Creating better library information systems: the road to FRBR-land

Tanja Merčun, Katarina Švab, Viktor Harej and Maja Žumer
Department of Library and Information Science and Book Studies,
Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia

Abstract

Introduction. To provide valuable services in the future, libraries will need to create better information systems and set up an infrastructure more in line with the current technologies. The Functional Requirements for Bibliographic Records conceptual model provides a basis for this transformation, but there are still a number of unresolved issues connected to its actual implementation.

Method. Using conceptual analysis, the paper investigates three main aspects of future FRBR-related research and developments.

Analysis. For each of the three areas, the paper describes its current position and identifies the questions that will need to be resolved in the broader library community.

Results. The library community needs to rethink three key aspects connected to the creation and sharing of bibliographic data: what bibliographic data users need in their information seeking process, how this data should be presented and interacted with in bibliographic information systems and, last but not least, how this data should be encoded and managed.
Conclusion. Implementation of the Functional Requirements for Bibliographic Records model is only at its beginning. The paper provides a roadmap for future work on FRBR implementation and the creation of better library information systems.

Introduction

For some time now, libraries have been losing their position in the fast-paced information environment. Despite incorporating some of the new technologies, one of library’s primary activities - the creation of catalogues that help users explore library collections as well as find resources and information - fundamentally still follows the tradition of card catalogues. Developed in another time, for a different information environment, a different type of users, and a different set of technologies, library information systems are facing a major change in the near future and the library community will need to rethink and improve the way it provides its information services. On the one hand, libraries need to create an infrastructure that would support exchange and reuse of their rich data beyond the library domain, providing information where the users are (Tonta 2008). On the other hand, they will also have to take better advantage of their high quality data as well as centuries of experience in order to bring users back to the library by offering (in an effective and user-friendly way) information and services other providers do not.

The conceptual model Functional Requirements for Bibliographic Records (FRBR) has been developed to overcome the drawbacks of current bibliographic data and consequently library information systems. With the gradual adoption of FRBR in the new cataloguing rules and bibliographic frameworks, the conceptual model is slowly being transformed into an actual implementation. However, these changes are only the beginning (and not the end) of the road ahead and the next steps are crucial for the success of the model and the transition towards more modern and useful bibliographic information systems. This paper brings forward three important “stops on the road” where a number of questions will need to be resolved before the model can be fully employed and thus before libraries can take full advantage of its potential: the bibliographic data libraries catalogue, the display and interaction with the data in user interfaces and, finally, the encoding and management of the data.

Functional Requirements for Bibliographic Records and its potentials

Functional Requirements for Bibliographic Records (IFLA 1998), also well known as FRBR, is a conceptual, entity-relationship model of the bibliographic universe developed with the primary purpose of improving bibliographic records, the process of cataloguing as well as online library catalogues (Carlyle 2006). The core of FRBR lies in Group 1 entities work, expression, manifestation, and item which represent different levels of abstraction of intellectual and artistic products. To give a more tangible demonstration of these entities, we can present FRBR using a concrete example: the novel Don Quixote is an intellectual creation by Cervantes – a work, which has been expressed by the author in Spanish language – an original expression, but has been later, for example, translated into English by Charles Jarvis – a new expression and adapted into an audio version read by George
Guidall - another new expression. An English translation by Jarvis published in 1912 by Oxford University Press presents a manifestation of that expression and was reprinted in 1999 and 2008, which means there are two more manifestations of the same expression. Our local library holds one copy of the 1999 edition and three copies – items of the 2008 edition. Each entity also has a set of attributes - characteristics that describe it, but the greatest value of FRBR lies in relationships it establishes between entities. They create the possibility to group different editions under a work, to alert users of works that are related to each other, to display the type of relationship that exists between two entities, or to display all works of an author. These features, as well as many others, show the important potential of FRBR for designing exploratory systems that could significantly improve current display of information which Rose (2012) describes as “bewildering list of disparate manifestations of various expressions ordered only by some obscure “relevancy” algorithm”.

Tillett (2005) reminds us that one of the basic beauties of FRBR is that it brings back the key objectives of a catalogue. The idea of grouping versions of the work, connecting related works or bringing together all works and endeavours of an author is in fact not a new concept, but can be traced back to Panizzi, Cutter, and the Paris Principles (Denton 2007). In the time of card catalogues, many of the features we hope to gain by implementing FRBR were already in use through filling arrangement but have then been lost in transition from card to computer catalogues. Transforming the traditional, manifestation-oriented approach into an entity-relationship approach where each entity type is equally important, FRBR also provides a basis that is better suited for today’s technologies. While FRBR describes how bibliographic databases could be structured and what functions they should fulfil, it is highly theoretical and does not prescribe what its implementation should look like. Becoming the basis of the new Resource Description and Access (RDA) cataloguing rules has been an important milestone, but to achieve a true FRBR-based information environment, the conceptual model, the cataloguing rules, the cataloguing practice, and the final implementation in information systems will all need to go hand in hand.

Towards an FRBR-based information environment

Identifying what users need

The FRBR model presents a major shift in cataloguing which defines a completely new structure of bibliographic data and takes user tasks as its foundation. However, in the process of creating the model, the authors of FRBR have not re-examined through (user) studies the attributes and relationships that should be included but have more or less adopted existing guidelines, rules and practice for bibliographic description. Data currently recorded in bibliographic records has remained basically the same over the last century and the question is whether this bibliographic data is really what library users are looking for or what library information systems need in order to support user’s information seeking process.

The library community has so far not taken a user-centred approach in the development of cataloguing standards and, as Hoffman (2009) states, users have been typically studied only in relation to existing systems and standards. Also the FRBR model did not “involve studies of how actual users approach and make use of bibliographic records” (Madison 2000) and Zhang and Salaba (2009) in their Delphi study confirmed the necessity to verify attributes and relationships through user studies. The need to re-examine some aspects of the model, particularly attributes, relationships, and tasks that need to be supported by bibliographic data, was recognized also by the authors of the FRBR report themselves who suggested that “the identification and definition of attributes for various types
of material could be extended through further review by experts and through user studies” (IFLA 1998: 5).

So far, research tells us that users intuitively understand the differences between works, expressions and manifestations (Pisanski and Žumer 2010; Pisanski and Žumer 2012) and perceive the differences between versions which could (not) be substituted with each other (Carlyle and Becker 2008). While some search for manifestations (i.e. particular editions) when they are particularly interested in the first or the latest edition or when they are looking for publications with additional materials, such as illustrations or commentaries, most users seek works, expressions, and groups of expressions (e.g. any edition of a work in a particular language) (Yee 1998; Leskovec 2005). In some contrast to those indications, current catalogue records describe manifestations in detail, whereas information about respective work(s) and expression(s) is not always evident and many important relationships and attributes are not recorded (Žumer 2011).

To find the selection of attributes and relationships required in our catalogues on the work, expression and manifestation level, a number of questions need to be answered, for example: Which attributes and relationships are missing in current catalogues and which are redundant? Which attributes and relationships are most important to users? Are they the same for all user groups, all types of materials or all information needs? Which attributes and relationships are essential for supporting user tasks find, identify, select, obtain, and explore - a task added to the original four by Functional Requirements for Subject Authority Data? (Zeng et al. 2011)

Different user groups, for example parents, children, and adults, have different information needs. To get a better idea of what bibliographic data is required to fulfil those needs, the information behaviour and information needs of each specific user group has to be studied in more detail. Our own research, for example, examined how five different user groups (parents of preschool children, high school students, school librarians, adults, and students) search and select fiction. Using not only existing but also fictitious bibliographic records in order to include a range of data not recorded or displayed in current bibliographic records, we have studied which attributes and relationships are important to specific user groups. Results have shown that in addition to current bibliographic elements (language, year of publication, extent of the carrier), users also need attributes that are not typically included in bibliographic systems (content, condition of the item, reading level, typical page of the book, font size, intended audience) for identifying and selecting the needed resources. The results also depended on the user group and the context of use: when selecting picture books, parents of preschool children were most interested in illustrations; for required reading in high school, the most important attributes were additional content in the book (biography, table of contents, introduction, preface…) and the reading level, while for leisure reading the most important attributes were a typical page of the book, font size, and binding of the book. Common to all contents and to all user groups were the aesthetical attributes of a book, such as the cover of the book, condition of the item, and the design of the book. Participants also recognized some relationships between works as important or very important: the information on sequels and prequels, transformation, and summarization.

Our preliminary results give same indication which attributes and relations are important to users and show that there is some difference between currently recorded attributes and relationships and user’s expectations. However, more user studies are needed in this area to really encompass the various scenarios as well as materials beyond fiction books.

Identifying which attributes, relationships, and entities are important to users at what point in their information seeking process, how they select resources, or which editions of the same work they find substitutable is important also for improving bibliographic information systems, not only for
displaying the data in the record display but also for designing user’s interaction and exploration of the library collection.

Creating a discovery environment

Knowing what bibliographic data are really needed in order to support users in their search is only a part of the journey; another question that has to be dealt with is how to make the best use of these entities, relationships, and attributes in user interfaces of bibliographic information systems (Tillett 2005).

Due to the lack of guidance as well as user research, FRBR implementations have so far been largely exploratory in nature in terms of user interface display, system features, and the way they applied the FRBR model (Zhang and Salaba 2012). The need to involve users in the evaluation and the design of FRBR-inspired catalogue displays has been emphasized also by Carlyle (2006). She stressed that user research on what kind of displays are most effective and what attributes most facilitate the use of the catalogue are highly desirable and needed to guide the decision-making process in the design of online catalogue displays that incorporate FRBR.

What little discussion in literature has been made in the last fifteen years on the presentation and interaction within FRBR-based information systems has comprised mainly of some general thoughts by individual researchers. Dickey (2008), for example, suggested that a tree organization was user-friendly and allowed users to maintain a visual sense of the organization they were encountering. Aalberg (2002) noted that a “complexity of the FRBR model calls for a user interface that will provide an overview of the large structure”, while Boston et al. (2005) emphasized the importance of considering appropriate screen labelling, terminology, and layout that would assist users to anticipate, understand, and fully exploit the delivery of the clustered FRBR results. Yee (2005), on the other hand, gave a few detailed scenarios of how entities could be presented in catalogues. She envisioned that for each author there should be a list of works and for each work separate categories for editions of the works, works about the given work, and related works. On the next level, categories would be used for distinguishing between complete works, selections, arrangements etc. and on the last level users would be able to rearrange expressions/manifestations of a chosen work by language, translator, editor, illustrator, edition statement, publisher, date, performer, format, or extent. However, these were just some ideas of what should be considered or how user interfaces could be designed and it is only recently that we have been witnessing some initial user-based research investigating the presentation of and interaction with FRBR entities (for example Zhang and Salaba 2012, Arastoopoor et al. 2011).

While the few FRBR-inspired catalogues implemented so far show improvements over traditional catalogues, they have not implemented the complete model or fulfilled the full potential of FRBR (Pisanski and Žumer 2007; McGrath and Bisko 2008). In most cases, they have created flat lists that only grouped records into work sets and, in some cases, enabled collocation of editions by language. It is true that due to the lack or the inappropriate form of essential bibliographic data, many of the missing features are difficult to implement using only frbrization of existing records, but future online catalogues based on FRBR should be able to create exploratory environments that move beyond the “list of manifestations” concept. The question, however, remains how such systems should be designed in order to better exploit the richness of the new bibliographic framework.

Building a prototype system to test the potential of information visualization techniques for presentation and interaction with fully frbrized data, we were faced with a number of design
questions, many of which could also be used as the framework for future discussions on FRBR-based displays within the library community, for example:

- How should entities be collocated? Should we create additional subcategories within a work grouping as indicated by Arastoopoor et al. (2011) or should we use only groupings on the work level and then embed other tools such as faceted navigation for browsing and narrowing down to the manifestation? What kind of subcategories would be most useful for users when researching a work or when researching an author?
- How should we deal with the discrepancy between complex work families or very prolific authors on the one hand and works with only one expression and one manifestation or authors with only one creation on the other?
- How should we present relationships between related works, between derivative expressions or between manifestations?
- How can we create the best overview of the bibliographic family and enable the user to explore the network of relationships that exist in the bibliographic universe?
- How should we form results list for keyword searches? Should we always present results on the work level even when the matching has been made only on the expression or manifestation level or on the combination of all three levels?
- How can we best bring together similar materials which are interchangeable for most users, but at the same time retain the detailed information that will allow users with specific needs to determine the differences between these similar materials?
- What kind of presentation method will enable us to show and interactively explore the hierarchical top-down, bottom-up as well as horizontal relationships between entities?

Knowing what kind of interaction and presentation of data should be provided in the newest generations of bibliographic information systems will not only provide the framework for design, but also for the identification of the most essential entities, attributes, and relationships that need to be catalogued in the future.

**Formatting the data**

To enable a full scale implementation of FRBR on all levels and create a richer information environment, it is not enough to revise the bibliographic data, cataloguing rules, and the conceptual design of user interfaces. All the changes need to be addressed also in a wider bibliographic framework and supported by the underlying data models and formats that encode the bibliographic data, enabling its exchange and processing. While the currently used MARC format can, to a certain extent, encode FRBR entities and relationships (Aalberg et al. 2011), the reality is that the format has been designed more than 50 years ago. It has been primarily intended for the exchange and display of records and the data was, to a large degree, structured for human interpretation and not for automated processing and retrieval as it is required today (Lee and Jacob 2011). Furthermore, MARC is also not able to efficiently support FRBR-born data that introduces an approach different from the traditional manifestation-based bibliographic record. For all these reasons, it is time for libraries to step forward and define a new format. As Picco and Ortiz Repiso (2012) stress, libraries have been pioneers in information organization, but clinging to “outdated methods and tools that are out of step with the technological reality” will make them lose their position in the information society.

The need to rethink the future of bibliographic control as well as the MARC format itself has been recently recognized also by the Library of Congress. In May 2011 it launched the Bibliographic Framework Transition Initiative where a major focus will be given to the transition from MARC 21
exchange format to more Web based, Linked Data standards (Library of Congress 2012). While the project is still in its early stage, it presents a globally oriented attempt to define a new data model that will guide the implementation of the FRBR-based bibliographic data. It has already proposed a BIBFRAME data model which however, as far as we can see, represents a deviation from the FRBR conceptual models since the expression entity has been left out.

One of the possible approaches to the formulation of a new format is, as indicated also by the Bibliographic Framework Initiative, to join the linked data “cloud”. The Semantic web platform provides the means for identification of entities and relationships and makes it possible to generate more appropriate ways of structuring bibliographic records that transcend the flat structure of the MARC format. In this new reality, relationships can be recorded explicitly which allows navigation between related resources and turns catalogues into true information networks that overcome the limitations of current catalogues (Picco and Ortiz Repiso, 2012).

However, the Semantic web technologies provide only the basic framework and not an “out-of-the-box” solution to all underlying problems. Technological solutions for publishing the data and making it searchable are already available for the Semantic web, but there are still a number of questions that libraries will need to address before they can move their data to the new platform. For example, how will the data be exchanged between libraries? Libraries will need to come up with some kind of an exchange library format and design procedures for maintaining the quality control of the data. Another issue will also be the global identification of entities and relationships. On the Semantic web, identification is achieved through the use of URIs (uniform resource identifiers), but some institutional structures would be needed to support the sharing of URIs created for entities on the Semantic Web so that the quality standards for bibliographic and authority control would be met (Yee 2009). A promising attempt towards global URIs for authority data is the VIAF project (OCLC 2013), but it presents only a starting point as the library community needs a similar mechanism for global identification of works, expressions, manifestations, places, subjects, and so on. For the presentation of bibliographic (or any other) domain on the Semantic web, an ontology encoded in one of the proposed ontology languages is also needed. The FRBR conceptual model can serve as the basis for the ontology, but more elaborate for that purpose is FRBRoo (International Working Group 2012). Which ontology will eventually become “the” ontology for bibliographic data on a global scale, if any, and which body will govern it are all questions that still need to be addressed in the future.

And last but not least, before moving towards the new platform, the FRBR model needs to be harmonized with the two additional models that have extended parts insufficiently addressed in the original FRBR model: Functional Requirements for Authority Data (FRAD) and Functional Requirements for Subject Authority Data (FRSAD). The so called FR family of models (FRBR, FRAD, and FRSAD) does not act as one model yet and more work is needed (Žumer et al. 2011; Riva et al. 2008) in order to create a strong platform for future work.

Some of the major questions that will need to be resolved in the near future if libraries wish to move into the 21st century information environment are therefore:

- How can we ensure the semantic and format interoperability of the new library data?
- How can we address the problems of identification and how will we satisfy the standards of authority and bibliographic control in this new, more open environment?
- How can we ensure a rapid and controlled exchange of bibliographic data that would support new ways of catalogization and how will this new process of catalogization even look like?
• How will the legacy bibliographic data be integrated into the new bibliographic framework and semantically enriched in a way that will enable the identification of all the needed entities and relationships?

Conclusion

With the adoption of the FRBR model, it seems that libraries are finally moving away from ground zero where they had to work within a modern information environment using their legacy data and restricting underlying formats. But the road to FRBR-land, where the full implementation of the model would enable libraries to use their data in more innovative ways and to create bibliographic information systems that would better support users’ needs and information seeking process, is really only beginning. The paper describes three important stops on the road where deliberations, decisions, and (user) studies will have to be made in order to pave the way towards a successful implementation: the bibliographic data libraries catalogue, the display and use of those bibliographic data in information systems, and data management that will enable re-use of data and the creation of a new generation of bibliographic information systems, but at the same time allow libraries to maintain a high level of bibliographic control. The paper provides a roadmap for future FRBR-based research and development, but time is of the essence: if guidelines, design concepts, data models, frameworks, and formats that complement each other are not established quickly enough, libraries might end up implementing only some aspects of the model, thereby again failing to use the full potential of bibliographic data.

References


• Rose, M.Z. (2012). The ship has sailed and we aren’t on it: how catalogers could support user tasks and why we won’t. *Journal of Library Metadata, 12*(2-3), 127-139.


How to cite this paper

Find other papers on this subject

[Scholar Search] [Google Search] [Bing]

Check for citations, using Google Scholar

© the authors, 2013.
Last updated: 9 August, 2013