

# Applying FRBR model to bibliographic works on Al-Quran

Yushiana Mansor\* and Sharifah Nur Amirah Sarif Abdullah

Department of Library and Information Science,

Kulliyyah of ICT, International Islamic University Malaysia,

P.O. Box 10, 50728 Kuala Lumpur, MALAYSIA

e-mail: \*yushiana@iium.edu.my (corresponding author); shnuramirah@iium.edu.my

## ABSTRACT

*This study explores the feasibility of applying the object-oriented Functional Requirements for Bibliographic Records (FRBR) model to MARC-based bibliographic records on the Al-Quran. Based on the content analysis of 127 MARC-based bibliographic records on Al-Quran from the International Islamic University Malaysia (IIUM) OPAC system, this paper reports on the process of mapping FRBR entities to set of works on Al-Quran. The attributes of the bibliographic works in the MARC records were identified and grouped according to the FRBR entities. The findings suggest that, overall, most of the MARC-based bibliographic records on Al-Quran were sufficient to represent the FRBR model. However, several issues were identified as affecting the process of creating entity-relationship model for "FRBRizing" bibliographic works on Al-Quran. These include inconsistencies in romanizing records in Arabic scripts, difficulties in identifying complex works, missing fields for subject headings, and missing fields for record-object relationship identification. Thus, a major conclusion drawn is that the quality of MARC records is an important aspect in ensuring the bibliographic records are having complete, correct, and reliable data for FRBRization process.*

**Keywords:** Functional Requirements for Bibliographic Records (FRBR); Cataloguing; OPAC; MARC (Machine Readable Catalogue); Al-Quran.

## INTRODUCTION

Cataloguing has been referred traditionally to as the process of describing collections in library and information institutions. A catalogue can be defined as a set of bibliographic records, where each record acts as a surrogate for a publication (Monch and Aalberg 2003). It is the main tool that enables users to search for items in a library using specific fields such as author, title, or subject. As the world moves into the era of digital libraries, metadata analysis has been identified as a new technique for dealing with networked resources. In order to clarify the process of metadata analysis, conceptual models such as the Functional Requirements for Bibliographic Records (FRBR) model was developed by the International Federation of Library Associations and Institutions (IFLA) in 1998. According to Mangunhas, Freire, and Borbinha (2010), the FRBR model offers a single container for entities that are shared among several bibliographic records (e.g. two works created by the same author, sharing the same subject, published by the same editor, etc.), which builds up a graph of entities. The FRBR model uses an entity-relationship model of metadata for information objects, instead of the single flat record concept underlying current cataloguing standards.

One of the great advantages of the FRBR-based OPAC displays as highlighted by Carlyle and Fusco (2006) is that long displays may be made much shorter, thereby enhancing the

intelligibility and browsability of results. Much research has been done to improve the current performance of online library catalogues or similar bibliographic databases. The Library of Congress has developed a FRBR mapping and display tool based on the MARC 21 structure (<http://www.loc.gov/marc/marc-functional-analysis/tool.html>). FRBRized OPAC display demonstrates better collocation of records as FRBR emphasizes relationships between bibliographic entities described in catalogues and organizes these entities according to logical hierarchies. Thus, FRBR potentially extends the scope of a single record to include information about multiple parts, multiple formats, publication and distribution, and creators and contributors in a logical way which in turn would allow users to find, identify, select and obtain information in one place.

The concern from which this study derives is to investigate the feasibility of applying the FRBR model to existing bibliographic database which is non-FRBR compliant, having noted the general acceptability of the model internationally. The FRBR model was introduced to allow for better arrangement, collocation, and navigation of bibliographic databases. This is being achieved by adapting FRBR to simplify record retrieval and aggregate result sets into manageable clusters. As such, testing the feasibility of implementing FRBR in a large bibliographic catalogue, and exploring how FRBR can assist in the integration of the traditional catalogue with the web environment makes it more relevant to library users. Previous research in FRBR has looked at several distinct datasets such as literature, fiction, Bible, and special formats. Not much has been done to understand the attributes of works on Al-Quran from the perspective of object-oriented model. It is hoped that FRBRized catalogues on Al-Quran will allow users to explore library holdings in more sensible groupings, with clusters of similar materials together. As most libraries in the world are holding MARC-based bibliographic records, mapping to the FRBR model will pose several challenges, some of which this study had revealed. Findings from this study will contribute to research on enhancing the retrieval of works on Al-Quran in online catalogues. Also, better understanding of the attributes of works on Al-Quran through the FRBR model will assist in improved design of online catalogues that support FRBR user-tasks, i.e. finding, identifying, selecting, and obtaining works on Al-Quran. Furthermore, findings from this study will assist in assessing how amenable the FRBR model is to bibliographic works on Al-Quran.

## **LITERATURE REVIEW**

Dickey (2008) summarizes the benefits of FRBR to the next generation library catalogues in various types of collections. It was illustrated that FRBR offers a display option in a revamped OPAC that is simpler than current result lists, and more elegant in its reflection of relatedness among items. Hence, each feature will better enable the users of the catalogue to find, select, and obtain appropriate resources, and will bring libraries into the next generation of cataloguing practice. Weng and Mi (2006) proposed how the FRBR model can be used to improve the access and discovery of digital cultural collections from the Digital Library Federation member sites in the USA. The study explores how the FRBR model can improve collocation of works and resolve major issues relating to public access to digital content. The findings show that applying the FRBR model to digital catalogues will help narrow down the search results and collocate related digital objects, thus resulting in an easier retrieval process. FRBR also supports manifestation level in which items in the same entities can be collocated under the same manifestations and automatically organize complex work with many manifestations into simple and clear format especially in books,

online video and literature collections (Greenberg, Trujillo and Mayer 2012). Chen and Chen (2004) shared how FRBR can be deployed as a logical framework for proceeding metadata analysis and developing metadata format based on their case study at the National Palace Museum of Taiwan. FRBR has been proven by the authors to be a useful and fundamental framework, particularly to media-centric and association rich contents. However, authors such as Beall (2006) had questioned the applicability of FRBR. According to him, FRBR has only been tested on selected datasets of larger size, such as works of Shakespeare, Bible, Hamlet, and Humphrey Clinker. Pietras and Robinson (2012) addressed the difficulties to apply FRBR model to serials collections due to difficulties and complexity of serials bibliographic records. Thus, it is still not convincing how FRBR can benefit smaller size bibliographic databases. He further argues that the interoperability of MARC records and crosswalk of MARC data to conform to FRBR model is another aspect that needs further examination. Aalberg and Zumer (2013) had looked at the limitations of existing MARC-based bibliographic records in expressing FRBR model. Consistent cataloguing practice and complete MARC records were found to be important in ensuring successful FRBRization.

## **OBJECTIVES AND METHOD**

This study aims at examining the issues associated with the conversion of a set of MARC-based bibliographic records of works on the Al-Quran to conform to FRBR model. In view of this, the underlying questions to be investigated are:

- a) What are the issues associated with creating an entity-relationship model for FRBRizing MARC-based bibliographic records on Al-Quran?
- b) To what extent is the information available in the bibliographic records sufficient to reliably identify the FRBR entities?

The procedure of mapping MARC-based bibliographic records to FRBR model is as described below:

### **Phase 1: Selection of Bibliographic Records**

A sample of 127 MARC-based bibliographic records of works on Al-Quran was selected from the International Islamic University (IIUM) library's OPAC. Since the library houses huge collections on Al-Quran, the initial observation by the researchers was difficulties in clustering the relevant records, thus the selection of the sample was done based on uniform title governed by the Library of Congress (LOC), that is, 'Koran'. The decision was made based on the assumptions that search query with 'Koran' would result into a smaller size of dataset, manageable for an academic research. It was also assumed that these records were more likely prepared by non-local catalogers, thus could provide insights into the issues of merging MARC records from various sources into local repository. All the samples were identified according to their work clusters and each record of the sample were examined through FRBR major entities. The sample of elements sets in the chosen MARC bibliographic records on Al-Quran were then identified in order to determine the different expressions and manifestations according to FRBR model.

### **Phase 2: Analysis of the Records**

- a) Identifying the Entities

The process of identifying entities in this study includes inspecting a MARC record of an item (or work) on Al-Quran to determine the entity being described in the record and to

find out the role this entity has in relationship with others. Each of the 127 records was analyzed to generate an understanding of the detailed characteristics of all the elements in MARC records on Al-Quran that will be further converted into FRBR model. The MARC fields reflect specific FRBR entities and the role of these entities. All the 127 records of Al-Quran in MARC format were assigned to the specific entities according to the FRBR conceptual model. The entities themselves were sorted into 3 groups as shown below:

1. Group 1 entity is a product of intellectual and artistic endeavour that are named or described in bibliographic records: Work, Expression, Manifestation, and Item (WEMI).
2. Group 2 are the entities responsible for the intellectual or artistic content, the physical production and dissemination or the custodianship of such products: person and corporate body.
3. Group 3 are the entities that serve as the subjects of intellectual or artistic endeavour: concept, object, event, place, and any of the Group 1 or Group 2 entities.

b) Establishing the Relationship

The interpretation of relationships between entities can either be based on the implicit roles of the entities occurring in a record, or it can be based on explicit information about roles and relationships found in indicators, relator codes or field linking subfields. This process was done to identify the kinds of relationships that may exist between entities. For each kind of relationship it is necessary to know the conditions for when a relationship can be identified as well as the conditions for determining what target entity the relationship points to. To achieve a final set of interrelated entities with a consistent set of relationships between all entities in the whole collection, the output from the previous interpretation was normalized.

### **Phase 3: Mapping MARC to FRBR**

In this phase, issues and challenges of mapping MARC based bibliographic records on Al-Quran were identified. This includes the problems related to the process of converting the records in manual method. Moreover, three classes of work according to FRBR model were assigned to each of the records. The types of the work were based on these three categories: Elemental work, Simple work and Complex work. The "Elemental" work is a work with a single expression and a single manifestation. "Simple" work means a work with a single expression but multiple manifestations, and "Complex" work describes a work with multiple expressions and multiple manifestations.

## **RESULTS**

### **Expression of Works**

Information on the language of work is important in order to determine the types of expression of work according to the FRBR model. The data element that represents language in MARC is tag 008.

Table 1 presents the distribution of languages in the MARC records on Al-Quran. Of the 127 records analyzed, 3.9 % (5) were in Arabic; 63 % (80) were in English; 7.9 % (10) were written in French language, 15% (19) were in German language; 3.1 (4) were written in

Latin language; 2.4 % (3) were from Russian and Turkish languages; and the remaining 2.4 % (1) were in Polish, Spanish and Tagalog languages.

Table 1: Expression of Works

Language	Frequency	Percent (%)
Arabic	5	3.9
English	80	63.0
French	10	7.9
German	19	15.0
Latin	4	3.1
Russian	3	2.4
Turkish	3	2.4
Polish	1	0.8
Spanish	1	0.8
Tagalog	1	0.8
<b>Total</b>	<b>127</b>	<b>100.0</b>

Table 2 illustrates the selected examples of language used in the existing MARC records on Al-Quran. The 008 identifier represents fixed-length fields with various kinds of coded information, and there are also specific data elements being defined according to the data position. The character position (000) provides coded information about the record as a language identifier of the item being catalogued. These coded data elements were potentially useful for retrieval and data management purposes.

Table 2: Identifying Language of Works

Language	Language Identification	Examples
Arabic	008 050804s1996 ua g 000 ara d	مختصر دستور الأخلاق في القرآن
French	008 960611s1990 fr b a001 0 fre d	Le Coran : essai de traduction de l'arabe annoté et suivi d'une étude exegetique
German	008 070925s1998 gw a001 0 ger u	Die Bedeutung de Qur'ans
Latin	008 800904t1893 gw000 0 lat c	Chrestomathia Qorani Arabica

### Manifestation of Works

Form of work in the MARC records was important in confirming the manifestation entities in the FRBR model. Table 3 summarizes the form of work distribution of the records. Findings show that 0.8% (1) of the MARC records was in a form of microform; 0.8% (1) was computer software; 0.8% (1) was a manuscript; and 97.6% (124) were books. It was observed that works reflected in the form of books were dominant among the records analyzed in this study.

Table 4 presents some of the examples for form of works in the existing MARC records on Al-Quran. This medium of works was detected in the tag 245, subfield (h), and tag 630, subfield (x).

Table 3 : Form of Works

Form of Work	Frequency	Percent (%)
Microform	1	0.8
Computer Software	1	0.8
Manuscript	1	0.8
Book	124	97.6
<b>Total</b>	<b>127</b>	<b>100.0</b>

Table 4 : Examples of Forms of Works

ID	Form of Work	Examples
027	Microform	<b>245</b> 14 \$aThe Koran. Daghestan, XIV-XVI centuries \$h <b>[microform]</b>
042	Computer File	<b>245</b> 03 \$aal-Qur'an al-Karim\$h <b>[computer file]</b> = The Holy Qur'an = Al-Qur'an yang suci = Kur'an-iKerimprogrami = Le Saint Coran = Der Heilige Koran
089	Manuscript	<b>630</b> 0 \$aKoran\$x <b>Manuscripts</b> \$xCatalogs

### Types of Works

Table 5 exhibits the types of work in the MARC records. Information on type of work is crucial in explaining multiple expressions and multiple manifestations. The result illustrates that most of the works on Al-Quran (81.9%) were elemental; 15.7% (20) were simple, and the remaining 2.4% (3) were complex works. Instances of every type of work in the MARC record is shown in Table 6.

### Subjects of Works

The subject added entry – uniform title for every MARC record was also observed in this study. Subject added entry – uniform title was important to find out the concepts and the aboutness of the work. The element that represents subject added entry – uniform title according to the MARC convention was tag 630.

Table 7 presents the distribution of subject added entry – uniform title in the MARC records on Al-Quran. As shown in the table, 15% (19) were translations in subject added entry – uniform title; 4.7% (6) were commentaries, 18.1% (23) were criticisms & interpretations; 2.4% (3) were controversial literatures; 7.9% (10) were glossaries & vocabularies; 7.1% (9) were quotations; and 33.9% (43) were those other than the aforementioned subject added entry- uniform title. However, 11% (14) of records did not have subject added entry- uniform titles. Examples of the subject added entry – uniform titles are shown in Table 8. Usually, the MARC tags for subject added entry – uniform title is tag 630, and the subfield (x) which indicates a general subdivision was also considered.

The result shows that most of the records (98%) were sufficient and have enough data to be converted. Only 2.4% (3) of the MARC records were not sufficient for FRBR conversion in this study.

Table 5 : Types of Works

Type of Work	Frequency	Percent (%)
Elemental	104	81.9
Simple	20	15.7
Complex	3	2.4
<b>Total</b>	<b>127</b>	<b>100.0</b>

Table 6: Examples of Types of Works

ID	Type of Work	Examples
006	Elemental	[W] William Montgomery Watt : Companion to the Qur'an: based on the Arberry translation [E1] The original English text [M1] Original ed. (English), 2008 (printed text)
089	Simple	[W] Arthur John Arberry: The Koran illuminated : a handlist of the Korans in the Chester Beatty Library [E1] the original English text [M1] original ed. (English), 1967 [printed manuscript] [M2] original ed. (English), 1990 [printed text]
013	Complex	[W] Toshihiko Izutsu: God and man in the Koran: semantics of the Koranic weltanschauung [E1] the original English text [M1] original ed. (English), 1984 [printed text] [M2] original ed. (English, 1984 [TIFF file] [E2] revised ed. (English) [M3] revised ed. (English), 1989 [printed text] [E3] 3rd ed. (English)

Table 7: Subject Added Entry – Uniform Title

Subject added entry – uniform title	Frequency	Percent (%)
Translations	19	15.0
Commentaries	6	4.7
Criticism & Interpretation	23	18.1
Controversial Literature	3	2.4
Glossaries & Vocabularies	10	7.9
Quotations	9	7.1
Others	43	33.9
Missing	14	11.0
<b>Total</b>	<b>127</b>	<b>100.0</b>

Table 8: Examples of Subject Added Entry – Uniform Title

ID	Subject	Example
007	Translation	The Holy Koran: an introduction with selections <b>630 0 \$aQur'an \$xTranslations, English</b>
006	Commentaries	Companion to the Qur'an: based on the Arberry translation <b>630 0 \$aQur'an \$xCommentaries</b>
004	Criticism & Interpretation	Le Coran <b>630 \$aQur'an \$xCriticism, interpretation, etc.</b>
110	Quotations	Gesetzim Koran : kultus und ritus <b>630 00 \$aQur'an \$xQuotations, Early</b>

## DISCUSSION AND CONCLUSIONS

The purpose of this study was to examine the feasibility of applying the object-oriented FRBR model to MARC-based bibliographic records on Al-Quran. The study attempted to analyze the dataset on selected bibliographic records on Al- Quran in the IUM library's OPAC. The entities were identified, and the relationship between MARC-based bibliographic records and FRBR model were established. The results were also normalized, and the issues were identified, using the three classes of work according to the FRBR model. Overall, the information on the bibliographic record on Al-Quran is sufficient to reliably identify the FRBR entities. Based on the findings, most of the records (98%) were sufficient and have enough data to be converted to the FRBR model. However, issues were found with works in Arabic scripts, complex work and missing of subject headings.

### Issues with Arabic Scripts

Based on the record analysis, the result shows that 3.6% (5) of the MARC language identifier (006) were from Arabic or non-Roman scripts work. Table 9 exemplifies five records in Arabic scripts which may likely affect the process of FRBRization. Most of the records displayed in the Arabic scripts were in the tags 245 and 100 which represent the title statement and main entry – personal name. Future research is needed to look at the complexities of frbrizing records in multi-lingual and multi-script environment.

### Issues with Complex Works

Identification of complex work was challenging and time consuming. The finding on the types of work in the MARC based bibliographic records on Al- Quran shows that 2.4% (3) of the records are complex work. Generally, a complex work in these records has multiple expressions, thus naturally has multiple manifestations as well. Among them are books that have been translated into several languages and published by several publishers, have several revisions, or have been published through various media. The usefulness of the FRBR model is demonstrated very effectively by such complex works. However, the challenge comes in identifying complex works and in correctly interpreting and mapping them according to FRBR model. This requires a means of identifying the existence of multiple expressions subsumed under a single work or the forms of variation that distinguish one expression from another. Hence, the task of determining the complex work involves longer process.

Table 9: Examples of Arabic Scripts MARC Records

ID No.	Example
032	دراسة لأخلاق النظرية = مختصر دستور الأخلاق القرآن \$a 245 10 \$ / والعملية القرآن الكر يم مقار نة بالنظر يات الأخلاقية القديمة والحديثة محمد عبد العظيم علي، محمد عبد الله دراز؛ إعداد المختصر، تأليف \$c 100 1 \$a محمد عبد الله
053	مشروع ترجمة معان القرآن الكر يم باللغة الألمانية \$a 245 10 \$
059	... كتاب جيز ييحتسبيرة النبي الأكرم القرآن الكر يم \$b : تاريخ القرآن \$a 245 10 \$ / تأليف أبو عبد الله الزنجاني؛ ومصدر بمقدمة لأحمد أمين \$c 100 1 \$a زنجاني، أبو عبد الله
062	تأليف محمد \$c / عرض تاريخي تحليلي مقارن \$b : مدخل إلى القرآن الكر يم \$a 245 10 \$ عبد اللدراز؛ ترجمة محمد عبد العظيم علي، مراجعة السيد محمد بدوي 100 1 \$a محمد عبد الله دراز، محمد عبد الله
111	تأليف أبو عمرو \$c / كتاب المقتطف من مصاحف المصارع مع كتاب النقط \$a 245 10 \$ عثمان بن سعيد الداني؛ بإعتناء أو توبرتزل 100 1 \$a داني، عثمان بن سعيد

### Issues with Subject Headings

Some of the records under study do not have the appropriate data recorded to reliably identify the work with the group 3 entities of FRBR. For example, 11% (14) of the 127 records displayed missing subject headings in the tag 630 and 650 of the MARC record. This could affect the recall of relevant items by the users as they will undoubtedly miss out or not be able to locate clusters of records on the aboutness of work through subject search.

In summary, although the findings show good results, there is still much work to be done in improving the overall process, especially in relation to the FRBRization of expressions of works, and the representation of subjects and complex work. The study suggests the need to improve the quality of bibliographic records and exercise normalization of their contents. This finding is quite consistent with the claim that “transforming from a list of isolated MARC records into a graph like structure (FRBR model) is not an easy task due to internal challenges of cataloguing practice” (Manguinhas, Freire and Borbinha 2010, p. 225).

The local cataloguing industry must be trained to be skilful in understanding the underlying principle of the FRBR model in relations to cataloguing practice. Further refinement should take into consideration the compliance with the Functional Requirements for Authority Data (FRAD), and Functional Requirements for Subject Authority Data (FRSAD) models. As such, the importance of authority files must not be overlooked as the quality of authority data affects the overall performance of bibliographic databases. As this study has only looked at a very small dataset, it is believed that embarking similar research on a bigger sample would generate more reliable and interesting findings. However, this would require deployment of sophisticated FRBRization tool/software, and findings of this study would be useful in providing information on the current state of local MARC records for FRBRization purposes.

## **ACKNOWLEDGEMENT**

We gratefully acknowledge funding (Endowment B Grant) provided by the Research Management Centre, International Islamic University Malaysia (IIUM).

## **REFERENCES**

- Aalberg, T. and Maja, Z. 2013. The value of MARC data, or, challenges of frbrisation. *Journal of Documentation*, Vol. 69, no. 6: 851 – 872.
- Beall, J. 2006. Some Reservations about FRBR. *Library Hi tech News*, Vol. 23, no.2: 15 - 16.
- Carlyle, A. and Fusco, M. L. 2007. Understanding FRBR as a conceptual model: FRBR and the bibliographic universe, *Bulletin of the Association for Information Science and Technology (ASIS&T)*. Vol. 33, no. 6: 12-16. Available at: [http://www.asis.org/Bulletin/Aug-07/carlyle\\_fusco.html](http://www.asis.org/Bulletin/Aug-07/carlyle_fusco.html).
- Chen, Y.N. and Chen, S.J. 2004. A metadata practice of the IFLA FRBR model: A case study for the National Museum in Taipei. *Journal of Documentation*, Vol. 60, no. 2: 128-143.
- Dickey, J. T. 2008. FRBRization of a library catalog: Better collocation of records, leading to enhanced search, retrieval, and display. *Information Technology and Libraries*, Vol.27, no. 1: 23-32.
- Greenberg, J., Trujillo, S. and Mayer, K. P. 2012. YouTube: Applying FRBR and exploring the multiple description coding compression model. *Cataloging & Classification Quarterly*, Vol.50, no. 5-7: 742-762.
- Manguinhas, M. A. H., Freire, M. A. N. and Borbinha, L.B. J. 2010. FRBRization of MARC records in multiple catalogs. *Proceedings of the 10th Annual Joint Conference On Digital Libraries*, 225-234. New York, USA: ACM.
- Monch, C. and Aalberg, T. 2003. Automatic conversion from MARC to FRBR. *Research and Advanced Technology for Digital Libraries*, Vol.2769: 405–411.
- Pietras, M. and Robinson, L. 2012. Three views of the “musical work”: Bibliographical control in the music domain. *Library Review*, Vol.61, no. 8/9: 551-560.
- Weng, C. and Mi, J. 2006. Towards accessibility to digital cultural materials: A FRBRized approach. *OCLC Systems and Services*, Vol.22, no. 3: 217-232.