

# A Quran Repository System for Ethics, Economy and Traveling

**Madihah Mohd Saudi<sup>1</sup>, Mohd Haizam Saudi<sup>2</sup>, Obsatar Sinaga<sup>3</sup>, Athira Abd Rahim<sup>4</sup>**

<sup>1</sup>Islamic Science Institute (ISI), Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

<sup>2</sup>Widyatama Universiti, Bandung, Indonesia

<sup>3</sup>Padjadjaran University, Indonesia

<sup>4</sup>Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia

\*madihah@usim.edu.my

## ABSTRACT

Quran is the main resource of references and guidance for Muslims. However, as Quran contains a vast amount of information and unstructured data, it is crucial to extract the encapsulated knowledge for ethics, economy and traveling in the Quran scientifically. The need to retrieve faster and accurate information focusing on ethics, economy and traveling is the urge to produce this paper. This paper aims to develop a Quran repository system for ethics, economy and traveling domain. Thus, this paper discussed how to extract the encapsulated knowledge and how to develop a secured Quran repository system that consists of information about traveling, ethics and economy using the Naïve string search algorithm. The tools used in this paper were based on open-source tools. For a secure system, Nikto has been used as the vulnerability scanner to scan this system, to ensure this system is free from any website security vulnerabilities issues. The system developed is beneficial to society in retrieving information about traveling, economy and ethics from Quran. This secured Quran repository system is easily accessible and provides comprehensive information about ethics, economy and traveling in the Quran.

## Keywords

Secure system, Naïve string, Quran repository, ethics, economy and traveling

*Article Received: 18 October 2020, Revised: 3 November 2020, Accepted: 24 December 2020*

## Introduction

It is very important to retrieve Islamic related information on the web and other storage mechanisms systematically. The information is usually organized in spiral or event-based such as in the Quran (Naseem Shazadi et al., 2011). Quran contains a rich repository of practical guidance on real-life situations, with injunctions regulating a vast field, from issues of international relations and matters of war and peace down to the habits of everyday life (Onder Bakircioglu, 2014). Quran contains 114 surahs, 77,000 words and 6,200 verses which have a vast amount of knowledge. The task to extract only certain knowledge is difficult due to inefficient technique in retrieving information. Therefore, this paper proposed to build a system of Quran repository for ethics, economy and traveling domain.

In software development, a repository is a central file storage location. A repository contains three primary elements which are a trunk, branches, and tags. The trunk contains the current version of a software project. Branches are used to store new versions of the program. Tags are used to save versions of a project but are not meant for active

development (Christensson, 2011). The term repository is from the Latin *repositorium*, a vessel or chamber in which things can be placed, and it can mean a place where things are collected. Depending on how the term is used, a repository may be directly accessible to users or maybe a place from which specific databases, files, or documents are obtained for further relocation or distribution in a network. A repository may be just the aggregation of data itself into some accessible place of storage or it may also imply some ability to selectively extract data (TechTarget, 2005).

This paper is organized as follows. Section II explains related works, Section III presents the method used in this paper, Section IV consists of the findings and evaluation, and Section V concludes the paper and makes suggestions for future work.

## Literature Review

The existing works that are related to extraction methods as follows in Table 1.

**Table 1.** List of previous work of extraction methods

Author	Methods
Maha Al-Yahy et al. (2010)	<ul style="list-style-type: none"> <li>- Used automatic extraction method to acquire ontology.</li> <li>- Focus mainly on salat or prayer concepts.</li> <li>- Another ontology development model used time nouns from the holy Quran to derive ontology structure.</li> </ul>
Ontology Research Team, UiTM (2010)	<ul style="list-style-type: none"> <li>- Automatically generate ontology instances from an unstructured document</li> <li>- Extracted concept and build the taxonomy of Islamic Knowledge.</li> <li>- The approach was the integration of ontology learning, ontology population and text mining framework for the extraction of information from Islamic knowledge sources.</li> </ul>
Sumayya Baqai et al. (2009)	<ul style="list-style-type: none"> <li>- Focused on bilingual (English/Arabic) comprehensive search tool using the keyword search.</li> <li>- Users need to insert the correct keyword</li> </ul>
Ivan Habernal et al. (2013)	<ul style="list-style-type: none"> <li>- Leveraging Semantic Web technology used for standardized knowledge modeling and retrieval of holy Quran and religious text.</li> <li>- The objectives of the system: for proving Quran knowledge and information sharing, storing, modeling, reasoning and retrieval from diverse Islamic domain sources.</li> <li>- Users need to insert the correct keyword</li> </ul>
Quratul Ain et al. (2011)	<ul style="list-style-type: none"> <li>- Based on keyword search which allows users to enter a keyword or</li> </ul>
	<ul style="list-style-type: none"> <li>navigate through the form.</li> <li>- It did not accept user Query with complex and long phrases.</li> </ul>
Ta'a et al. (2018)	<ul style="list-style-type: none"> <li>- The proposed approach consists of two sub-stages: the development of the ontology of Al-Quran and the development of a method with semantic search.</li> <li>- Islamic ontology, based on the stability, homogeneity and peace of the ordinary social system, as a solution to the problems.</li> </ul>
Pane et al. (2018)	<ul style="list-style-type: none"> <li>- Multi-label classification varies from single-label classification, so this study provided a new classifier model to manage multi-label classification.</li> <li>- With multiple stages of pre-processing data such as case folding, tokenization, and stemming, the framework was designed using Multinomial Naïve Bayes. The framework also used bags of words as a means of extracting functions.</li> </ul>
Radzid et al. (2018)	<ul style="list-style-type: none"> <li>- In-text line segmentation, this paper proposed a new technique to recognize overlaps between adjacent text lines and segment each line with precision.</li> <li>- Line segmentation could be difficult because the text row can be overlapped by diacritical marks or strokes of the Arabic word between adjacent text lines.</li> </ul>
Al-Zoghby et al. (2018)	<ul style="list-style-type: none"> <li>- Presented a comprehensive study of Arabic Semantic Relation Extraction and Arabic Ontology Learning Research Areas.</li> <li>- In particular, though</li> </ul>

- |                          |   |                           |  |
|--------------------------|---|---------------------------|--|
|                          | <p>concentrating on Arabic Semantic Relationship Extraction, this paper studied Arabic Ontology learning in general as the most important, but challenging activity in the learning process of Ontology.</p>  |                           | <p>word synonyms and combined with Quran ontology to get the relationships between concepts within the expanded query.</p>   |
| Alhawarat et al. (2015)  | <ul style="list-style-type: none"> <li>- In this paper, the holy Quran text was pre-processed, and then different text mining operations were applied to it to reveal simple facts about the terms of the holy Quran.</li> <li>- The results were based on term frequencies that were calculated using both Term Frequency (TF) and Term Frequency-Inverse Document Frequency (TF-IDF) methods.</li> </ul>  | Safee et al. (2018)       | <ul style="list-style-type: none"> <li>- Development of ontology for the Medical and Health Science domain in the Quran and implementation of ontology-based search method to answer related queries in the Quran.</li> <li>- The development of ontology adopted the ontology 101 approach which was later evaluated by Quran experts.</li> <li>- A hybrid search tool was developed that encompassed semantic-based and keyword-based technique to answer a user query.</li> <li>- The search tool was evaluated using the Recall and Precision measurements.</li> </ul> |
| Zouaoui and Rezeg (2021) | <ul style="list-style-type: none"> <li>- Designed a semantic search engine based on ontology as an index.</li> <li>- Focused on creating a new ontology for Quranic documents based on a set of useful words extracted from the Quranic E-Arab book with grammatical functions that serve as concepts.</li> <li>- Created links between the existing Quranic words in the same verse, which will be used with the user's query to find the desired verses.</li> </ul> | Mohamed and Shokry (2020) | <ul style="list-style-type: none"> <li>- Presented a concept-based searching tool (QSST) for the Holy Quran. It consisted of four phases.</li> <li>- The performance of the proposed QSST is measured by comparing results against Mushaf Al-Tajweed.</li> <li>- Then, precision, recall, and F-score are computed, and their percentages were 76.91%, 72.23% 69.28% respectively</li> </ul>   |
| Yusuf et al. (2020)      | <ul style="list-style-type: none"> <li>- Proposed a new hybrid query expansion approach to improve Quran search results using semantic search and Lucene ranking.</li> <li>- A novel semantic search for Quran search is first presented, in which, Quran search.</li> <li>- Queries are expanded with</li> </ul>   |                           |  |

However, there were slight differences between Sumayya Baqai et al. (2009) and Ivan Habernal et al. (2013) which were based on the ability of the system to answer a query with complex phrases. The users were allowed to ask any query in natural language by inserting how many words or sentences wanted. While other existing works were a lack in domains of the ethic, economy and traveling. Hence, we are proposing this paper to

fill in the existing gap on these domains by creating a system based on keyword search and focusing on ethics, economy and traveling domain only. A user could search for any keywords related to these three domains.

Methodology

Figure 1 depicted the overall overview of the methods used for this paper.

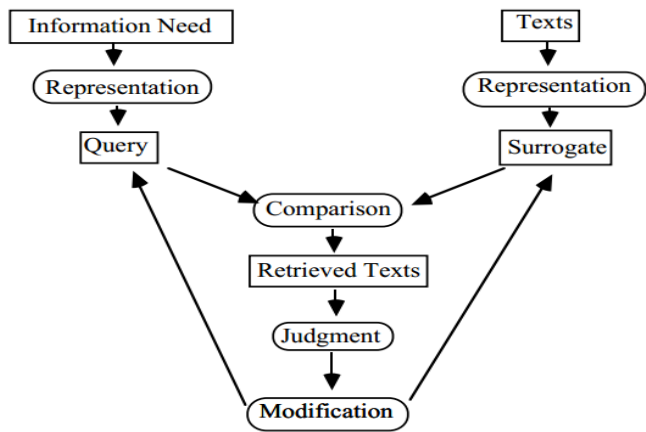


Figure 1. A model of the standard view of the proposed system

Figure 2 illustrated the system architecture of the Quranic Search Engine system that is derived from the design phase discussed previously. It consists of 3 phases, which are the interface layer, abstract layer and physical storage layer. The user interface design of the system is delivered on the interface layer, while the process of the system is processing on the abstract layer and the database storage is on the physical storage layer respectively. To make the system running properly, the two layers of the system structure must have integration within it.

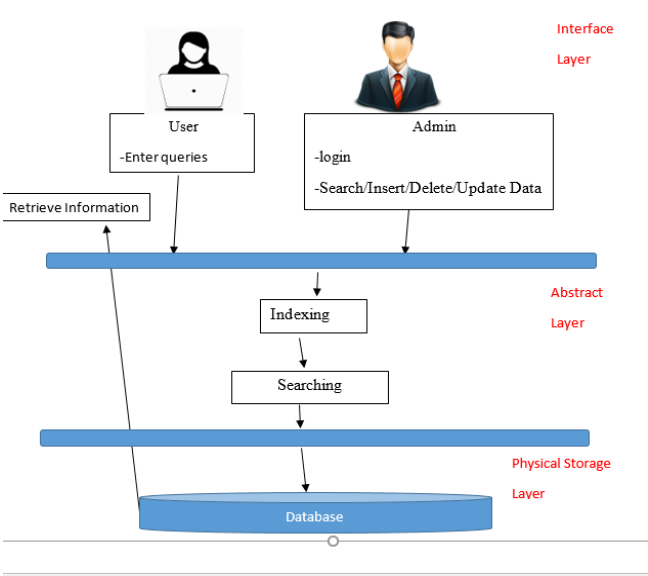


Figure 2. Quranic search engine system structure

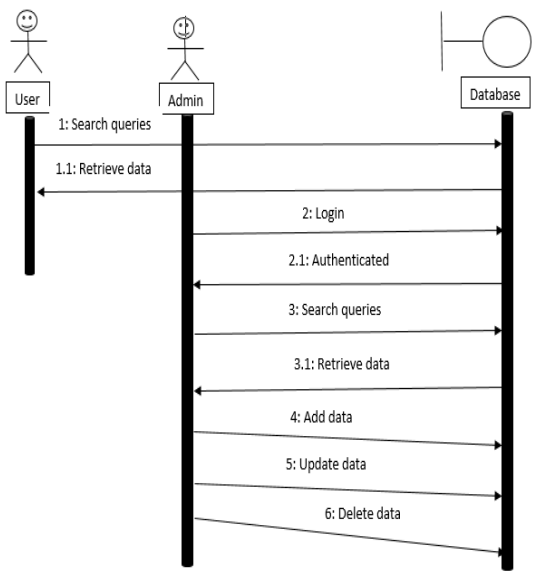


Figure 3. Sequence diagram of Quranic search engine

For the search string, we used the Naïve algorithm. This algorithm consists of trying to match any substring of length  $m$  in the text with the pattern. The expected number of text pattern comparisons performed by the naive or brute force algorithm when searching with a pattern of length  $m$  in a text of length  $n$  ( $n \geq m$ ) as displayed in Equation (1).

$$\bar{C}_n = \frac{c}{c-1} \left( 1 - \frac{1}{c^m} \right) (n - m + 1) + O(1) \tag{1}$$

The naïve solution is compared to the character by character of the text. It returns all valid shifts found. The time complexity of this algorithm is in its worst-case which is  $O(M \times N)$ . For example, if the pattern to search is am and the text is an, then M comparison is needed for comparison by shift. For all text,  $(N-M+1) \times M$  operation is needed but M generally is very small compared to N, that why the complexity can be considered as  $O(M \times N)$ .

For secure system development, we used Nikto to scan the vulnerabilities in our proposed system. We were focusing on the Content Security Policy(CSP) which consists of risks on policy misconfiguration and permissive policy. The screenshot of our security analysis is depicted in Figure 4. Based on our analysis, our system is secured from any possible web attacks. But still updating must be done from time to time to ensure its safety.

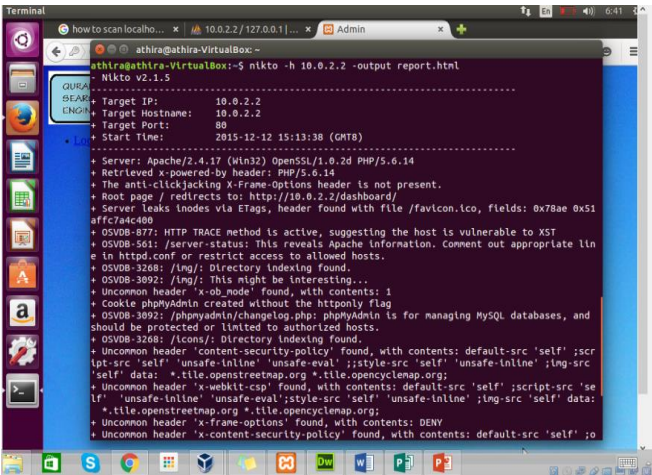


Figure 4. Nikto analysis report

Results and Discussion

The main interface of the proposed system is depicted in Figure 5.

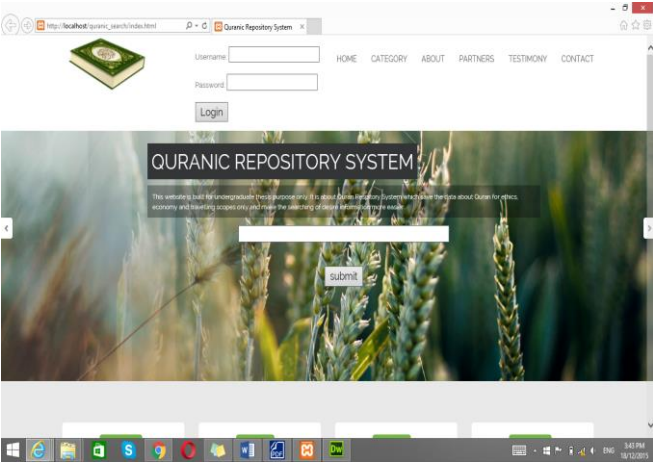


Figure 5. Main interface

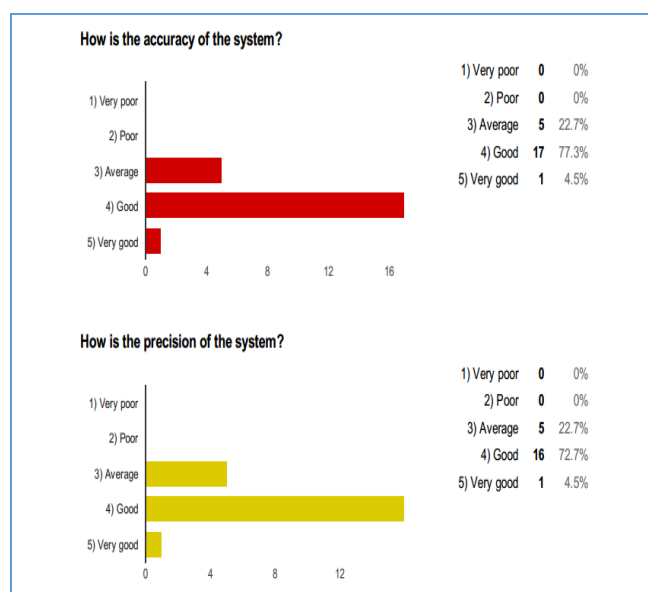
While, the following Table 2 showing our findings and comparison with the benchmark works.

Table 2. Summarization comparison with the benchmark works

Quran Repository System (Our Proposed System)	Quran Arabic WorldNet model was proposed by Manal AlMaayah and Majdi Sawalha ( 2014)	SearchQuran.com
Consists of 52 chapter and 237 verses only.	Has 77,430 words, 114 chapter and 6,243 verses.	Consists of a large amount of information.
Only based on English documents.	Compare Arabic and English documents.	Compare English and Arabic/Urdu documents.
Return searched results with an explanation and law for the ayah.	The complex structure, contradiction and similarities in the synonyms.	Search results returned Quranic verses/hadith and translations only.
Focus on ethics, travel and economy only.	Focus on one whole Quran.	Focus on the whole Quran and some Hadith.



Table 2 shows that Quran Repository System focuses specifically on the economy, traveling and ethics domain which differentiates it from other existing systems that focus on the whole Quran because this system is narrowing down the scopes to these three major domains to make it easy for the user to search the related information. Figure 6 showed our evaluation findings and evaluation of the accuracy and precision of the system. It showed that overall, users were satisfied with the proposed system in terms of accuracy and precision.



**Figure 6.** Evaluation summary of the proposed system

## Conclusion

For future work, this system could be upgraded by applying semantic search for better information retrieval and system performance improvement. This paper contributes to collect related information on ethics, economy and traveling in one repository system which will help users to search desired information and also get the explanations on searched results. The system developed is beneficial in specific domains of economy, travel, and ethics.

## Acknowledgement

The authors would like to express their gratitude to Widyatama University and Universiti Sains Islam Malaysia (USIM) for the support and facilities provided.

## References

- [1] Aliyu Rufai Yauri, Rabiah A. Kadir, Azreen Azman, Masrah Azrifah Azmi Murad. (2013). Quranic verse extraction base on concepts using OWL-DL ontologies. *Research Journal of Applied sciences, Engineering and Technology*, 6(23), 4492-4498.
- [2] R. Radzid, M. S. Azmi, I. E. A. Jalil, N. A. Arbain, A. K. Draman Muda, and A. Tahir, (2018). "Text line segmentation for mushaf al-Quran using hybrid projection based neighbouring properties," *J. Telecommun. Electron. Comput. Eng.*, vol. 10, no. 2-7, pp. 53-57.
- [3] Al-Zoghby A.M., Elshawi A., Atwan A. (2018). Semantic relations extraction and ontology learning from Arabic texts—A survey. In Shaalan K., Hassanien A., Tolba F. (Eds.), *Intelligent Natural Language Processing: Trends and Applications. Studies in Computational Intelligence*. Springer, Cham, pp. 199-225.
- [4] Alhawarat, M., Hegazi, M., & Hilal, A. (2015). Processing the text of the Holy Quran: a text mining study. *Int. J. Adv. Comput. Sci. Appl*, 6(2), 262-267.
- [5] Farooqui N.K. and Dr. Mohammad Fauzan Noordin. (2015). Knowledge Exploration: Selected Works on Quran Ontology Development, *Journal of Theoretical and Applied Information Technology*, 72(3).
- [6] Ivan Habernal, Miloslav Komopik. (2012). SWSNL: Semantic Web Search Using natural language. *Expert Systems with Application*, 40, 3649-3664.
- [7] Maha Al-Yahy, Hend Al-Khalifa, Hend Al-Khalifa, Alia Bahanshal, Iman Al-Odah, Nawal Al-Helwah. (2010). An Ontological Model for Representing Semantic Lexicons: An Application on Time Nouns in the Holy Quran. *Arabian Journal for Science and Engineering*, 35(2), 21.
- [8] Mohamed, E. H., & Shokry, E. M. (2020). QSST: A Quranic Semantic Search Tool based on word embedding. *Journal of King*

Saud University-Computer and Information Sciences, 2020, 1-12.

- [9] Shahzadi, N., & Shaheen, A. (2011). Semantic Network-based Semantic Search of religious Repository. *International Journal of Computer Applications*, 36(9), 1-5.
- [10] R. A. Pane, M. S. Mubarak, N. S. Huda and Adiwijaya, (2018) "A Multi-Label Classification on Topics of Quranic Verses in English Translation Using Multinomial Naive Bayes," 6th International Conference on Information and Communication Technology, pp. 481-484,
- [11] Safee, M. A. M., Saudi, M. M., Pitchay, S. A., Ridzuan, F., Basir, N., Saadan, K., & Nabila, F. (2018). Hybrid search approach for retrieving Medical and Health Science knowledge from Quran. *International Journal of Engineering and Technology (UAE)*, 7(4.15), 69-74.
- [12] Ta'a, A., Abed, Q. A., & Ahmad, M. (2017) "Al-Quran Ontology Based on Knowledge Themes." *Journal of Fundamental and Applied Sciences*, 9(5S), 800-817.
- [13] Yusuf, N., Yunus, M. A. M., Wahid, N., Nawi, N. M., Samsudin, N. A., & Arbaiy, N. (2020). Query expansion method for quran search using semantic search and lucene ranking. *Journal of Engineering Science and Technology*, 15(1), 675–692.
- [14] Zouaoui, S., & Rezeg, K. (2021). A Novel Quranic Search Engine Using an Ontology-Based Semantic Indexing. *Arabian Journal for Science and Engineering*, 2021, 1–22.