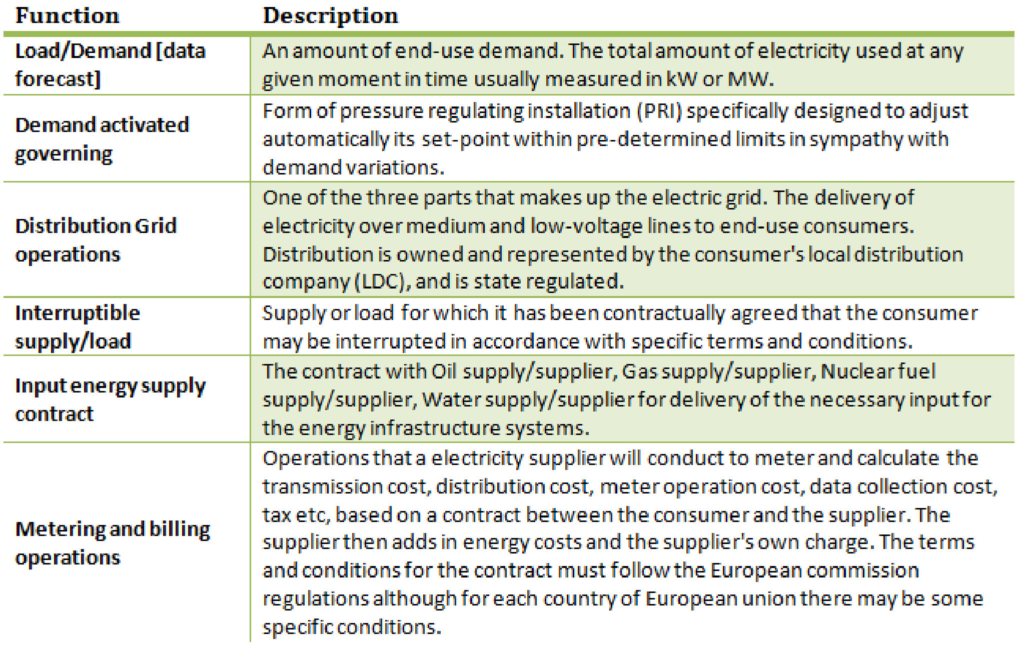
Globally, more than 100 references, of scientific, technical and regulatory nature, have been identified and systematically reviewed. After analysis, a portion of the sources turned out not to be useful for the ontologies description. At the end 62 documents have been used to develop the final 22 sub-ontologies (Table 1). Through this literature review, for each type of infrastructure a list of physical assets and functions has been derived. All the listed assets are classified according to a common classification scheme developed during the analysis, accompanied with a standardised physical description (Figure 4). Similarly, a list of functions with associated descriptions is given for each infrastructure sector (Figure 5).

Figure 4: List of assets and their description - Electricity sector example



Assets	Description
Meter	A device used to measure the amount of el electricity flowing through a point on the system.
Smart Meter	An advanced electric meter that records consumption in intervals of an hour or less and communicates that information back to the utility for monitoring and billing purposes.
Power Generation Plant	A power generation plant is a source of electricity. It is most likely fossil fuel–powered (coal, fuel oil, or natural gas) but could also be powered by nuclear, hydroelectric, a wind farm, or some other alternative power source.
Generator	Technically, the generator is the part of the power plant that converts the mechanical power of a spinning shaft to electricity.
Gas Turbine	High speed rotating machine in which fuel is burned continuously in a combustion chamber at high pressure and the combustion products are expanded through the turbine to produce shaft horsepower.
Steam Turbine	Is a device that extracts thermal energy from pressurized steam and uses it to do mechanical work on a rotating output shaft.

Figure 5: List of functions and their description - Electricity sector example



Function	Description
Load/Demand [data forecast]	An amount of end-use demand. The total amount of electricity used at any given moment in time usually measured in kW or MW.
Demand activated governing	Form of pressure regulating installation (PRI) specifically designed to adjust automatically its set-point within pre-determined limits in sympathy with demand variations.
Distribution Grid operations	One of the three parts that makes up the electric grid. The delivery of electricity over medium and low-voltage lines to end-use consumers. Distribution is owned and represented by the consumer's local distribution company (LDC), and is state regulated.
Interruptible supply/load	Supply or load for which it has been contractually agreed that the consumer may be interrupted in accordance with specific terms and conditions.
Input energy supply contract	The contract with Oil supply/supplier, Gas supply/supplier, Nuclear fuel supply/supplier, Water supply/supplier for delivery of the necessary input for the energy infrastructure systems.
Metering and billing operations	Operations that a electricity supplier will conduct to meter and calculate the transmission cost, distribution cost, meter operation cost, data collection cost, tax etc, based on a contract between the consumer and the supplier. The supplier then adds in energy costs and the supplier's own charge. The terms and conditions for the contract must follow the European commission regulations although for each country of European union there may be some specific conditions.

The physical arrangement of a generic infrastructure has been specified using class hierarchy in OWL language and implemented in Protégé software (Figure 6). As for the design of the Functional Ontology of CI, the aim was to cover all operations phases from the acquisition of resources (supply side) to the final service delivery to end users (demand side). Therefore, all the functional sub-ontologies have been organized with reference to a standardized functional representation of a general service delivery process (Figure 7). Accordingly, the highest level of the functional ontology has been organized into five main phases:

- sourcing;
- source stock;
- service generation;
- service stock;
- service delivery.

The lower layers of each functional sub-ontology contain more detailed functions specific for each CI sector and sub-sector..

Figure 6: Portion of the Physical Asset sub-ontology for Electricity Infrastructure

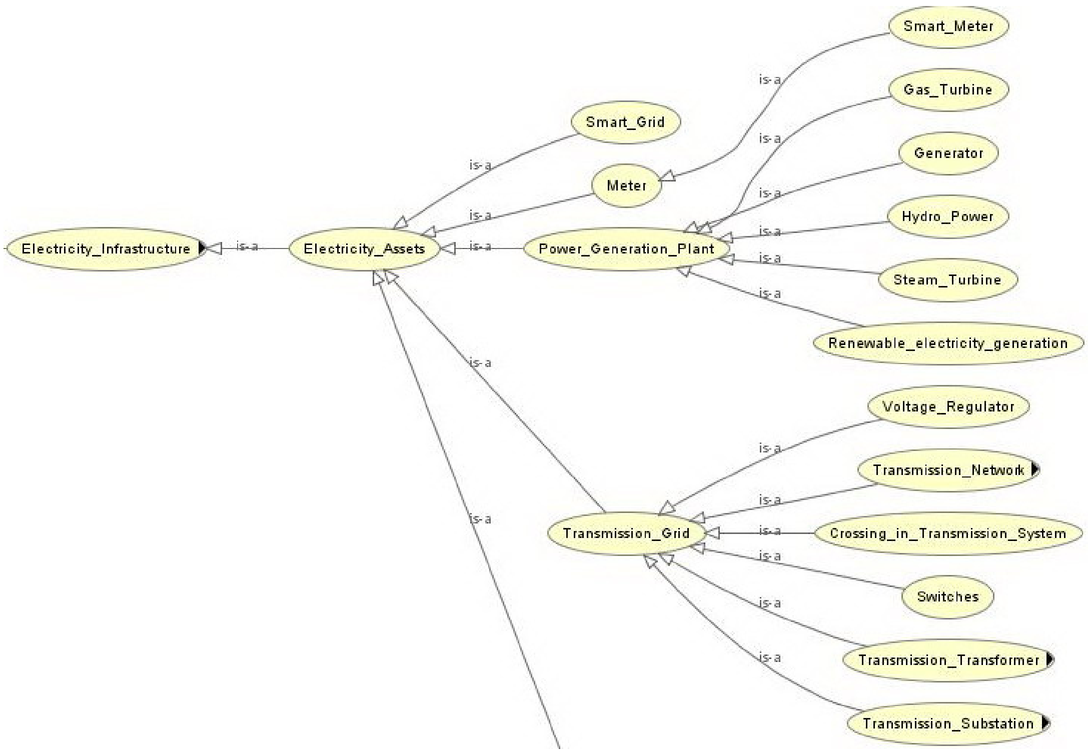
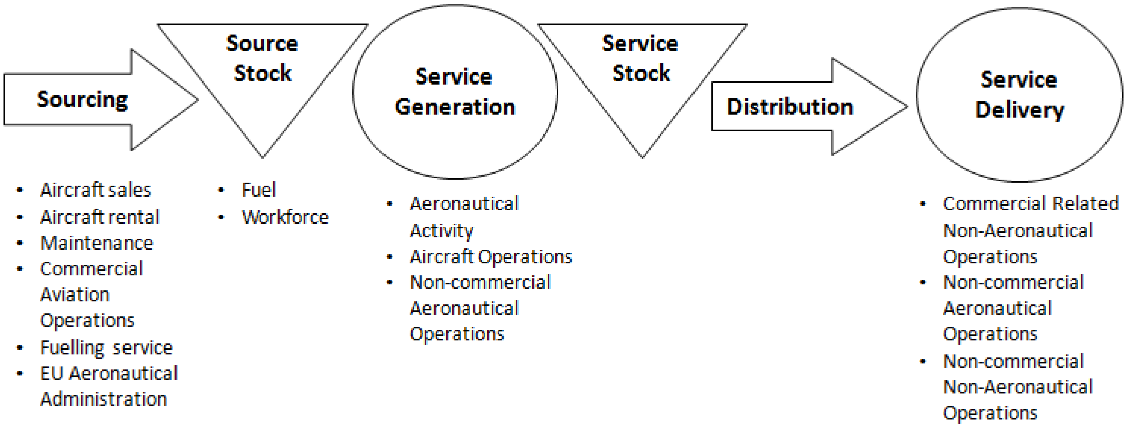


Figure 7: General Service delivery process (Air transport functions used as an example)


Sourcing
Source Stock
Service Generation
Service Stock
Distribution
Service Delivery
• Aircraft sales
• Aircraft rental
• Maintenance
• Commercial Aviation Operations
• Fuelling service
• EU Aeronautical Administration
• Fuel
• Workforce
• Aeronautical Activity
• Aircraft Operations
• Non-commercial Aeronautical Operations
• Commercial Related Non-Aeronautical Operations
• Non-commercial Aeronautical Operations
• Non-commercial Non-Aeronautical Operations

The relations (links) between Assets and/or Functions have been defined adopting the Integration Definition Function Modeling (IDEF∅) standard, generally used for developing structured representations of a system or enterprise. This formalism allows the construction of models comprising system functions (activities, actions, processes, operations), functional relationships, and data (information or objects) that support systems integration [6]. More specifically, a ‘function model’ is a structured representation of the functions, activities or processes within the modelled system or subject area. It relates classes to its inputs (e.g. requirements, materials), mechanisms (e.g.

resources, assets), controls (e.g. plans, legislations, monitoring) and outputs (e.g. functions, services) – Figure 8.

In the context of THREVI2 the IDEF∅ nomenclature is used to specify relations among classes, belonging to both the physical and functional ontologies, in terms of:

- input: entity X *is input to* entity Y (or, entity Y *has input* entity X), e.g. the output of a function or an operations activity is the input of another one;
- mechanism: entity X *is mechanism of* entity Y (or, entity Y *has mechanism* X);
- control: entity X *controls* entity Y (or, entity Y *is controlled by* entity X);
- output: entity X *delivers* entity Y (or, entity Y *is delivered by* entity X).

Figure 8: An IDEF Meta Model [6]

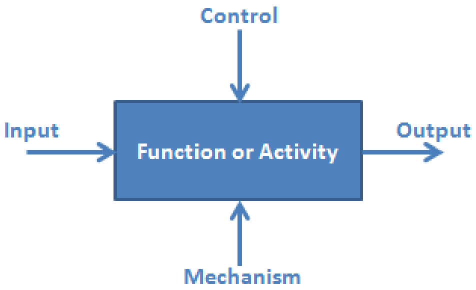
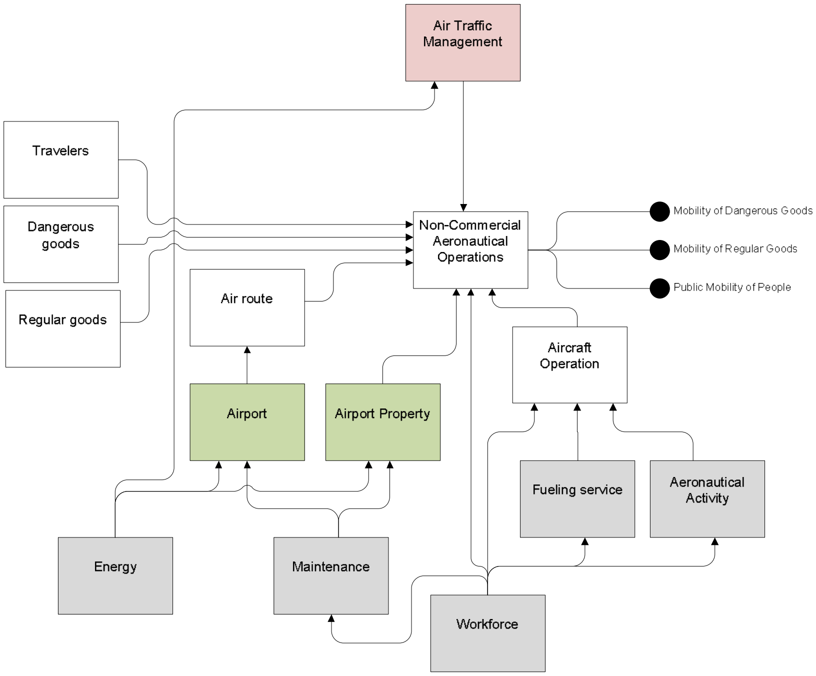


Figure 9 represents an example of links (relations) between Assets and/or Functions within the service delivery process, where *Service Generation* stage within Air transport sector has been used as an example.

Figure 9: Service generation in Air transport sector



3.3. Hazards and Threats Ontology

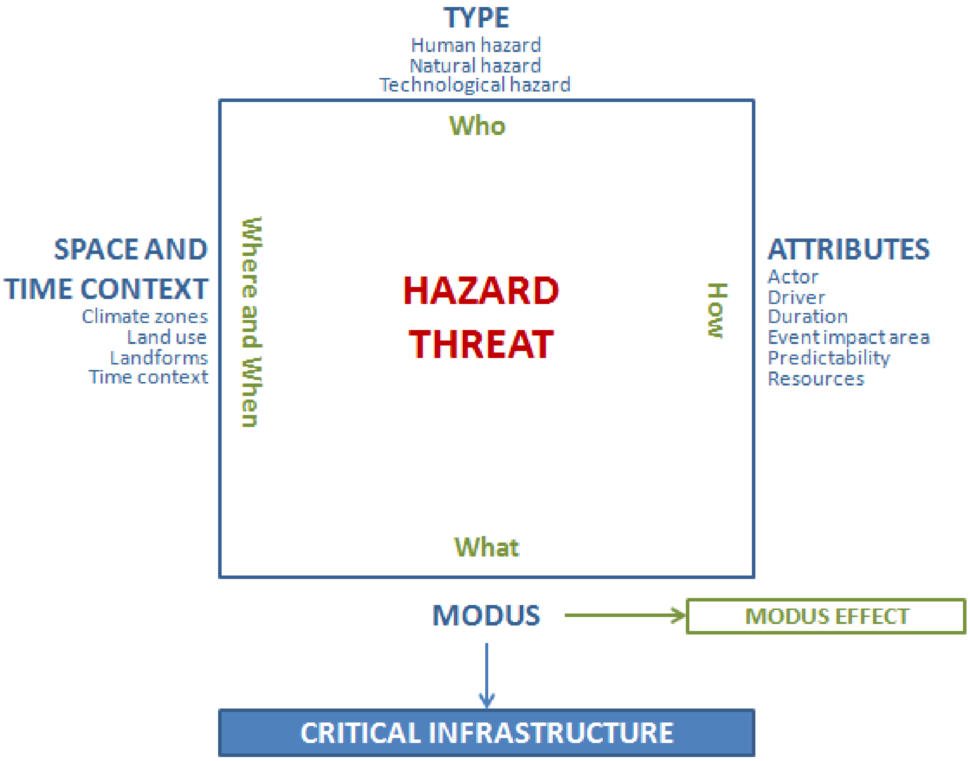
As for the Hazards and Threats (H&T) Ontology, there is no standardized form of data collection. However, countries have collected a lot of information on hazards in various databases at international, country, regional and local levels. This task seeks to describe the available body of knowledge about phenomenological behavior of natural and technological hazard and modus-operandi related to intentional acts. It will describe original causes that can trigger the manifestation of hazard (e.g. the extreme meteorological event can originate a landslide) or motivation malicious human behavior (e.g. religious motivation or political strategy based on the generation of public terror). Terroristic attacks aim at selected targets deliberately chosen on the base of the weaknesses of defenses and the level of preparedness and imply a wide range of means and methods. Our society presents an almost infinite array of potential targets that can be attacked through a variety of methods. The full set of information about the hazards and threats considered by the project therefore needs to be precisely defined and justified.

Hazard and Threats (H&T) Ontology aims to systematically characterise different typologies of events to be used for the development of the Pathfinder tool that allows first-order recognition of:

- all possible threats that can affect or destroy a generic critical infrastructure;
- all possible infrastructure that can be affected or destroyed by a specific threats.

The ontology is based on a hierarchical structure (classes and sub-classes), and is developed considering the possibility to reuse available literature on threat classification. For each class, a set of features (duration, impacted area, etc.) are assigned to better characterize the classes and to allow flexible navigation of the user within the ontology.

Figure 10: Hazard & Threat ontology

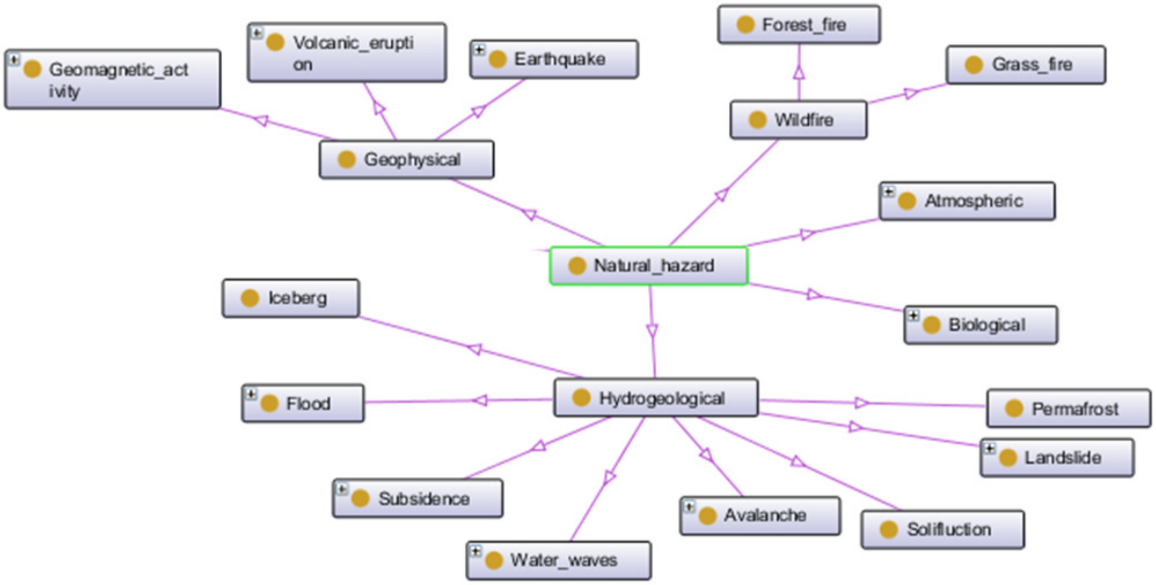


The overall Hazards & Threats Ontology framework is organized in four interconnected sub-ontologies, each one responding to a simple question (Figure 10):

1. Who is the hazard? *Events Type sub-ontology*. Potential events sub-ontology is created in the form of a hierarchical taxonomy. At the first level of this hierarchical taxonomy identifies the considered hazards have been classified as *Natural* (e.g. flood, landslide, etc.), *Technological* (e.g. dysfunction of equipment or system components) and *Human* (e.g. malicious act). Partial view of Natural hazards taxonomy is given as an example in Figure 11.
2. How the hazard can occur? *Hazard attributes sub-ontology*. A sub-ontology of hazard attributes is specified. Hazard attributes sub-ontology includes *Duration, Resource, Event impact area, Actor, Driver and Predictability*
3. What action are (can be) triggered by the hazard? "*Modus*" and "*Modus effect*" concepts are introduced in order to describe the actions/processes (impact mechanism) through which the Critical infrastructures can be impacted and the relevant effects.
4. When and Where the hazard can occur? *Spatial and temporal attributes sub-ontology*.

In addition to the common attribute it was decided to introduce a new feature to define how a given event, classifiable in terms of threat or hazard, occurs and can affect the CI. At this attribute has been given the name of Modus. Modus is useful because it allows simplifying the complexity of the ways in which an event can occur because, by approximation, events with different origins (natural, technological and human) may have the same modus (E.g.: both a landslide and a manifestation can create a road obstruction. In this case "Obstruction" is the one of the possible modus that characterised two hazards with different origin). Each modus, in turn, can create one or more effects on infrastructure (Modus effect). According to our description level, Modus is used as the link with the ontology of the CI Assets.

Figure 11: Partial view of Natural hazards (first and second levels)



4. VALIDATION PROCESS

In the first phase, about 25 experts have been invited to review the Critical Infrastructure ontologies on assets and functions. Experts have been provided with an evaluation template (available on request) where they have been able to propose:

- doubts on the clarity of, or different nomenclature and description for assets/functions;
- missing relevant assets and/or functions;
- not relevant assets and/or functions (candidates to be removed).

They have also reviewed topologies (the service delivery process) where they have validated connections (and their types) between assets and/or functions, indicating missing or wrong links.
Based on the comments and recommendations received by the experts, some of the assets and functions have not been used in the final topology implementation. It is either due to a high level of detail – assumed not to be relevant to describe the effects of threats to service delivery process -, or to an activity that is not being carried out on regular bases (and thus not relevant in the standard service delivery process). We still keep these assets/functions inside the catalogue in order to assure the completeness of the ontology and a comprehensive description of CI.

In the subsequent phase experts will validate integration of CIs and H&T ontologies which will be connected through different types of interdependencies. This part of the validation will be carried out through face-to-face interviews with technical managers and experts for each specific infrastructure sector.

5. FURTHER DEVELOPMENTS

On the modeling side the main task to be completed is the modeling of interdependencies within the already existing federation of ontologies.

Geographic interdependency occurs if a local environmental event creates state changes in infrastructure [7]. Geographic interdependencies will be defined as links between modus (created by H&T) and CI assets. Impact of modus on CI functions will be taken into account through possible unavailability of asset needed to execute the function.

Functional (or Physical) interdependency is a physical reliance on material flow from one infrastructure to another [7]. Within CI topologies (service delivery process) *mechanism, control* and *material/resource* inputs have been defined for each single function, as well as material flow between related functions – covering both dependencies within and between infrastructure sectors (e.g. see Figure 9).

Cyber interdependency occurs if the state of an infrastructure depends on information transfer between infrastructures (e.g. SCADA, communications, monitoring, controlling) [7]. The information is normally transmitted through the information infrastructure. Cyber interdependencies will be modelled by connecting *information* (a common resource that has been identified as an input to functions) and *telecommunication assets.*

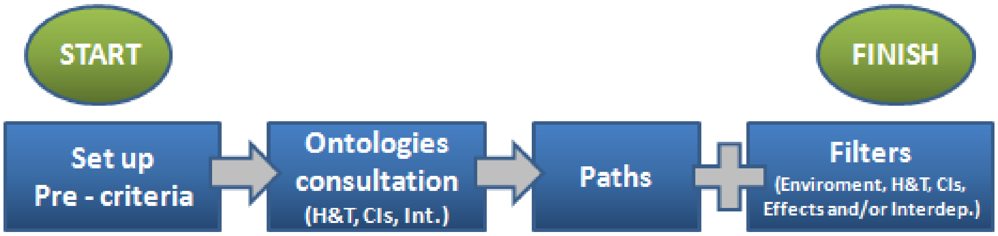
An open issue is how to model logical interdependencies. Two or more infrastructures are logically interdependent if the state of each depends upon the state of the other via some mechanism that is not a physical, cyber, or geographic connection [7]. This category can contain policy, legal or regulatory regimes; economic systems and trends; social and human factors, etc. – making it very complex and uneasy to properly cover.

On the implementation side the remaining work consists of the PATHFINDER Tool development, as the main project output. It is principally planned for two standard types of users – *Critical Infrastructures managers/operators* and *National Authorities* (*Civil Protection operators).* The general approach for identifying vulnerability paths is given in Figure 12, but may vary according to different users' needs.

It will be possible to use the tool in two main modes:

- *WIZARD mode* – the order of setting the criteria (filters) necessaries to delineate your paths is defined. In this use mode a minimal cut set of selection criteria is defined.
- *FREE mode* – the ontologies can be interrogated in free order without limitation (for expert users).

Figure 12: PATHFINDER tool utilisation process



5. CONCLUSIONS

THREVI2 project aims at assisting Authorities and Operators to get comprehensive information of all potential disruption scenarios relevant for CIP/R. It is achieved through creating a comprehensive and multi-dimensional catalogue for CI – linking and characterizing the relationships between hazards threatening CI and various infrastructure systems. It will enable evaluation of vulnerabilities and resilience capacities of the main assets and system components. In the first phase *Hazards & Threats* and *CI Systems* ontologies have been elaborated in order to capture and organize logically the available information. The main outcome of this phase consists of standardised and validated nomenclature of assets and key functions for 11 different CI systems, all hazard catalogue for CI and federation of ontologies to be implemented in a software tool for disruption scenario generation. The advantage of an ontological approach to disruption scenario generation for CI is a systematic specification and classification of concepts in the domains of interest. It embraces all-hazards approach, allowing detection of complex cascading effect mechanisms difficult to be identified and directly elicited by experts.

The main limitations of an ontology-based approach are mainly related to the limited geo specification and description of CI (partial modeling of geo interdependencies), and to poor integration of logical interdependencies (generated by organisational, economical or social related drivers). Forthcoming steps will deal with modeling of interdependencies within the already existing federation of ontologies, on one side, and with the software tool development on the other. The PATHFINDER Tool will be finalised and tested in two pilot applications developed in collaboration with at least one EC Member State and one relevant CI operator.

Acknowledgments

THREVI2 research project has been co-funded by DG Home Affairs of the European Commission, under CIPS/ISEC Work Programme. The financial support is gratefully acknowledged.

References

[1] M. Grüninger and M. S. Fox. "*Methodology for the design and evaluation of ontologies*", Technical Report, University of Toronto, 1995, Toronto (Canada).
[2] T. R. Gruber. "*A translation approach to portable ontologies*" Knowledge Acquisition, 5(2), pp. 199-220, (1993). http://www-ksl.stanford.edu/kst/what-is-an-ontology.html
[3] Protégé official website: http://protege.stanford.edu/
[4] M. Horridge, "*A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools*", Edition 1.3, The University Of Manchester, 2011, Manchester (UK).
[5] N. F. Noy and D. L. McGuinness, "*Ontology Development 101: A Guide to Creating Your First Ontology*", SMI technical report SMI-2001-0880, Stanford University, 2001, Stanford (CA).
[6] *Integration Definition for Function Modeling (IDEF∅) Standard*, Draft Federal Information Processing Standards Publication 183, 1993.
[7] S.M. Rinaldi, J.P. Peerenboom and T.K. Kelly. "*Identifying, Understanding and Analyzing Critical Infrastructure Interdependencies*" IEEE Control Systems Mag. 21, pp. 11-25, (2001).