ONTOLOGY-BASED APPROACH FOR RETRIEVING KNOWLEDGE IN AL-QURAN

QUSAY ABDULLAH ABED

MASTER OF SCIENCE (INFORMATION TECHNOLOGY)

SCHOOL OF COMPUTING
COLLEGE OF ARTS AND SCIENCES
UNIVERSITI UTARA MALAYSIA

2015

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Capaian maklumat maklumat bergantung kepada data berkaitan yang diperolehi dari satu set sumber pengetahuan, seperti Al-Quran. Pencarian boleh berdasarkan metadata, pengindeksan teks penuh, atau lain-lain yang berasaskan kandungan. Al-Quran adalah kitab yang paling banyak dibaca di dunia dan mengautomasikan kaedah pencarian pengetahuan dari kesusasteraan agama telah mendorong minat para penyelidik dan ianya amat mencabar. Ini telah membawa kepada pembangunan beberapa aplikasi carian, yang boleh membuat carian pengetahuan berdasarkan kata kunci atau ayat. Mendapatkan pengetahuan daripada ontologi Al-Quran berdepan dengan beberapa masalah asas, iaitu satu daripadanya adalah ketepatan. Dalam kebanyakan kes, pencarian yang tidak boleh mengambil konsep yang berkaitan sesuai dengan ayat-ayat yang berkaitan. Pendekatan semasa menggunakan kaedah konvensional seperti taksonomi, haraki, atau struktur pokok yang hanya menyediakan takrif konsep tema tanpa dikaitkan dengan konsep ilmu yang betul daripada Al-Quran. Tujuan utama kajian ini adalah untuk membangunkan ontologi Al-Quran berdasarkan klasifikasi tematik. Pendekatan kajian yang baru terdiri dari dua peringkat. Peringkat pertama: melibatkan pembangunan ontologi Al-Quran berdasarkan bahasa RDF/OWL melalui alatan Protégé-OWL. Peringkat kedua: melibatkan pembangunan kaedah carian dengan menggunakan rangka kerja JENA yang berasaskan bahasa pengaturcaraan Java. Kaedah carian membolehkan pemprosesan ontologi, dan dilakukan pencarian menggunakan kata kunci yang diberikan dan mendapatkan pengetahuan yang berkaitan dengan kata kunci. Pendekatan carian adalah, dinilai menggunakan ukuran Recall dan Precision yang menunjukkan ketepatan yang tinggi dalam carian pengetahuan ontologi Al-Quran. Tambahan pula, klasifikasi ontologi telah dinilai oleh dua orang pakar dalam bidang pengajian Islam. Kajian ini menyumbang kepada kemudahan pembelajaran dan kefahaman Al-Quran kepada semua orang di semua peringkat umur.

Kata Kunci: Ontologi, Capaian Maklumat, Carian Semantik, Pengurusan Pengetahuan, Ilmu Al-Quran.

Abstract

Information retrieval relies on obtaining relevant data from a set of knowledge resources, such as Al-Quran. Searching can be based on metadata, indexing, or other content-based. Al-Quran is the most widely read book in the world and automating knowledge retrieval from this of religious literature is very challenging. This has led to the development of a number of search applications, which can retrieve knowledge based on keywords. Retrieving the knowledge of Al-Quran ontology includes several fundamental problems, one of which is the lack of accuracy. In most cases, the searching cannot retrieve the relevant concept of knowledge and verses. Current approaches use conventional methods such as taxonomy, hierarchy, or tree structure, which only provide the definition of the concept of themes without linking to the correct knowledge concept of Al-Quran. The main aim of this study is to design a method that uses the ontology approach to search and retrieve relevant verses in Al-Quran. The new approach consists of two stages. The first stage: involves the Al-Quran ontology development based on thematic classification which was implemented using Protégé-OWL. The second stage: involves the development of a search method by using the Jena framework which is based on Java programming languages. The search method allows ontology processing, and performed the searching using the given keywords and retrieve the knowledge pertaining to the keyword. The search approach was evaluated using the Recall and Precision measurements, which shows a high accuracy in retrieving the knowledge of Al-Quran. Furthermore, the ontology classification was evaluated by two experts in Islamic Studies field. This study contributes to the ease of learning and understanding Al-Quran by people of all ages.

Keywords: Ontology, Information Retrieval, Semantic Search, Knowledge Management, Al-Quran Knowledge.

Acknowledgement

First and foremost, I would like to thank Allah (S.W.T), for having made everything possible by giving me strength, confidence, and courage to accomplish this project.

I wish to express my sincere gratitude to my supervisor, Dr. Azman Ta'a, for his guidance and direction in this work. He provided me with many interesting, valuable, and sincere feedbacks, throughout his supervision. I greatly benefited from his detailed comments and insights, which helped me clarify the ideas presented in "Ontology-based Approach for Retrieving Knowledge in Al-Quran". I am also grateful to my second supervisor, Assoc. Prof. Dr. Mohd. Syazwan Abdullah. I am thankful to them for their support and motivation.

I wish to express my sincere gratitude and thankfulness to the Al-Quran experts at the Islamic Center, University Utara Malaysia (UUM), Assoc. Prof. Dr. Ismail Ishak, Director of the Islamic Center, and Ustaz Mohd. Yusnaidi Sabri, to help me and giving me of their precious time to evaluate some aspects of the dissertation.

I sincerely thank to my evaluator Dr. Yuhanis Yusof and Dr. Hazaruddin Harun, As well as, Dr. Norliza Katuk for spending time reviewing this research and giving valuable suggestions and comments on my work. I would also like to say a big thanks all University Utara Malaysia (UMM) lecturers and staff at the School of Computing who were kind enough to give me their precious time and assistance, without which I would not have been able to complete this Masters Dissertation. I will always remember it as one of the best phases of my life. I am indebted and thankful to the Chancellor of University Utara Malaysia (UMM), who referred me to valuable e-resources at the Sultanah Bahiyah Library.

I wish to thank the Ministry of Higher Education and Scientific Research of Iraq for the financial support awarded to me, and all the staff and employees of the Technical Institute of Karbala, Iraq.

Last, but not least, words cannot express my gratitude to my family, especially my beloved father, my loving mother, and my brothers and sisters, who always gave me the encouragement to acquire the education which I wanted. Special thanks to my faithful wife, and to my two children, Siraj and Ali. Words cannot describe the importance of their constant love, care, concern, patience, and direction, in every aspect of my life during my the two years of my study abroad. I will be forever thankful, grateful, and indebted to them.

Thank you UUM.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

API	Application Programming Interface
DDL	Data Definition Language
GUI	Graphical User Interface
HTML	Hyper Text Markup Language
IDE	Integrated Development Environment
IS	Information System
JSP	Java Server Pages
OWL	Ontology Web Language
RDBMS	Relational Database Management System
RDF	Resource Description Framework
SQL	Structured Query Language
UML	Unified Modeling Language
URL	Uniform Resource Locator
XML	Extensible Markup Language

CHAPTER ONE
INTRODUCTION

1.1 Introduction

This chapter provides an overview of this research. It includes a background of the study, which focuses on retrieving knowledge in Al-Quran using the ontology approach. Then, the research problem, research questions and research objectives are discussed. This is followed by the scope and significance of this research, at the end.

1.2 Background of the Study

Islamic scholars have described Al-Quran as the holy book of Muslims that teaches morals, purification, and good deeds. Al-Quran provides guidance to mankind, promotes justice between one another, and provides guidance on how to live on earth with neighbors (Ahmad et al., 2013; Yauri, Kadir, Azman, & Murad, 2012). A related study described Al-Quran as the source of information on any subject matter concerning the world and the hereafter (Shoaib, Nadeem Yasin, Hikmat, Saeed, & Khiyal, 2009). In other words, knowledge gained from Al-Quran cannot be compared with scientific books because the former provides real and deep discussions on matters under examination unlike the latter (Ahmad et al., 2013; Shoaib et al., 2009).

The study of Khan, Saqlain, Shoaib, and Sher (2013) emphasized that searches and retrieval of knowledge in Al-Quran sometimes lacks clarity and accuracy due to the non-implementation of sophisticated and dynamic ways for retrieving knowledge or

The contents of the thesis is for internal user only

REFERENCES

Ahmad, O., Hyder, I., Iqbal, R., Murad, M. A. A., Mustapha, A., Sharef, N. M., & Mansoor, M. (2013). A Survey of Searching and Information Extraction on a Classical Text Using Ontology-Based Semantics Modeling: A Case of Al-Quran. *Life Science Journal, 10*(4), 1370-1377.

Al-Saif, A., & Markert, K. (2010). *The Leeds Arabic Discourse Treebank: Annotating Discourse Connectives for Arabic*. Paper Presented at the Proceedings of Language Resources and Evaluation Conference, Valetta, Malta.

Aleksovski, Z. (2008). *Using Background Knowledge in Ontology Matching.* Vrije Universiteit, The Dutch Research School for Information and Knowledge Systems. (No. 2008-24).

Ameen, A., Khan, K. U. R., & Rani, B. P. (2012). Creation of Ontology in Education Domain. *Life Science Journal*, pp. 237-238.

Ardakan, M. A., & Mohajeri, K. (2009). Applying Design Research Method to IT Performance Management: Forming a New Solution. *Journal of Applied Sciences, 9*(7), 1227-1237.

Atwell, E., Brierley, C., Dukes, K., Sawalha, M., & Sharaf, A.-B. (2011). *An Artificial Intelligence Approach to Arabic and Islamic Content on the Internet.* Paper presented at the Proceedings of NITS 3rd National Information Technology Symposium (pp.1-8).

Atwell, E., Habash, N., Louw, B., Abu Shawar, B., McEnery, T., Zaghouani, W., & El-Haj, M. (2010). Understanding the Al-Quran: A new Grand Challenge for Computer Science and Artificial Intelligence. *ACM-BCS Visions of Computer Science 2010*, pp.1-13.

Baqai, S., Basharat, A., Khalid, H., Hassan, A., & Zafar, S. (2009). *Leveraging Semantic Web Technologies for Standardized Knowledge Modeling and Retrieval from the Holy Qur'an and Religious Texts.* Paper Presented at the Proceedings of the 7th International Conference on Frontiers of Information Technology (pp. 42).

Beck, H., & Pinto, H. S. (2002). Overview of Approach, Methodologies, Standards, and Tools for Ontologies. *Draft Paper, The Agricultural Ontology Service, UN FAO*, pp.3-8.

Beseiso, M., Ahmad, A. R., & Ismail, R. (2010). A Survey of Arabic Language Support in Semantic Web. *International Journal of Computer Applications, 2*, pp.1-6.

Bilal, M., & Khan, S. (2008). *Ontology-Driven Relevance Reasoning Architecture for Data Integration Techniques.* Paper Presented at the Intelligent Systems, 2008. IS'08. 4th International IEEE Conference (pp. 8-22).

Bryman, A., & Bell, E. (2011). *Business Research Methods*. New York: Oxford University Press.

Churchill, G. A., Brown, T. J., & Suter, T. A. (2010). *Basic Marketing Research* (7th ed.). London: ISBN.

Colton, D., & Covert, R. W. (2007). *Designing and Constructing Instruments for Social Research and Evaluation* (1st ed.). United States of America: John Wiley & Sons, Inc.

Decker, S., Erdmann, M., Fensel, D., & Studer, R. (1999). *Ontobroker: Ontology Based Access to Distributed and Semi-Structured Information. (pp. 351-369)*: Springer.

Du, Z. Q., Hu, J., Yi, H. X., & Hu, J. Z. (2007). *The Research of the Semantic Search Engine Based on the Ontology*. Paper presented at the Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007. International Conference on (pp. 5403-5406). IEEE.

Fang, W. D., Zhang, L., Wang, Y. X., & Dong, S. B. (2005). *Toward a Semantic Search Engine Based on Ontologies.* Paper Presented at the Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on (pp. 1913-1918), Guang Zhou.

Fensel, D., McGuiness, D., Schulten, E., Ng, W. K., Lim, G. P., & Yan, G. (2001). Ontologies and Electronic Commerce. *Intelligent Systems, IEEE, 16*(1), 8-14.

Geyer, M. (2005). OWL Bible Ontology. *International Journal of Information Technology, 5*(2), 199-220.

Gruber, T. R. (1993). A Translation Approach to Portable Ontology Specifications. *International Journal of Human-Computer Studies, 5*(2), 199-220.

Gruber, T. R. (1995). Toward Principles for the Design of Ontologies Used for Knowledge Sharing? *International Journal of Human-Computer Studies, 43*(5), 907-928.

Harvey, F., Kuhn, W., Pundt, H., Bishr, Y., & Riedemann, C. (1999). Semantic Interoperability: A central Issue for Sharing Geographic Information. *The Annals of Regional Science, 33*(2), 213-232.

Hislop, D. (2013). *Knowledge Management in Organizations: A critical Introduction*. Oxford City, England: Oxford University Press.

Hoffer, J. A. (2004). *Modern Database Management*: Pearson Education India.

Jones, M. V., Coviello, N., & Tang, Y. K. (2011). International Entrepreneurship Research (1989–2009): A domain Ontology and Thematic Analysis. *Journal of Business Venturing, 26*(6), 632-659.

Kaner, C. (2003). What is a Good Test Case? *Software Testing Analysis & Review Conference (STAR East). May 2003*, pp. 1-17.

Khalid, R. (2011). *Classification of Al-Qur'an's Contents (original format)* (1st ed.). Perth, Australia: Education Ministry.

Khan, H. U., Saqlain, S. M., Shoaib, M., & Sher, M. (2013). Ontology Based Semantic Search in Holy Quran. *International Journal of Future Computer and Communication, 2*(6), 570-575.

Klein, M. (2001). *Combining and Relating Ontologies: An analysis of Problems and Solutions.* Paper Presented at the IJCAI-2001 Workshop on Ontologies and Information Sharing (pp. 53-62).

Knublauch, H., Horridge, M., Musen, M. A., Rector, A. L., Stevens, R., Drummond, N., . . . Wang, H. (2005). The Protege OWL Experience. *Workshop on OWL: Experiences and Directions, Fourth International Semantic Web Conference (ISWC2005),Galway, Ireland.*

Kuechler, B., & Vaishnavi, V. (2008). On Theory Development in Design Science Research: Anatomy of a Research Project. *European Journal of Information Systems, 17*(5), 489-504.

Kuhn, T. (2009). *Controlled English for Knowledge Representation.* (Doctoral Thesis, University of Zurich), Dissertation Abstracts International , 63(12), 230.

Leiyu, S. (2008). *Health Services Research Methods* (2nd ed.). USA: Delmar Learning.

Lynn, S. C., & Ronald, R. P. (2010). *Basic Research Methods for Librarians* (5th ed.). Greenwood Publishing Group: USA.

Mika, P. (2005). Ontologies are us: A unified Model of Social Networks and Semantics *The Semantic Web–ISWC 2005* (Vol. 3729, pp. 522-536). Heidelberg: Springer.

Nguyen, N. T., & Rusin, M. (2006). *A consensus-Based Approach for Ontology Integration.* Paper Presented at the Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (pp. 514-517).

Noor. (2009). QUR'AN. Retrieved April 20th, 2014, from http://quran.com/.

Nyame-Asiamah, F., & Patel, N. (2009). Research Methods and Methodologies for Studying Organisational Learning. *Intelligent Systems, IEEE, 3*, pp. 1-15.

Qawaqneh, Z., & Kayed, A. (2007). *New Method for Ranking Arabic Web Sites Using Ontology Concepts.* Paper Presented at the Digital Information Management, 2007. ICDIM'07. 2nd International Conference on (Vol. 2, pp. 649-656).

Qurany. (2009). AL-QURAN : The Guidance for Humanity. Retrieved April 20th , 2014, from http://quran.net/

Ranjit, K. (2011). *Research Methodology: A Step-By-Step Guide for Beginners* (3rd ed.). London: Sage Publication Limited.

Saad, S., Salim, N., & Zainal, H. (2009). *Islamic Knowledge Ontology Creation.* Paper Presented at the Internet Technology and Secured Transactions, 2009. ICITST 2009. International Conference on (pp. 1-6).

Saad, S., Salim, N., Zainal, H., & Muda, Z. (2011). *A process for Building Domain Ontology: An Experience in Developing Solat Ontology.* Paper Presented at the Electrical Engineering and Informatics (ICEEI), 2011 International Conference on (pp. 1-5), Bandung, Indonesia.

Sherri, J. (2011). *Research Methods and Statistics: A critical Thinking Approach* (4th ed.). USA: Cengage Learning.

Shoaib, M., Nadeem Yasin, M., Hikmat, U., Saeed, M. I., & Khiyal, M. S. H. (2009). *Relational WordNet Model for Semantic Search in Holy Quran.* Paper Presented at the Emerging Technologies, 2009. ICET 2009. International Conference on) pp. 29-34).

Sminia, T., & Stuckenschmidt, H. (2002). Ontology-Based Information Sharing in Weakly Structured Environments. *SIKS Dissertation Series No. 2003-01*, 195.

Sridaran, R., Padmavathi, G., & Iyakutti, K. (2009). A Survey of Design Pattern Based Web Applications. *Journal of Object Technology, 8*(2), 61-70.

Staab, S., & Studer, R. (2010). *Handbook on Ontologies*. Bandung, Indonesia: Springer.

Studer, R., Benjamins, V. R., & Fensel, D. (1998). Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, *25*(1), 161-197.

Syuhada, N., & Ta'a, A. (2013). *An ontology Approach for Al-Quran Knowledge Representation and Classification.* (Unpublished Master's Project), Universiti Uatra Malaysia.

Ul Ain, Q., & Basharat, A. (2011). *Ontology Driven Information Extraction from the Holy Qur'an Related Documents.* Paper presented at the Proceeding of the 26th IEEEP Students' Seminar (pp. 17-23).

Uschold, M. (1998). Knowledge Level Modelling: Concepts and Terminology. *The Knowledge Engineering Review*, *13*(01), 5-29.

Villa, F., Athanasiadis, I. N., & Rizzoli, A. E. (2009). Modelling with Knowledge: A review of Emerging Semantic Approaches to Environmental Modelling. *Environmental Modelling & Software, 24*(5), 577-587.

Visser, U., Stuckenschmidt, H., & Schlieder, C. (2002). *Interoperability in GIS-enabling Technologies.* Paper Presented at the Proceedings of the 5th AGILE Conference on Geographic Information Science (p. 291).

Wimalasuriya, D. C., & Dou, D. (2010). Ontology-Based Information Extraction: An Introduction and a Survey of Current Approaches. *Journal of Information Science, 36*(3), 306-323.

Yang, J. (2005). Knowledge Integration and Innovation: Securing New Product Advantage in High Technology Industry. *The Journal of High Technology Management Research, 16*(1), 121-135.

Yauri, A. R., Kadir, R. A., Azman, A., & Murad, M. A. A. (2012). *Quranic-Based Concepts: Verse Relations Extraction Using Manchester OWL Syntax.* Paper Presented at the Information Retrieval & Knowledge Management (CAMP), 2012, International Conference on (pp. 317-321).

Yauri, A. R., Kadir, R. A., Azman, A., & Murad, M. A. A. (2013). Quranic Verse Extraction Base on Concepts Using OWL-DL Ontology. *Research Journal of Applied Sciences, Engineering and Technology, 6*(23), 4492-4498.

Zekr. (2009). ZKER QUR'AN. Retrieved April 20th, 2014, from http://zekr.org/quran/quran-for-windows.