

FRBR: Past, present, and future

LIS 415 Essay One

Paul Hoffman <hoffman@simmons.edu>

Introduction

In 1992, the International Federation of Library Associations (IFLA) formed a Study Group on the Functional Requirements for Bibliographic Records. The group's work lasted from 1993 to 1997 and its findings, approved in final form in 1997 and formally published in 1998, are known collectively as FRBR.

The FRBR group was charged with exploring the ways in which bibliographic records are used in order to formulate 'a clear, precisely stated, and commonly shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs' (IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998). The group was also asked to recommend the basic functional and data requirements of records in national bibliographic utilities, setting aside preconceived notions of how records should be structured and used. This paper explores the essential aspects of their findings, subsequent efforts to refine the FRBR model and to implement aspects of it, FRBR's expected impact on the future of cataloging, and the historical backdrop against which this reformulation of bibliographic records and practices is being played out.

The FRBR model

FRBR is 'a conceptual model for the bibliographic universe' (Tillett, 2004) that attempts to describe and classify the entities that belong in a bibliographic record (publications, authors, subjects, and so on), to identify the relationships among them, and to match these entities and relationships to the tasks for which people turn to bibliographic records. The FRBR Study Group identified four general user tasks: (1) to *find* resources that meet users' search criteria; (2) to *identify* resources – that is, to confirm the relevance of found resources and to distinguish them from each other; (3) to *select* resources that meet users' needs; and (4) to *obtain* access to these resources (IFLA Study Group, 1998).

To support these tasks, the study group identified the entities of interest to users of bibliographic records and divided them into three conceptually constructed groups. Group 1 consists of the products of intellectual or artistic endeavor. Group 2 comprises the people and institutions responsible for the creation, production, and distribution of those products. Group 3 consists of other entities that may be the subjects of the entities in group 1 (IFLA Study Group, 1998). FRBR follows the entity-relationship model formulated by relational database theory: entities have attributes and are related by formal relations. This paper will focus primarily on entities.

Foremost among the entities that FRBR identifies are the four that constitute the first group: *works*, *expressions*, *manifestations*, and *items*. Of these, the last two are most familiar to today's catalogers: an item is the physical or digital object that a user reads or listens to or otherwise makes use of, while a manifestation – represented in today's catalogs by a single MARC record built according to AACR2 or other ISBD-based principles – is the set of all items with identical content and form, ignoring slight differences (such as bindings or authors' inscriptions or missing pages) that have no bearing on items' intellectual or artistic content *per se*.

The *work*, in contrast, is a wholly abstract notion: a work is 'a distinct intellectual or artistic creation' independent of any physical or digital object (IFLA Study Group, 1998). Conceptually, a work may stand on its own; in practice, however, a work has no significance in the bibliographic world unless it is embodied in a physical or digital object.

The least familiar entity modeled in FRBR – and the one that seems most difficult to define and delimit – is the *expression*.¹ Expressions are defined as much by what they are not as by what they are: they are not works, and they are not the physical embodiment of works; rather, they are the intellectual or artistic realization of works. A book in print form and the same book reproduced in microform are two manifestations of a single expression of a work, because their intellectual content is identical; Aaron Copland's *Appalachian spring suite* for orchestra and a 1985 arrangement for two pianos are distinct expressions of the same work, because their content, though not identical, express the same underlying intellectual or artistic endeavor.

Explicitly modeling a work and its expressions is a leap forward from current bibliographic practice because it bestows a formal existence upon concepts central to user tasks, concepts that are second-hand citizens in today's catalogs.

The practical value that this entails is illustrated by Denton (2003), who posits a scenario in which a library user, wanting to read a particular book, is frustrated by the mismatch between her mental model of what she wants and the structure of her library's catalog:

Brigid wanted to read Raymond Chandler's *The Big Sleep*. She remembered seeing the movie a few years ago, the one with Humphrey Bogart, and when some friends talked about the book the other day and [sic] she decided it was time to finally read it. She went to the library, but their copy was not on the shelf. Someone had already borrowed it. Interlibrary loan within the city was very fast, so she went to the computer, pulled up the online catalogue, searched for the title 'big sleep' – and got a list [of] ten books, all named *The*

¹ The difficulty in delimiting works and expressions is due, in part, to the interdependence of the entities in the FRBR model: it's not possible to define (or understand) *expression* without defining (and understanding) *work* and *manifestation* as well. This is an essential aspect of FRBR, not a flaw.

Big Sleep or some slight variation. The top book on the list was by Chandler and had 73 copies in the system. Brigid choose it from the menu – and got a list of 46 libraries, all holding at least one copy of the book.

Brigid got her wish, obtaining a copy through interlibrary loan, but she had to follow an annoying sequence of steps first because the catalog failed to formally represent the thing she wanted: a copy – *any* copy – of *The big sleep* by Raymond Chandler. She didn't care who the publisher was, nor the date of publication; she merely wanted to read the book. In this case, the catalog complicated, rather than simplified, her efforts to *select* (FRBR's third task) the thing she wanted.

Reworking library catalogs to use the FRBR model will achieve much more than the division of bibliographic records into the separate concerns of work, expression, and manifestation; it will also result in bibliographic records containing information currently found only in authorities records (Le Boeuf, 2001). A personal name authorities record today, for example, would most closely be expressed in FRBR not by a string of letters in a field, but by a full-fledged author record existing on a par with records for works and the other entities of group 1. A full realization of FRBR may also encourage a radical reassessment of the way in which subjects are represented in the catalog, especially since those subjects that fall under FRBR's rubric (works and authors as subjects, for example) will necessarily be modeled as distinct entities with first-class bibliographic status, and the benefit of giving such a status to subjects will be hard to ignore.

The roots of FRBR

While FRBR is only eight years old, its roots go back much further. The concept of the work, which lies at the heart of the FRBR model, is a thread running throughout the last 150 years of advances in cataloging. Anthony Panizzi led the way in the mid-nineteenth century, asserting in his *Rules for the compilation of the catalogue* that the library catalog should serve as more than a mere inventory of books: 'A reader may know the work he requires; he cannot be expected to know all the peculiarities of different editions; and this information he has a right to expect from the catalogues' (Panizzi, 1841). Charles Cutter's (1876) 'Objects' for library catalogs advanced the notion that 'catalogs not only should point the way to an individual publication but should also assemble and organize literary units' (Taylor, 2004). The revision of cataloging codes in the next several decades, however, were more an elaboration of complicated, and often inconsistent, rules – with numerous unprincipled exceptions – than an expression of these ideas. In the 1950s, Seymour Lubetzky revitalized interest in the concept of the work, but his ideas – though met with widespread approval and codified in the 1961 Paris Principles – were given short shrift until the late 1970s, which saw the development of AACR2, 'the triumph of the logic and analysis of Seymour Lubetzky' (Gorman, 2000).

The IFLA study group describes FRBR in terms of the relational model, the rigorous mathematical foundation upon which modern-day database management systems such as Oracle and MySQL depend. The relational model is well understood in the field of information technology and widely implemented in software.

Recent developments

There has been some success in identifying works in large catalogs of MARC bibliographic records constructed according to AACR2 rules, but the distinction between work and expression has been much harder to draw – either in an automated fashion or by manual examination of records (Bennett, Lavoie, and O’Neil, 2003; IFLA Cataloguing Section’s FRBR Review Group, 2004). One possible response to this difficulty is to ignore the distinction altogether, something which RLG has done in their RedLightGreen system (RLG, 2005). Others have suggested clarifying the boundaries between works and expressions by dividing the latter into two or more subtypes (Le Boeuf, 2001).

The first large-scale effort to implement the FRBR model was AustLit, ‘a non-profit collaboration between eight Australian Universities and the National Library of Australia providing authoritative information on hundreds of thousands of creative and critical Australian literature works...’ (AustLit Gateway Consortium, 2004a). AustLit uses XML records to represent works, expressions (which they call versions), manifestations (which they call publications), and agents; items are not represented in AustLit (AustLit Gateway Consortium, 2004b).

In 2003 RLG created RedLightGreen, a FRBResque version of its union catalog, containing 130 million records; the system is ‘built on an IBM DB2 database containing XML data’ (RLG, 2005). Despite the impressive numbers, RedLightGreen can’t be described as a faithful implementation of FRBR, because it reduces the four entities of group 1 to just two: works and manifestations.

Other progress on FRBR includes a mapping from ISBD elements to FRBR (Delsey, 2004) and an algorithm, implemented as an XSL transformation on XML-encoded MARC records, to automate the conversion of existing bibliographic databases to data structures that implement – at least partially – the FRBR model (OCLC, 2003).

The publication of a successor to AACR2 is scheduled for 2007 (Joint Steering Committee for the Revision of Anglo-American Cataloguing Rules, 2004). This new revision, known as AACR3, will adopt the FRBR terms *work*, *expression*, and *manifestation*, and will change its use of the term *item* to accord with the definition used in FRBR. It’s not clear, however, how closely AACR3 will mirror the principles of FRBR.

ILS vendors have been slow to adopt the FRBR model, but this isn’t surprising since there are still so many unknowns. As of this writing, the only commercial vendor to release an ILS based on the FRBR model is VTLS, whose Virtua ILS first gained FRBR support in 2002 (Virtua ILS Is First to Support FRBR Model, 2002).

Virtua supports catalogs containing records of both types (FRBR and non-FRBR); this practical solution may give VTLS an advantage over other vendors, but it's not clear how well such hybrid catalogs will meet users' needs.

The advent of bibliographic catalogs that take full advantage of the FRBR model may mark the beginning of the end for MARC bibliographic records; the latter's linear structure, based on the requirements of card-producing systems before the advent of online catalogs and the personal computer, is unwieldy and inflexible, a poor match for the rich web of relations found in the FRBR data model. Although it is in theory possible to adapt MARC to this purpose – say, by defining a bevy of new linking fields and subfields and strictly constraining their values – the result would be nearly impossible to put into practical use. However, the transition from MARC to an XML bibliographic schema – or whatever takes the place of MARC – will take a long time. Already, nearly 8 years have passed since IFLA approved the final report on FRBR; a similar span of time will certainly elapse before the FRBR model is widely implemented. Retrospective data conversion projects of the size we're talking about just don't happen overnight.

FRBR and everyday cataloging practices

The impact of FRBR on everyday cataloging will be big, but how big? Some aspects of original cataloging practice will change dramatically. Catalogers will have to adopt a new mental model; they can no longer operate on the principle that the piece they have in hand will result in a single bibliographic record (and perhaps an authorities record or two). They will have to draw distinctions between expressions and works, and between one expression and another. This is not to say that all the changes will be painful; for example, the elimination of the 'rule of three' in all its guises will undoubtedly be a cause for celebration.

To imagine how an implementation of the FRBR model might affect the day-to-day work of catalogers, it may be useful to consider the work of two catalogers given a newly acquired book to catalog. One cataloger (let's call him Jack) works at an institution that follows the old practice of AACR2 and MARC 21, while the institution that the other cataloger works in (call her Jane) uses an environment that takes full advantage of the FRBR model. Both catalogers have the same book in hand – not the same item, of course, but copies of the same manifestation – a curiously heavy volume titled *Rutabagas and Rosicrucians: a Freudian synthesis* by a certain Ulysses Karl Fishwick. The book was published in 2001 by Odds and Ends Press in Oshkosh, Wisconsin.

The two catalogers' first task is to determine whether the book needs original cataloging. We'll set aside the possibility that this determination has already been made and force the two unsuspecting catalogers to make it themselves.

Jack pulls up OCLC Connexion and does a quick derived name/title search for 'fish,ruta' but no matching records are found. A derived name search for 'fish,uly,k'

pulls up another book – *Travels with my id* – written by Ulysses K. Fishwick and published in 1996. Jack concludes that the two Fishwicks are one and the same.

Jane searches the personal name index in Ednecat, the FRBRized national utility her institution subscribes to; this yields about 25 *person* records for various Fishwicks, among them one for a Ulysses K. Fishwick. A single entry appears in the list of works and expressions by the author: the same *Travels with my id* that Jack found.

Jack presses a function key on his keyboard; in a couple of seconds, the macro assigned to it has done its work and a template for a MARC monographic record awaits him, many of its fixed and variable fields already populated with reasonable defaults.

Jane clicks on a link in the author's record to bring up a very different template. The display is divided into three parts labeled 'Work/Expr/Mani,' 'Creators,' and 'Subjects'. The first part occupies the left half of the window, and consists mostly of empty fields; the right half is divided between the other two. The second part indicates that the creator, Ulysses K. Fishwick – identified by a non-editable 'Name' field whose contents are linked in by the value in a 'Creator no.' field – has the role of author; the subject part is empty.

Jack types quickly into the MARC record template, entering the appropriate fixed field codes – which he long ago memorized – and transcribing the author and title information from the title page into the fields and subfields he knows so well. A quick flip to the back cover and he enters the ISBN one-handed from the numeric keypad; adding a 100 field, he enters the form of the author's name that he finds in the name authority file.

Jane follows much the same procedure. The record template she fills in is constructed for works with a single expression and manifestation, easily the most common type of record that she creates during the course of a normal day. The fields in the 'Work/Expr/Mani' section are typeface-coded: bold text for work-level information, normal text for expression attributes, and italic text for manifestation-level information. Jane quickly modifies the 'Stmt of Resp' field (in italics) from 'Fishwick, Ulysses K.' to 'Fishwick, Ulysses Karl'; she knows the name will be indexed both ways in the work/expression/manifestation title index and that the latter form will be added to the author's record as a 'Variant Name' attribute.

I'll stop here, because it should be clear from what I've already laid out that the essential skills used by the two catalogers, and to a lesser extent the labors they undertake, are not very different, even though the conceptual models that underlie AACR2/MARC and FRBR are dissimilar. Both catalogers use similar search techniques and strategies to attempt to relate the item in hand to other bibliographic entities – namely, to the author and to books with the same, or similar, title – and both transcribe information taken from the item. Both will go on to assign subject headings and one or more classification codes. Their approaches differ in important ways: Jane is more likely to link information entities explicitly and without any possibility of ambiguity using record numbers, while John must rely on strict

adherence to rules and authorities information, as well as highly accurate typing, to ensure that the headings he assigns will be collocated correctly in his library's OPAC so that users searching and viewing bibliographic records will find what they need. The specifics of how information is encoded under the two regimes differ widely, but the essential expertise that the two catalogers draw upon is still the same, and they share the same aim.

This doesn't mean that the transition to FRBRized catalogs will be painless to library staff; on the contrary, it is likely to cause a great deal of consternation as catalogers and other staff revise their mental models of bibliographic structures and relations and learn new interfaces and encodings. Efficiency is likely to suffer in the short term, which is a grave concern in the face of financial constraints; the long-term impact on cataloger productivity can only be guessed, though certainly there is the potential for cost savings as original catalogers focus less on the transcription of information, which serves users' needs more obliquely, and more on efforts that more directly support the functions of bibliographic records and the tasks to which users put them.

Conclusion

The transition from cataloging based on AACR2 and MARC to cataloging based on the FRBR model will be difficult and require a great deal of time and effort. Although there have been other large-scale changes in cataloging in the past – for example, from AACR to AACR2 and from card-based catalogs to OPACs – none combined a reformulation of data models with a dramatic rethinking of the conceptual foundations of cataloging, at least not on the scale that a wholesale adoption of the FRBR model will entail.

The FRBR model is such an improvement on the card catalog mentality that lives on in our current cataloging practices that only three things can really stop it – political division and infighting among interested parties, the financial costs of the transition from current practices, and implementations that are so poorly done that FRBR is rejected as the wrong solution at the wrong time. The first possibility seems remote given the general goodwill felt toward FRBR not only by the library world but also among those interested primarily in organizing the Web. The prospect of an expensive transition to the FRBR model, however, is daunting. The challenge of reducing the costs of cataloging as currently practiced was an important impetus for IFLA's FRBR study (IFLA Study Group, 1998), but it's not clear whether the potential long-term savings of a cataloging paradigm that more directly addresses users' needs will offset the short-term costs associated with full implementation of the FRBR model; the greatest danger I see is the possibility of poorly done half-implementations that echo the failures of the original AACR.

References

- AustLit Gateway Consortium. (2004a). *About AustLit*. Retrieved April 25, 2005, from <http://www.austlit.edu.au/about>.
- AustLit Gateway Consortium. (2004b). *AustLit data models*. Retrieved April 25, 2005, from <http://www.austlit.edu.au/about>.
- Bennett, R., Lavoie, B. F., & O'Neill, E. T. (2003). The concept of a work in WorldCat: an application of FRBR. *Library Collections, Acquisitions, and Technical Services* 27, 45–59.
- Cutter, C. A. (1876). *Rules for a printed dictionary catalog*. Washington, D.C.: Government Printing Office.
- Delsey, T. (2004). *Mapping ISBD elements to FRBR entity attributes and relationships*. Retrieved April 20, 2005, from <http://www.ifla.org/VII/s13/pubs/ISBD-FRBR-mappingFinal.pdf>.
- Denton, W. (2003, May.) *FRBR and fundamental cataloguing rules*. Retrieved April 23, 2005, from <http://www.miskatonic.org/library/frbr.html>.
- Gorman, M. (2000). Seymour Lubetzky, man of principles. In T. H. Connell & R. L. Maxwell (Eds.), *The future of cataloging: insights from the Lubetzky Symposium*. Chicago: American Library Association.
- IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). *Functional requirements for bibliographic records: Final report* (UCBIM Publication – New Series, 19). München: K. G. Saur. Retrieved April 17, 2005, from <http://www.ifla.org/VII/s13/frbr/frbr.pdf>.
- Le Boeuf, P. (2001). FRBR and further. *Cataloging and Classification Quarterly*, 32, 15–52.
- Joint Steering Committee for the Revision of Anglo-American Cataloguing Rules. (2004). Tom Delsey appointed as AACR3 Editor. Retrieved April 23, 2005, from <http://www.collectionscanada.ca/jsc/aacr3editor.html>.
- OCLC. (2003). *FRBR work-set algorithm*. Retrieved April 25, 2005, from <http://www.oclc.org/research/software/frbr/>.
- Panizzi, A. (1841). Rules for the compilation of the catalogue. In A. Panizzi, *The catalogue of printed books in the British museum*. London: British Museum.
- RLG. *Under the hood: RLG's RedLightGreen*. (2005). Retrieved April 20, 2005, from http://www.rlg.org/en/page.php?Page_ID=4721.
- Taylor, A. G. (2004). *Wynar's introduction to cataloging and classification* (revised 9th ed.). Westport, Conn.: Libraries Unlimited.
- Tillett, B. B. (2004). What is FRBR? Retrieved April 20, 2005, from <http://www.loc.gov/cds/downloads/FRBR.PDF>.
- Virtua ILS Is First to Support FRBR Model. (2002, September). *Computers in Libraries*, 22(8). Retrieved April 17, 2005, from <http://www.infotoday.com/cilmag/sep02/newsline.htm>.