World Applied Sciences Journal 10 (8): 980-985, 2010
ISSN 1818-4952


# A Conceptual Architecture of SAHARA - A Semantic Disaster Management System

[1]*Qazi Mudassar Ilyas and* [2]*Ijaz Ahmad*

[1]Department of Information Systems, King Faisal Univesity, Al-Ahsa, Kingdom of Saudi Arabia
[2]Department of Computer Science, COMSATS Institute of Information Technology, Abbottabad, Pakistan

**Abstract:** Disasters, whether natural or human inflicted, have always affected mankind severely. Information and Communication Technologies (ICT) solutions are best suited to solve problems involving collection, analysis, dissemination and integration of data. Conventional Disaster Management Systems (DMS), however, depend upon significant human intervention for collaboration among individuals and organizations. One of the most important aspects in a DMS is timeliness of data. The utility of conventional DMSs is limited because they do not process real time data. In this paper we propose Semantic Disaster Management System. The basic idea is to make use of the great amount of data shared by people during disasters. This data is shared on Faecbook, Twitter, discussion forums, news sites, blogs, Wikiwiki pages and RSS feeds etc. The proposed system uses backend knowledge in the form of ontologies to collect data from these resources automatically. This data, instance meta data and the ontologies form the knowledge base. This knowledge base can later be used for performing reasoning. This reasoning process results in generation of implicit knowledge. This combination of explicit and implicit knowledge helps for decision making and information dissemination for better decision making that results in more effective disaster management.

**Key words:** Disaster Management · Ontology · Knowledge Management · Information Extraction

## INTRODUCTION

Disasters are part and parcel of life. Earthquakes, landslides, tsunamis, cyclones, tornados, hurricanes, heavy snow and rainfalls, floods, volcano eruptions and draughts are the major natural disasters. These natural disasters are further augmented with human-inflicted disasters including war-fares, terrorist attacks, bombings, political and social strikes and agitations, radiological accidents, large-scale vehicular collisions and refugee emigration etc. All disasters require mitigation measures. There are three main tasks at the heart of a typical disaster management system, namely data gathering and analysis, interconnectivity and integration. Traditional ICT technologies can contribute in all these three tasks to some extent and facilitate collaboration among the individual volunteers, national agencies and Non Governmental Organizations. Semantic Web based ICT solutions provide next level of sophistication for data collection, analysis, decision making, integration and dissemination. It was extremely difficult to incorporate description and reasoning capabilities into Web applications a decade ago but now because of the standardization and maturity of languages, tools and frameworks (OWL [1], RACER [2], GATE [3], Jena [4]) and awareness of the research community, these applications are becoming a reality. These capabilities can be exploited for automatic data gathering and analysis and decision making. Current Disaster Management Systems (DMS) are dependant upon manual data entry, analysis and decision making. Our idea is to use Semantic Web technologies to make this whole process automatic. This automated data collection is extremely helpful in case of a disaster as sometimes it is not possible to collect data from disaster struck area. Furthermore, data collection is not the only objective of using Semantic Web technologies, but it will also prove helpful in reasoning and decision making. We will also make use of existing state-of-the-art and well-tested DMSs by providing integration facilities with them.

Rest of the paper is organized as follows. Section 2 discusses the related projects and identifies their weaknesses and limitations. Section 3 gives an overview of the proposed system. Section 4 elaborates the related tools and techniques required for the implementation of the proposed framework. Section 5 highlights potential challenges and problems in the research work. The paper is concluded in section 6.

**Corresponding Author:** Qazi Mudassar Ilyas, Department of Information Systems, King Faisal Univesity, Al-Ahsa, Kingdom of Saudi Arabia, Tel. +(966)552019591, E-mail: qilyas@kfu.edu.sa.

**Related Work:** Realizing the need for disaster management, many organizations and nations have taken initiatives to mitigate the effects of disasters. There is, however, a dearth of tangible solutions as most initiatives in this regard are limited to providing strategies for planning, preparedness and rescue response team development ([5-7]). The few available tangible DMSs along with their strengths and limitations are discussed below.

**Existing Disaster Management Systems:** To the best knowledge of authors, the following are the only two Disaster Management Systems that have actually been successfully deployed for disaster management.

**Sahana:** Sahana [8] was developed by Centre of National Operations (CNO), Sri Lanka when a tsunami hit the country in the year 2004. Due to its utility and global acceptance, it was later funded by the Swedish International Development Agency (SIDA). Afterwards, it was successfully deployed and used in a number of disasters throughout the world including Sri Lanka, Pakistan, Philippines, Indonesia, New York City, Peru and China. It has won significant number of awards. The key strength of Sahana is its well structured modules, namely – Missing Person Registry, Organization Registry, Shelter Registry, Request/Aid Management System, Volunteer Coordination System and Situation Awareness. Other strengths include its free and open source format, easy customization and localization and its interoperability with other applications and datasets. The key weakness is that it requires manual data entry that is a laborious task and is not always possible in case of an emergency. Another weakness is its inability to generate dynamic inferences due to lack of formalism.

**Disaster Management Information System (DMIS):** DMIS [9] was an initiative taken by Civil Society for Research and Initiatives for Sustainable Technologies and Institutions (SRISTI), India in the year 2002. The motivation behind the initiative was earthquake in Gujrat in the year 2001. The major strength of DMIS is that it provides coordination among volunteers and strategies to deal with disasters. Moreover, it also provides literature regarding disaster management in the form of pamphlets, booklets and brochures etc. It is interactive in nature where affectees can directly interact with the system. It also creates awareness among its users with expert advice on structural engineering, rehabilitation planning, trauma and rescue activities. The key weakness of DMIS is that is just provides a platform for coordination and collaboration but does not provide any facility to store, retrieve, or process real-time data to help the field workers and decision makers in disaster management.

**Disaster Ontologies:** Some efforts have been made by the Semantic Web community to address the issue of disaster management. Their efforts are, however, only limited to the engineering of disaster ontologies. The notable efforts are discussed below:

In 2003 Matheus *et al.* [10] proposed an ontology for situation awareness. This ontology defines the core concepts necessary to describe the situation that leads to the domain specific ontologies. The authors also discuss the different design alternatives to represent situation awareness. A battlefield scenario is used to illustrate the concept.

In 2006 Wen-Yu Liu and Kwoting Fang [11] developed an ontology for debris flow by using the approach of Petri nets. They divided the task of ontology construction in three layers namely – lexical, conceptual and symbol layer. Lexical layer comprises of an analysis of domain and gathering relevant data from various resources. The conceptual layer identifies the possible relations among the collected entities. The symbol layer transforms the collected concepts and relations into some formal model such as structured graph.

Joshi *et al.* [12] in 2007 proposed an ontology for mitigation and disaster planning. The ontology considers infrastructure (IS) as an important concept of mitigation plan. The authors also propose high level disaster mitigation and modeling system. In 2005 Hoogendoorn *et al.* [13] devised a strategy to formally represent and compare different disaster plans by using organizational structure.

The strength of the above mentioned work is approaches proposed for ontology construction but there is no comprehensive ontology that covers all aspects of the disaster management domain. A critical comparison between this research and existing disaster management systems reveals that the granularity of the ontologies constructed is too course to make their integration with the existing DMSs possible. This existing artwork has, however, helped us in understanding the domain of disaster management, defining system requirements, identifying potential concepts and relations in the domain.

**Proposed Solution:** After thoroughly studying the architecture and services provided by the existing DMSs and disaster ontologies, we have proposed SAHARA – a Semantic Web based DMS - that elegantly solves the discussed problems by automatic data collection and deriving inferences. The data collection and retrieval process of SAHARA is automated and intelligent. The system would have the capability to integrate with traditional DMSs and thus exploit their capabilities.

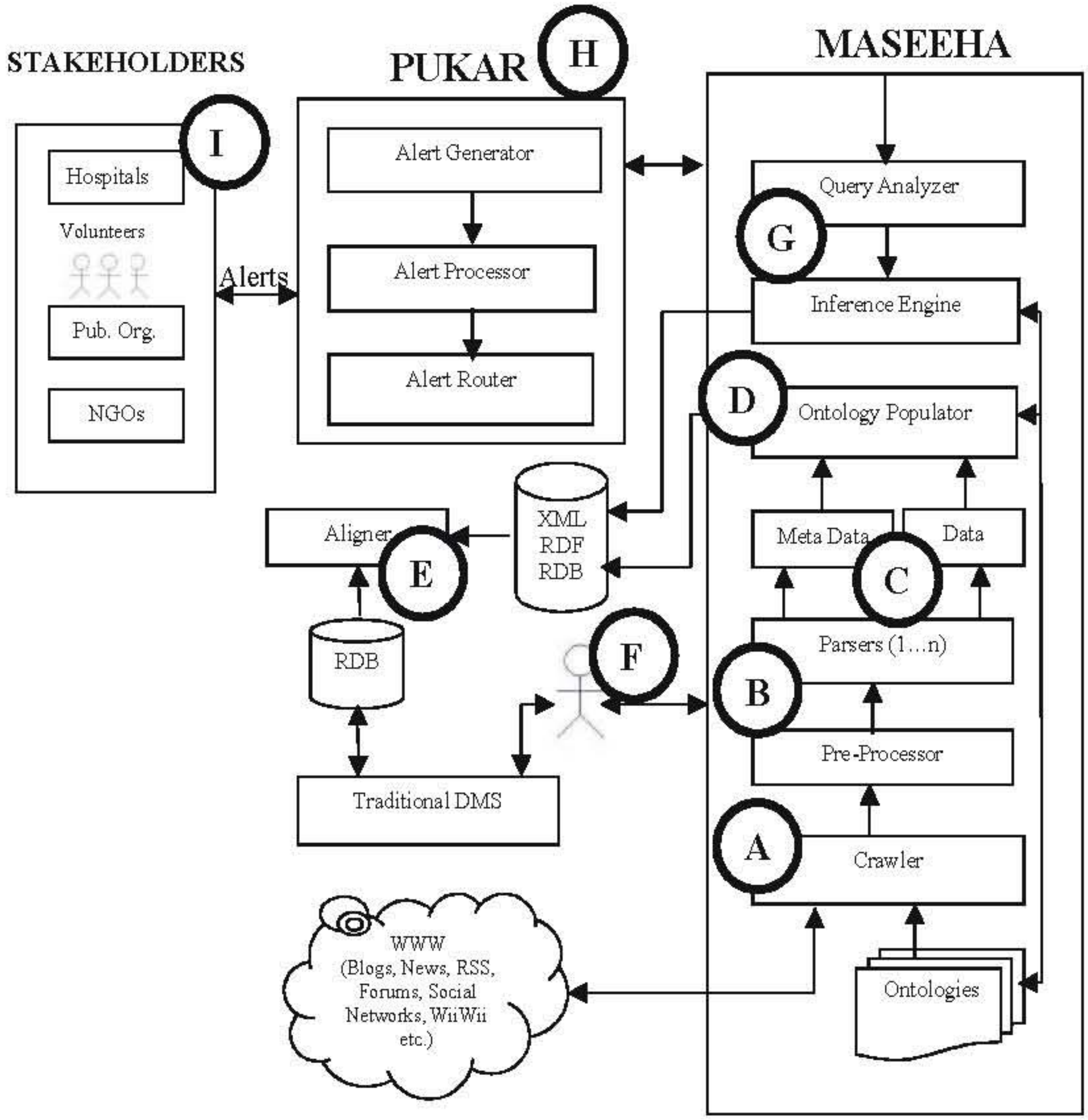


Fig. 1: A Conceptual Architecture of SAHARA

**The Conceptual Framework of SAHARA:** Figure 1 gives a conceptual architecture of the proposed system with all essential components and the data flow between them. The system works as follows:

**A:** The crawler crawls the Web to find the required data by using ontology as background knowledge. As a conventional crawler crawls the World Wide Web to find relevant resources, it looks merely for keywords. As it does not have contextual information, it is easily confused by synonyms and other complex grammatical structures. However, when the ontology is used, this backend knowledge guides the crawler in finding more related resources and ignoring those less relevant.

**B:** The crawled documents are sent to the pre-processor that classifies the documents into Blogs, News, RSS feeds and Wikiwiki pages etc. Every document is then sent to the relevant parser for further processing. Pre-processing does the dirty work of checking format, ignoring multimedia and other unrelated content, extracting data from frames and separating text and links in typical web documents.

**C:** Every parser parses a typical type of document sent to it by the pre-processor. Parsers are also required to separate data and metadata (if any). This results in identifying key entities in the processed document. Any repeated mentions of an entity are removed. New entities are created and mentions to old entities are established.

Finally the clean and processed set of entities to ontology populator for dumping them into the database.

**D:** Ontology populator uses this data, metadata and rules from the ontology to create instances and stores them in a database. These instances will then be used to build a meaningful knowledge base. This knowledge base can be used to perform reasoning that would help in decision making process in disaster management.

**E:** The proposed system also supports any conventional DMS to be used along with it. The "Aligner" converts the existing data of these systems into compatible data format and dumps it in the central database which is already populated with the instances from the crawled and processed resources.

**F:** The user interacts with the system through the query.

**G:** The query is passed to Query Analyzer that sends it to Inference Engine. Inference Engine answers the query by performing reasoning using instance data and ontology descriptions. This reasoning process can help to infer a huge amount of implicit knowledge from the knowledge explicitly expressed in the form of ontology and instances.

**H:** As the actual disaster management and mitigation tasks are actually carried by the rescue teams, government organization, NGOs and volunteers, we need to generate alerts to keep all these people aware of on-site situation, the exact requirements, communication infrastructure and any warning etc. "Pukar" is responsible for generating all kinds of alerts and sending them to the concerned stakeholders.

**I:** Finally these alerts are sent to the concerned stakeholders for decision making resulting in effective mitigation and management of disaster.

**Tools and Techniques:** At the heart of Semanic Web are two XML based technologies declared recommendations by World Wide Web Consortium (W3C) namely - Resource Description Framework (RDF) [14] and Web Ontology Language (OWL). RDF is used to describe the resources with the help of taxonomical structure and relations. Web Ontology Language (OWL) is the next layer of formalism that helps in creating more logically refined descriptions of the concepts. The ontologies in our proposed system would be created in OWL and the instances are generated in RDF. Protégé provides an easy to use GUI for creating ontologies. General Architecture for Text Engineering (GATE) is a widely used toolkit for text mining that would help SAHARA in parsing and classifying the documents. RACER and FaCT++ [15] are stable and efficient reasoners that provide reasoning support and may be used as inference engine. Jena is an open source framework that includes APIs for RDF and OWL storage and retrieval, processing and querying. The whole application of SAHARA will be built using Jena API.

Common Alerting Protocol (CAP) [16] can be used to generate alerts inferred by the system that can be propagated to the stakeholders through any means such as RSS feeds, mobile messages, or a website. CAP can also be used effectively in Emergency Data Exchange Language (EDXL) declared a standard by OASIS in June 2006. EDXL has three major components namely - Distribution Element (EXDL-DE) [17], Hospital Availability Exchange (EXDL-HAVE) [18] and Resource Message (EXDL-RM) [19]. EXDL-DE is used as a container to send CAP messages including the information about emergency type, geographical location, damage caused and information about sender and receiver of the message. As hospitals play a vital role in case of an emergency, hence OASIS has developed EXDL-HAVE for providing information about the availability of hospitals resources such as types of services available, number of available beds, medicine, blood, doctors and support staff etc. EXDL-RM is used to communicate with the hospitals, volunteers, public and Non Governmental Organizations for the resources request during an emergency. As time is always very critical in an emergency, hence EXDL-RM provides solace to affectees in acquiring the required resources such as blood, ambulance and rescue teams etc.

**Challenges, Issues and Possible Solutions:** There are numerous potential research problems and challenges that need to be explored. The first and foremost problem is the selection of sources for data gathering during a disaster because of conflicting, inaccurate or different information on various resources. On a whole a crawler would manage all this activity. The crawler would appreciate well known reliable sources of information e.g., Web sites of famous local and international media, Web sites of public and Non Governmental Organizations etc. and depreciate otherwise. A seed page and an ontology would be provided to the crawler and it would find the related Web pages, rank them and store for easy retrieval and future reference.

The second most important problem is the conversion of unstructured data into structured data. Natural Language Processing (NLP) tools such as GATE can be used to extract entities and relationships from the source and the extracted information would be stored in RDF format. Such general purpose NLP extractors would extract only general entities but we need to add another layer of precision that would give us only entities related with disaster domain. The constructed ontology of the disaster domain would educate this layer about the disaster domain. We have explored other NLP extractors such as Open Calais [20], T-REX [21] and TOPO [22] but we have a consensus that extra precision layer is necessary to manage this task. We would also exploit RSS feeds to extract information. It is relatively easier to extract information from RSS feeds as compared to unstructured formats e.g., HTML pages.

The third important issue is the representation of disaster domain i.e., ontology. A very detailed study is required to understand this domain and then map the concepts learned into ontology. Another important issue is the integration of extracted information with the existing DMS. We would need to make the data compatible with existing DMS such as conversion from relational database into XML, XML into RDF and XML into OWL.

The final challenge is the dissemination of information to the concerned stake holders by considering the limitation of media such as mobile phones.

## CONCLUSIONS

In this paper we have proposed a conceptual model of the semantic DMS and elaborated all the necessary components of the model along with their interaction and data flow. We have proposed the tools, technologies and frameworks necessary to build these components. We have also conceived the potential research issues and challenges involved and their possible solutions. Our future work would concentrate on materializing the conceptual model presented in this paper.

## REFERENCES

1. Schneider, P. and H. Horrocks, 2004. OWL Web Ontology Language Semantics and Abstract Syntax. W3C Recommendation.
2. Racer Systems Gmb, H. and K.G. Co, 2007. Racer Pro User's Guide.
3. General Architecture for Text Engineering (GATE), http://gate.ac.uk/
4. Jena, http://jena.sourceforge.net/
5. Ministry of Food and Disaster Management Bangladesh, 2006. Disaster Management Framework Disaster Management Framework of Bangladesh. In: Workshop on Regional Cooperation among BIMSTEC Countries for Disaster Risk Reduction and Management, New Delhi, India.
6. Singh, R.K., 2004. National System for Disaster Management. In: India-United States Conference on Space Science, Appllications and Commerce, Bangalore, India.
7. Venkatachary, K.V., B. Manikiam and S.K. Srivastava, 2004. Harnessing Information and Technology for Disaster Management. In: India-United States Conference on Space Science, Appllications and Commerce, Bangalore, India.
8. Careem, M., *et al.* 2006. Sahana: Overview of a Disaster Management System. In: International Conference on Information and Automation, Colombo, Sri Lanka.
9. SRISTI - Society for Research and Initiatives for Sustainable Technologies and Institutions, http://Www.Sristi.Org/Cms/
10. Matheus, C.J., M.M. Kokar and K. Baclawski, 2003. A Core Ontology for Situation Awareness. In: Proceedings of the Sixth International Conference of Information Fusion, pp: 545-552. International Society of Information Fusion, Queensland, Australia.
11. Liu, W.Y. and K. Fang, 2006. Using IDEF0/Petri net for Ontology-Based Task Knowledge Analysis: The Case of Emergency Response for Debris-Flow. In Proceedings of the 39th Hawaii International Conference on System Sciences, pp: 1: 10, Hwaii.
12. Joshi, H., *et al.* 2007. Ontology for Disaster Mitigation and Planning. In: Proceedings of the 2007 Summer Computer Simulation Conference, Society for Computer Simulation International San Diego, California, USA.
13. Hoogendoorn, *et al.* 2005. Formal Modelling and Comparing of Disaster Plans. In Proceedings of the Second International ISCRAM Conference, pp: 1: 11, Brussels, Belgium.
14. Beckett, D., 2004. RDF/XML Syntax Specification (Revised). W3C Recommendation.
15. FaCT++, http://owl.man.ac.uk/factplusplus/
16. Jones, E. and A. Botterell, 2005. Common Alerting Protocol, pp: 1: 1.

17. Raymond, M., S. Webb and P.I. Aymond, 2006. Emergency Data Exchange Language (EDXL) Distribution Element, pp: 1.0.
18. Dwarkanath, S., 2008. Emergency Data Exchange Language (EDXL) Hospital AVailability Exchange (HAVE) Version 1.0.
19. Aymond, P., *et al.* 2008. Emergency Data Exchange Language Resource Messaging (EDXL-RM) 1.0.
20. OpenCalais, http://www.opencalais.com/
21. Albanese, M. and V.S. Subrahmanian, 2007. T-REX: A System for Automated Cultural Information Extraction. In Proceedings of International Conference on Computational Cultural Dynamics, USA.
22. Valero1, A.T., M.M. Gómez and L.V. Pineda, 2006. A Machine Learning Approach to Information Extraction. In Computational Linguistics and Intelligent Text Processing, Volume 3406, pp: 539: 547, Springer, Heidelberg.