SEMANTIC CATALOGUE TO MANAGE DATA SOURCES IN DISASTER MANAGEMENT SYSTEM

Claire PRUDHOMME
i3mainz, University of Applied Sciences, Mainz, Germany
claire.prudhomme@hs-mainz.de
Ana ROXIN
Laboratoire Le2i, UMR CNRS 6306, Dijon, France
ana-maria.roxin@u-bourgogne.fr
Christophe CRUZ
Laboratoire Le2i, UMR CNRS 6306, Dijon, France
christophe.cruz@u-bourgogne.fr
Frank BOOCHS
i3mainz, University of Applied Sciences, Mainz, Germany
frank.boochs@hs-mainz.de

Abstract. *With the climate change, disasters occur more frequently and the need for efficient disaster management systems becomes highly recommended to save lives. This paper deals with a study of existing systems, with the intention of determining the main recent improvement in the domain. The heterogeneous data integration process is a major central point. Thus, a semantic system with three main components is proposed as a new position in the disaster management systems. These three components are a knowledge base, a reasoner and a semantic catalogue. The knowledge base provides a controlled vocabulary and allows storing information retrieval. The semantic catalogue facilitates the access to data sources adapted to user's and agent's needs. The reasoner analyzes the information in the knowledge base thus replying to the user queries. In addition, the reasoner aims at adding automatically new data sources in the semantic catalogue. The fast access to a great number of data sources is of benefit for decision-making systems such as disaster management systems.*

Keywords: Disaster management, Semantic catalogue, Semantic web.
JEL classification: Q54 Climate; Natural Disasters and Their Management; Global Warming, C88 Other Computer Software.

1. Introduction

In literature, disaster management is described as a four-step cycle [1]: Mitigation, Preparedness, Response, and Recovery. The Mitigation step aims at assessing the risk of different disasters, according to their likelihood and their aftermaths. This assessment first allows increasing the resilience of a risk area in order to mitigate disaster aftermaths and thus decreasing the risk. Then, this assessment is used for the next step which is Preparedness. The preparation of public and rescue organizations is required for reducing life damages related to disasters. It needs to plan response activities and assign them organizations and responsible persons. This step also consists in training population and rescue teams, along with planning materials and equipment needs to response activities. A good preparation allows a more efficient response. During the Response phase, the aim is to minimize casualties, loss of life and property damages. Coordination and collaboration among agencies along with time are main related issues. The last step is Recovery. It requires to recover what it has been lost in order to come back to a *normal* situation. It must be used to learn from eventual errors or mistakes and rebuild with improvements. Disaster management systems' implementation and improvement strongly rely on information technologies. Several such systems have been developed. These systems are presented in the next part of related work.

Each step of disaster management requires a set of data which can be very diversified. The risk assessment of the Mitigation step is often based on information about the previous disasters. This type of information can be obtained by combining many types of data, such as, for example, geographic data of disaster areas or the number of occurrence according to the type of disaster. During Preparedness, one of tasks being the assessment of the humanitarian resource number, it requires data about the location of warehouse and potential shelters, but also about the humanitarian resource stock for each warehouse in order to increase them, if there are not enough to respond to a disaster. For the Response phase, many data types are required: geographic data about the location of rescue teams, casualties, or equipments, data of sensor in order to follow the evolution of the disaster, data about the policy of response according to the disaster, etc. In order to manage the Recovery step, other data are also used such as satellite data to evaluate damages, or financial data (donations, government funds, etc.). Consequently, our goal aims at integrating heterogeneous data in a semantic Geographic Information System (GIS) in the context of disaster management in an efficient way. The semantic has been used in several systems to permit a better flexibility and for integrating heterogeneous data. However, this purpose is not yet complete. Hence, in our opinion a good solution could be the integration of a semantic catalogue in the system. In order to understand the context of disaster management, the next section introduces the domain by presenting existing systems. Section 3 describes our approach using semantic catalogues in a disaster management system. The last section concludes the paper.

2. Related work

The Response step being the most critical, the majority of the developed systems is created in order to improve this step. Different types of systems exist with different aims. One of them is Sahana [2], an open source software supporting the deployment of Humanitarian Response management. It is composed of several systems (localization system, geographic information system, etc.), libraries and API, several modules for each functionality (shelter registry, organization registry, etc.), and some other optional modules (voluntary coordinate system, data importation module, etc.). This software is managed by a database. Other systems are not based on the database but on the semantic technologies such as UICDS [3]. This system uses an ontology and a rule-based reasoning systems to automatically find the most adapted organization in order to respond to a need or request. In Firegrid [4], Siadex [5] and IsyCri [6], semantic technologies are used to plan automatically activities which are adapted to situation needs. Workpad [7] is a system peer-to-peer, which aims at allowing the exchange of data and information between the different actors of response. A European project Cobacore [8] has the same goal as Workpad but focuses on the Recovery phase. The affected communities can use the platform in order to save their needs (medical needs as bandage for example, service needs, transportation needs, etc.). According to these needs and a set of information about the situation, the platform suggests some information which could respond to the user needs. This platform defines a plan of recovery, in order to support the decision-making of professional responding.

As a common feature, all disaster management systems require integration and retrieval processes. However, the main issue is the heterogeneity of data sources. These systems require an efficient way to integrate different types of data such as sensors. In [9], Soknos integrates sensor web services. This system contains a database called web service repository which contains the description of web services. It uses concepts which are in the user query to search in description of web services what are the adapted web services. A sensor ontology is then, used to incorporate information as assertions. Our approach focuses on the improvement of the information retrieval process to increase the diversity of integrated services. Hence, we propose the use of a semantic catalogue to link ontology concepts to data sources. The advantage of a semantic catalogue is to obtain the better and larger number of information using principles such as vector expansions in queries [10].

3. Using the semantic catalogue for disaster management

The aim of disaster management system (Fig. 1.) is to provide information according to the user queries. The reply process to a user query begins with the interpretation of this query.

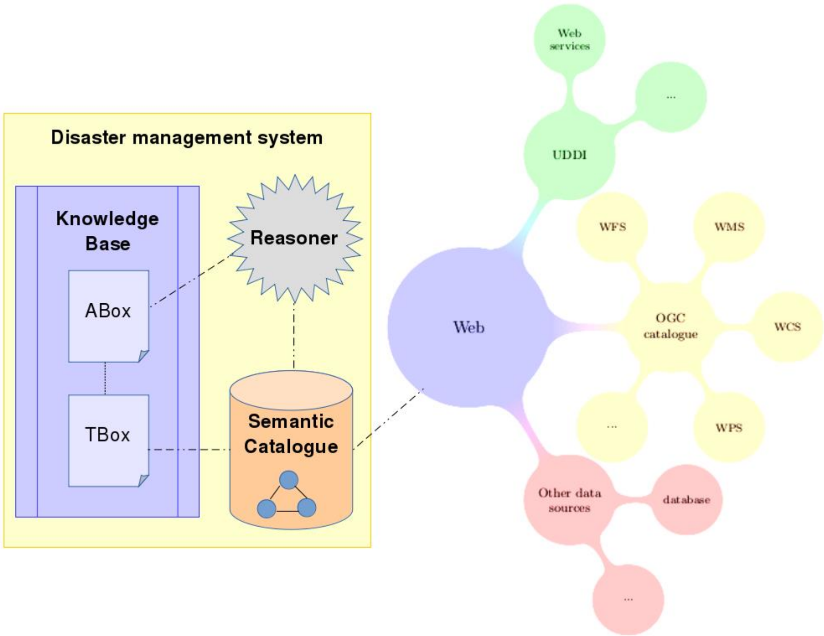

Figure 1. Structure of disaster management system using a semantic catalogue

This query is analyzed in order to identify the concepts of the query which can be linked with the controlled vocabulary in the knowledge base (Tbox). With the identified concepts and the usage of the semantic catalogue, the system retrieves data sources and services which are linked to the identified concepts. Thanks to information which are in the semantic catalogue, the system access to data of services and data sources. The information present in data is added to the knowledge base as individual in the Assertional Box (Abox) according to the controlled vocabulary of the Terminological Box (Tbox). The reasoner analyzes new information related to the user query and provides an answer. This answer can be a display of all information to support decision-making or a plan of activities or capacities which can answer to a need.

According to [11], a semantic catalogue needs a set of ontologies which can be gathered in four groups: general ontologies, specialized ontologies, spatial ontologies and temporal ontologies. In our case, specialized ontologies correspond to different domains whose information helps in the disaster management as meteorology, topography, infrastructure, etc. Spatial ontologies contain the vocabulary to manage geographic data with its diversity of format and diversity of information. Temporal ontologies allow the management of temporal data and the management of temporal information as for example the order or the duration of response activities. The general ontologies are ontologies which gather different elements of different domains. It needs to have an ontology to describe links, constraints between the different domains in order to explicit the meaning of each element in the controlled vocabulary (Tbox). The first general ontology is an ontology to describe the disaster management. This ontology needs to describe links between actors, needs, capacities and the disaster event. Another general ontology can be an ontology to describe data sources or an ontology to describe a sensor as in [12]. When all ontologies are defined, it is needed to do a mapping between the concepts in a same ontology and between the different ontologies according to their measures of similarity [11].

Thanks to the controlled vocabulary offered by Tbox, different facts, information can be saved in the knowledge base. These facts are gathered in the Abox. During the design of the knowledge base, some elements are added in the Abox. When they are added the designer has to take care of the consistency of the facts. However, facts can also be added during the reasoning thanks to the inference process. In this case, two approaches can be possible. The first approach is to save the new facts in another place than the knowledge base. This approach is the monotonic inference; it guarantees the preservation of the consistency of the knowledge base. The second approach is the non-monotonic inference. It consists to add new facts in the knowledge base. The risk is to obtain an inconsistent knowledge base. However, in [13], it is possible to find ways to manage this problem of non-monotonic inference and to benefit thus, of the advantage of a non-monotonic inference. The advantage of the non-monotonic inference is that it doesn't restrict the inference power contrary to the monotonic inference [14].

The semantic catalogue is a triple store RDF where are stored the metadata of services and other data sources [15]. The adding of metadata in the triple store is the first step of new data source integration process. The next step is the analyze of these metadata to determine concepts which will be linked to the knowledge base concepts. The result of this analyze is an ontology mapping which allows the retrieval of data sources corresponding to knowledge base concepts [11]. The reasoner is used in the two following processes: data analysis for replying and integration of new data sources.

- The data analysis corresponds to infer the information added to the knowledge base from data sources. The reasoner searches in the set of knowledge facts including new facts what are the facts linked to user query concepts. According to the choice of the user, the reasoner can return elements linked to user query concepts or apply some rules to plan activities or match capacities and needs to support the decision-making of the user.
- When a user query is formulated, the system uses an ontology mapping to retrieve data sources and their data thanks to the semantic catalogue. But the issue is the access of data sources limited to the content of semantic catalogue. In order to resolve this limit, we can create an automatic extension of the semantic catalogue by integrating new data sources. The majority of web services is described and references in OGC (Open Geospatial Consortium) catalogue or in a UDDI (Universal Description, Discovery, and Integration). Therefore, when a concept found in the user query is not linked to a data source, a process of research in OGC catalogue and UDDI begins with the concept. For the research in OGC catalogue, a calculation of semantic similarities can be applied between the concept and the content of metadata [11] to retrieve web services related to the concept. For the research in UDDI, a rule-based method in [16] could be used to retrieve related web services. At the end, these sets of data sources will be added in the semantic catalogue as explained previously.

4. Conclusion

Our researches focus on disaster management systems and how to improve them. The main point of improvement corresponds to integrate heterogeneous data based on semantic technologies. The system is composed of three main components: a knowledge base which provides a common vocabulary for managing different elements and contains heterogeneous information, a semantic catalogue, which allows the integration of different data sources, and a reasoner in order to analyze data and add new data sources in the semantic catalogue. The dynamic extension of the semantic catalogue with new data sources offers a diversity of information which allows a better situation assessment. The situation assessment is a key element in disaster management. It supports the decision-making of the user. The decision-making support can be reinforced by the rule usage in order to provide a possible response to the situation according to the policy of responders.

References

[1] D. P. Coppola, *Introduction to International Disaster Management*. Elsevier, 2011.
[2] M. Careem, C. De Silva, R. De Silva, L. Raschid, and S. Weerawarana, “Sahana: Overview of a disaster management system.” Information and Automation, ICIA International Conference on. IEEE, 2006.
[3] B. Shafiq, S. Ae Chun, V. Atluri, J. Vaidya, and G. Nabi “Resource sharing using UICDS™ framework for incident management.” Transforming Government: People, Process and Policy 6.1, 2012, pp. 41-61.
[4] L. Han, S. Potter, G. Beckett, G. Pringle, S. Welch, S. H. Koo, G. Wickler, A. Usmani, J. Torero, and A. Tate, “FireGrid: an e-infrastructure for next-generation emergency response support,” Journal of Parallel and Distributed Computing, 70, 11 , pp.1128-1141, 2010.
[5] M. de la Asunción, L. Castillo, J. Fdez-Olivares, Ó. García-Pérez, A. González, and F. Palao, “SIADEX: An interactive knowledge-based planner for decision support in forest fire fighting.” AI Communications 18.4, pp. 257-268, 2005.
[6] F. Bénaben, C. Hanachi, M. Lauras, P. Couget, and V. Chapurlat, “A metamodel and its ontology to guide crisis characterization and its collaborative management,” in Proc. The 5th International Conference on Information Systems for Crisis Response and Management (ISCRAM), Washington, DC, USA, May 2008, pp. 4-7.
[7] T. Catarci, M. de Leoni, A. Marrella, M. Mecella, M. Bortenschlager, and R. Steinmann, “The WORKPAD project experience: Improving the disaster response through process management and geo collaboration,” in Proc. The 7th International Conference on Information Systems for Crisis Response and Management, ISCRAM, May 2010.
[8] K. van Dongen, J. Willem Streefkerk, K. Anema, K. Meesters, and R. Pieneman. D5.2 Evaluation report on the intermediate evaluation. WP5 Experimentation and Evaluation. September 2014.
[9] G. Babitski, S. Bergweiler, O. Grebner, D. Oberle, H. Paulheim, and F. Probst, “SoKNOS – Using Semantic Technologies in Disaster Management Software,” in Proc. The Semantic Web: Research and Applications: 8th Extended Semantic Web Conference (ESWC), Part II, Heraklion, Crete, 2011, pp. 183–197.
[10] D. Werner, and C. Cruz, “Precision Difference Management using a Common Sub-vector to Extend the Extended VSM Method,” Procedia Computer Science, vol. 18, pp. 1179-1188, 2013, ISSN 1877-0509, http://dx.doi.org/10.1016/j.procs.2013.05.284.
[11] J. Vizcarra, and C. Cruz. “Multi-Domain Retrieval of Geospatial Data Sources Implementing a Semantic Catalogue.” in Proc. The 14th International Conference on Informatics in Economy, Bucarest, Romania, April 2015.
[12] G. Babitski, S. Bergweiler, J. Hoffmann, D. Schön, C. Stasch, and A. C. Walkowski, “Ontology-based integration of sensor web services in disaster management.”
[13] K. Frankish. Non-monotonic Inference. *The Encyclopedia of Language and Linguistics, 2nd edition*. K. Brown. Oxford: Elsevier, 2005.

[14] S. Lange, and T. Zeugmann. Monotonic versus non-monotonic language learning. *Non monotonic and Inductive Logic*. Springer Berlin Heidelberg, 1991. pp.254-269.

[15] H. Arenas, B. Harbelot, and C. Cruz. "Implementing a Semantic Catalogue of Geospatial Data." in Proc. The Tenth International Conference on Web Information Systems and Technologies, Barcelona, Spain, April 2014.

[16] K. Tamilarasi and M. Ramakrishnan, "Semantic cluster based Search in UDDI for Health Care Domain," Indian Journal of Science and Technology, [S.l.], ISSN 0974 -5645, March 2016. Available at: <http://www.indjst.org/index.>.