

A Cloud-Based Cross Language Search Engine for Quranic Application

Zameer Ahmed Adhoni*, Husam Al Hamad*, Abdul Ahad Siddiqi**, Zaheer Ahmed Adhoni***

*Information Technology Department, Qassim University, Qassim 51477, Saudi Arabia.

**Computer Science Department, Taibah University, Madinah, Saudi Arabia

***Senior Software Engineer, Huawei Technology, Bangalore, India.

zameeradhoni@yahoo.co.in, hhamad@qu.edu.sa, asiddiqi@tabahu.edu.sa, zaheerkadhoni@huawei.com

Abstract— In this paper, we describe a Cloud-Based Cross LANGUAGE Search engine, which aims at providing a unifying framework for searching the Quranic verses in multi-language. It uses Drupal technology for the implementation. Includes all necessary components for searching, building mobile applications and social networks around the central theme of Qur'an. A salient feature of the proposed API is the ability to retrieve Qur'anic verses in Arabic using search phrases in languages other than Arabic. The other part is design of a Mushaf recitation which is also cloud based and multi-language., we describe the core components and design patterns of the proposed API with emphasis on key design criteria.

Keywords— Qur'an API, Drupal.

I. INTRODUCTION

Many Qur'an related applications and software have been developed to satisfy the needs of online and offline users. All these are designed and developed based on technology and requirement. With the common use of Internet User prefer tools to be develop on internet base as it is most easy and common way to utilize .Many websites are built based on these requirement. In this work, we present a Quran Related search Based on Cross language which makes the user learn and search more conveniently and easily, the user can access by website and internet enabled this is develop by programming language Drupal. This creates more interest for user to learn, search and develop their knowledge

II. QUR'AN SEARCH

Quran Search is being designed and developed using search API module of Drupal 7 software, which enables us to store the indexing in DB or Apache Solr. We have currently used DB based Indexing. Other related projects work as a) Cross language information retrieval (CLIR)'. It presents semantic technique on queries for retrieving more relevant results in CLIR, that concentrate on the Arabic, Malay or English query(s) translation (a dictionary based method) to retrieve documents according to query(s) translation. According to this study, the semantic ontology significantly improves and expands the single query itself with more synonym and related words [1]. b) Concentrates on the evaluation of the existing visualization techniques to find ways to improve the view,

access and retrieve the content of Surah Al Baqarah effectively. The work focuses on five categories, namely muamalat, judiciary, Islamic criminal law, Islamic family law and jihad. In the prototype of this system, the paper describes several visualization techniques such as overview and detail, hierarchical, dynamic queries, and iconic representation of the surah [2]. c) DataQuest', an efficient framework for modelling and retrieving knowledge, from distributed knowledge sources primarily related to the Holy Quran related scholarly texts, with the use of Semantic Web, Information Extraction and Natural Language Processing techniques. The documents were annotated using the domain ontology and then a semantic based intelligent search engine let the user query that filtered and concise knowledge [3].

d) development made in the design and development of a complete and comprehensive online cloud-based Qur'an portal. The designed portal makes all the reading and resource sections accessible to the audience whether they are using laptops, PCs, mobile, tablet, or personal digital assistants. In its completed form, the portal covers almost all themes related to Qur'an, namely 'Al-Mus'haf', 'cross language information', 'reciter and bookmarks', 'translation and transliteration', 'memorizer', 'search', 'study materials', 'videos of scholars on Tajweed', 'personalization or self-evaluation', 'Quran for kids', and much more. The application designed, provides a specific widget interface on a mobile phone, enabling quick access and implementing the specific multimedia features of a mobile phone. A Qur'an authentic translation for non-Arabic users is also available, and transliteration of Qur'an in Urdu for people who cannot read Urdu but understand it, has also been included [4].

A. How does it work?

Search for any word in the Quran and all the Ayahs with those words shows up. The search can filter by Surah (or any additional parameters defined by us). For example if a user wants to get only the results from a particular surah, he can do it so by clicking on the Surah name. The data is added to the site systematically, Ayah by Ayah with proper information architecture and metadata's. It makes retrieving data in complex ways very easy and possible. Apart from that the data is indexed by Drupal and can be useful for full text search.

For instant using the relationships and metadata's we can pull an ayah from a surah, with the keyword given and from a particular Juz'. This is made possible by the search function of Drupal & the index. The search being on native API exposed by Drupal also makes it possible to later use solutions such as 'Apache Solr' for performance in case the site traffic grows. So the method is future centric. All the coding in the search is done in PHP and using the Drupal coding standards and API's.

B. Design of Qur'an Search API

The API for searching the Qur'anic text is designed to provide the functionalities:

- Search Qur'an by a word or an ayah (full text search).
- Search by using filters, like in selected Surahs (chapters of the Qur'an) or within some ranges of ayah (verse) range.
- Search by translations and transliterations.
- Search for commentaries, articles and translation of Qur'an.
- Search ayah based on topics.

Qur'an search API has a pluggable architecture. Translations or transliterations of the Qur'an text can be added with this API with ease using an interface to import. Right now the format followed by Tanzil [5] is supported. Once various translations/ transliterations are imported into the Qur'an database, the search data can be re-indexed to include the newly imported data. The search box used in our system is intelligent and multilingual. People can search in the language of their choice provided that the translation of that language is added to the site already. There are no separate selection or input based on different language.

The API facilitates the users to search for Ayahs in Arabic using their own language as shown in Figure 1. To the best of our knowledge, such cross-lingual search is the first of its kind for Qur'anic text. For example, people don't need to know the Arabic word امي (means 'my mother') to search about 'mom' in the Qur'an text. They could search 'mom' in English or the equivalent word in Urdu or Arabic, or even in Urdu transliteration. They will be instantly presented with the results in Arabic. The same is true vice versa; an Arabic search could bring up English results. The most interesting part of this is that, a new language can be added and all the features work with that language instantly. This feature can also help users to search for Qur'an audio in their preferred language.

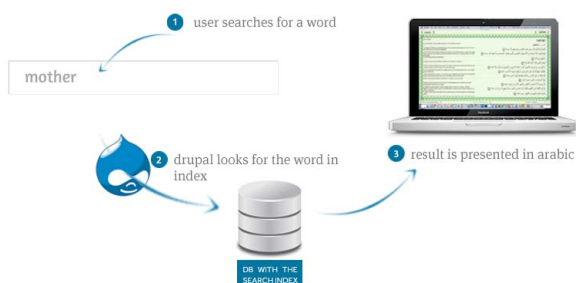


Figure 1. Visual illustration of Qur'an search process)

A robust feature of our AP is the ability to search in the translation/transliteration space. This search process is illustrated in figure 4. The basic idea is to maintain a central index for the Ayahs and Surahs of Qur'an. For each ayah, one or more translations (either in native language or in transliterated form) are stored. Such indices of the Qur'anic text create the central database where any search word (as shown in Figure 2, search string is 'father') is referred to. The search is then carried out in the indices using the preferred language (the language of the search string). The resulting Ayahs are then retrieved from API, along with the desired translations or explanations.

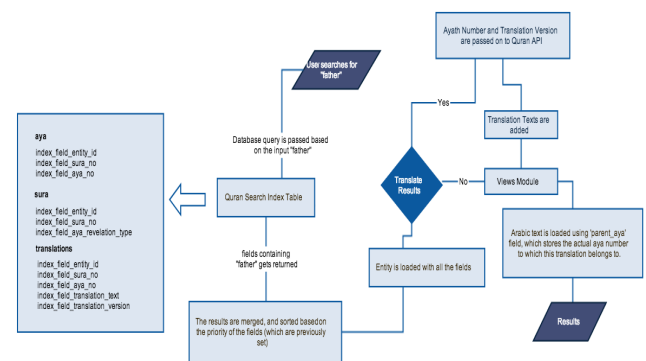


Figure 2. Illustration of different phases in searching for Qur'anic text in the translation space3

Such a framework for searching Qur'anic text offers several advantages. *First*, Qur'anic applications can be developed in any languages desired with a minimal effort on creating the indices. For a new language, all needed is to amend the indices with translations of the Ayahs in that particular language. The core search API remains same and can be used for searching in any language. *Second*, the development time for Qur'anic applications can be greatly reduced, as the main focus will now be the implementation of the interface, rather than the core search API. *Third*, applications developed using the proposed framework can be useful to users of various languages, as the core functionalities of the application will remain same for various languages, with a minimal change in interface for each language.

In the following section, we describe some implementation details of the proposed search framework for Qur'anic text. The implementation is done with state of the art tools, keeping in mind the ease-of-use and compatibility of the API for most of the developers

A. Search Scenario

Let's try out a scene by the scenario of search if a user wants to search for a word in particular, for example word "cow".

1. The user will access the search url '/search' which is being provided by the view module, that integrates with the search API, and acts as a middleman who receives and retrieves index data for based on user inputs.
2. An SQL query will be generated by the searchapi module that searches for words, in our case "cow"; in all of the fields of the particular search index we use. Figure 3 of

this appendix shows the result from the 'field_translation' field which has cow in it.

item_id	word	score
20360	cow	1
20361	cow	2
20363	cow	1
20364	cow	1
20366	cow	1

Figure 3. Search scenario of word 'cow'

- Each result will have an item_id, which is the entity id for a particular entity type. In this case, the searched component looked at all the fields of all entity types we have opted to index, and since we have also opted entity translation, we will also get results from it. It is based on this item_id, that a full resulting object is created, which is given to the view module to handle.
- View module takes the data and builds html objects, based on the setting provided in the view configuration. The result might look like as shown in Figure 4 of this appendix.

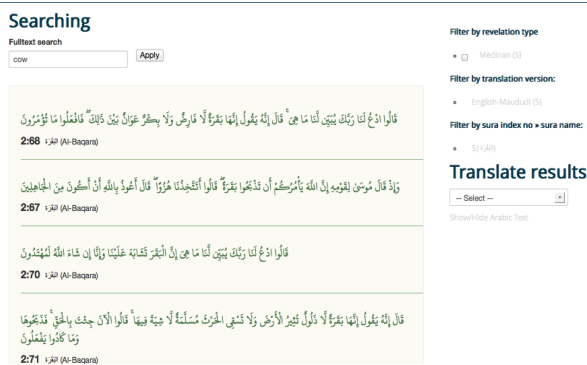


Figure 4. Search result of word 'cow'

- When a result is displayed, the system also provides some additional HTML tags that show a unique entity ids along with it, so when a user for example wants to see the results in a translation, and when he clicks on a particular translation, an Ajax function that requests a particular translation of Ayath gets performed. An instance of this feature can be accessed at:

http://quran.ziyr.com/api/trans_switch/18/12894+12894+12893+12893+12896+12896+12897+12897+12899+12899

It will further return a Json object with the translated text, as shown in Figure 5 of this appendix.

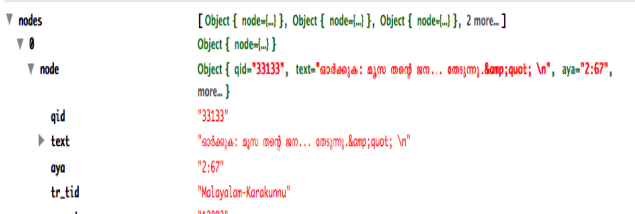


Figure 5. A snapshot of returned JSON object

6. Fetched data is then, appended to the existing listings by the JavaScript code. This option is stored in the cookie, so that the same function can be performed on pagination or future search, unless the user wishes to remove it. Figure 6 shows a flow chart about how the word 'cow' is searched by the system using various components of Search API module and other modules.

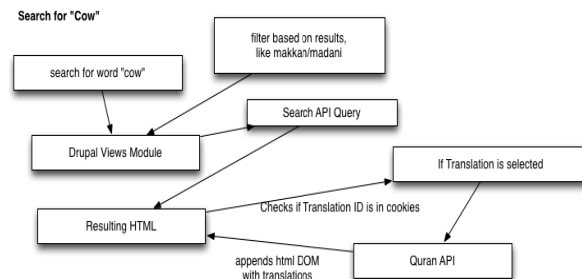


Figure 6. Flow chart of how word 'cow' is searched using Search API

III. THE RESULTS

A. Result of search by Arabic and Urdu transliteration

Figure 7 shows result of search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Urdu transliteration.

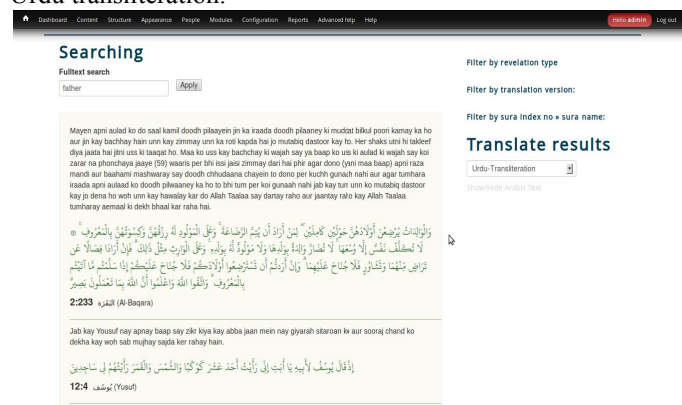


Figure 7. Search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Urdu transliteration

B. Result of search by cross language option

Figure 8 shows result of search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Malayalam language.

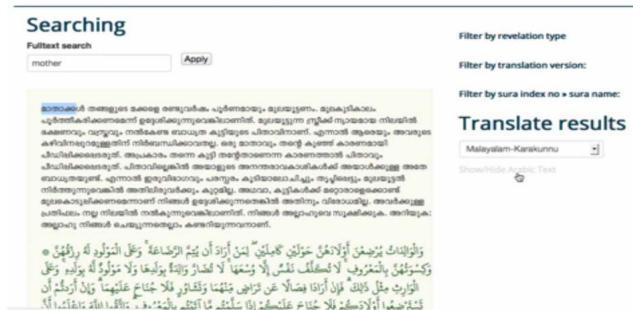


Figure 8. Search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Malayalam language

C. Result of search by Ayahs and its meaning in Urdu

Figure 9 show result of search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Urdu language.

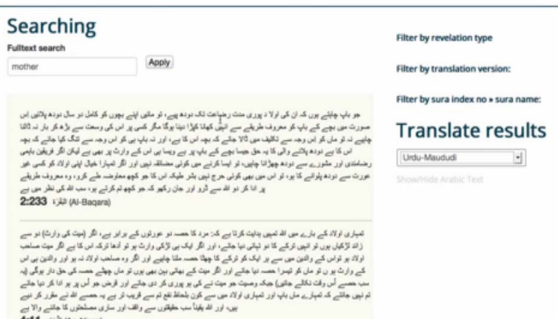


Figure 9. Search using a word in English that results in Arabic Ayahs with the meaning of the Ayahs in Urdu language

IV. CONCLUSIONS

The Qur'an Search & API mainly focus to make search of Qur'an more compelling and rich. Instead focusing on search based on roots and other such details and dependency on one language, this search application enables a unique cross language search which enables more results .its cross language search features help the user to increase their knowledge.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support provided by the IT Research Centre for Holy Qur'an (NOOR), Madina Saudi Arabia ,under research project reference number, NRC1-1, entitled, 'Developing a Complete Online Qur'an Portal and Mobile Friendly Qur'an

REFERENCES

- [1] Mohd Amin Mohd Yunus, Roziati Zainuddin, and Noorhidawati Abdullah, "Semantic Method for Query Translation", the International Arab Journal of Information Technology, Vol. 10, No. 3, May 2013.
- [2] Robiah binti Hamzah, "Visualizing Surah Al Baqarah: The New Innovation Of Reciting Al Quran", Library of Malaysian Institute Of Information Technology (MIIT), University of Kuala Lumpur Malaysia, 2000.
- [3] Qurat ul Ain and Amna Basharat, "Ontology driven Information Extraction from the Holy Qur'an related Documents", 26th IEEE Students' Seminar 2011, NUST, Pakistan.
- [4] Zameer Ahmed Adhoni, Husam Al Hamad, Abdul Ahad Siddiqi and Latifa El Mortaji, "Towards a Comprehensive Online Portal and Mobile Friendly Qur'an Application", NOORIC 1435/2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences, Madinah, Saudi Arabia, 22-25 December 2013.
- [5] Tanzil - Quran Navigator. http://tanzil.net/wiki/Tanzil_Project



Dr. Zameer Ahmed Adhoni Assistant Professor Qassim University Dept of IT ,Qassim,Saudi Arabia
B.E M.Tech ,Ph.d .His research information security



Dr. Husam Al Hamad, Assistant Professor of Information Technology at Qassim University, in Qassim, Saudi Arabia. He has got the BS, MS, and Ph.D. in the field of computer science and Information Technology in 2001, 2003, and 2008 respectively. His research interests are data Engendering, pattern recognition, image processing, artificial intelligence, and neural networks.



Abdul Ahad Siddiqi received a PhD and a MSc in Artificial Intelligence in year 1997, and 1992 respectively from University of Essex, U.K. He also holds a bachelor degree in Computer Systems Engineering from NED University of Engineering and Technology, Presently he is an Associate Professor at College of Computer Science and Engineering at Taibah University, Madina, Saudi Arabia. He has over two dozen research publications to his credit. He has received research grants from various funding agencies, notably from Pakistan Telecom, and Deanship of Research at Taibah University for research in areas of Intelligent Information Systems, Information Technology, and digital Quran applications.



Zaheer Ahmed Adhoni is a Senior Software engineer in Huawei Technologies. His area of interest is software testing.