STRATEGIES FOR COLLECTING ELECTRONIC RESOURCES ON THE QUR'ANIC RESEARCHES

Amar Arbaoui,
Yasser M. Alginahi &
Mohamed Menacer

ABSTRACT

Electronic dissemination of digital information has benefited from recent advancements in Information and Communication Technologies (ICTs) and dissemination of Islamic digital content in different formats has taken advantage of this rapid progress in technology. Information is available in many formats and on many different applications and devices. However, the online availability of such resources on Qur'an and its sciences is very limited. In this work, digital Qur'an resources available from authentic sources will be used to formulate strategies on how to extract, format, present and make these resources available for researchers in a centralized knowledgebase. The methodology starts with collecting these authentic electronic Islamic resources. Next, resources are classified into their main categories, sub-categories and related categories. Then, the list of metadata for each type of resource is developed. Following this, strategies for efficient dissemination for the different types of digital resources are developed; depending on the format of the files, the metadata is extracted automatically, if possible, otherwise data is entered manually into the knowledgebase. Therefore, the manual entry of data into the knowledgebase for resources having poor or lacking metadata would efficiently help researchers to easily locate the resources they are searching for. Developing these strategies for each type of format could efficiently help in the search process. From the data collected so far, the amount of data gathered shows that books contribute to the highest number of resources available in this area followed by post graduate dissertations, conference papers and journal papers respectively. Finally, efficient dissemination of electronic information will help researchers access, explore and disseminate quality research in Qur'anic studies.

Keywords: *Data collection, metadata, Qur'an classification, Qur'an research studies and resources.* .

INTRODUCTION

The Holy Qur'an has been and is continuing to be an inspiration to many scholars and scientists to embark on scientific research and innovation (Shaikh Mohd Salleh & Azizan binti Baharuddin, 2011). The last decade has seen vast technological advancements in digital technologies in the areas of publishing, indexing, searching and multimedia; hence, the research in the area of Qur'an and its sciences has noticed a vast increase in the last few years. With this continual increase of digital Islamic publications, ways to ease the gathering and organizing of information is very essential to researchers. In this work, strategies for data collection and data entry will be developed to aid in the dissemination of Qur'an resources.

The information on the Internet has been exponentially increasing and no doubt that Qur'anic content is also increasing. The search results of books on Qur'an from the online bookseller Amazon showed 7,852 items which include books in the following languages English, French, Spanish, German, Italian, Arabic, Hebrew, Urdu, Hindi … etc. With the largest number of books available in English then Arabic followed by few books available in other languages. In addition, a search on the keyword "Islam" produced 74089 results in January, 2013 compared to 50,436 items in 2009 (Idrees and Mahmood, 2009). This shows an increase of at least 54% in Islamic publications in a period less than 4 years.

It is clear that the scattered Qur'an resources in different formats, printed and digital, which are available from different sources make it difficult to collect data. In addition, no uniform or standardized formats are noticed in the existing information on the Internet. This makes it even difficult to gather and organize in one place. The metadata for many digital resources is not available and many files are posted as Word, PDF, different Image formats, databases…etc., which makes it even more difficult to classify and make available in a central repository. Therefore, human intervention is needed in order to help in the process of data entry to organize these resources in a proper way to ease the search process for researchers.

From the literature survey, it is noticed that the number of journals, magazines or other research resources available in a language other than Arabic is considered very low and no doubt if a search is done for other languages than English less and less number of resources would be found. In addition, there could be many other journals publishing work related to Qur'an and its sciences; however, most of such work is not noticed by many researchers especially those who do not know any other

language besides Arabic or English. Hence, most of the work on Qur'an and its sciences is not published in specialized Qur'an journals. After conducting extensive research on Islamic journals and to the best knowledge of the authors it can be concluded that many journals published in Arabic language or those languages using Arabic script such as Persian and Urdu may not appear as results on search engines. In addition, to the best knowledge of the authors, journals in other languages such as Urdu and Persian may not be indexed in specialized database in those language and only few databases in Arabic language are available or have emerged in the last few years. Consequently, many Arabic resources are available only in specific libraries where such information may not be accessible by all researchers in the area of Qur'an or Islamic studies. Similarly, Qur'an related conference proceedings may not be available in standard digital formats and/or properly disseminated which may not be easily obtainable by the majority of researchers and therefore making it difficult to gather all Qur'an related conference papers for the purpose of building a Qur'an knowledgebase. Most conferences on Qur'anic/Islamic studies do not make conference research papers available online or indexed in research databases, i.e., the distribution of research papers are mainly on CDs in a local environment. However, such culture is changing and more and more conferences are seen to be indexed especially with the emergence of Arabic research databases. For example, the database IslamicInfo (2013) provided by Dar Almandumah contains a list of 243 conferences; however, only two conferences are on Qur'an related topics.

There are many Qur'an resources, Ancient Qur'anic manuscripts, Qur'an tafsir, books related to Qur'an, conference research papers, magazines, journal papers …etc.) available in many different places around the world. However, no centralized body has decided to collect all of the information which exists mainly in paper format to be available electronically under one umbrella. Few digital resources exist on the Internet through different organizations, thus what is available is only a drop from an ocean of what could be found in many places around the world. For example, the city of Timbuktu, in Mali holds many ancient Islamic manuscripts. According to the Library of Congress, "The texts and documents included in **Islamic Manuscripts from Mali** are the products of a tradition of book production reaching back almost 1,000 years " (2012). However, these manuscripts are at risk because of war.

This paper is organized as follows: section 2 provides the literature survey, section 3 discusses library classification systems, section 4 explains the data collection process, section 5 presents the methodology

and data collection strategies for Islamic and Qur'an resources, section 6 discusses the results, and finally section 7 concludes this paper.

Literature Survey:

The main sources of information which are considered up-to-date for researchers includes: journals, magazines, conference papers, and post-graduates dissertations. The existing number of journals or conferences presenting Qur'anic research is not known, as mentioned above such information is scattered and/or not easy to find. In an attempt to search for the number of journals related to Qur'an and/or Islamic studies, the following information was collected:

• From 12214 journals listed by Thomson Reuters/ ISI Web of Science List: Science (2012), the number of journals written in English which were found by using the keywords search "Islam", "Islamic" or "Muslim" are only six. No journal is found on Qur'an and its sciences.

• In addition, the Access to Mideast and Islamic Resources (AMIR) website listed 475 titles of open access journals in Middle Eastern Studies as of January, 2013, however the number of Islamic journals is around 55 journals. The journals are published in different languages which include: Arabic, English, Turkish, Urdu, French, Korean, etc. With no specific journal on Qur'an, however, some of the listed journals may contain topics/papers on Qur'an and its related sciences.

• The king Abdulaziz Foundation for Research and Archives publishes a directory on scientific peer reviewed journals published in Saudi Arabia (2012). According to Ahal Altafsir Forum this directory contains 64 journals from which there are only 3 journals on Qur'an and all are published in Arabic.

AskZad the first and largest Arabic digital library offers an extensive referential, cultural and academic database which contains Pan-Arab Academic Journal Index (PAJI) with full Arabic language indices of more than 700 Middle Eastern university-published journals and approximately 350 organization-published journals (Arabic Digital Library, 2012).

Another main resource in the research community is dissertations published by graduate students. There are many dissertations on Qur'an and its sciences available through university libraries across the Middle East, Islamic world, Western and Eastern universities which offer programs on Arabic and/or Islamic fields, however, many of them are not

accessible to all researchers. Thus, since many of these dissertations are not in digital format it is difficult for a researcher to visit universities in order to get the reference they are looking for especially if inter-library loan services are not available through those universities. Therefore, the AskZad database provides the Pan-Arab Dissertations (PAD) index which contains almost 7000 dissertations published by graduate students in the Middle East in any language. Currently, in Saudi Arabia universities are converting all dissertations to digital format in an attempt to make resources available to all researchers from all around the world.

With regard to conferences, forums, symposiums and workshops on Qur'an and its sciences, the last few years have seen a surge in their numbers. Even though the number of such events is increasing, more conferences are needed to address research and studies in the different sciences of the Holy Qur'an in order to raise awareness on the importance of such areas of research such as: Qur'an recitations, miracles of the Qur'an, Tafsir, information technology (IT) and Qur'an, … etc.

IT has influenced many things in our lives and IT resources for Qur'an has seen an increase in the last few years, especially with the increase use of smart phones and tablets. In a survey on mobile Qur'an applications by Khan and Alginahi, as of December 2012, 209 smart phone apps were found; however, since the main source of information is the Internet some applications may have been missed and/or developed before 2005, making it difficult to have a complete comprehensive data on all software/applications developed for the Qur'an. This survey showed that there are many Qur'an apps in different Operating Systems and formats with about 15% of them designed in different languages other than Arabic or English. The data collected shows that these apps are regularly updated and that many have several versions (Khan & Alginahi, 2013). Thus, this opens the door for new areas of research on Qur'an and its sciences where technology could be involved, such areas could be: the effect of technology on dissemination of Qur'an software, the authentication of software applications, the expansion of Qur'an and its sciences classification by adding IT related fields, … etc.   Finally, Table 1 shows the update year for the apps surveyed.

Table: 1: Mobile Apps Update Timetable (Khan & Alginahi, 2013)

Number of Apps	Update year
150	2012
35	2011

16	2010
3	2009
1	2008
2	2007
1	2006
1	2005

In conclusion, the research in the area of Qur'an and its sciences has seen an interest to many researchers from all over the world. Hence, the need to make resources available under one umbrella is becoming vital with more online visibility.

Library Classification Systems:

Library classification systems are used to catalog resources such as books, periodicals, films … etc. The two main standard library classification systems available and widely used are the Dewey Decimal System (DDS) and the Library of Congress (LOC) Classification System. In addition, there are other classification systems which are developed for specific fields and/or organizations such as the Colon Classification, Harvard-Yenching Classification: An English classification system for Chinese language materials and V-LIB 1.2, this is just to name a few. In addition, there are other universal classification systems in other languages such as: New Classification Scheme for Chinese Libraries, Nippon Decimal Classification (NDC), Chinese Library Classification (CLC), Korean Decimal Classification (KDC) and Library-Bibliographic Classification (BBK) from Russia (Decimal System vs. Library of Congress, 2012).

The DDS was developed in the second half of the 19th century as a library cataloging system to organize all knowledge which relies on a simple framework that starts with ten subject classes: religion, sciences, etc. These classes are then broken down into ten divisions, which are then broken down into ten subdivisions. Resources are assigned numeric call numbers based on where content within them falls in this taxonomy of knowledge. On the other hand, the LOC classification system which was developed at the turn of the 20th century differs in its design from the DDS. It was created to categorize books and other items held in the Library of Congress. It features 21 subject categories with resources being identified by a combination of both letters and numbers. The

number of categorization classes is not restricted, nor are the numerous subclasses included in the system (Decimal System vs. Library of Congress, 2012).

Each system has its shortcoming, for example, since the DDS was developed in the 19th century it may not be able to add new fields such as Computers since it was not accounted for under the ten subject category headings. While the system has been updated over time, a closed taxonomy has forced computers and other technology topics to be shoehorned into a category labeled 'General.' However, the LOC has 'Technology' as a subject heading. Consequently, "due to the flaws seen in the standard library classification systems, librarians think that libraries should follow practices which are best to their respective collections". "While some librarians and other bibliophiles have a strong preference for either Dewey or the LOC system, many others concede that both systems have flaws and that libraries should follow practices that are best for their respective collections. Many public libraries, for example, continue to use Dewey while some academic libraries have made the switch to LOC to allow for greater specialization in identifying resources (Decimal System vs. Library of Congress, 2012)."

Most of the libraries in the Islamic World use either the DDS or the LOC classification system; however, there are some organizations which developed classification schemes designed to suite their own collection. For example, the Center of Studies and Qur'anic Information at the Institute of Imam Alshatibi in Jeddah, Saudi Arabia, devised a classification system based on the Dewey system for Qur'anic Studies which includes five main divisions from which different sub-divisions are branched and so on (Database for Qur'anic Information Resources, 2012). In another example, from a visit to the Imam Ibn Alqayim Library in Riyadh, Saudi Arabic the Chief Librarian informed the authors that the library devised its own classification system which suites its needs since all its books are on Islamic and Qur'anic Studies.

On the other hand, it is also noticed that websites or search engines such as Google (Categories of Topics, 2013), Yahoo (Computers and Internet, 2013), etc., use their own classification/directory scheme which suites the way their information is organized.

In the work of Idrees, 2012, the author concluded that neither the standard classification systems, nor indigenous expansions or schemes are fulfilling the purpose for classifying Islamic resources. In response to shortcomings of the standard systems, different practices have been adopted. Organizations/libraries have developed their own systems

without following or developing any standards, e.g., International Islamic University, Islamabad, (Idrees, 2007). In other cases some organizations have developed expansions in the standard systems (Database for Qur'anic Information Resources, 2012). "Efforts were made to get such expansions formally incorporated in the original schemes, but, such efforts could not succeed. Subsequently, there have been very different approaches in the expansions of even same standard systems and no uniformity is found in this regard. Thus, the same kind of knowledge could be seen organized differently at different places (Idrees, 2012)." Also, from a visit to the International Islamic University of Malaysia in March 2013 the Chief Librarian stated that the library is in the process of devising its own Islamic classification extension to the LOC which will be available by the end of 2013 or early 2014. Idrees, 2012, proposes the development of a new, independent and comprehensive system that covers all the related and possible aspects of Islamic knowledge and the materials being produced on the associated topics.

 For IT and computing classification systems the Association for Computing Machinery (ACM), developed the 2012 ACM Computing Classification System (CCS) which replaces its traditional 1998 version of the ACM-CCS. It is being integrated into the search capabilities and visual topic displays of the ACM Digital Library. It reflects the state of the art of the computing discipline and is receptive to structural change as it evolves in the future.   Since, IT is noticed to be entering all disciplines a modified system has to be developed for Qur'an related IT classification.

In conclusion, due to the many classification schemes universally used in libraries or specific to organizations and websites, and the unavailability of a universal Arabic and/or Islamic classification system, the best way to classify Islamic research material would be to devise a new comprehensive expandable classification system which will allow the inclusion of all Qur'anic, Islamic and related IT resources. The current standard library classification systems and ACM-CCS could be used as the reference to design such an expandable Islamic library classification system compatible with international standards.   However, this is out of the scope of this work and the classification used in the process of data entry of gathered data was based on classification according to the main categories, sub-categories and related categories which were related to the different sciences of the Qur'an.

Data Collection

The data collection of resources on Qur'an and its sciences were mainly collected from authentic Islamic websites and visits to some organizations. Lists of different websites such as university libraries, Islamic/Qur'anic centers, Islamic/Qur'anic conferences, Islamic/Qur'anic journals, … etc., in addition visits were made to different organizations. In this work, the authors visited different organizations located in Saudi Arabia. The purpose of the visits was to collect as much research resources as possible on Qur'an and its sciences. Table 2 shows the list of organizations visited by the authors.

Table 2: List of organizations visited in Saudi Arabia

Organizations visited	City
Islamic University Central Library	Madinah
Taibah University Central Library	Madinah
The agency for the Muslim World League	Makkah
Umm Alqura University Central Library	Makkah
The custodian of the two holy mosques institute for	Makkah
Institute of Islamic research and revival of Islamic	Makkah
Imam Ibn Alqayim Library	Riyadh
King Fahd National Library	Riyadh
The Saudi Society for Qur'an and its Sciences, Imam Mohammed Bin Saud University	Riyadh
"www.islamhouse.com" project, Riyadh	Riyadh
Institute of Imam Alshatibi, Jeddah	Jeddah

From the lists of websites collected and organizations visited, it is noticed that the information collected comes in different formats and file sizes which cause difficulties during the data entry process. Therefore, a strategy is highly recommended to organize the data collected during the data entry process. It is noticed that most of the data gathered either from Internet resources or visiting different organizations in Saudi Arabic do not provide the metadata for the files posted on the Internet or distributed on CDs. This is because of the fact that the information is not organized in databases and therefore the metadata available is either minimal or nil. The aim is not only to collect digital resources related to Qur'an and its sciences, but also to classify and formulate strategies to ease the data collection and data entry which are the main stages of this project. Data collection is not that difficult if the essential means,

strategies, are set to guide the collector on how to deal with different formats of information. No doubt that with the existence of large number of Islamic organizations and the availability of enormous information on the web it is not that easy to find the information needed from one source. In addition, each organization may have its information in different formats than others which definitely may cause some problems for people who may not have the suitable software to deal with such file formats.

With the Internet being the main source of data collection, clearly the metadata is the essential requirement during a search process, otherwise, it may be difficult to search for any given item. Metadata provides information about the content of digital documents, files, such as text, images, audio, video … etc.

The objective of producing metadata for each Islamic resource available in digital format helps researchers identify and search for items very easily. The main objectives of metadata (Understanding Metadata, 2004) are:

1- Resource discovery.
2- Organizing e-resources.
3- Digital Identification – by using ISBN, ISSN, file name, URL, DOI (Digital Object Identifier)
4- Archiving and preservation in order to track the resources and their physical characteristics

For example, a text document metadata may contain the following data: title, author, date, size of document (no. of pages, no. of words.. etc.), abstract or summary, … etc. The example in Figure 1 below shows a case where there is a minimal metadata for a book available on a library website. Figure 1(a) shows the list of books available, 3 books only. Here, the title of the book and type of the resource are provided. Clicking on the download arrow the download process starts, then the open file window appears, Figure 1(b). Finally, when clicking on open file a window comes with the message: "File extension is unknown," Figure 1(c). This is just an example of many Internet links which cannot be accessed and therefore such information may not be available from another source.

Figures 2(a) and (b) show two examples of complete metadata for book search. Figure 2(a), (English Resources, 2013), is obtained by clicking on the books section of the website, then choosing the book of interest the detail is shown providing the metadata available on the book. Similarly, the book details in Figure 2(b) were obtained in the same way (Directory

of English Publications, 2013). This shows that the metadata is available through the organization and it may be difficult to obtain it from them to be migrated to another database or indexing system since they are the sole owners of the information. The purpose of this work is to provide researchers and students with the means they need in the research process.
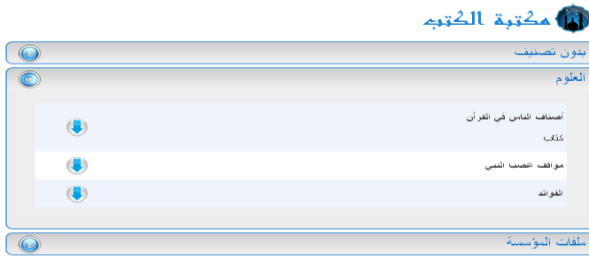
(a)A limited list of books available on a website.
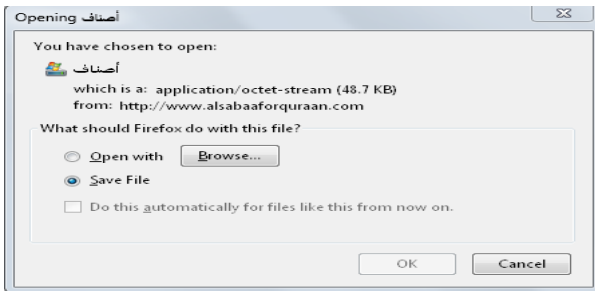
(b) After downloading file, option is given to open it
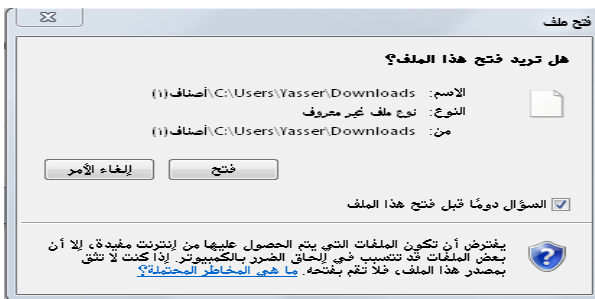
(c) A message with "unknow file extension"

Figure 1: An example of a Resource with minimal metadata (Library Books, 2013)

(a)Example with complete metadata

(b) Example with complete metadata

Figure 2: Examples of websites providing complete metadata for resources available

Therefore, in order to gather metadata on resources available from different organizations collaboration is required in order to obtain such information. In doing so, the metadata could be entered by the owner of information automatically with a written script provided to them in order to migrate the fields available on their database into the proposed knowledgebase. Otherwise, data has to be entered semi-automatically by cutting/pasting the data available into a data entry form, with referencing the source by providing the link to the source from where the information was obtained.

Another example of a database is provided by Umm-Alqura University, Makkah, Saudi Arabia. This database provides detail on Master and Doctorate dissertations. The database search window provides the ability to search by a keyword in the title or the name of the author, student, Figure 3(a). Following this, a list of the number of results with the titles of the dissertations and names of authors is provided, from this the specific item is chosen, Figure (b) shows an example of a chosen item, the complete metadata is available; however each metadata is provided in terms of numbers. These numbers mean nothing to the investigator,

researcher, unless the corresponding information associated with it is available.

In other cases, the information is available but not organized in a simple way, i.e. to ease the search process. In other words the metadata is available however, it is not organized in a format that makes it easy to find.

(a)   Search window.

(b)   A sample from the search on the topic of Qur'an

Figure 3: Example of Thesis database with metadata

Methodology for Data Collection Strategies for Islamic and Qur'anic Resources

The methodology implemented follows the following steps which will be explained in details below: In the initial stage of this project, data was gathered from different sources (Internet, libraries, visiting some organizations … etc.) and in different formats, printed and digital. Then the different formats of the collected material were studied to formulate a strategy for the data collection and data entry stages of the project. The

strategy formulated for the data collection and data entry starts by first studying the overall content of data in order to classify the material into their main categories. Here, the data collected is studied and depending on the file titles and content it is categorized into main, sub and related categories. The type of resources available can be any of the following: books, journals/magazines, conferences/workshops/forums and dissertations.   These are the most essential resources needed by researchers. Figure 4 presents the most essential data resources needed by researchers.
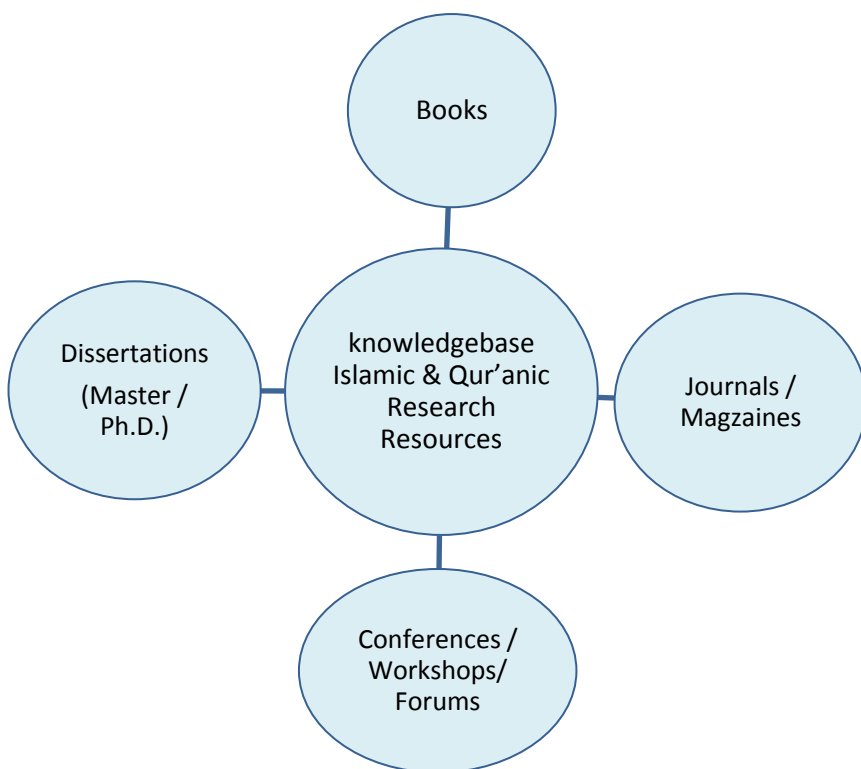

Figure 4: The most essential resources needed by researchers-

Then, the different file formats are separated according to these classifications. For example, let's assume the files are organized as follows: first we start with a folder which contains all the files available, then for each category we have different sub-category (sub-folders), next each sub-category folder contains five folders containing the three different formats found from the data collected, these are: text files

(word, pdf, txt, …. etc.), pdf created from images such as ancient manuscript files and files with different image formats, database files (simple or complex), audio and video files. Ancient manuscripts, audio and video files are not the target of the project at this stage. The databases sub-folder may contain two folders due to the fact that some databases could be simple not containing enough metadata and others could be complex, rich in metadata. The Proposed organization structure of data is shown in Figure 5.
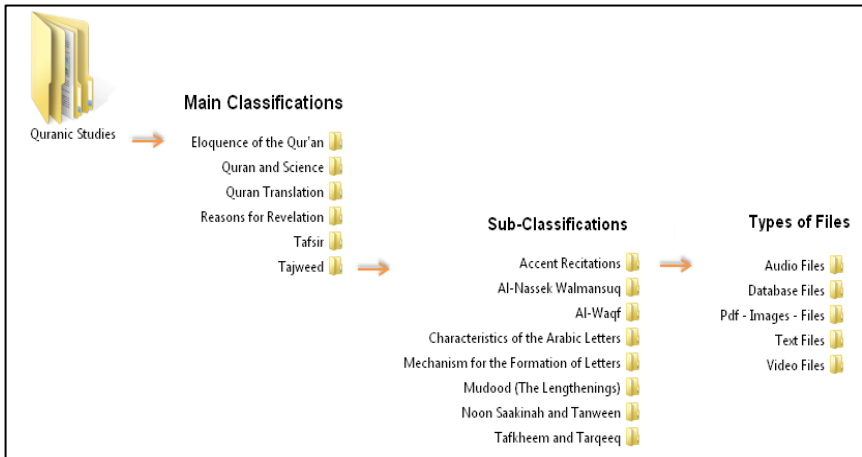

Figure 5: Organization Structure of Files

The main metadata fields needed from the different types of resources are specified in Table 3. From the table, it is very clear that some metadata fields may not be available for all resources, for example, books do not contain ISSN numbers, issue number … etc. Finally, depending on the format (printed or digital) the data entry process to create the metadata for each resource (book, article, research paper, dissertations, ….etc.) is divided into automatic, semi-automatic or manual. The data entry process for paper formats, text files (word, txt, … etc.), pdf created from images, simple database files and rich database files are processed manually, semi-automatic, manually, semi-automatic and automatic respectively.

Table 3: List of Metadata Required for the Different Types of Digital Resources

Books	**Journal Papers**	**Conference Papers**	**Theses**
Resource Type	Resource Type	Resource Type	Resource Type
Reference ID - **Book**	Reference ID – **Paper**	Reference ID – **Paper**	Reference ID - **Thesis**
Title	Paper Title	Paper Title	Thesis Title
Author	Author	Author	Author
Main Category	Main Category	Main Category	Main Category
Sub-Category	Sub-Category	Sub-Category	Sub-Category (Specialization)
Related Category	Related Category	Related Category	Related Category
Keywords	Keywords	Keywords	Keywords
Language	Language	Language	Language
Year of Publication	Year of Publication	Year of Publication	Year of Publication
Source	Source	Source	Source
Source URL	Source URL	Source URL	Source URL
Description (Abstract)	Description (Abstract)	Description (Abstract)	Description (Abstract)
File ID	File ID	File ID	File ID
Cover page	Cover page	Cover page	Cover page
Extra_data	Extra_data	Extra_data	Extra_data
Source_comments	Source_comments	Source_comments	Source_comments
Publisher Name	Publisher	Publisher	University Name
Place of Publishing	Publisher's Address	Pages	College Name
No. of Pages	Pages	Volume No.	Dept. Name

No. of Volumes	Volume No.	ISSN	Degree (MSc./Ph.D.)
Edition	Issue No.	E-ISSN	Supervisor's Name
Series Name	Country	Conference Name	
ISBN	ISSN (Printed)		
	E-ISSN (Electronic)		
	Journal Name		

Finally, in designing the knowledgebase metadata is considered the essential requirement for all resources available for entry into the knowledgebase. The burden of entering metadata should be reduced and only essential (minimum) fields for the different types of resources should be entered. The design of the data entry form should be friendly, easy to use, and should include all types of fields which could be needed for any type of resource. It should use drop-down lists to choose from, ensures meaningful field names and provided help text where needed.

Results and Discussion

In this stage of the project, most of the data collection so far has been done in Saudi Arabia where the data collection process started by visiting some organizations and collecting information from Islamic websites. The resources received/gathered were in different formats: Sample books, booklets, CDs, directories (book format), large amount of files with different formats copied on external hard disks … etc. As explained in the methodology section of this paper for different data formats different data entry approaches were used. Data entry interfaces were designed for each type of resource to help in the manual and semi-automated data entry process, Figure 6 shows a screenshot of a portion of the interface for books. The data entered is organized in a database; Figure 7 shows an example of few items saved in the database. In this work, the data entry process has just started with limited entries that require more human resources. So far, the intermediate results are shown in Table 4, presenting the amount of data collected. From approximately 737

sources, websites/journals/conferences/organizations (libraries, universities, societies, etc.), mainly from Saudi institutions and organizations, as of June, 2013, the number of validated resources (items) entered into the knowledgebase is 47,233, as shown in Table 4. Meanwhile, the collection process is still progressing according to the plan setup for this project.

On the other hand, other alternatives are being investigated depending on the data formats being collected since too many resources are not properly cataloged, classified or disseminated or/and contain poor amount of metadata which makes searching for Islamic and Qur'anic resources difficult to reach at best and lost at worst.

Table 4: Quran resources collected so far

No.	Resource Name	No. of items collected
1	Books	38,354
2	Theses	5,273
3	Conference Papers	2,943
4	Journal Papers	663
	Total	47,233

The metadata field: main category, sub-category and related category fields link all resources together in the knowledgebase. With such a link different resources could be located according to the categories. In addition, the metadata contains a field named extra data which may contain comments, sub-categories and other information which may be suitable to build the knowledgebase.

Figure 6: Manual Data Entry Interface for books

Figure 7: Sample of data organization in the database after using the manual data entry form.

Conclusion

This work so far concentrated on the collection of authentic resources for Qur'an and its related sciences, as well as, the design of a dedicated data entry system. The process of data collection and data entry was found to be a very tedious process, but a very important one to build a proper foundation for efficient collection, classification and dissemination of research findings and results on Qur'an and its related sciences. Therefore, data classification has been designed to facilitate the integration of such important resources into a knowledgebase. It can be expanded as necessary, compared to other classification systems which are not efficient and flexible enough for Qur'an related resources, hence, a dedicated and expandable classification for Qur'an and its sciences. Finally, this work contributes an attempt to organize, classify and catalog, as well as, raise awareness about the importance of developing unified classification/standards for related Islamic and Qur'an research resources across the Arab Muslim world.

Acknowledgement

References:

Alphabetical List of Open Access Journals in Middle Eastern Studies. (2013). Access to Mideast and Islamic Resources (AMIR). Retrieved from
http://amirmideast.blogspot.com/search?q=islamic+studies.
Peer Reviewed Journals Published in Saudi Arabia. (2012). Ahal Altafsir Forum. Retrieved from http://www.tafsir.net/vb/tafsir24164/.
Library Books. (2013). Alsabaa International Establishment. Retrieved from http://www.alsabaaforquraan.com/Library.aspx?libID=19.
Arabic Digital Library. (2012). Askzad. Retrieved from
  http://askzad.com/e_genpages/AboutUS.aspx.
The 2012 ACM Computing Classification System. (2012). Association for Computing Machinery. Retrieved from
http://www.acm.org/about/class/2012.
IslamicInfo database. (2013). Dar Almandumah. Retrieved from
  http://www.mandumah.com/islamicinfoconf.

Decimal System vs. Library of Congress. (2010). Education Insider. Retrieved from

http://education-portal.com/articles/Dewey_Decimal_System_vs_Library_of_Congress_Whats_the_Difference.html

Categories of Topics. (2013). Gooogle Guide. Retrieved from http://www.googleguide.com/directory.html.

Haroon, I. (2012). Organization of Islamic Knowledge in Libraries: The Role of Classification Systems, Library Philosophy and Practice - Libraries at University of Nebraska-Lincoln, pp. 1 – 14, Retrieved from http://digitalcommons.unl.edu/libphilprac/.

Haroon, I. (2007). User Relationship Management, Dr. Muhammad Hamidullah Library, Islamic Research Institute. *Pakistan Library & Information Science Journal*, 38(3), pp. 25-31.

Haroon, I and Khalid, M. (2009). Devising a Classification Scheme for Islam: Opinions of LIS and Islamic Studies Scholars. *Library Philosophy and Practice (e-journal) - Libraries at University of Nebraska-Lincoln*, pp. 1 – 15.

Database for Qur'anic Information Resources. (2012). Imam Alshatibi Institute. Retrieved from http://www.Qur'an-c.com/.

English Resources. (2013). Islam House. Retrieved from http://www.islamhouse.com/s/9661.

Directory of English Publications. (2013). Islamic Research and Training Institute. Retrieved from

http://www.irtipms.org/PubDetE.asp?pub=54.

List of Books. (2012). King Abdulaziz Foundation for Research and Archives. Retrieved from http://www.darah.info/WebTrBooks.aspx.

Islamic Manuscripts from Mali. (2012). Library of Congress. Retrieved from http://international.loc.gov/intldl/malihtml/islam.html.

Muhammad, K. K & Yasser, M. A. (2013). The Holy Qur'an Digitization: Challenges and Concerns. *Life Science Journal*. 10(2), pp. 156-164.

Understanding Metadata. (2004). National Information Standard Organization. pp. 1 - 16. Retrieved from,

http://www.niso.org/publications/press/UnderstandingMetadata.pdf.

Shaikh M. S. & Azizan B. B. (2011). Significance of Science and Scientific Thought from the Islamic Perspective, *International Journal on Qur'anic Research*, Vol. 1, No. 1. pp. 73 – 87.

Thomson Reuters-ISI Web of Science List: Science. (2012). Support Services. Retrieved from http://supportservices.ufs.ac.za/dl/userfiles/Documents/00002/1828_eng.pdf,

Center for Scientific Research and Islamic Heritage. (2013). Umm Alqura University. Makkah, Saudi Arabia. Retrieved from http://uqu.edu.sa/isr/islamic_culture_cen.html

Library Classification. (2012). Wikipedia. Retrieved from http://en.wikipedia.org/wiki/Library_classification.

Computers and Internet. (2013). Yahoo directory. Retrieved from http://dir.yahoo.com/computers_and_internet/

About the Authors:

Amar Arbaoui, an Associate Professor at Taibah University, College of Computer Science and Engineering, Madinah, Saudi Arabia. Email: aarbaoui@taibahu.edu.sa

Yasser M. Alginahi, an Assistant Professor at the Dept. of Computer Science, College of Computer Science and Engineering, Taibah University. Email: yginahi@taibahu.edu.sa

Mohamed Menacer, an Assistant Professor at Taibah University, College of Computer Science and Engineering, Madinah, Saudi Arabia. Email: mmenacer@taibahu.edu.sa