A SEMANTIC WEB-BASED ONTOLOGY FOR DISASTER TRAIL MANAGEMENT IN PAKISTAN

BY

ASHFAQ AHMAD

A thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy in Information Technology

Kulliyyah of Information and Communication Technology
International Islamic University Malaysia

JANUARY 2021

ABSTRACT

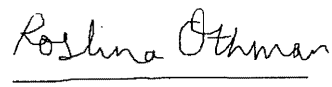
Disasters whether natural or human-made, leave a lasting impact on human lives and require mitigation measures. In the past, millions of human beings lost their lives and properties in disasters. Information and Communication Technology provides many solutions. The issue of so far developed DMSs is their inefficiency in semantics that causes failure in producing dynamic inferences. Here comes the role of semantic web technology that helps to retrieve useful information. Semantic web-based intelligent and self-administered framework utilizes XML, RDF, and ontologies for a semantic presentation of data. The ontology establishes fundamental rules for data searching from the unstructured world, i.e., the World Wide Web. Afterward, these rules are utilized for data extraction and reasoning purposes. Many disaster-related ontologies have been studied; however, none conceptualizes the domain comprehensively. Some of the domain ontologies intend for the precise end goal like the disaster plans. Others have been developed for the emergency operation center or the recognition and characterization of the objects in a calamity scene. A few ontologies depend on upper ontologies that are excessively abstract and are exceptionally difficult to grasp by the individuals who are not conversant with theories of the upper ontologies. The present developed semantic web-based disaster trail management ontology almost covers all vital facets of disasters like disaster type, disaster location, disaster time, misfortunes including the causalities and the infrastructure loss, services, service providers, relief items, and so forth. The objectives of this research were to identify the requirements of a disaster ontology, to construct the ontology, and to evaluate the ontology developed for Disaster Trail Management. The ontology editor applied by this research is Protégé version 5.2.0, which utilizes OWL as ontology language. The ontology consists of 6969 axioms, 84 concepts, 103 properties, and 726 individuals. The ontology was assessed efficaciously via competency questions; externally by the domain experts and internally with the help of SPARQL queries. The ontology was assessed by a software tool and found 100% accurate concerning its structure and overall 97% perfect as evaluated by the domain experts.

خلاصة البحث

الكوارث، سواء كانت طبيعية أو من صنع الإنسان، تترك أثراً دائماً على حياة البشر وتتطلب بعض التدابير للتخفيف من أثرها. في الماضي، فقد ملايين من البشر أرواحهم وممتلكاتهم بسبب الكوارث. توفر تقنية المعلومات والاتصالات العديد من الحلول. مسألة DMSs المطورة حتى الآن هي عدم كفاءتها في الدلالات التي تسبب الفشل في إنتاج استنتاجات ديناميكية. هنا يأتي دور تكنولوجيا الويب الدلالية semantic web technology التي تساعد على استرجاع المعلومات المفيدة. شبكة الدلالة الذكية والإطار الذاتي، تستخدمان XML، RDF، وأنطولوجيا العرض دلالي للبيانات. تضع الأنطولوجيا قواعد أساسية للبحث عن البيانات من العالم غير المهيكل، أي شبكة الويب العالمية. بعد ذلك، يتم استخدام هذه القواعد لأغراض استخراج البيانات والمنطق. تمت دراسة العديد من الأنطولوجيا المتعلقة بالكوارث؛ ومع ذلك، لا شيء تصور المجال بشكل شامل. بعض الأنطولوجيا المجال تهدف إلى الهدف النهائي الدقيق مثل خطط الكوارث. وقد تم تطوير البعض الآخر لمركز عمليات الطوارئ أو التعرف على الكائنات وتوصيفها في مشهد الكوارث. تعتمد بعض الأنطولوجيا على الأنطولوجيا العلوية المجردة بشكل مفرط والتي يصعب فهمها من قبل الأفراد الذين لا يعرفون نظريات الأنطولوجيا العليا. يغطي علم الوجود الحالي المطوّر المستندة إلى الويب لإدارة آثار الكوارث تقريبًا جميع الجوانب الحيوية للكوارث مثل نوع الكوارث وموقع الكوارث ووقت الكوارث والمصائب بما في ذلك الخسائر وفقدان البنية التحتية والخدمات ومقدمو الخدمات ومواد الإغاثة وما إلى ذلك. يهدف هذا البحث إلى تحديد متطلبات علم الوجود في حالات الكوارث، وبناء علم الوجود، وتقييم علم الوجود الذي تم تطويره لإدارة الكوارث. محرر الأنطولوجيا المطبق من خلال هذا البحث هو Protégé version 5.2.0 والذي يستخدم OWL كلغة أنطولوجيا. تتكون الأنطولوجيا من 6969 بديهية و84 مفهومًا و103 خاصية و726 فردًا. تم تقييم الأنطولوجيا بفعالية من خلال أسئلة الكفاءة؛ خارجيا من قبل خبراء المجال وداخليا بمساعدة استعلامات SPARQL. تم تقييم الأنطولوجيا بواسطة أداة برمجية ووجدت أنها دقيقة بنسبة 100٪ فيما يتعلق بهيكلها وإجمالي 97٪ مثالية كما تم تقييمها من قبل خبراء المجال.

APPROVAL PAGE

The thesis of Ashfaq Ahmad has been approved by the following:



Roslina Othman
Supervisor

Mohamad Fauzan
Co-Supervisor

Asadullah Shah
Internal Examiner

Tengku Mohd bin Tengku Sembok
External Examiner

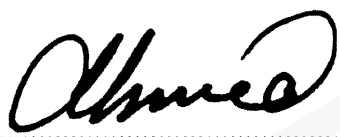
Datin Noor Habibah HjArshad
External Examiner

Ma'an Fahmi Rashid Al-Khatib
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Ashfaq Ahmad

Signature  Date 07-01-2021

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

A SEMANTIC WEB-BASED ONTOLOGY FOR DISASTER TRAIL MANAGEMENT IN PAKISTAN

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

By signing this form, I acknowledged that I have read and understand the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Ashfaq Ahmad

.....
Signature

07-01-2021
.....
Date

ACKNOWLEDGEMENTS

First and foremost, all praises and gratitude be to Almighty Allah, the Lord, the source of knowledge and guidance. His blessings made me able to complete this research work successfully.

I am incredibly grateful to my father for his true love and prayers for my future. My heartfelt thanks to my loving, caring and supportive better half whose constant encouragement and motivation always work as a driving force especially during ordeals and trying circumstances. My heart muses to pay sincere thanks to all my near and bosom ones who supported me directly or indirectly to accomplish this research with their precious time and guidance. Also, my bona fide appreciation goes to my brother in law Tahir Nazeer and my nephew Asif Javed for their untiring efforts and assistance in data acquisition.

My humble thanks ought to be due for those who provided their time, support and effort to accomplish this project. To the high esteemed committee members who translated my effort into reality.

Finally, my deep and sincere gratitude goes to my supervisor Professor Dr Roslina Othman for her leadership, motivation, enthusiasm, knowledge sharing and invaluable guidance throughout this research. Her advice helped me in the time span of research and writing of this thesis. I would extend my special thanks to my co-supervisor Professor Dr Mohamad Fauzan for his support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 RESEARCH BACKGROUND

A large number of human beings are being affected by disasters every year. In the year 2016, 569.4 million people were affected by disasters (Guha-Sapir, Hoyois, Wallemacq & Below, 2016) and 110.3 million in 2015 (Guha-Sapir, Hoyois & Below, 2016). The collected data shows that from 2005 to 2014, on an average, 196.3 million people were affected per year (Guha-Sapir, Hoyois & Below, 2016). In 2016, 342 disasters were reported, causing 8733 human causalities along with financial damage worth US$154 billion (Guha-Sapir, Hoyois, Wallemacq & Below, 2016). Pakistan was on the list of top 10 affected countries in Annual Disaster Statistical Review 2015 and 2016 prepared by Centre for Research on the Epidemiology of Disasters (CRED) (Guha-Sapir, Hoyois & Below, 2016 and Guha-Sapir, Hoyois, Wallemacq & Below, 2016).

In disaster-affected areas, the survivors suffer a lot due to an interruption in the essential services like health care, communication, transportation, and so forth. Infrastructural damages can also affect food and water supply. Although the man has made considerable progress in the field of science, engineering, and technology, yet he is unable to control the occurrence of disasters. All his efforts, so far, aim at managing hazards, mitigation and to reduce the impact of disasters. Due to the devastating effects of disasters on human lives, catastrophes and crises management have always been given vital importance.

Disaster management is planning, arrangement, and deployment of resources with a precise aim of reducing disaster's damaging effects. Socio-economic conditions of the affected area and existence of effective information system regarding the

occurrence of emergency are the significant factors that influence this management. Timely information plays a vital role in reducing disaster impact up to a certain level. The arrangements and organization of resources and efforts for mitigation majorly depend upon disastrous areas' situation and effects of the disaster on the local population. Efforts are made for gathering, organizing and disseminating factual information to various stakeholders taking part in the mitigation process. Efficiency in the deployment of resources is one of the significant concerns in disaster management as it can minimize the disastrous aftereffects to a great extent.

1.1.1 Overview

The word 'disaster' itself shows that it is something troublesome that needs to be avoided or requires mitigation to bring down its outcomes if it ever happens again. Disaster mitigation focuses on long-term measures for diminishing risks. These measures can be structural or non-structural. Developing technological solutions and training of key personnel are examples of structural measures whereas legislation and communicating potential threats to the public are considered as non-structural measures. Disaster mitigation or management process can be divided into three major phases:

- The data collection and analysis: Data is collected through observation techniques of the data collection and visualization, knowledge modeling, event forecasting, and information management after critical analysis.
- Data communication: Data communication or interconnectivity involves the mode of communication for information sharing among stakeholders.

- Data integration: Data integration phase involves combining the data from several disparate sources into meaningful and valuable information and providing a unified view to users.

Traditional Information and Communication Technology (ICT) can contribute significantly to all the three phases discussed above. For collecting and compiling data with a view to its dispersal and assimilation, semantic web-based ICT solutions provide all these levels in a plausible way.

"The word 'disaster' comes from ancient Greek words *dis* means 'bad' and *aster* means 'star'. The astrological sense of disaster based on calamity blamed on star positions." (ewonago.wordpress.com)

Disasters are divided majorly into two types; natural and human-made. These could further be categorized into slight and extraneous emergencies. The major catastrophe result in great devastation. The gigantic natural catastrophic paradigms are cyclones, drought, earthquake, flood, tornadoes, and tsunami; however, life-threatening human-made tragedies are epidemic, deforestation, chemical pollution, wars and holocaust in the shape of terrorism (Raza & Kandhro, 2015; Van de Walle & Comes, 2015). Minor natural disasters include thunderstorms and cold/hot waves whereas traffic accidents, rampage, botulism, and factory disasters are known as minor human-made disasters (Raza & Kandhro, 2015; Van de Walle & Comes, 2015).

Ontology is getting importance for providing clear and definite search by focusing the concepts in documents collection and data sources. Ontologies are designed to help improve the communication whether it is between human and machine or is between computers. The use of ontologies can support in obtaining accurate and well-organized information in a better way (Iqbal, 2012; Suganya, G., & Porkodi, R., 2018). In other words, ontology helps in managing knowledge.

An ontology primarily comprises of concepts (classes), properties (attributes) and possible relationships (slots) among concepts. There may exist some constraints (facets) on slots, or cardinalities on relationships among concepts. Collectively, the components and instances (individuals) form the knowledge base that helps in reasoning.

There are various types of ontologies which have been defined or discussed by multiple researchers including:

- Upper ontology: very general concepts familiar to numerous domains for supporting the development of an ontology
- Domain ontology: domain-specific concepts
- Interface ontology: concepts relevant to the juncture of two disciplines
- Process ontology: knowledge domain of processes

Domain ontology is one of the classifications of ontology. Ontological design can better accumulate the knowledge of diverse nature as the disaster domain does. Hence, there is a need for a specific domain ontology to conceptualize the disaster trail management (DTM).

The primary purpose of designing ontologies is to formulate sketches for disaster plans, operational centers set up for emergency may be aided with task ontology like study and planning of objects in disastrous scenes. Upper ontologies are enigmatic and abstract hence hamper understanding for those who are alien to its application. The developed DTM ontology covers nearly all sovereign states like nature of the hazard, occurrence date, damages like mislaying including the loss in infrastructure, refugee camps and facilities available or required in the refugee camp, rehabilitation tasks associated with the contributor, location, and relief index, and so forth.

1.1.2 Issues of Current Web and Need of Semantic Web

ICT has developed new concepts of information sharing. These concepts are being utilized in different fields of life. ICT is playing an essential role in disaster management. The internet has become the most common and affordable medium of fastest information sharing. The evolution of the World Wide Web (WWW) has introduced new ways of communication among people. WWW refers to a collection of documents and other web resources interlinked by hypertext links. It is a massive information space with an almost unlimited amount of data which is being flooded with an enormous amount of data every moment. Millions of websites share these data with billions of internet users. It appeals to the research community to think about how they can use this source of data for management. The limitation of WWW space is that it is a collection of documents without a mechanism of interpretation of information.

WWW is in the form of data source, not a knowledge base. It means that the available data is not in a machine-readable format. Hence, it is not understandable by the machines. An example of people speaking different languages who want to communicate among themselves can elaborate on this issue. The people will not have the capacity to explain themselves adequately through their own spoken words. They need an interpreter to make good communication. With the increase of people with different languages, the problem gets more demanding because they will need more translators!

Some effective mechanism for interpretation is required to make this data more useful. In other words, the information should be uploaded in such a well-defined format that can make it easy to formalize to be machine readable. Therefore, a solution that can be proposed for this problem is the use of the semantic web. The semantic web is a

structure of tagged and interlinked information that makes the information easily understandable by a machine.

The erection of semantic web is dependent upon universal resource identifier (URI) descriptions used for the data representation. The data representation normally enshrines in the resource description framework (RDF). The foundation upon which the edifice of semantic web stands is ontology (Gruber, 1993). The application of ontologies in the diversity of fields for systemizing data in a well-organized form has a vibrant role to play (Davies, Studer & Warren, 2006).

Disaster management in any country is the responsibility of the state government. After the occurrence of a disaster, people need immediate help. In severe disasters, non-government organizations (NGOs) and volunteers also come forward to work with government agencies in helping the victims. The rescuers need accurate information; where and what help is required. What is the urgency of relief and how to reach the place; because sometimes due to infrastructural damages, regular routes are not available to enter the site.

Social media is one of the fastest sources of current information. People share much information through social media websites. Other than entertainment, social media has become a source of breaking news or even used to create a pressure group against a specific social issue. Sometimes, even the news channels broadcast news based on trends and news from social media websites (Westlund, 2103; Chen & Sakamoto, 2014; McClendon & Robinson, 2013). Twitter is one of the favorite and widely used social media channels. Twitter news sometimes become channel headlines, which shows that Twitter as a vital source of information that can be used for collecting disaster-related information.