

AL-QURAN THEMES CLASSIFICATION USING ONTOLOGY

Azman Ta'a¹, Syuhada Zainal Abidin², Mohd Syazwan Abdullah³, Abdul
Bashah B Mat Ali⁴, and Muhammad Ahmad⁵

¹Universiti Utara Malaysia, Malaysia, azman@uum.edu.my

²Universiti Utara Malaysia, Malaysia, s808229@students.uum.edu.my

³Universiti Utara Malaysia, Malaysia, syazwan@uum.edu.my

⁴Universiti Utara Malaysia, Malaysia, bashah@uum.edu.my

⁵Universiti Utara Malaysia, Malaysia, md.ahmad@uum.edu.my

ABSTRACT. Knowledge representations of Al-Quran involved the classification of contents for understanding the readers. In information technology environment, the representation and classification approaches must be understood by human and computer in order to ensure the correct semantics of Al-Quran knowledge. Current approaches used conventional methods such as taxonomy, hierarchy or tree structure, which only provides a concept definition without linked to another possibility of knowledge explanation. This research aims to explore the approach to represent and classify the Al-Quran knowledge by using ontology. The ontology model for Al-Quran is developed according to the themes of Al-Quran knowledge as described in *Syammil Al-Quran Miracle the Reference*. The ontology approach has shown that the Al-Quran knowledge can be classified and presented to the readers systematically. Moreover, the ontology structure that representing the relationships among the theme concepts in Al-Quran was reviewed and validated by the domain experts in Al-Quran.

Keywords: Knowledge representation, knowledge classification, Al-Quran knowledge, Ontology, Al-Quran application

INTRODUCTION

Al-Quran is the eternal miracle of Islam. It is the complete and best guide to living one's life and seeking Allah's pleasure. Al-Quran consisted of 30 divisions (Juz), 114 chapters (surah), 6236 verses (ayah), and less than 80,000 words. The verse may contain one or more sentences to form the verse. A group of verses will form a chapter with the given names (e.g., Al-Baqarah, Ali-Imran, An-Nisa). The concept of Al-Quran knowledge is defined by using the Quran hierarchy as determined by the sequence of divisions, chapters, and verses. These are the most important components for presenting the verse as final references of the particular Al-Quran concepts in a systematic manner. Understanding the Al-Quran knowledge required appropriate teachers, which is sometime difficult to be available at all times. However, computer technology can support the learners to understand Al-Quran easily, especially in the web and mobile-based environment. Mostly, the approaches to presents the Al-Quran knowledge is using tree-structure hierarchy, which makes it limited to identify the actual meaning of the concept in the Al-Quran (Shohib et al., 2010). Therefore, this research used an ontology approach to represent Al-Quran concepts that can be classified and organized according to specific themes.

RELATED WORKS

Understanding the Al-Quran knowledge is a major challenge for Computer Science and Artificial Intelligence research area (Atwell et al., 2010). Knowledge representation is a systematic approach to learn and understand a science based on a number of classification methods (Kuhn, 2010). Some methods of classification adopted by researchers such as taxonomies, hierarchical, tree structure, paradigm and constrained facet analysis in the delegate, discover and construct new knowledge (Kuhn, 2010; Villa, Athanasiadis & Rizzoli, 2009). Currently, the researchers using the ontology to represent knowledge in a wider context, specifically in the domain of Al-Quran. It is important to present the Al-Quran knowledge in a systematic manner because the Quran knowledge is categories into *muhkamat* (i.e., semantic sentence is clear) and *mutashabihat* (i.e., semantic sentence is not clear) that require a high level of complex explanations.

An effort to create information systems that enable Al-Quran knowledge to be accessed online has been implemented by many researchers and system developers (Abbas, 2009). Among them are Atwell, Brierley, Dukes, Sawalha and Sharaf (2011) who have built the Quran ontology-based information system in common Arab language and Arabic corpus. The ontological structure is focused on the relationship of concept in Al-Quran taken from Arabic language. Saad and Salim (2007) have developed ontology that makes Al-Quran and Hadith as the source for determining the related concept of knowledge in Islam. Saad, Salim, Zainal, and Noah (2010) have developed knowledge of Islam based on the Al-Quran by using ontology. Al-Yahya, Al-Khalifa, Bahanshal, Al-Odah, and Al-Helwah (2010) study focused only on the ontology for semantic lexicon for a concept in the Al-Quran, in particular the concept of time.

All studies still infancy and much more work needs to be done to complete the ontology of Al-Quran. However, the application of the Al-Quran still being developed based on traditional methods (i.e., using entity-relationship modeling). Moreover, most Al-Quran ontology is developed based on the English and Arabic languages, and does not have a complete application that allows searching and learning with a semantic based method.

DEVELOPMENT OF AL-QURAN ONTOLOGY

Classification of Knowledge Concept

Al-Quran verses can be more than one sentence or part of sentences. A group of verses will form a chapter with a given unique name for explaining about the content of the chapter. All chapters and verses can be identified by numbers such as 2:255 refers to chapter 2 (Al-Baqarah) and verse 255 (known as Al-Kursi verses). The concept of Al-Quran knowledge is defined by the structure of content hierarchy of divisions, chapters and verses, which can be classified by the appropriate themes. The theme (i.e., thematic) approach is a way of teaching and learning in such a way that many areas of the knowledge are integrated and connected within a theme, therefore it allows learning to be less fragmented and more natural (Hislop, 2004). The number of verses and chapters may share similar themes such as *Iman* and *Akhlaq*. Therefore, the definition of concepts will be based on classification of themes in *Syammil Al-Quran Miracle the Reference* (Shohib, Enang, Syatibi & Sidqi, 2010), and referred to the correct verses by using a hierarchy structure of Al-Quran contents as illustrated in Figure 1.

The Al-Quran Ontology Model

The ontology is used to present the content of Al-Quran in a systematic and structured manner by mapping the meaning of themes toward the appropriate verses of Al-Quran, which

provides the correct semantics of the Al-Quran knowledge. The model of Al-Quran ontology is developed and illustrated as shown in Figure 2.

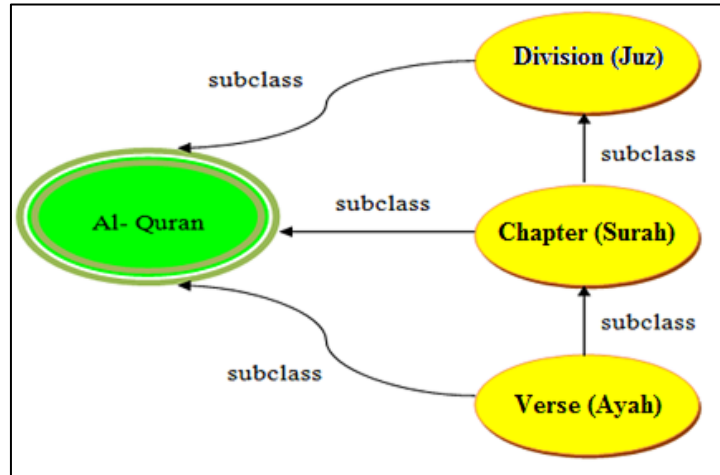


Figure 1. The Al-Quran Content Hierarchy

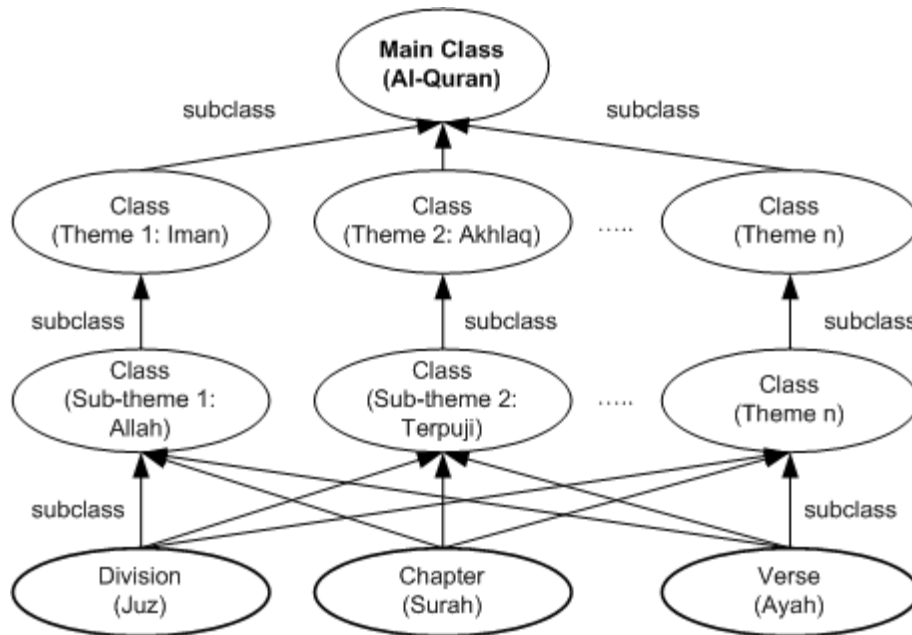


Figure 2. Al-Quran Ontology Model

Based on the model in Figure 2, the Al-Quran ontology structure such as class (the themes), property (the relationship between the themes) and individual (the verses) were defined in specification as presented in Table 1.

Table 1. The Al-Quran Ontology Structure

Terms	Ontology Structure	Al-Quran Knowledge Themes
Root	Main class	Al-Quran
Theme(s)	Class / subclass	Iman, Akhlaq
Sub-theme(s)	Class / subclass	Iman (Allah, Hari Akhir, Kitab-kitab Allah, Malaikat, Masalah Ghaib, Orang-orang

		Beriman, Para Nabi dan Rasul) Akhlak (Akhlak yang tercela, Akhlak yang terpuji)
Division, Chapter, and Verse	Class / subclass	Division (Juz), Chapter (Surah), and Verse (Ayah)
Relationship	Property	Link between classes to classes, classes to subclasses and, sub-classes to sub-classes. Example: <i>Iman</i> (class) <i>wajib_kepada</i> (property) <i>Hari Akhir</i> (subclass)

The ontology is defined based on themes in *Syammil Al-Quran Miracle the Reference*, which has been endorsed by the experts in the Al-Quran knowledge. Therefore, the model of Al-Quran ontology can be defined as following form:

$$\text{Al-Quran Ontology Model} = (\text{CT}, \text{CS}, \text{P}, \text{I}_1, \text{I}_2, \text{I}_3 \dots \text{I}_n)$$

Where:

- CT – Class for main themes in Al-Quran
- CS – Class (sub-class) for sub-themes in Al-Quran
- P – Properties for relationship between class to another class/sub-class
- $\text{I}_1, \text{I}_2, \dots, \text{I}_n$ – Individual for verses that referred by the themes

Given the Al-Quran knowledge themes, the ontology of Al-Quran is constructed by defining the CT, CS, P and $\text{I}_1, \text{I}_2, \dots, \text{I}_n$. The class is created for each primitive concept of themes and sub-themes. Whereby, property P is created for domain of P that set to be the class corresponding to the primitive concept of CT and CS. Therefore, each relation of P is a defined class is created. The created class is defined in the form: $C_p \equiv C_1 \sqcap C_2 \sqcap \dots C_n$, where C is the class representing the concept mapped to P. Each class C_i is derived from the themes or sub-themes definition.

The Al-Quran Ontology Development

The identification of Al-Quran themes, scope and terms was collected through reading the *Syammil Al-Quran Miracle the Reference* and interviewing with several peoples who expert in Al-Quran knowledge. *Iman* (faith) and *Akhlak* (deed) main themes were chosen as the research scope for constructing the ontology. Interviews have been conducted with experts (lead by Ustazah Nur Fadlina Mat Dali) from Al-Quran Study, Kolej Universiti Insaniah (KUIN). She has explained in detail about the relationship among divisions, chapters and verses. The ontology is developed by using Protégé-OWL and Malay language as medium of concepts and semantics to help users understand and learn the Al-Quran.

The first step is to create a class hierarchy for *Iman* and *Akhlak* as main themes, and other concepts will be defined as sub-themes. This is showed in Figure 3 and Figure 4. From the sub-themes, the next level of sub-themes are derived. The next task is to define the properties (object properties and data properties) that present the relationship between two classes. Figure 5 shows both properties in relations with the classes. Individuals represent the verses have been identified relevant to the main themes, sub-themes and the next levels of themes. The individual value is annotated by text string, but most values are constrained by data literal in a particular format (e.g., date, float, and integer), prior versions of the ontology, and Universal Resource Indicator (URI) referring to other resources (e.g., another ontology). The individual annotations were set according to the classes and properties as shown in Figure 6.

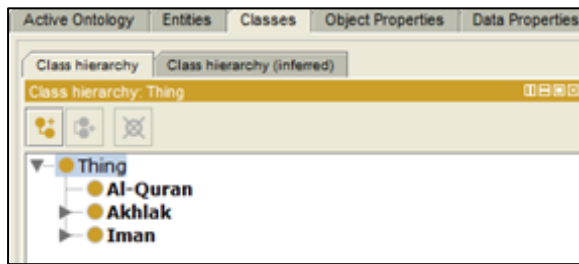


Figure 3. Main Class Hierarchy



Figure 4. Sub-Themes Class Hierarchy

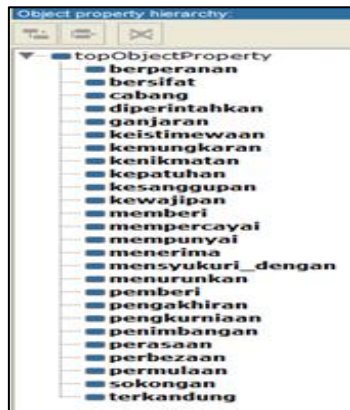


Figure 5. Object Property

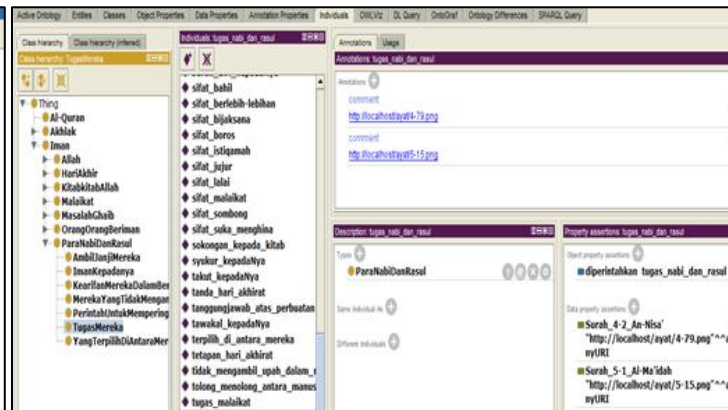


Figure 6. Individual Annotations from Classes and Properties

VALIDATION

The validation is performed to ensure the correctness of ontology according to the themes, sub-themes and the next level of sub-themes that presents the actual definition of the Al-Quran. This process has been carried out by using an expert review technique, which seven Al-Quran experts in Kulliyah of Al-Quran, KUIN involved. The focus group session was conducted and validity elements of the ontology model have been asked. Examples of the questions are: i) does the Al-Quran discusses only one topic for each verse in a chapter? ii) how to find verse on a certain theme? iii) what are the relationships between each theme according to Al-Quran knowledge?).

The results show that the proposed ontology model generally has strong support for presenting the Al-Quran knowledge using the theme-based approach (i.e., thematic) in the ontology structure. The Al-Quran ontology can be illustrated in graph as shown in Figure 7.

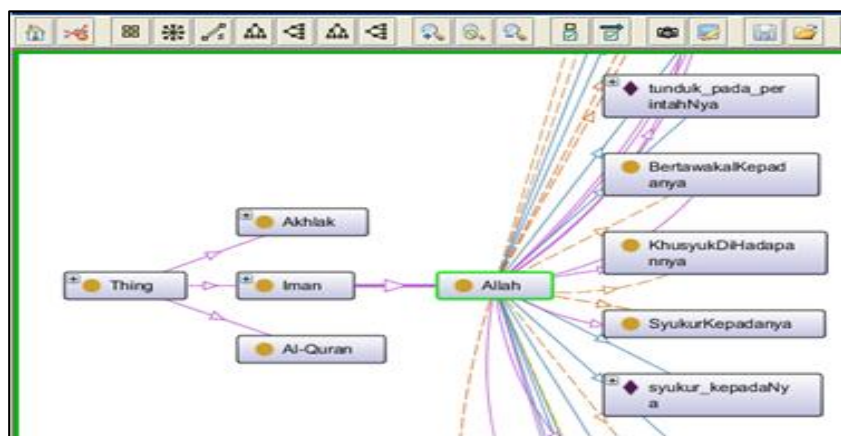


Figure 7. Al-Quran Ontology

CONCLUSION

This research aims to present the Al-Quran knowledge in theme-based approach by using ontology. The thematic approach is believed easier for users to understand the Al-Quran knowledge in a systematic way. In information technology environment, the representation and classification approaches must be understood by human and machine in order to ensure the correctness of Al-Quran semantics. Ontology approach was headed the use of conventional methods such as taxonomy, hierarchy or tree structure to present and classify the Al-Quran knowledge for developing a smart community of Muslims. Further works will explore the application for benefiting the ontology of Al-Quran to the users.

REFERENCES

- Abbas, N. H. (2009). *Quran 'Search for a Concept' Tool and Website*. Unpublished Dissertation, Leeds.
- Al-Yahya, M., Al-Khalifa, H., Bahanshal, A., Al-Odah, I., & Al-Helwah, N. (2010). An Ontological Model for Representing Semantic Lexicons: An Application on Time Nouns in the Holy Quran. *The Arabian Journal for Science and Engineering*, 35(2C), 21-35.
- 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 National Information Technology Symposium (NITS), Riyadh, Saudi Arabia.
- Atwell, E., Dukes, K., Sharaf, A.-B., Louw, N. H. B., Shawar, B. A., McEnery, T., et al. (2010). *Understanding the Quran: A new Grand Challenge for Computer Science and Artificial Intelligence*. Paper presented at the British Computer Society Workshop, Edinburgh.
- Guarino, N. (2009). Formal ontology, conceptual analysis and knowledge representation. *Special issue on Formal Ontology, Conceptual Analysis and Knowledge Representation*.
- Hislop, D. (2004). *Knowledge Management In Organizations: A Critical Introduction*. London: Oxford University Press.
- Kuhn, T. (2010). *Controlled English for Knowledge Representation*. Unpublished PhD, University of Zurich.
- Noy, N. F., & Musen, M. A. (2000). PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment. Paper presented at the AAAI'00.

-
- Saad, S., & Salim, N. (2007). Building Islamic Ontology Based on Ontology Learning. *Paper presented at the Postgraduate Annual Research Seminar, FSKSM.*
- Saad, S., Salim, N., Zainal, H., & Noah, S. A. M. (2010). A Framework for Islamic Knowledge via Ontology Representation. *Paper presented at the Information Retrieval & Knowledge Management (CAMP), Kuala Lumpur.*
- Shohib M., Enang H., Syatibi H.M., Sidqi A.Z. (2010). *Syammil Al-Quran Miracle the Reference.* Religious Department, Jakarta, Indonesia: Sygma Publishing.
- Villa, F., Athanasiadis, I. N., & Rizzoli, A. E. (2009). Modeling with knowledge: A review of emerging semantic approaches to environmental modeling. *Environmental Modeling & Software*, 24(5), 577-587.