

FRBRizing MARC records with the FRBR Display Tool

by Jodi Schneider

May 16, 2008

Introduction

Who Should Read this Report

This paper is aimed at three audiences:

- Administrators who need to understand what FRBR is, how it benefits library users, and why trends towards increased digitization are making FRBRization even more important
- Researchers interested in automatic methods for FRBRizing MARC records
- Users of the FRBR Display Tool

What is FRBR? A Conceptual Model for Library Catalogs

A Conceptual Model Based on User Needs

FRBR, the Functional Requirements for Bibliographic Records, is a conceptual model, establishing the purposes, uses, and requirements for bibliographic records. FRBR sets out “a clearly defined, structured framework for relating the data that are recorded in bibliographic records to the needs of the users of those records.”¹ We need such a framework because “the catalog of the future, and even the catalog of the present, isn't like the catalog of the past.”² Automation, changes in user expectations, increasing opportunities for national and international collaboration, and economic pressures are driving changes in cataloging practices. FRBR can be seen as part of the ongoing evolution of cataloging.³

The FRBR User Tasks

The group preparing the FRBR model considered “a broad spectrum [of catalog users], including not only library clients and staff, but also publishers, distributors, retailers, and the providers and users of information services outside traditional library settings.”⁴ The FRBR Report, published in 1997 by the Standing Committee of the IFLA Section on Cataloguing, establishes 4 user tasks:

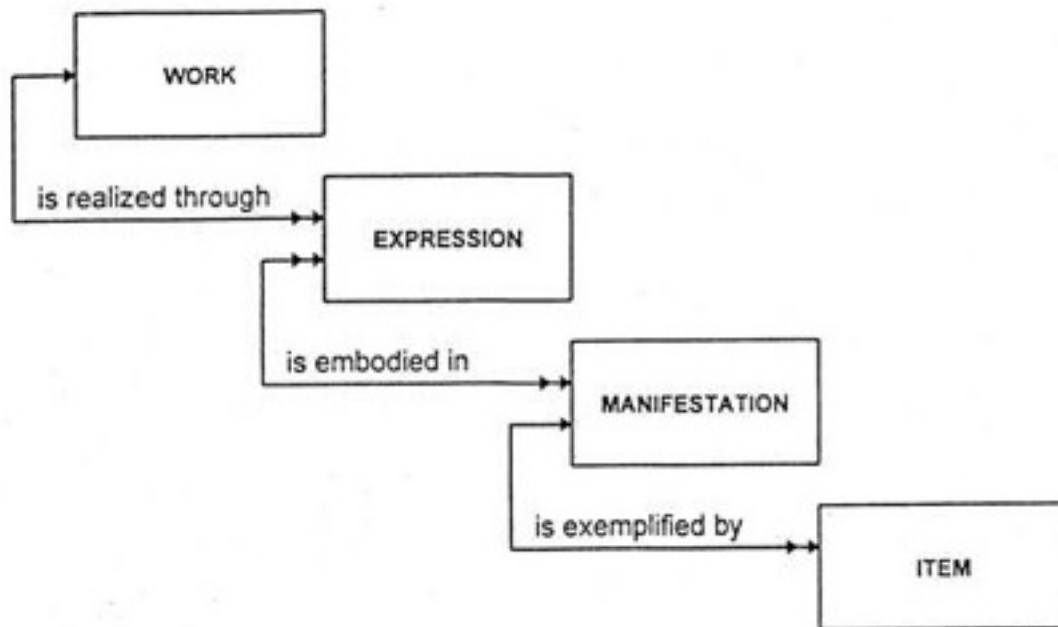
- To *find* entities that correspond to the user's stated search criteria (i.e., to locate either a single entity or a set of entities in a file or database as the result of a search using an attribute or relationship of the entity);
- To *identify* an entity (i.e., to confirm that the entity described corresponds to the entity sought or to distinguish between two or more entities with similar characteristics);

- To *select* an entity that is appropriate to the user's needs (i.e., to choose an entity that meets the user's requirements with respect to content, physical format, etc., or to reject an entity as being inappropriate to the user's needs);
- To *acquire or obtain access* to the entity described (i.e., to acquire an entity through purchase, loan, etc., or to access an entity electronically through an online connection to a remote computer).⁵

A fifth task, navigation, is added by some experts: "To *navigate* a bibliographic database (that is, to find works related to a given work by generalization, association, and aggregation; to find attributes related by equivalence, association, and hierarchy)".⁶

FRBR: An Entity-Relationship Model

FRBR presents an entity-relationship model of bibliographic records. The Group 1 Entities, which represent "the products of intellectual or artistic endeavor", are central for our purpose.⁷ These entities are Work, Expression, Manifestation, and Item:



They are well-described by Pat Riva:

These entities are the focus of the bibliographic record and the heart of the model. The first two are entirely abstract and reflect intellectual or artistic content:

- Work (a distinct intellectual or artistic creation)
- Expression (the intellectual or artistic realization of a Work)

The next two are (more or less) concrete and reflect physical form (although physical should not be taken too literally, as it includes remote electronic resources):

- Manifestation (the physical embodiment of an expression of a Work)
- Item (a single exemplar of a manifestation).⁷

What Library Users Want

Users want to find “stuff”. They use particular physical or digital Items, but in many cases related Items, Expressions, or Manifestations would also serve their needs. Here's an example: “A library user may ask a question like ‘Do you have Seamus Heaney’s translation of *Beowulf*?’ (a request for an expression) or ‘Do you have Stephen Hawking’s *A Brief History of Time*?’ (a request for a work). Upon further questioning, many such users do not have a particular item or even a particular manifestation in mind. What they are interested in are abstractions – the content, either at the expression or the work level”.⁸ FRBR provides a model for grouping related content.

Library catalogs typically have entries at the Manifestation level, because for most purposes, any copy (Item) of a book is acceptable. Additional collation can help both users who may accept alternatives, as well as users with specific requirements for a particular format, edition, language, or copy. Collating non-identical Manifestations would bring together, for example, a VHS tape and a DVD with the same content. Collating Expressions would bring translations together with the original. Collating Works would bring together related creative endeavors, such as adaptations of a Work into a movie or a children's book.

Studies of how people conceive of and group materials^{9 10} suggest that FRBR is a promising approach to summarizing library holdings. Grouping materials according to the FRBR Group 1 Entities may help users *find*, *identify*, and *select* materials. A library catalog ought to collate similar materials which are likely to be interchangeable for most users, while retaining detailed information to allow users with more particular needs to determine the differences between these similar materials.

Improving Library Catalogs

FRBRized catalogs would allow users to explore library holdings in sensible groupings, which cluster similar materials. “One of the great advantages of FRBR-based displays is that long displays may be made much shorter, enhancing the intelligibility and browsability of results.”¹¹ In catalogs which primarily collate Items, the *identification* and *selection* tasks are compromised by a lack of collation; FRBR might bring improvements. For instance, FRBRized displays might be more intuitive when they are closer to user-conceived groupings.

FRBRization is particularly beneficial in some areas, such as music. Dickey quotes a study of musical scores¹² that found that “more than 94.8% [of the scores in the sample] exhibited at least one bibliographic relationship with another entity in the collection.”¹³ Making these relationships explicit can help users *find*, *identify*, and *select* materials more easily in a FRBR-based catalog. While collation is important, exploration at the Item level will remain important for some purposes. For example, while early printed books may be grouped into Expressions and Manifestations, individual copies retain a striking amount of uniqueness, and their differences are important to scholars.

FRBRizing the Catalog

What does FRBRizing the Catalog Mean?

FRBRizing the catalog involves collating MARC records of similar materials. FRBRization brings together sets of Works, Expressions, or Manifestations, rather than just sets of Items. Researchers in several countries, including Korea^{14 15}, Norway^{16 17}, and Portugal¹⁸, are experimenting with FRBRizing their national catalogs on a large scale. I have been experimenting with an existing tool aimed at creating an access/display layer for MARC21 records, the North American flavor of MARC.

What is the FRBR Display Tool?

The FRBR Display Tool¹⁹ is an experimental tool distributed by the Network Development and MARC Standards Office of the Library of Congress. The tool generates FRBRized MODS records from MARC records. MODS is a natural choice for the output format it was designed both to accommodate MARC21 and to facilitate a FRBRized record structure. MODS uses current standard data storage and exchange idioms, such as XML data exchange and semantic metadata tag names.²⁰

Display is a major focus of the FRBR Display Tool: it displays materials sorted by Work, form, and edition with links into the Library of Congress OPAC for further detail. While MODS can store any MARC data field, the Tool does not transfer the entire MARC record into MODS but rather extracts pertinent details needed for creating FRBRized Work sets and for grouping and displaying records.

Overview of How the FRBR Display Tool Works

The FRBR Display Tool is primarily an XSLT processing system. The FRBR Display Tool uses MARC4J, an opensource Java toolkit, to convert MARC records stored in ISO2709 binary format to MARCXML. (MARCXML is another possible starting point for toolkit users.) Next the MARC data is transformed through a sequence of 4 XSLT transformations, which are based in part on the MARCXML toolkit. In the fourth transformation, FRBRized MODS XML is styled into HTML. The final output is two files: this HTML file and the XML file from which it was generated.

The Pipeline

The FRBR Display Tool consists of a shell script or a Windows batch file and supporting files. Here is a shell script, which can be obtained from the most recent version of the toolkit or at <http://jodischneider.com/pubs/supinfo/2008MayFRBRizingMARC/marc2frbr.sh>

```
#!/bin/sh
if [ -z $1 ] || [ -z $2 ]
then
echo "Usage: $0 marcfilerecord outputstem"
echo "marcfile.mrc will be transformed into 2 files:"
echo "  outputstem.xml"
echo "      Will contain the FRBR xml representing the marc records"
echo "  outputstem.html"
```

```

    echo "      Will contain the HTML display of the FRBR xml"
    exit
fi

echo Transforming $1 to MARCXML ...
java -cp marcxml-20080423.jar:marc4j.jar gov.loc.marcxml.MARC2MARC21slim $1
    slimfrbr.xml
echo Transforming the MARCXML into FRBR XML and saving to $2.xml ...
java -jar saxon9.jar -warnings:silent -u -o clean.xml slimfrbr.xml http://
www.loc.gov/standards/marcxml/frbr/v2/clean.xsl
java -jar saxon9.jar -warnings:silent -u -o match.xml clean.xml http://
www.loc.gov/standards/marcxml/frbr/v2/match.xsl
java -jar saxon9.jar -warnings:silent -u -o $2.xml match.xml http://
www.loc.gov/standards/marcxml/frbr/v2/FRBRize.xsl
echo Transforming the FRBR XML into HTML and saving to $2.html ...
java -jar saxon9.jar -warnings:silent -a -o $2.html $2.xml
echo Complete

```

The process can be summarized as: MARC --> MARC XML --> clean --> match --> FRBRized MODS --> html display

The output from each step: (input.mrc) --> slimfrbr.xml --> clean.xml --> match.xml --> output.xml --> output.html

The XSLT from each step: (marc4j.jar) --> <http://www.loc.gov/standards/marcxml/frbr/v2/clean.xsl> --> <http://www.loc.gov/standards/marcxml/frbr/v2/match.xsl> --> <http://www.loc.gov/standards/marcxml/frbr/v2/FRBRize.xsl> --> <http://www.loc.gov/standards/marcxml/frbr/frbr2html.xsl>

This final transformation, which serializes to an html display format, is called inside the FRBRize.xsl stylesheet, as an xsl:processing-instruction. Throughout the process, the MARCXML utility transformations <http://www.loc.gov/marcxml/xslt/MARC21slimUtils.xsl> are used via an xsl:import.

Versions of the FRBR Display Tool

Version 1 - Released by July 2003,²¹ the tool was created based on Thomas Delsey's functional analysis of MARC 21²². Version 1, was still current as of 2005.²³

Version 2 - By May 2007, when Denton documented his experience with the tool, version 2 was available.

New version (in progress)- As a result of my query to the MARC FRBR listserv, the tool was updated to take advantage of character set processing from the latest version of MARC4J, and a UNIX shell script was added²⁴. A temporary version is available at <http://memory.loc.gov/natlif/cred/marcFRBR.tar.gz> while the update makes its way through the LC's formal publishing process to the official page²⁵.

Sample files

Sample files created from the new version of the tool can be downloaded from my examples directory <http://jodischneider.com/pubs/suppinfo/2008MayFRBRizingMARC/> .

Running the tool on mahler.marc produces results in html and XML (see mahler.html and mahler.xml in the examples directory) as well as three intermediate files slimfrbr.xml, clean.xml, and match.xml, which are normally deleted at runtime.

Prior Publications about the FRBR Display Tool

Official Publications Relating to the FRBR Display Tool

Several publications of the MARC Standards Office relate to the tool. The most important of these is a paper from the MARC Standards Office describing the tool.²⁵ An earlier paper describes general techniques for manipulating MARCXML.²⁶ Finally, a presentation by the Chief of the MARC Standards Office gives further details about the tool, including explanations of several types of collation failures.²⁷

Publications by users of the FRBR Display Tool

To date, I have found 1 master's dissertation,²⁸ 1 paper,²⁹ 1 presentation,³⁰ of a student project,³¹ and 1 series of blog posts focusing on the FRBR Display Tool.³²

Denton's blog posts served as my introduction to the FRBR Display Tool and influenced my work throughout the project. Along with publications of the MARC Standards Office, Denton's series gives the most detailed information. Therefore, I summarize his work here.

In May 2008, William Denton reported on his experiments with the FRBR Display Tool in a series of 9 blog posts at the FRBR Blog³², named after the primary example, *Pride and Prejudice*. The experiment started with a single manifestation of *Pride and Prejudice*, with ISBN 0192833553. Denton queried thingISBN and xISBN services to retrieve 792 ISBN for related Works. He then used a ruby script to search Z39.50 servers for MARC records for these ISBNs, and managed to retrieve 383 records from open servers. According to the marclint from the MARC/Perl library, these records contained 249 errors. Next he wrote a shell script modifying the FRBR Display Tool (version 2) to run on his linux box, and ran the script on the sample mahler.mrc data from the Library of Congress.

Next Denton tried running the script on the Pride and Prejudice MARC records pride-and-prejudice.mrc. Data errors caused the FRBR Display Tool script to fail, and Denton set to hacking on the records to get them to run through the tool. In order to delete a particular record, he first wanted to parse results using MARC/Perl, but then Denton discovered character encoding errors, which had come from a bug in the ruby script he'd used to collect records. He continued working with the data he had while re-retrieving records from Z39.50 with an updated ruby script. In the meantime, he deleted problem records and removed some fields of remaining records. At this point, he was able to get through the first step of the FRBR Display Tool pipeline, and had intermediate results in MARCXML.

Unfortunately, these intermediate results had two undefined entities, `` and ``, which caused processing problems. After removing these entities, Denton was able to get the first XSLT transformation to run. However, running the second one gave a stack trace with Saxon 7 and he switched to Saxon 8.9. At this point, the final XSLT still failed because in some records a space appeared in the 245 Title Statement Field³³ first or second indicators where a number was expected. Finally, after editing the XSLT, Denton was able to get output, shown in <http://www.frbr.org/eg/pp/pp.html>. The series concluded with a discussion of running the FRBR Display Tool on Harry Potter records with a smoother, not yet documented process, which required removing characters but not wiping fields. He also comments that “The Library of Congress’s FRBR Display Tool needs to be robustified so it can withstand heinous MARC records, but it’s still a fine piece of work.”³⁴

Experiments with the Library of Congress's FRBR Display Tool, based on Denton's work, formed the practical work for this paper.

My work with the FRBR Display Tool

My original goal was to repeat William Denton's experiments. As I used the FRBR Display Tool, I discovered problems, both with the records I started with and with the usability of version 2 of the tool. My interest expanded to how others have automated the MARC to FRBR conversion.

Problems, Challenges, & Discoveries

Charsets

MARC may be encoded in one of two character sets, MARC-8 or UTF-8, as expressed in the Leader 9. While UTF-8 is widely supported, MARC-8^{35 36} is almost entirely unknown outside the library programming community. Version 2 of the FRBR Display Tool used an early version of MARC4J to transfer records into MARCXML; newer versions of MARC4J have improved charset handling.

Version 2 of the Tool worked smoothly on Windows. Shell scripts from William Denton, based on the Windows batch file, returned results along with character encoding errors on the Mac OS X and linux machines I used. Here's a typical error, which comes from the old version of MARC4J used:

```
Transforming nahler.mrc to MARCXML ...
Character not found: cb in Code Table: 45
Character not found: 2d8 in Code Table: 45
Character not found: c8 in Code Table: 45
Character not found: cc in Code Table: 45
Character not found: cb in Code Table: 45
Character not found: cf in Code Table: 45
Character not found: 2d8 in Code Table: 45
Character not found: 2c6 in Code Table: 45
Character not found: 2c6 in Code Table: 45
Character not found: 2830 in Code Table: 45
Character not found: 2830 in Code Table: 45
Character not found: 2830 in Code Table: 45
Character not found: cc in Code Table: 45
Character not found: cc in Code Table: 45
Character not found: 2c6 in Code Table: 45
Character not found: 2830 in Code Table: 45
Character not found: 2830 in Code Table: 45
Transforming the MARCXML into FRBR XML and saving to examples/nahler.xml ...
Error on line 1 column 281 of file:///Users/jodi/Desktop/2008%20Spring%20LIS590DPL%20document%20processing%20in%20LIS/FRBRproject/frbrfolder/elinfrbr.xml:
  Error reported by XML parser: The prefix "xsi" for attribute "xsi:schemaLocation" associated with an element type "collection" is not bound.
Transformation failed: Run-time errors were reported
Transforming the FRBR XML into HTML and saving to examples/nahler.html ...
Complete
```

Characters are not found in the Extended Latin Codetable and an xsi namespace was not bound to its proper URI. The FRBR Display Tool was updated in April in response to my question on the MARCFRBR listserv regarding these problems. The newest version includes UNIX shell scripts as well as Windows batch files, properly binds the namespace and correctly handles both MARC-8 and UTF-8 input.

Data consistency and data quality

The FRBR Display Tool is optimized for Library of Congress records. For instance, the tool is based on MARC21 and assumes a certain amount of authority control. Some normalization routines are also based on the Library of Congress source data; for example, "[from old catalog]" is removed from certain fields in the clean XSLT. Collections harvested from multiple sources, such as Denton's pride-and-prejudice.marc collection, may prove more difficult to process due to variations in the MARC tags or field values used. For instance, Hegna and Murtomma³⁶ list differences between Norwegian and Finnish MARC formats, NORMARC and FINMARC.

Besides differences in the MARC tags used, there may also be differences in the rulesets (e.g. AACR2) applied. Local changes are also possible, so there may be inconsistencies in the use of certain fields. For instance, the 700 field in many catalogs becomes a miscellany, since those fields are typically indexed for search, retrieval, and display. Finally, there are non-intentional differences. Every catalog has spelling errors, and most seem to have differences in authority headings, since the headings are updated more often (for example to add birth and death dates) than many catalog records.

Data Quality Example

Sometimes obvious errors lead one to discover less obvious errors. For example, running the new FRBR Display Tool on `pride-and-prejudice-cleaned-chars.marc`, fails because the `slimfrbr.xml` contains undefined entity references, `` and ``; which XML flags as invalid in this context. The error message says “Character reference ‘’ is an invalid XML character.”:

```
jodi-schneiders-computer:zzz jodi$ ./marc2frbr.sh pride-and-prejudice-cleaned-chars.marc examples/pp
Transforming pride-and-prejudice-cleaned-chars.marc to MARCXML ...
Transforming the MARCXML into FRBR XML and saving to examples/pp.xml ...
Error on line 15881 column 46 of file:/Users/jodi/Desktop/2008%20Spring%20LIS5900PL%20Document%20Processing%20in%
20LIS/FRBRproject/zzz/slimfrbr.xml:
  SXP0003: Error reported by XML parser: Character reference "&#31;" is an invalid XML character.
Transformation failed: Run-time errors were reported
Transforming the FRBR XML into HTML and saving to examples/pp.html ...
Complete
```

Looking at `slimfrbr.xml`, we find invalid characters:

```
<datafield tag="080" ind1=" " ind2="&#31;">
  <subfield code="a">Austen, Jane</subfield>
  <subfield code="a">Pride and prejudice</subfield>
  <subfield code="h">&amp;lt;Ljudupptagning&amp;gt; /</subfield>
  <subfield code="c">Jane Austen</subfield>
  <subfield code="a">Ashland, Or.</subfield>
  <subfield code="b">Blackstone Audiobooks</subfield>
  <subfield code="c">2005</subfield>
  <subfield code="a">Huddinge</subfield>
</datafield>
```

As it turns out, a number of records from that source are bad. Tag 080 holds an “n” in indicator 1 and a “d” in indicator 2:

```
<datafield tag="080" ind1="n" ind2="d">
  <subfield code=" " >prejudice /</subfield>
  <subfield code="c">Jane Austen ; with illustrations by George Allen</subfield>
  <subfield code="a">Ware, Hertfordshire</subfield>
  <subfield code="b">Wordsworth</subfield>
  <subfield code="c">1994</subfield>
  <subfield code="a">Wordsworth classics</subfield>
  <subfield code="a"> land</subfield>
</datafield>
```

However, the indicators should be " " to indicate they're undefined. And in fact, tag 080 should hold a call number, rather than publication information!³⁷ Here is a more typical value:

```
<datafield tag="080" ind1=" " ind2=" " >
  <subfield code="a">821.111-3</subfield>
</datafield>
```

To get the script to run at this point, various solutions are possible. We could declare the entities, remove the (incorrect) 080 indicators, or even the entire fields, which happen to not be used in the processing or display of the FRBR Display Tool. Another solution would be to work on validation and error handling. Knowing the data, and how it is used, will help!

Simplifying the Process

One way to simplify the process would be to start with MARCXML records, which would eliminate encoding problems and might improve data quality. Working with a single catalog could also increase the uniformity in the data, which would bring improvements. Data cleaning and data validation could also become a preprocessing step. Preliminary work with the MARCXML validation stylesheet suggest that flagging MARC errors is promising. Further possibilities are given in the future work section which follows.

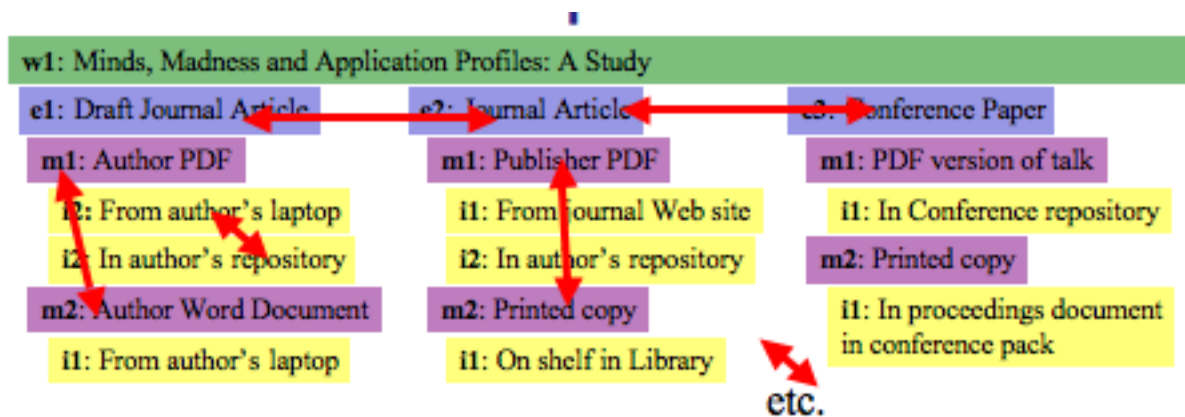
Ideas for Future Work

- Clean and validate data, using existing tools such as the MARC DTD,³⁸ MARCXML Validation Stylesheet,³⁹ or other validation tools from the Library of Congress,⁴⁰ and other institutions with significant MARC processing apparatus (e.g. OCLC, UC System, Library of Congress, other national libraries).
- Scale the FRBR Display Tool to run on a typical recordset, and analyze results.
- Compare FRBRizing algorithms and methods on well-understood test data.
- Put data into a Storage, Search and Retrieval Environment such as Solr⁴¹ or a Solr-based next-generation catalog such as Blacklight⁴², VuFind⁴³, or the fac-back-opac⁴⁴. (Consider modifying the MARCImporter⁴⁵ to import and index MODS. Determine what XSLT processing models are possible using the XSLTResponseWriter⁴⁶ or the forthcoming (in version 1.3) DataImportHandler⁴⁷.)
- Put data into a navigation-focused environment using Topic Maps (XTM) or consider a tool that primarily improves display such as Exhibit (or other SIMILE tools)⁴⁸ or Flamenco⁴⁹, or a custom XML to HTML stylesheet.
- Improve output display. Consider a tree structure, as shown in Tillett⁵⁰ or a concept map structure, as shown in LeBoeuf⁵¹
- Do user testing of FRBRized catalogs
- Incorporate related tasks such as record collection. While the FRBR Display Tool might be used by a single institution already in possession of records, adding record collection at the beginning of the pipeline might be useful for some users. For instance, chain with existing Z39.50 and SRU/SRW tools, for instance with Yaz.⁵²

Related Considerations for the Future of the Catalog

FRBR in the Networked Era

FRBR becomes even more important in the digital networked environment. As we digitize items, we create surrogates that are “the same” for certain purposes. The networked digital environment also increases the amount of data available from an individual's desktop. And as data proliferates, we commonly find multiple surrogates for the same artifact. So in the digital environment, substitutions of similar items is even more likely. Peter Brantley talks about substituting one photo of a place for another: “One of the things not extensively discussed in the library and archival communities, yet which is increasingly implicit in the larger digital environment, is the possibility that even if an original, source object is not re-locatable or re-discoverable... often an acceptable surrogate can be found. Further, it is increasingly likely that there is no real ‘original’, but rather a stream of mildly differing replicas. These are surrogates of type, and form.”⁵³ Consideration of the Items, Manifestations, and Expressions associated with a single journal article illustrates⁵⁴ this point:



Users may, for instance, substitute a preprint in the author's repository for a publisher PDF behind a paywall, usually (but not always) without loss of information.

The digital networked environment also gives users more entries into bibliographic universe, for example through booksellers and search engines. While those environments do not always provide exceptional bibliographic control, they are typically easy to use, and in a full-text environment, the cost (in time as well as money) of *obtaining* an item may be low. Depending on the environment, the other user tasks may be harder or easier than in a typical catalog. For example, *finding* something acceptable is often (though not always) easy in a search engine such as Google, while an information-rich environment aids Amazon users in *selecting* the Item desired. Improved, readily-available alternatives lend urgency to the desire to improve library catalogs.

Identifiers

Identifiers are extremely important for network-level services. As Kristin Antelman says: “Documents do not need to be described to be referenced in a networked world; they must be identified. An inherently descriptive element, such as title, cannot meet the requirements of a network identifier.”⁵⁵

Linked data such as the new LCSH linked data service⁵⁶ will become increasingly important as we seek identifiers not only for Works but also for (persistently linkable) subject authorities. Identifiers would allow a multi-level record structure to function as a unified record.⁵⁷

More Research is Needed

Although FRBR has been studied for over 10 years, most prototypes are still small-scale experimental systems, or experiments with single-country national catalogs. FRBR has not been sufficiently tested, especially with regard to user studies of working systems. FRBRization of catalogs is still a research area, rather than an area of practice.

For a survey of FRBRized prototype and production systems, see the recent article, *From a Conceptual Model to Application and System Development*, especially its tables.⁵⁸ IFLA's bibliography of FRBR-related publications⁵⁹ is also useful for surveying the field.

The Library of Congress Report on the Future of Bibliographic Control devotes Section 4.2 to FRBR. The section stresses that FRBR is still only a conceptual model which has not been implemented on a large scale: "The library community is basing its future cataloging rules on a framework that it has only barely begun to explore. Until carefully tested as a model for bibliographic data formation for all formats, FRBR must be seen as a theoretical model whose practical implementation and its attendant costs are still unknown." The current environment does not adequately support FRBR: "the impact of the FRBR model on cataloging practice and on the machine-readable bibliographic record has not been extensively explored. There is no standard way to exchange Work-based data, and no cataloging rules that yet support the creation of records using the FRBR model." Further attention to sharing our data outside the library community is also needed, according to the report: "LC and the library community need to find ways of 'releasing the value' of their rich historic investment in semantic data onto the Web."⁶⁰

As new cataloging rules (RDA) are developed, it is particularly important to consider how to best fulfill users' needs. "More research and development is needed to explore the application and implementation of FRBR for future retrieval systems that better support information seeking."⁶¹ User studies will be particularly important in determining the extent to which systems meet the FRBR user objectives: *find, identify, select, obtain, and navigate*.

FRBRized catalogs

FRBRizing catalogs could bring improvements to the overly-long, flat displays currently in use. FRBR may be useful particularly for reformatted materials and taming content proliferation of popular Works. Further research about FRBRizing MARC records, and particularly user testing of prototype systems, is needed.

Acknowledgements

Thanks to Stephen Linhart, Wendell Piez, Clay Redding, Jon Wheeler, Ryan Wick, Karen Wickett, and to UIUC classmates in spring 2008 sections of LIS590ROL and LIS590DPL.

Endnotes

- [1] Section 2.1: Objectives of the Study, IFLA Study Group on the Functional Requirements for Bibliographic Records. 2008. Functional requirements for bibliographic records: final report, amended and corrected for 2008. http://www.ifla.org/VII/s13/frbr/frbr_current2.htm
- [2] Wendell Piez, personal communication, May 10, 2008.
- [3] William Denton. 2007. FRBR and the History of Cataloging. In: Understanding FRBR: What It Is and How It Will Affect Our Retrieval. Ed. Arlene G. Taylor. Westport, Connecticut: Libraries Unlimited. Denton traces the history of 4 ideas in the evolution of cataloging, which prepare the way for FRBR: logical principles, user needs, the idea of the work, and standardization and internationalization. <http://hdl.handle.net/10315/1250>
- [4] Section 1.2: Approach, IFLA Study Group on the Functional Requirements for Bibliographic Records. 2008. Functional requirements for bibliographic records: final report, amended and corrected for 2008. http://www.ifla.org/VII/s13/frbr/frbr_current1.htm
- [5] Section 6.1: Mapping Attributes and Relationships to User Tasks, IFLA Study Group on the Functional Requirements for Bibliographic Records. 2008. Functional requirements for bibliographic records: final report, amended and corrected for 2008. http://www.ifla.org/VII/s13/frbr/frbr_current6.htm
- [6] Page 20, Elaine Svenonius. 2007. The Intellectual Foundation of Information Organization. Boston: MIT Press.
- [7] Page 8, Pat Riva. 2007. Introducing the Functional Requirements for Bibliographic Records and Related IFLA Developments. Bulletin of the American Society for Information Science and Technology 33(6): 7-11. <http://www.asis.org/Bulletin/Aug-07/riva.html>
- [8] Page 11-12, Allyson Carlyle and Lisa M. Fusco. 2007. Understanding FRBR as a Conceptual Model: FRBR and the Bibliographic Universe. Bulletin of the American Society for Information Science and Technology. 33(6): 264-273. http://www.asis.org/Bulletin/Aug-07/carlyle_fusco.html
- [9] Page 8, Allyson Carlyle. 1999. User categorisation of works: toward improved organisation of online catalogue displays. Journal of Documentation 55(2). 184-208. <http://www.emeraldinsight.com.proxy2.library.uiuc.edu/Insight/ViewContentServlet?Filename=Published/EmeraldAbstractOnlyArticle/Articles/2780550204.html>
- [10] Maja Žumer. 2005 August 12. FRBR as User's Model. Presentation at Bibliotheca Universalis: How to organize chaos? Satellite Meeting to the 71st World Library and Information Congress, Järvenpää, Finland, organized by the Finnish Library Association. <http://www.kaapeli.fi/%7Efla/frbr05/ŽumerFRBR2005.pdf>
- [11] Page 14, Allyson Carlyle and Lisa M. Fusco. 2007. Understanding FRBR as a Conceptual Model: FRBR and the Bibliographic Universe. Bulletin of the American Society for Information Science and Technology. 33(6): 264-273. http://www.asis.org/Bulletin/Aug-07/carlyle_fusco.html
- [12] Page 238, 251, Sherry L. Velluci. 1997. Bibliographic Relationships in Music Catalogs. Lanham, Md: Scarecrow.

- [13] Page 23, Timothy J. Dickey. 2008 March. FRBRization of a Library Catalog: Better Collocation of Records, Leading to Enhanced Search, Retrieval, and Display. 27(1): Information Technology and Libraries.
- [14] Sam Oh. 2008. MARC, FRBR and RDA. [Presentation]. Topic Maps 2008. http://topicmaps.com/tmc/presentation.jsp?conf=TM2008&id=TM2008~Sam_Oh~MARC_FRBR_and_RDA
- [15] Jane Cho. 2006. Library Collections, Acquisitions & Technical Services. 30:202-213. <http://dx.doi.org/10.1016/j.lcats.2006.08.003>
- [16] Knut Hegna and Eeva Murtomaa. Data mining MARC to find: FRBR? Paper presented at the 68th IFLA General Conference and Council, Glasgow. <http://folk.uio.no/knuthe/dok/frbr/datamining.pdf>
- [17] Trond Aalberg. 2006. A Process and Tool for the Conversion of MARC Records to a Normalized FRBR Implementation. ICADL. LNCS 4312, 283-292.
- [18] Nuno Freire, José Borbinha, and Pável Calado. 2007. ICADL. LNCS 4822, 267-276.
- [19] Library of Congress Network Development and MARC Standards Office. Functional Analysis of the MARC 21 Bibliographic and Holdings Formats: FRBR Display Tool. Version 2.0 <http://www.loc.gov/marc/marc-functional-analysis/tool.html>
- [20] Page 85, Sally H. McCallum. 2004. An introduction to the Metadata Object Description Schema (MODS). Library Hi Tech. 22(1): 82-88. MODS has 19 top-level elements. 17 can be organized by FRBR Group 1 Entities, while the remaining elements provide information about the record itself and allow for extensibility without recourse to an application profile.
- [21] Jacqueline Radebaugh. 2003 July 23. Email from jrad@loc.gov to XML4J. Subject: FRBR Display Tool. Message-ID: <sf1e96eb.042@loc.gov> <http://lists.webjunction.org/wjlists/xml4lib/2003-July.txt>
- [22] Thomas Delsey. 2001. Functional Analysis of the MARC 21 Bibliographic and Holdings Format. Network Development and MARC Standards Office, Library of Congress. http://www.loc.gov/marc/marc-functional-analysis/original_frbr.html
- [23] Corey Keith and Jacqueline Radebaugh. 2005. FRBR display tool. Cataloging & Classification Quarterly. 39(3/4): 271-283.
- [24] 2008 April 23. Clay Redding. Email to the MARCFRBR listserv. <http://www.loc.gov/marc/marc-functional-analysis/marcfrbr.html>
- [25] FRBR Display Tool (Official Webpage) <http://www.loc.gov/marc/marc-functional-analysis/tool.html>
- [26] Corey Keith 2004. Using XSLT to manipulate MARC metadata. Library Hi Tech. 22(2): 122-130.
- [27] Sally McCallum. 11-12 August 2005. The FRBR Tool of the Library of Congress. Presentation at Bibliotheca Universalis: How to organize chaos? Satellite Meeting to the 71st World Library and

Information Congress, Järvenpää, Finland, organized by the Finnish Library Association. http://www.fla.fi/frbr05/McCallumHelsinki_frbrtool.pdf

[28] Fernanda P. Moreno. 2006. FRBR - Requisitos Funcionais para Registros Bibliográficos: um estudo no catálogo da Rede Bibliodata. Master's degree dissertation, Information Sciences, Universidade de Brasília (Brasil). (in Portuguese). English Abstract: In the context of library studies and Information Science, descriptive representation a series of specialist meetings trying to establish standards to bibliographic description. Originated by one of international meetings, FRBR - Functional Requirements for Bibliographic Records, published by International Federation of Library Association and Institutions IFLA in 1998, it started a new interpretation of the bibliographic records was introduced, reorganizing their elements and presenting concepts and definitions of entities, attributes and relationships. The study became reflexes of the model FRBR in an electronic catalog, on-line, nationwide, through the examination of bibliographic records in format Machine Readable Cataloging – MARC, in order to systematize the relationship between the elements. Using appropriate literature and characteristics regarding potentials to realize the study, had been selected registers according to Catálogo Coletivo da Rede Bibliodata (Bibliodata Network Public Collection Catalog), and characterizing a study case. To illustrated the reflex of the model on the records, was used the conversional tool, FRBR Display Tool, from Library of Congress (LC), and besides the use of auxiliary tools when was made necessary. In order to complete the study, a few issues encountered in the literature were brought back as a manner to reach the goal of systematizing the relationships, given the relative failure of the tool under this aspect. Analytics stages comprise sample registers, notoriously descriptive and of the register converted in document that contains register model. In this phase, that is impossible to discuss every case, some typical cases for each type of entitle purpose in FRBR. The results conducting to a necessity of register normalization and it is suggested tags MARC linking a better visualization of the model. Absence of discussions due to the use of models and international standards is considering as a potential trouble and there are paths indicated for the ways of future researches that would be realize, and one contribution to construct theorist corpus of the area. <http://eprints.rclis.org/archive/00006330/>

[29] Bernal Rajapatirana and Roxanne Missingham. 2005 September. The Australian National Bibliographic Database and the Functional Requirements for the Bibliographic Database (FRBR). The Australian Library Journal 54.1 <http://alia.org.au/publishing/alj/54.1/full.text/rajapatirana.missingham.html>

[30] Miao Jin & Nashaat I. Sayed. FRBR-izing library collections. 2004. Draft of Presentation for Mississippi Library Association Conference Poster Session. Jackson, Mississippi. http://orca.st.usm.edu/~nsayed/temp_review/articles/frbr/Poster/FRBR%20presentation-draft3.ppt

[31] Nashaat I. Sayed. 2004. Project Plan: FRBR-izing library collections. http://orca.st.usm.edu/%7Eensayed/temp_review/stuff/frbr/research/A%20Conceptual%20Look%20at%20the%20OCLC%20FRBR%20Display%20Tool.doc

[32] William Denton. 2008 May. FRBR Blog. Pride and Prejudice [blog posts]. <http://www.frbr.org/categories/examples/pride-and-prejudice>

[33] Library of Congress Network Development and MARC Standards Office. MARC Standards: 245 - Title Statement (NR). <http://www.loc.gov/marc/bibliographic/concise/bd245.html>

- [34] William Denton. 2008 May. FRBR Blog. Pride and Prejudice 8: Comments [blog post]. <http://www.frbr.org/2007/05/18/pride-and-prejudice-8-comments>
- [35] Library of Congress. December 2007. MARC 21 Specifications for Record Structure, Character Sets, and Exchange Media CHARACTER SETS AND ENCODING OPTIONS: Part 2. MARC-8 Encoding Environment <http://www.loc.gov/marc/specifications/speccharmac8.html>
- [36] Library of Congress. Codetables for MARC-8. <http://www.loc.gov/marc/specifications/codetables.xml>
- [37] Library of Congress Network Development and MARC Standards Office. MARC Standards: 080 - Universal Decimal Classification Number (R) <http://www.loc.gov/marc/bibliographic/bd080.html>
- [38] Library of Congress Network Development and MARC Standards Office. MARC in XML. <http://www.loc.gov/marc/marxml.html#marcdtd>
- [39] Library of Congress Network Development and MARC Standards Office. MARCXML Validation Stylesheet. <http://www.loc.gov/standards/marcxml/validation/validate.xsl>
- [40] Library of Congress Network Development and MARC Standards Office. [Directory of Validation Tools] <http://www.loc.gov/standards/marcxml/validation/>
- [41] Solr Lucene opensource search engine <http://lucene.apache.org/solr/>
- [42] Blacklight: Faceted browsing with Solr and Ruby on Rails <http://blacklight.rubyforge.org/>
- [43] VuFind: library resource portal with Solr and PHP <http://vufind.org/>
- [44] Fac-back-OPAC: Faceted Backup OPAC using Solr and Django <http://code.google.com/p/fac-back-opac>
- [45] Blacklight MARCImporter http://blacklight.rubyforge.org/svn/blacklight_importer/
- [46] XSLTResponseWriter for Solr <http://wiki.apache.org/solr/XsltResponseWriter>
- [47] DataImportHandler for Solr <http://wiki.apache.org/solr/DataImportHandler>
- [48] Exhibit: SIMILE's Javascript Publishing Framework <http://simile.mit.edu/exhibit>
- [49] Flamenco python-based search interface framework <http://flamenco.berkeley.edu>
- [50] Barbara B. Tillett. 2004. What is FRBR? (Functional Requirements for Bibliographic Records). Presentation at Back to the Future: Understanding the Functional Requirements of Bibliographic Records Model (FRBR) and its Impact on Users, OPACS, and Knowledge Organization Preconference, Presented by the ALCTS Cataloging and Classification Section's (CCS) Committee on Cataloging: Description & Access (CC:DA) at American Library Association. <http://www.ala.org/ala/alcts/education/presentations/Tillett.pdf>
- [51] Patrick LeBoeuf. 2005. Is it possible to organise all information? A library view. Presentation at Presentation at Bibliotheca Universalis: How to organize chaos? Satellite Meeting to the 71st World Library and Information Congress, Järvenpää, Finland, organized by the Finnish Library Association. http://www.kaapeli.fi/~fla/frbr05/2005_Järvenpää_LeBoeuf.pdf

- [52] Yaz. <http://www.indexdata.dk/yaz/>
- [53] Peter Brantley. 2008 March 19. Digital Steel [post on shimenawa blog]. http://blogs.lib.berkeley.edu/shimenawa.php/2008/03/19/digital_steel
- [54] Peter Cliff. 2008. FRBR and Metadata Application Profiles. Presentation at Version Identification Workshop, London, UK, April 2008. <http://www.ukoln.ac.uk/ukoln/staff/p.d.cliff/presentations/vif2008-apfrbr.ppt>
- [55] Kristin Antelman. 2004. Identifying the Serial Work as a Bibliographic Entry. Library Resources & Technical Services, 48(4): 238-55 as cited by David Mimno, Gregory Crane, and Alison Jones. 2005. Hierarchical Catalog Records: Implementing a FRBR Catalog. DLIB 11(10). <http://www.dlib.org/dlib/october05/crane/10crane.html>
- [56] Ed Summers. 2008. LCSH Linked Data <http://lcsch.info>
- [57] Karen Coyle. 2004. Future considerations: the functional library systems record. Library Hi Tech. 22(2): 166-174.
- [58] Page 14, Athena Salaba and Yin Zhang. 2007. From a Conceptual Model to Application and System Development. Bulletin of the American Society for Information Science and Technology 33(6): 17-23. http://www.asis.org/Bulletin/Aug-07/salaba_zhang.html
- [59] International Federation of Library Associations and Institutions (IFLA). FRBR Review Group, IFLA, Division IV, Cataloguing Section. FRBR Bibliography. http://www.ifla.org/VII/s13/wgfrbr/FRBR_bibliography.pdf
- [60] Section 4.2, Page 32-3, The Library of Congress Working Group on the Future of Bibliographic Control. 2008. On the Record [Final Report]. <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>
- [61] Page 6, Yin Zhang. 2007. Introduction to the Special issue on FRBR. Bulletin of the American Society for Information Science and Technology 33(6): 6. <http://www.asis.org/Bulletin/Aug-07/zhang.html>