

# An Ontological Approach towards Dialogue-based Information Visualization System: Quran Corpus for Juz' Amma

*Mohd Amin Mohd Yunus*<sup>1,\*</sup>, *Aida Mustapha*<sup>1</sup>, *Rizwan Iqbal*<sup>2</sup> and *Noor Azah Samsudin*<sup>1</sup>

<sup>1</sup>Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Batu Pahat, Johor, Malaysia

<sup>2</sup>Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

**Abstract.** This paper presents a corpus that offers rich and authentic knowledge for Juz' Amma using the ontological engineering approach. The reason of the study conducted is because of lack of understanding of Juz' Amma in the form of listing result. At the same time, the listing result does not provide pictorial tidings for clustering the listing result. Therefore, the corpus covers knowledge from various authentic sources, including translation, annotation of keywords, tafsir, and related hadiths for every verse in English and Malay. The ontology is engineered and designed to effectively utilize the authentic knowledge extracted from the corpus intended to serve a dialogue-based information visualization system for Quranic text. A standard ontology engineering methodology was applied for developing the ontology. After creation, the ontology was evaluated and was found to fulfill the goals for which it was developed. The corpus and ontology are hoped to aid various semantic applications related to Quranic text in the future.

## 1 Introduction

With regards to studies in the Quranic text, most religious scholars agree that contextual information is considered mandatory for correct interpretation of the Quranic verses. In the contextual aspect, Quranic exegeses (tafsir) greatly and facilitate understanding by providing correct interpretation and meaning of the verses. Nonetheless, the tafsir is not verse-based of word-to-word translation. It is a lengthy and detailed explanation of the verses that may span from a few sentences to a full paragraph. In the same way, hadith may also become a potential source for providing keen understanding to the true meaning of some verses. Hadiths are the sayings or oral traditions related to the words and acts by Prophet Muhammad (pbuh).

---

<sup>1</sup> Corresponding author: [amin@uthm.edu.my](mailto:amin@uthm.edu.my)

AQILAH approach is an ambitious effort to qualitatively enhance human experience in navigating the Quranic text using human-machine dialogues. The use of natural language dialogues via conversational agents offers higher understanding of the presented information [1]. The dialogue-based architecture of AQILAH facilitates human-computer interaction by using the ontology to rearrange the Quranic content based on need-to-know basis, as opposed to sequential order as in recitals. Hence, an ontological engineering approach to its core knowledge base is essential so the knowledge for Juz' Amma is accessible in a non-sequential form. Therefore, the objective of this study is to design and develop a Quran corpus and ontology that present rich, authentic knowledge for Juz' Amma of the Quran that encapsulates information revolving around the translations, tafsir, and hadiths. The ontology will be derived from the corpus using a standard ontology engineering methodology by Gruninger and Fox [4].

## 2 Related Work

Ontologies have gained considerable focus in the last two decades. At present day, they are extensively used in different areas like knowledge engineering, artificial intelligence, natural language processing, e-commerce, intelligent information integration, information retrieval, database design and integration, bioinformatics and expert systems. An ontology engineering methodology caters the methodological aspect of ontology development. It gives a set of guidelines and activities to develop a particular ontology. Among the major ontology engineering methodologies proposed to date include the methodology by Uschold and King [10] and Gruninger and Fox [4].

In the Quranic domain, researchers have worked on encapsulating the knowledge of Quran using ontologies [8,9, 11], which focus on ontology creation for Islamic concept such as prayers sourced from the Quranic text. Ontology developed by Dukes (2012) covers about 300 linked concepts with 350 relations, which primarily acts as a means of mapping different concepts to related verses in the entire Quran. However, all the ontologies discussed do not provide any contextual information related to the verses and do not report on the opted ontology engineering methodology.

## 3 Experimental Approach

This section discusses the details of the Quran ontology developed based on the corpus of Juz' Amma. The ontology design chose the methodology proposed by Gruninger and Fox (1995) because the motivation scenarios presented in the initial phase of this methodology facilitates the development team to identify intuitively possible solutions to the scenario problems [3,4].

### 3.1 Ontology design

The Gruninger and Fox's methodology is inspired by the development of knowledge-based systems using first-order logic [4]. The methodology proposes to develop the main scenarios, which describe the set of ontology's requirements that the ontology should satisfy after implementation. These scenarios provide a starting point to visualize the intended semantics of the objects and relations that will later become part of the

computable ontology.

Subsequently, natural language questions or also known as the informal competency questions are derived from the motivation scenarios to help determine the scope of ontology. After formally implementing the ontology, the competency questions will be transformed into an equivalent formal form for evaluation purposes. This is to ensure that the ontology developed fulfills the requirements for which it is made. Figure 1 shows the processes involved in the methodology by Gruninger and Fox [4].

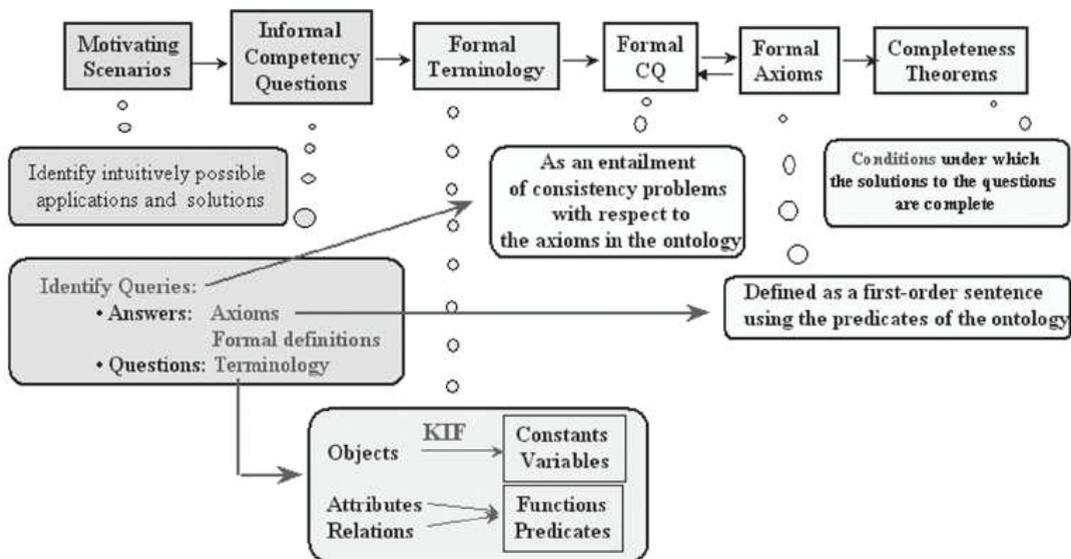


Fig. 1. Processes in the methodology (Source: Gruninger and Fox 1995 [4])

Based on Figure 1, the process begins by identifying the motivation scenarios. Such scenarios are related to the applications that will utilize the ontology. The scenarios define the set of requirements which the designed ontology should satisfy. They intuitively provide possible solutions to the scenario problems. The next process is defining the informal competency questions. They are derived from the motivation scenarios, specifying questions in natural language which the ontology should answer. The third process starts when the ontology engineer extracts the terminologies such as concepts, attributes, and relations in first order logic from the competency questions. Answers to these competency questions will guide the ontology engineer to extract the knowledge to be included in the formal definitions of the concepts, relations and formal axioms used.

Next, the competency questions are translated into their formal form followed by definitions of the ontology terms and constraints in their interpretation using axioms. The final process is to specify the completeness theorems, meaning defining the conditions under which the solution to the questions is complete. In this paper, evaluation of the Quran ontology will be carried out by running Description Logic (DL) queries and getting correct answers for it, which serve as the criteria for satisfying the completeness theorems.

### 3.2 Development of ontology

The Quran ontology for Juz' Amma has the following features which was distinguished from the existing Quran ontologies already mentioned in Section 2. The ontology is developed by extracting authentic knowledge from the Quran corpus. Gruninger and Fox methodology [4] is opted for developing the ontology, which is well-founded for building and evaluating ontologies [3]. The ontology covers the knowledge of Quran in the form of relationships between Juz', Surah and verse (Ayat). In order to provide contextual information, the verses of Quran have been linked with Quranic exegeses (tafsir), hadith and dual translation in English and Malay for correct interpretation of the verses. The ontology is engineered in a manner that it can be used as core knowledge base for a dialogue-based information visualization system and other related applications in the future. The design of

Quran ontology is flexible; it can be extended and modified as per need. The ontology development began by identifying the motivation scenarios. Motivation scenarios were created to discover the set of requirements for the Quran ontology that has to be formally implemented. It is often useful to use a template as a representational artefact to share the motivation scenarios. Use case templates in object-oriented methodology are useful for this purpose since it shows meaningful information in the most concise way [2]. Table 1 shows a sample of use case representing the motivation scenario developed for this ontology and it describes a user can input formal query in the query tab, provided in the Protege IDE.

**Table 1.** Use case for the Quran ontology.

<b>Brief Description</b>	<b>The use case describes that how a user can input formal query in the query tab, provided in the Protege IDE.</b>
Actors	User & Protege IDE
Preconditions	The Protege IDE is properly installed. The Qur'an ontology is loaded in the environment. The user is familiar with the formal query language, like SPARQL
Basic flow of events	The user initializes the Protege IDE. The user inputs a query in formal language (Like SPARQL) in the query tab. After inputting the query user presses the execute button. The system searches the Qur'an ontology for the appropriate answer, based on a triple (RDF) searching format. The answer/answers are populated on the screen If the user wants to run another query; he erases the previously written query and repeats the above procedure from step 2.
Alternative flow	Incorrect query syntax: If the user inputs the formal query in an incorrect syntax in step 2, then no results will be displayed. It is out of scope query, if the user inputs the query in correct syntax in step 2 but it is out of scope of the Qur'an ontology, then no results will be displayed.
Key scenarios	The system does not respond to the query and no results are displayed.
Post conditions	Successful query execution: The user successfully receives the correct answer for his query from the system. For failure condition, the user do not receive any response/answer for his query from the system.
Special	The user should be familiar with the formal language (SPARQL) and Protege
Requirements	IDE

After identifying the motivation scenarios, a set of informal competency questions are defined based on the motivation scenarios. These questions are prepared in natural language to identify the scope of ontology under development. The set of informal competency questions help the ontologist to extract content for the ontology, which will later become part of the computable ontology. Some informal competency questions in natural language

defined for the Quran ontology may include which Surahs belong to Juz' 30, which verses (Ayat) belong to Surah 113, which Surahs were revealed in Makkah and have verses (Ayat) greater than 15, which Surahs belong to Juz' 30 and were revealed in Medina, what is the English Translation for verse 4, Surah 111, what is the tafsir (Quranic exegeses) of verse 2, Surah 114 and what is the English name for Surah 114.

Informal competency questions ease the work of the ontologist to identify concepts, attributes and relations to be included in the ontology. Furthermore, answers in natural language to the competency questions help the ontologist to extract knowledge to be included in the formal definitions for the ontology content [3]. Table 2 presents some relations and inverse relations for the Quran ontology. Some concepts, instances and attributes identified for the Quran ontology are presented in Table 3.

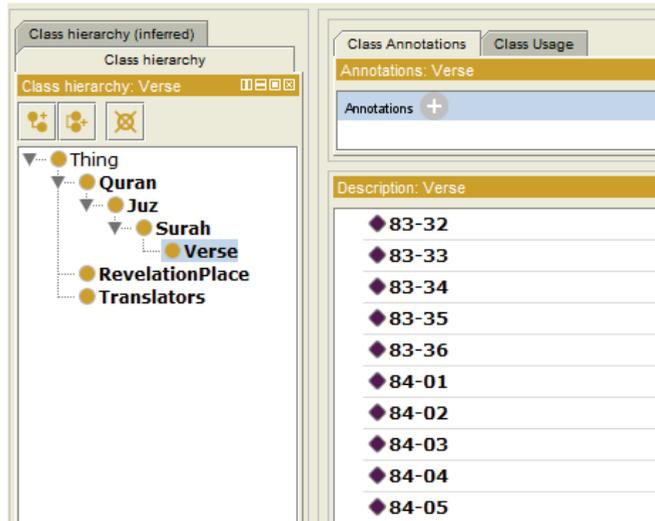
**Table 2.** Relations and inverse relations.

Relation name	Source concept	Source cardinality	Target concept	Inverse relation
Belong to Juz	Surah	(1...*,1...*)	Juz	Has Surah
Belong to Surah	Verse	(1,1)	Surah	Has verse
Has Surah	Juz	(1...*,1...*)	Surah	Belong to Juz
Has verse	Verse	(1...*, 1)	Surah	Belong to Surah

**Table 3.** Concepts, instances and their attributes.

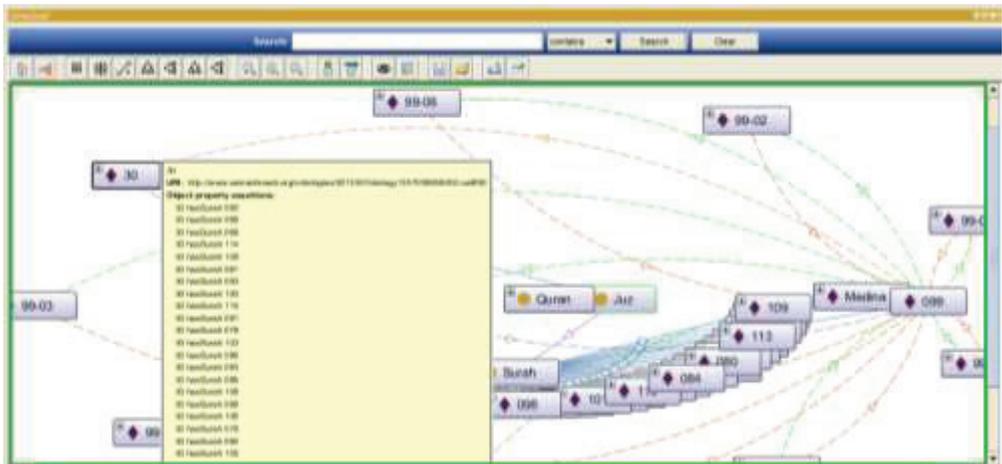
Concept name	Instances	Instance Attributes
Juz	78-Juz 99-Juz 80-Juz	Name of Juz' in Arabic
Surah	An-Naba At-Takwir	Surah name in Arabic Surah name in English No of verses
Verses	114-1 114-2 114-3	Belong to Surah Has translation Has Tafsir
Revelation Places	Makkah Medina	Not available
Translator	Hilali-Khan Abdel Haleem	Not available

Once all the ontological content is finalized, the ontology is transferred into a computable form using an ontology editor, which is the Protege IDE [7]. Protege is an open-source ontology editor that offers a wide range of useful plugins. Figure 2 shows a snap shot of the computable Quran ontology implemented in Protege. The wide range of plugins supported in Protege allows the ontologist and the entire development team to develop a clear understanding of the implemented ontology.



**Fig. 2.** View of class hierarchy of Quran ontology in Protege.

Besides many plugins available in Protege, OntoGraf was found to be very [6] effective in interactively navigating the Quran ontology. OntoGraf provides support for interactively navigating the relationships of the ontology using multiple layouts to automatically organize the structure of the ontology. It allows the user to create the desired view by filtering relationships [6]. Figure 3 illustrates a view of OntoGraf showing the relationships between classes and instances in the Quran ontology.



**Fig. 3.** OntoGraf showing relationships between classes and instances in the Quran ontology.

## 4 Result and Discussion

Table 1 shows the formula to calculate the percentage of precision and recall. After the ontology is implemented, it is important to ensure that the computable ontology fulfils the requirements and scope for which it is made. In order to perform evaluation, natural language competency questions are transformed into their machine level equivalent formal query language. For this research, competency questions for Quran ontology were converted into their equivalent queries in the form of Description Logic.

Description Logic (DL) is a family of formalisms that belongs to logic-based knowledge representation. DL empowers OWL ontology to provide precise meaning to the description and also helps to develop reasoning algorithms, which help to provide answers to queries about the domain [5]. Protege IDE provides a DL-Query Tab to run DL queries to retrieve knowledge from the loaded ontology. Table 4 represents some informal competency questions transformed in to their DL query equivalent, for evaluating the Quran ontology.

**Table 4.** Informal competency questions and their equivalent DL queries.

Informal competency questions	Equivalent DL query
Which Surahs were revealed in Makkah and have verses (Ayat) greater than 15?	Revealed in value Makkah and no of Verses some int [ $> 15$ ]
Which Surahs belong to Juz 30?	Belong to Juz value 30
Which Surahs belong to Juz 30 and were revealed in Medina?	Belong to Juz value 30 and revealed in value Medina
Surahs which have verses greater than 20 but less than 35 belong to Juz 30 and revealed in Makkah?	noOfVerses some int [ $> 15, < 35$ ] and belong to Juz value 30 and Revealed in value Makkah

Next, the Quran ontology was evaluated by running DL queries, using the DL-Query tab in Protege. It was found that the computable Quran ontology returned correct answers for the DL queries submitted, using the DL-Query tab. This manifests that the Quran ontology successfully fulfilled the requirements for which it was made and it can be effectively used for extracting the desired Quran related knowledge. Figure 4 shows a snapshot of a DL query running in Protege.



**Fig. 4.** A snapshot of Protege running a DL query on Quran ontology.

## 5 Conclusions

The corpus developed for Juz' Amma is based on individual verse or Ayat. Each verse in this corpus is annotated with literal translation, tafsir, and supporting hadith. All the knowledge in this corpus is based on authentic knowledge extracted from various authentic sources. The Quran ontology is developed by extracting knowledge from this authentic Quran corpus. The developed ontology encapsulates authentic knowledge of Quran in a machine understandable format, which makes it suitable to be used as the core knowledge base for a dialogue-based information visualization system such as AQILAH. The Quran ontology is developed using a standard ontology engineering methodology, which makes it authentic, robust and reliable. The ontology has been evaluated to ensure that it fulfils the requirements for which it was developed. Furthermore, the design of this ontology is flexible so it can be easily extended and modified in the future. It is hoped that this ontology will be able to serve AQILAH in aiding mental visualization when studying the Quran by enabling the users to communicate the content of Juz' Amma with clarity, precision, and efficiency. Both the corpus and ontology carry considerable potential to benefit upcoming Quran related semantic applications in the future.

## Acknowledgement

This work was supported in part by a grant from the Ministry of Education under the Research Acculturation Grant Scheme (R045) and in part by a grant from Research Gates IT Solution Sdn. Bhd.

## References

1. R. J. Beun, E. Vos, C. Witteman, Embodied conversational agents: Effects on memory performance and anthropomorphisation. *Intelligent Virtual Agents*, pp. 315-319 (2013)
2. G. Brusa, L. Caliusco, O. Chiotti, Towards ontological engineering: A process for building a domain ontology from scratch in public administration, *Expert Systems* **25**(5), 484-503 DOI 10.1111/j.1468-0394.2008.00471.x (2008)
3. A. Gomez-Perez, M. Fernandez-Lopez, O. Corcho, Ontological engineering: with examples from the areas of knowledge management, e-Commerce and the Semantic Web, Machine Learning and Data Mining for Computer Security, *Methods and Applications*, Springer-Verlag. <http://fs1.bib.tiera.ru/content/dvd44/Maloof> M.A (2005)
4. M. Gruninger, M. S. Fox, Methodology for the design and evaluation of ontologies, *Procs. Workshop on Basic Ontological Issues in Knowledge Sharing*, IJCAI, **95**, pp. 26-37 (1995)
5. I. Horrocks, Semantic web: the story so far, *Procs. 2007 Int. Cross-Disciplinary Conf. Web Accessibility (W4A)*, ACM. pp. 120-125 (2007)
6. OntoGraf Plugin (2013). <http://protegewiki.stanford.edu/wiki/OntoGraf> (2016)
7. Protege IDE (2013). <http://protege.stanford.edu/> (2016)
8. S. Saad, N. Salim, H. Zainal, Islamic knowledge ontology creation, *Procs. Internet Technology and Secured Transactions (ICITST)*, pp. 1-6 (2009)
9. S. Saad, N. Salim, H. Zainal, S. A. M. Noah, A framework for Islamic knowledge via ontology representation, *Procs. Int. Conf. Information Retrieval and Knowledge*

- Management (CAMP)*, IEEE. pp. 310-314 (2010)
10. M. Uschold, M. King, Towards a methodology for building ontologies, *Procs. Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada, **74**, pp. 23-42 (1995)
  11. F. E. Zahari, Ontology application for the Quran, Master's thesis, Universiti Utara Malaysia (2009)