



Breaking barriers between old practices and new demands: the price of hesitation

**Maja Žumer
Jan Pisanski
Polona Vilar
Viktor Harej
Tanja Merčun
Katarina Švab**

Department of Library and Information Studies and Book Science, Faculty of Arts
University of Ljubljana
Ljubljana, Slovenia

Meeting:

Cataloguing Section

Abstract

In recent years, cataloguing has been going through a number of quite fundamental changes. This paper offers an evaluation of what the cataloguing community has accomplished so far and a suggestion of what it ought to focus on in the future. Catalogues should provide users with a more efficient way of accessing and using information as libraries have the tools to expose users to a broader context that is absent from the catalogues today. We do have FRBR (as well as FRAD and FRSAD) which could and should be used as the background for cataloguing, as indicated by the findings in two of our studies. Some cataloguing codes based on these documents, as well as on the International Cataloguing Principles, have been developed so far (perhaps most importantly RDA) and they present an important step forward. However, we can identify several key issues that the library community needs to focus on in order to complete the process: harmonisation (of various initiatives); frbrisation, identification and presentation. For each of these areas, the purpose of cataloguing needs to be kept in the centre and the decisions must be based on user studies, which have been sorely lacking so far. The library community has been too slow in this transformation, which is why these issues must be quickly and efficiently resolved if libraries wish to retain their role and importance in today's information world.

The FRBR-based revolution

In recent years, cataloguing has been going through a number of revolutionary changes driven by developments in technology, the need to change current rigid practices that have somewhat lost their purpose in the digital world, as well as by the recent findings about information needs and searching patterns of users. Undoubtedly, FRBR - Functional Requirements for

Bibliographic Records (1998), the first formal model of the bibliographic universe, represents a bold step into a new direction for the library community and its cataloguing practices and outcomes.

The benefits of FRBR for users as well as librarians have been predicted shortly after the publication of IFLA study at an ELAG workshop (Noerr, Goossens, Matei, Otten, Peruginelli and Witt, 1998): users benefit by easier searching, more focused results, clustering at the work-level, better navigation, and better utility of bibliographic references, while librarians profit by better placement of metadata in records, easier copy-cataloguing and sharing of records, and adding new metadata (e.g. rights management).

When it comes to the provision of information, we are becoming painfully aware that the library community has not been active enough. While undoubtedly hosting the richest source of information about resources, the majority of library catalogues is still based on the model that worked incredibly well for card catalogues but has been obsolete ever since computers were first introduced. As Yee (2005) put it, FRBR enables the development of online public finding lists into true online public catalogues.

Is FRBR intuitive?

While FRBR is clearly user oriented, it was not based on dedicated user studies (Madison, 2005). Therefore many have called for user verification of FRBR, perhaps most noticeably Library of Congress Working Group on the Future of Bibliographic Control (2008) and experts on FRBR in a Delphi study by Zhang and Salaba (2009a). This void prompted our research of intuitiveness of FRBR. So far, we have performed two user studies of mental models of the bibliographic universe. In both cases we focused on printed books.

The first study, which was performed in 2007-2008, explored mental models of non-librarians using three different approaches: card sorting, concept mapping and comparison task. While the details are presented elsewhere (Pisanski, Žumer, 2010a, 2010b), the results indicate that the participants of the study found FRBR to be intuitive.

The second study from 2010 was conducted on 60 students, 10 each from the fields of computer science, design, economics, history, medicine, and social studies. We took six graphs based on the models observed in the first study and asked participants to choose the one they felt best represented the bibliographic entities listed and their relationships. Results show that the FRBR model was chosen by the majority of the participants (33 out of 60) and there was no clear alternative.

While agreement with FRBR was overwhelming in both studies, we found that the positioning of the original expression in FRBR was questioned by some participants, as they viewed it as a crucial link between the work and the rest of the bibliographic universe. This distinction is not reflected in FRBR, but is present in the FRBRoo model (International Working Group, 2010). We will be continuing our research with a slightly enhanced extension of the second study.

A somewhat similar study (Arastoopoor, Fattahi, Parirokh, 2010) considered the views of experts in a particular field on collocating related entities. It, too, found that FRBR was understood by the participants, but additionally they used other, more specific, criteria for collocation.

The results of these studies suggest that FRBR is an appropriate user-oriented conceptual model for at least parts of the bibliographic universe.

From theory to practice

Recently, there have been a number of developments based on FRBR, such as the new International Cataloguing Principles (2009) and even the first attempts at FRBR based cataloguing rules in the form of Italian Cataloguing rules (Trombone and Canepa, 2009, Petruccianni, 2009) and RDA (2010). On the other hand, the completion of the two other conceptual models from the FRBR-family, FRAD (2009) and FRSAD (2010), enables further insights into various parts of the bibliographic universe. Another important area is the on-going development of namespaces by IFLA Namespaces Task Group.

While we have these major building blocks for the new generation of catalogues in place, the debate has continued on some of the major issues that still need to be resolved (see, e.g., Yee (2005), Žumer (2006, 2007 a and b), Dickey (2007), Salaba and Zhang (2007), Zhang and Salaba (2007, 2009a, 2011a and b), Carlyle (2006), Carlyle and Fusco (2007)). We will try to give accounts of some of these issues in our paper.

Thirteen years after the publication of FRBR, we believe that the time has come for the library community to embrace the thought that FRBR could and should be used to develop a new approach to cataloguing. Yet it seems that librarians still do not recognize the full potential of a networked library environment and want to hold on to some tools and practices that have lost their purpose with library automation. In this sense, initiatives that allow continuation of current practices will not help.

Experiences of recent years and of other information providers tell us that, in order to survive, catalogues should provide users with a more efficient way of accessing and using information, as well as better data. Libraries should start offering the tools which would provide users with a broader context than is available in catalogues today. By also opening their tools and data to the outside world, libraries could have a major influence on the development of Semantic Web. Indeed, we pursue the same main goal: improved accessibility and organisation of information. We should therefore actively engage in interaction with the Semantic Web community.

Key Issues

There are several key issues that we feel the library community needs to focus on, always keeping the purpose of cataloguing in mind.

1. Development of the model
 - a) Harmonisation (of various initiatives)

While both FRAD and FRSAD are based on FRBR and follow the same method of modelling, the developers have made different modelling decisions, resulting in partially incompatible solutions. FRBR is a conceptual model of the bibliographic universe and defines all entities and relationships, but focuses particularly on Group 1. FRAD expands the model

in the area of authority data for Group 2 entities and *works*; FRSAD deals with the subject relationship.

Main differences between FRBR and FRAD are summarised as:

- Two user tasks, »Contextualise« and »Justify« are added in FRAD
- Three appellation entities, *name*, *identifier*, and *controlled access point* are added
- 'has as subject' relationship in FRBR is replaced by subject as attribute of *work* in FRAD
- *Rules* and *Agency* as new entities in FRAD

The main differences between FRBR and FRSAD are:

- The addition of the “Explore” task;
- *Thema* is introduced as a superclass of all entities that can be subjects of a *work*;
- No entities are explicitly predefined in Group 3;
- *Nomen* is introduced and is defined as a separate entity instead of an attribute.

All these differences discourage wide implementation of FRBR, therefore the harmonisation of the FRBR family is essential and urgent. FRBR Review Group has started the process, partially in conjunction with the Namespaces project, but the real work has not started yet.

On the other hand, the library community has been relatively successful in developing a common model with the museum community: FRBRoo is the result of the harmonisation of FRBR and CIDOC CRM and can serve as the basis for development of common tools and services for the cultural heritage documentation.

b) Attributes and relationships

In spite of all the above mentioned initiatives, we still have little research-based evidence on exactly what attributes and relationships different user groups require. FRBR model took the relationships and attributes mainly from the current cataloguing practice: *International Standard Bibliographic Descriptions* (ISBDs), the *Guidelines for Authority and Reference Entries* (GARE), the *Guidelines for Subject Authority and Reference Entries* (GSARE), and the *UNIMARC Manual*. The designers of the FRBR model state that: “The identification and definition of attributes for various types of material could be extended through further review by experts and through user studies” (FRBR, 1998). However, newer initiatives, such as RDA and the new version of ISBD, still basically rely on attributes and relationships as specified in the original FRBR study.

A small study by Leskovec (2005) confirms that the attributes and relationships recorded in current catalogues do not always correspond to user needs. She analysed user requests in a public library and found that most users search for expressions, groups of expressions (e.g. any edition of a work in a particular language) and sometimes even works in general. Some users search for manifestations (i.e. particular editions) when they are particularly interested in the first or latest edition or when they are looking for publications with additional materials, such as illustrations or commentaries. However, while catalogue records describe manifestations in detail, information about respective work(s) and expression(s) is not always evident and many important relationships and attributes are not recorded (e.g., whether text is integral or abridged, information about sequels, etc.).

2. Frbrisation

There are huge amounts of legacy data that will somehow have to coexist with born-FRBR data. The best solution for that is frbrisation. Frbrisation is the process of extracting FRBR entities from the existing bibliographic data. As recataloguing huge amounts of existing data is not a viable option, frbrisation is of crucial importance for transparent joint use of old and new data.

At IFLA congress two years ago, we reported on an experiment on data from Slovenian national bibliography, the Norwegian Bibsys database and the Swedish Burk database (Pisanski, Žumer, Aalberg, 2009). We found that existing data can be frbrised in a satisfactory manner, but the quality of the results depends on various important factors, such as the quality and consistency of bibliographic records, local cataloguing practices, cataloguing rules and formats. These inconsistencies lead to misidentification of entities, which may overshadow the benefits of frbrisation for the end-user. Perhaps most importantly, our experiment clearly showed that frbrisation is not trivial and that it requires customisation to allow for the differences in individual cataloguing practices. Reports on previous frbrisation attempts, such as Hegna and Murtomaa (2002) and Hickey and O'Neill (2005), clearly deliver the exact same message.

There are many FRBR-inspired initiatives in existence, clearly showing the acknowledgement that users would benefit from better structured information. However, none of the attempts, except some small scale experimentation, actually follow FRBR completely, as expression level is often omitted or reduced to language-based groups of expressions, mainly due to difficulties with identifying expressions during frbrisation. However, as already mentioned, Leskovec (2005) found that users do sometimes look for a particular expression (and not just in terms of a particular language) and would definitely benefit from such information.

There is no simple way of recording all the important information (including relationships, which, after all should be the focal point of next-generation catalogues) in current bibliographic records in a structured manner using current tools. This is particularly true for complex cases such as aggregates of different kinds.

3. Identification

At last year's IFLA Congress, we presented a paper on the importance of identification in the area of the bibliographic universe, with a particular emphasis on the necessity of identification systems in multilingual and multicultural environments (Pisanski, Žumer, Aalberg, 2010). There are a number of identification systems in place, however even the most widespread systems, such as ISBN, are still underused. Also, in many cases there are different understandings of what entity an identifier should identify, largely based on the different expectations that different communities place on identifiers. Additionally, many identification systems are intended to be used for a particular type of materials (audio, books, serials, etc.).

As any other platform, Semantic Web technologies also have demands, and even shortcomings (see Yee, 2009). One step in our struggle to give data semantic meaning is the Linked data initiative. Technically, Linked data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets (Bizer, Heath, Berners-Lee,

2009). One of architectural prerequisites of linked data initiative is to use URIs (Uniform Resource Identifiers) as names for things, since URIs identify any kind of object or concept (Berners-Lee, 2009).

Yee (2009) explains how this would work in FRBR context: »[...] we would share in the creation of URIs for works, expressions, manifestations, persons, corporate bodies, places, subjects, and so on.« But as Yee (2009) continues: »Very much in the air is the question of what institutional structures would support the sharing of the creation of URIs for entities on the Semantic Web. For the data to be reliable, we would need to have a way to ensure that the system would be under the control of the people who had been educated about the value of clean and accurate entity definition, the value of choosing most commonly known preferred forms (for display in lists of multiple different entities), and the value of providing access. « In other words, if we want to expose our data on the Semantic Web, we need to uniquely identify each entity (and relationships, as well).

The agent part of the problem is targeted by the VIAF project (<http://viaf.org>) and one option would be to use their URIs for agent identification, at least for those agents which are in their database.

How to address identification of all FRBR Group 1 entities is even more difficult to imagine. One approach would be to achieve identification on national level (a complicated task on its own) and have an international coordinating body.

Some would probably argue that just publishing national (or even library specific) URIs for entities on the web would be better than nothing. In this scenario, librarians would leave the »power« to make connections to the enthusiasts or companies, and this is exactly what Yee (2009) is afraid of.

Whatever our approach will be, we need to be aware that nothing will be done just by adopting Semantic Web platform. Identifiers need to be defined manually and also data will not connect itself. Connections need to be explicit and this connecting should better be done by us. While some of this can be done automatically or semi-automatically, the results are not always reliable, as frbrisation which relies on identification has already proven. Partial solutions in this area will be neither cost- nor time-saving.

4. Presentation

When discussing changes in cataloguing and catalogues, one area that has often been overlooked is the data presentation aspect, which deals with the creation of user interfaces that would efficiently support user tasks and adequately present data. Carlyle (1997) summoned the library community to provide relationship-based displays and stressed that such change “is long overdue”, but 13 years later, the situation is still basically the same.

The classical display of data that has been traditionally used in library catalogues is not capable of showing all the possibilities provided by the various new models for bibliographic records which build on relationships. Critiques on the inability of our information systems to efficiently collocate records and present relationships are not new by any means, but the issue has not really been in the forefront, also because our existing cataloguing practice, rules and formats did not really provide a good base for implementation of those ideas.

As Zhang and Salaba (2009b) point out, the FRBR model offers great potential for developing effective and user-friendly system interfaces and displays by assembling records into interrelated clusters and displaying the hierarchy and relationships based on the FRBR model. So far, most FRBR-inspired initiatives have only attempted the work-(expression)-manifestation collocation and presentation, but have not yet focused on work-to-work and other relationships discussed in the model.

Due to the lack of investigations into presentation techniques for frbrised records and only very basic implementations of the FRBR ideas, our research group decided to explore the prospects of information visualization for presentation of and interaction with frbrised records, as it could enable collocation of records as well as a better presentation of various relationships. With the help of a prototype system, three different visualization techniques and interactions will be put to the test. As visualizations need to work equally well for very simple as well as very large data sets, the test examples will be constructed in a way that will represent various possible structural as well as relational complexities.

To our knowledge this will be the first experiment that will test possible display techniques for FRBR-based data. As such, it will not only tell us whether information visualization has the potential to support user's information exploration and discovery process, but will also provide other valuable information. We hope that in the future, more studies will be done in this area; not only because they could offer a great insight into what data is actually needed in our records, but also because it will soon become essential to adjust user interfaces for the new generation of library data. Without that, much of our work will be lost to the users. We also believe that a platform which will provide support for the implementation of enhanced library data can be a good incentive for libraries and cataloguers to quickly and more consistently take up the new practices.

User studies

While users have been the focus of libraries worldwide, all our decisions must be evidence based. We feel that users and usage must be systematically and continuously studied in all of the above mentioned areas. While our group has performed its small share of studies and there have been other sporadic attempts, we would like to encourage the research community to join in these efforts.

Conclusion

We have shown some results of our research and the issues that need to be addressed in order to make the most out of the recent promising developments in cataloguing and beyond. However, libraries are quickly running out of time for experimentation and comfortable, slow-paced approach to changes. Also, the library community is trying to adapt new ideas to the old practices, which may not prove to be efficient, as current catalogues are not in line with the information behaviour of users (UCL, 2008, OCLC, 2009). Users find catalogues hard to use and not intuitive enough. Additionally, most of them are not willing to invest their time and resources into learning the details of any system, even if it could help them achieve better results. Libraries do have potentially world class tools, however, they need to be adapted to the requirements of the end-users. The goal should be clear: to provide a solution that users will explicitly find beneficial. The price of hesitation in implementing the necessary changes might just be too high. In the time when users seek quick access to information,

context, content, exploration and discovery, libraries' failure to provide that will cost them their users as well as the position of leaders in the information world.

References

Arastoopoor, S., Fattahi, R. and Parirokh, M. (2011). Developing user-centred displays for literary works in digital libraries: Integrating bibliographic families, *FRBR and users*. *2nd International Conference on Asian Special Libraries*.

Berners-Lee (2009). *Linked data*. <http://www.w3.org/DesignIssues/LinkedData.html>

Bizer, C., Heath, T. and Berners-Lee, T. (2009). Linked data – the story so far. *International Journal on Semantic Web and Information Systems*. 5 (3), 1-22.

Carlyle, A. (1997). Fulfilling the second objective in the online catalog: Schemes for organizing author and work records into usable displays. *Library Resources and Technical Services*, 41 (2), 79–100.

Carlyle, A. (2006). Understanding FRBR as a conceptual model: FRBR and the bibliographic universe. *Library Resources & Technical Services*, 50 (4), 264-273

Carlyle A. and Fusco, L. (2007). Understanding FRBR as a conceptual model: FRBR and the bibliographic universe. *Bulletin of the American Society for Information Science and Technology*, 33 (6), 12-16.

Dickey, T. (2007). FRBRization of a library catalog: better collocation of records, leading to enhanced search, retrieval and display. *Information Technology and Libraries*, March 2008, 23-32.

Hegna, K. and Murtomaa, E. (2002). *Data mining MARC to find: FRBR?*. <http://folk.uio.no/knuthe/dok/frbr/datamining.pdf>

Hickey, T. and O'Neill, E. (2005). FRBRizing OCLC's WorldCat. *Cataloging & Classification Quarterly*, 39 (3/4), 239-251.

International Federation of Library Associations and Institutions. Working Group on Functional Requirements for Subject Authority Records. (2010). *Functional Requirements for Subject Authority Data (FRSAD): a conceptual model*. June 2010.

<http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>

International Federation of Library Associations and Institutions. *Statement of international cataloguing principles*. http://www.ifla.org/files/cataloguing/icp/icp_2009-en.pdf

International Federation of Library Associations and Institutions. Study Group on the Functional Requirements for Bibliographic Records (1998). *Functional Requirements for Bibliographic Records: final report*. Munich, Germany: KG Saur

International Federation of Library Associations and Institutions. Working Group on Functional Requirements and Numbering of Authority Records. (2009). *Functional Requirements for Authority Data: a conceptual model*. Munich, Germany: KG Saur.

International Working Group on FRBR and CIDOC CRM Harmonisation. (2010). *FRBR: object-oriented definition and mapping to FRBR_{ER} (version 1.0.1)*. http://www.cidoc-crm.org/docs/frbr_oo/frbr_docs/FRBRoo_V1.0.1.pdf

Joint Steering Committee for Development of RDA. (2010). *RDA: Resource Description and Access*. <http://www.rda-jsc.org/rda.html>

Leskovec, M. (2005). *Delo, izrazna oblika, pojavna oblika : kaj uporabniki res iščejo?*. Diplomsko delo. Ljubljana: Univerza v Ljubljani, Filozofska fakulteta.

Library of Congress. Working Group on the Future of Bibliographic Control. (2008). *On the record: report of the Library of Congress Working Group on the Future of Bibliographic Control*. <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>

Madison, O. (2005). The origins of the IFLA study on Functional Requirements for Bibliographic Records. *Cataloging & Classification Quarterly*. 39 (3/4). 15-37.

Noerr, P., Goossens, P., Matei, D., Otten, P., Peruginelli, S., and Witt, M. (1998). User benefits from a new bibliographic model: Follow-up of the IFLA Functional Requirements study. *International Cataloguing and Bibliographic Control: Quarterly Bulletin of the IFLA UBCIM Programme*, 28 (3), 80-81.

OCLC (2009). *Online catalogs : what users and librarians want : an OCLC report*. 68 p.

Petruciani, A. (2009). Every reader his work, every work its title (& author). *World Library and Information Congress: 75th IFLA General Conference and Council, 23-27 August 2009, Milan, Italy*. <http://www.ifla.org/files/hq/papers/ifla75/107-petruciani-en.pdf>

Pisanski, J. and Žumer, M. (2007). Functional Requirements for Bibliographic Records: an investigation of two prototypes. *Program*, 41 (4), 400-417

Pisanski, J. and Žumer, M. (2010a). Mental models of the bibliographic universe. Part 1: Mental models of descriptions. *Journal of Documentation*, 66 (5), 643-667

Pisanski, J. and Žumer, M. (2010b). Mental models of the bibliographic universe. Part 2: Comparison task and conclusions. *Journal of Documentation*, 66 (5), 668-680

Pisanski, J., Žumer, M. and Aalberg, T. (2009). Frbrisation: towards a bright new future for national bibliographies. *World Library and Information Congress: 75th IFLA General Conference and Council, 23-27 August 2009, Milan, Italy*. <http://www.ifla.org/files/hq/papers/ifla75/77-pisanski-en.pdf>

Pisanski, J., Žumer, M. and Aalberg, T. (2010). Identifiers: bridging language barriers. *World Library and Information Congress: 76th IFLA General Conference and Assembly, 10-15 August 2010, Gothenburg, Sweden*. <http://www.ifla.org/files/hq/papers/ifla76/93-pisanski-en.pdf>

Salaba, A. and Zhang, Y. (2007). From a conceptual model to application and system development. *Bulletin of the American Society for Information Science and Technology*, 33 (6), 17-23.

Trombone, A. and Canepa, F. (2009). The new Italian cataloguing rules and the UNIMARC standard. *World Library and Information Congress: 75th IFLA General Conference and Council 23-27 August 2009, Milan, Italy*. <http://www.ifla.org/files/hq/papers/ifla75/135-trombone-en.pdf>

UCL (2008). *Information behaviour of the researcher of the future*. http://www.jisc.ac.uk/media/documents/programmes/reppres/gg_final_keynote_11012008.pdf

Yee, M. (2005). FRBRization: A method for turning online public finding lists into online public catalogs. *Information Technology and Libraries*, 24 (3), 77-95.

Yee, M. (2009). Can bibliographic data be put directly onto the semantic web? *Information Technology and Libraries*. 28 (2), 55-80.

Zhang, Y. and Salaba, A. (2007). Critical issues and challenges facing FRBR research and practice. *Bulletin of the American Society for Information Science and Technology*, 33 (6), 30-31.

Zhang, Y. and Salaba, A. (2009a). What is next for FRBR? A Delphi study. *The Library Quarterly*, 79 (2), 233-255.

Zhang, Y. and Salaba, A. (2009b). *Implementing FRBR in libraries: Key issues and future directions*. New York: Neal-Schuman.

Zhang, Y. and Salaba, A. (2011a). FRBRizing existing MARC record at expression and manifestation level. *ALA Midwinter Conference, January 7, 2011*. <http://frbr.slis.kent.edu/presentations/ALA2011-midwinter-FRBR.pdf>

Zhang, Y. and Salaba, A. (2011b). Developing an FRBR-based system to support user tasks. *NOR-ASIST, April 4, 2011*. http://frbr.slis.kent.edu/presentations/NOR-ASIST_2011-04-04.pdf

Žumer, M. (2005). Functional Requirements for Bibliographic Records as users' model. *Satellite meeting to the 71st World Library and Information Congress, IFLA 11-12 August 2005 : Biblioteca Universalis - How to Organise Chaos? FRBR, a new effort to organise content, and some practical applications*. Järvenpää, Finland. <http://www.fla.fi/frbr05/zumerFRBR2005.pdf>

Žumer, M. (2007a). FRBR: The end of the road or a new beginning? *Bulletin of the American Society for Information Science and Technology*, 33 (6), 27-29.

Žumer, M. (2007b) FRBR: is this the beginning of a beautiful friendship? *Ontology-driven interoperability for cultural heritage objects : working notes : DELOS - MultiMatch workshop 15 February 2007*. Tirrenia: DELOS, 2007, 5-6.

This paper is authored by members of a research group at the Department of Library and Information Studies and Book Science, Faculty of Arts, University of Ljubljana, Slovenia. The main research focus of the group is on modelling of bibliographic information systems.

Members of the group are: Maja Žumer (Professor), Jan Pisanski (Assistant Professor), Polona Vilar (Assistant Professor), Viktor Harej (Doctoral Student), Tanja Merčun (Doctoral Student) and Katarina Švab (Doctoral Student).