The Functional Library Systems Record

Karen Coyle
August, 2003

Abstract

The author performs a thought experiment on the concept of a record based on FRBR and library system functions, and concludes that if we want to develop a functional bibliographic record we need to do it within the context of a flexible, functional library systems record structure. The article suggests a new way to look at the library systems record that would allow libraries to move forward in terms of technology but also in terms of serving library users.

Introduction

The library card catalog performed a suite of functions with a single technology: the card. Today's library automation systems have integrated a much larger number of functions into a single system. These include the functions of discovery and location that were performed by the card catalog, but expand to other library management functions like acquisitions, serials control, and circulation. The library system is also being asked to expand beyond these functions. We want it to provide interaction with outside user services such as full text, to enhance catalog entries with images and sound, and to allow users to search a variety of local and remote databases with a single search.

When we contemplate how our bibliographic record should be structured in the future, and what data elements it should contain, we need to look at more than just the MARC record but also the context in which it is used, which is the library system. Changing the MARC record without taking this holistic system view would be a grave mistake. It would also be a mistake to assume that the library system of today is a finite and fixed context; instead, our systems are in a constant state of evolution, as is are all computer systems, and they are part of a larger context of networked information resources.

At the same time that those of us in the library systems area are contemplating our next record structure, catalogers in our profession are looking at the bibliographic record from a conceptual and functional point of view. The bibliographic view of what is functional and the systems view of functional are not currently being discussed in concert. Bringing these two reform movements together will be a better formula for success than either of them will have on its own. This article proposes one way to think about those two changes and how they might work together.

AACR::MARC ...

The MARC record was created as a digital mirror image of the cataloging rules of its time, which were not so different from the cataloging rules of our time. Those cataloging rules were originally designed to produce cards for library catalogs, and they still reflect that heritage with their main entry headings, inverted forms of names, and the grouping of data elements into paragraph-like segments.

A library catalog's cards served a variety of functions. They carried the descriptive catalog for works owned by the library; they were the discovery mechanism for users of the library; they provided users with the shelf location of the items; and for the library administration the card catalog was an inventory of the library's holdings.

The data elements for this library card were the original focus of the MARC record, and the first use of the MARC record was to print traditional catalog cards in an era of computer-driven typographic machinery. The creation of the

first online catalogs began a transformation of library catalogs that was not anticipated by either the cataloging rules nor the machine-readable record that served them.

Most notably, the online catalog made a radical change in the discovery function of the library catalog. Discovery in the card catalog had been an entirely linear affair. Each designated access heading in the catalog record was an entry point in an alphabetical list of headings. Users searched for their desired author, title, subject or series in this alphabetical list. In the online catalog, discovery could be linear, but it could also take place as a keyword or string search within the access headings. Not only could the records be retrieved by words in the headings rather than the entire heading, the boundaries between headings could be broken down. A single search could be performed against more than one heading, for example a search could include words from all subject headings in the record, or could combine keywords from both author and title fields. It could also go beyond the designated access headings and allow searching in fields that were previously unavailable for discovery, such as notes, identifying numbers, and tables of contents.

As new forms of discovery were presented in online catalogs, the MARC record began to respond to this environment. Fields were added to the MARC record that did not arise from the cataloging rules. Fixed field coding for various item formats became increasingly detailed so that searchers could limit their retrievals to specific physical formats such as to videotapes in VHS format, or to music on cassette tape. A field was added for a coded form of the mathematical data carried in map records that was normalized for machine manipulation. Other coded fields served the retrieval of music records by composition and number of instruments. None of this was conceivable in the era of the card catalog.

Discovery wasn't the only library catalog function that has changed in this era of automated library systems; the concept of location has made significant changes. Networking, and in particular the Internet, means that the library catalog is no longer a closed system that only refers to items inside the library. The location function of the catalog has changed from that of identifying a shelf location in a library to pointing to a networked location anywhere in the world. Location is increasingly becoming a dynamic concept that refers less to a fixed position in space and more with networked functions like the OpenURL and the DOI, which resolve to a means to obtain the item or a service that can be offered related to the item.

The inventory function has changed as well. The descriptive record is no longer the primary record of the library's inventory. Inventory, as well as acquisitions and licensing, have their own functional segments of the integrated library system. Although coupled with the descriptive record, these modules are themselves sophisticated accounting and control systems. Part of a library's inventory is now virtual; licensed resources that are neither owned by the library nor possessed by it must be accounted for in terms of resources that the library is making available to its users.

Since the automation of the library catalog, the most radical change to the MARC record was the creation of a separate record for the very complex functions relating to holdings and locations. The MARC Format For Holdings was the first — and so far only — time that a new MARC record format was developed to fulfill the requirements of library systems.1 The Holdings record was needed in particular to express complex serials holdings patterns for the system functions that support check-in and receipt prediction. The creation of a holdings-level record that is linked to the MARC bibliographic record gives us a direction for further developments toward a multi-level, multi-functional library systems record. The data structure of the Holdings record, however, is the same as that of the bibliographic record, and that is based on a standard developed in the mid-1960's and shares the structural limitations.

There are many reasons to contemplate a more modern replacement for the MARC record. Already there is a movement to transform the 1960's record structure of MARC (Format for Information Exchange, ISO 2709) to a variety of XML formats. These changes aim at allowing greater extensibility in the record structure and better integration with mainstream computing. But they do not necessarily encourage any modification of the fundamental content of the MARC record. Challenging the content is more difficult than challenging the structure — after all, the content of the record has a legacy of over a century of library cataloging rules. Also, the challenges to the MARC record are coming from the technologists in the library field, and following the division of labor in the library profession, technologists work with the structure but the content is entirely the responsibility of catalogers. A new development in the cataloging community, however, may give us the opportunity to work on both sides of the library record format.

It is an opportune time to move on from the AACR/MARC model that is based on the card catalog to a cataloging philosophy and a machine-readable record that are grounded in database management capabilities and networked information resources. And this brings us to one of our profession's more recent acronyms: FRBR.

... as FRBR::?

There is a great deal of buzz in the library world today over the Functional Requirements for Bibliographic Records (FRBR)2. FRBR is indeed a new way of thinking about bibliographic description because it places emphasis on the context of textual works and their relationship to each other rather than merely on the description of individual publications. It moves us toward a view of a universe of inter-linked publications where users eventually will not need to be concerned with differences in formats or the vagaries of nearly identical printings of the same works.3

FRBR is not itself a record structure, but it speaks conceptually of four bibliographic levels. The most general level is that of the work, the fundamental intellectual product. That is followed by an individual expression of that work, which is generally considered to be the specific content or edition. The next level is the manifestation. Manifestation is the level where the work is "productized," that is a particular publication or production. The final level is the item itself, the copy or physical package that is handled by the library.

These four levels are the aspect of FRBR that most librarians are aware of, to the extent that they are even aware of the FRBR movement. But the FRBR document also describes a number of entity-relationship elements that are pertinent to the intellectual work, such as authorship and topic assignment. These entity-relationships essentially define the relationship of key data elements, like creator and topic, to the four bibliographic levels.

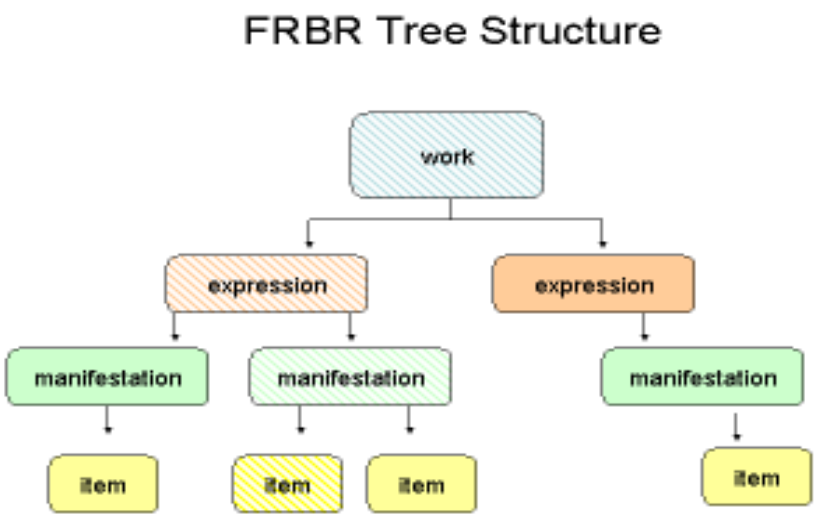
Although some of the concepts of FRBR have caught on in the library profession like a mild fever, how these concepts might affect both cataloging and library data development is still fairly unclear. One of the dilemmas we face when thinking about FRBR is that it is so clearly incompatible with our current data structure, the MARC record. Attempts are being made to "FRBR-ize" collections of MARC21 records, but in these cases we are trying to imitate some FRBR concepts with records created using pre-FRBR cataloging rules and pre-FRBR record structure. Because MARC records are what we have today, any early experimentation with the FRBR concepts in library systems will have to use the MARC data, but if we want to move forward to a library systems record that is based on the FRBR concepts we need we need to invent a record structure that supports that experimentation.

As a preliminary step to creating that record structure we can sketch out the basic functions and relationships that such a record will need to have. Whether or not this turns out to be the library world's next record structure, having a logical map should help us perform the necessary gedanken experiments to determine if this multi-level, multi-functional record structure fulfills the needs of the library catalog of the new millennium.

A Record Structure

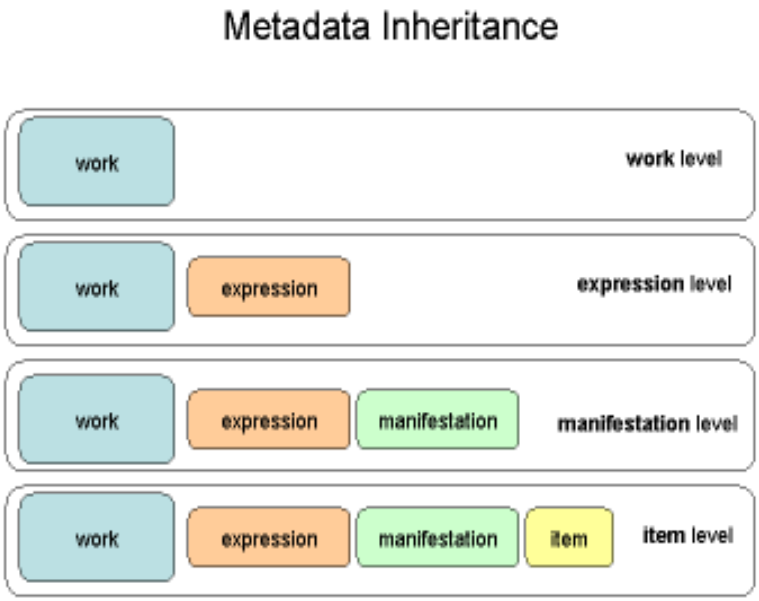
Today we have a two-level record using MARC Bibliographic and MARC Holdings. If we take FRBR at its face value we can assume that we will have a four-level record: work, expression, manifestation, item and that any individual document has aspects in each of those levels. In fact, we may have more or fewer levels, but the exact number of levels is not important.

Fig 1



The creators of the FRBR concept state that they were inspired by relational database design4. Viewing the FRBR design as object-oriented, however, allows us to make use of the concept of "inheritance" in which qualities and data elements of higher levels are inherited by the linked levels below them. In this way, every expression inherits the elements associated with the work, and manifestations inherit those of the expression and the work.

Fig. 2



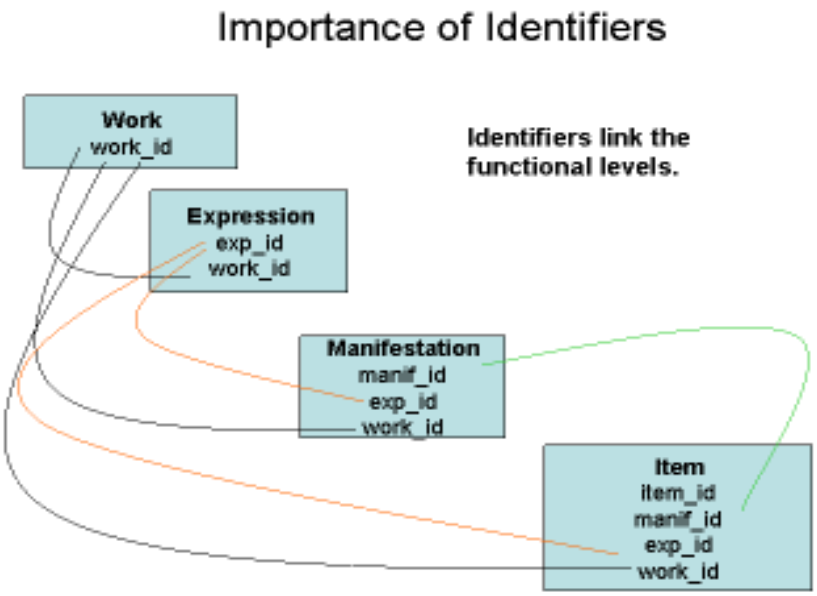
The object-oriented design also helps solve one of the conceptual difficulties of FRBR, which is that the upper levels may be insufficient or incomplete on their own. With object-orientation we can consider upper levels such as work and expression to be abstract in nature, and therefore not standing alone without at least a manifestation entry. At the same time the abstract levels can carry data elements absolutely essential to the meaning of the manifestation or item levels and can sometimes fulfill functions on their own, such as some user-oriented displays.

The Importance of Identifiers

Identifiers are important for linking records or parts of records that actually relate to the same item, or for identifying those parts from either inside or outside of the basic bibliographic record. Although information for a single item can transmitted in a single record, the nature of the flexibility of a multi-level record means that within systems records may not be unitary, as the MARC21 bibliographic record is, but could consist of parts that may be accessed separately during different system operations. The exact nature of these operations is not predictable, and we should not attempt to constrain them through a record structure. Instead, any part of the record with its own coherent structure needs to identify itself in a way that maintains its relationship to the whole set of data elements related to a bibliographic item. This means that the record structure will rely heavily on identifiers and that standard identifier schemes for the levels will need to be defined.

Assuming a record that uses the FRBR levels as its primary structure, we have a similar structure for identifiers of the various levels:

fig. 3



The creation of identifier schemes is one of the more difficult aspects of metadata development. For the sake of the readability of the examples below I have used some identifiers that we are already familiar with, but this is not meant to suggest that those would necessarily be the appropriate identifiers for our future record.

Systems and Functionality

Once we have a bibliographic record that is flexible and extensible, we can begin to look at the functions that we want the record to perform.

The "F" in FRBR is "Functional" and the FRBR movement, if it can be called such, can open up a discussion not only of the functional roles of traditional cataloging elements but could allow us to look at overall library systems functionality and how that functionality can integrate with the core record for bibliographic description. In the FRBR analysis the bibliographic record contains data elements for the complete range of functions that the catalog record performs today, including subject analysis, inventory, circulation and preservation. I will suggest that we can take "functional" even further and can break out some, if not all, of the uses of the library system record that all are anchored by a core bibliographic description.

Library Systems Functions

Below are some examples of key library systems functions. This is not a complete list; in fact, one of the points I wish to make is that the design of library system records needs to be so flexible that no definitive list of functions is appropriate. Instead, we need to be open to the possibility that our record will be of a plastic nature, and yet we can still expect to have standards that we can count on to make our systems accurate and strong.

These library systems functions are not the same as the functions that are defined in the FRBR document, although there is some overlap. The FRBR analysis limits itself to the traditional functions of a library catalog (not a library *system*): find, identify and select records based on bibliographic characteristics. In our re-thinking of the bibliographic record we should think broadly about a wide range of functions that have been supported by the bibliographic record and those functions that we support outside of that record in our current library systems, and see if we can't find a more holistic approach to the system functions and the record structures that support them.

Description

Let's begin our list of functions with the one with which we are the most familiar: *description*. Description will be the central function of any library systems record because it is the focus of the system itself; it defines the information item. Description is the purview of the cataloging community and the future of bibliographic description may be informed by the principles expressed in FRBR.

Description records the title of the item, the author or authors, the publisher or producer, the date of issue or publication, place of publication, and the extent of the item. It may also place the item in relation to other bibliographic entities such as a series. And it may list parts, such as chapters, in some publications.

Description, however, has never existed as a fully separate function from *discovery*, and in the traditional catalog record the same fields perform both of these functions, as well as other functions like display. Although it may not be possible draw a firm line between description and discovery, a functional approach to library systems records gives us an opportunity to rethink some discovery-related fields such as subject headings and variant titles that are intended primarily for discovery purposes.

Discovery

Library catalogs exist not simply as inventories of library holdings but as a means to discover what the library has. Although *discovery* is the key function served by library catalogs today, many data elements used for discovery also serve the description function. The data elements serving this dual function are based on the card catalog where the

primary means of discovery was based on the ordering of cards, not database searching. We still create a record replete with inverted entries (Doe, Jane; Mary, Blessed Virgin) intended to be filed sequentially. Very little of the record is designed to facilitate the keyword indexing that is the mainstay of discovery in the 21st century. The ability to carry additional keywords, based perhaps on indices or even the full text of the item, are essential to modern discovery methods.

None of this means that catalogers have to create separate data elements to describe and discover the author of a work. Making multiple uses of a single incoming field is something that computer systems do quite well. Using the appropriate data structures, programs can derive a variety of displays and discovery elements from a single field.

Location

One of the primary functions of the library catalog entry has been to give users the shelf *location* of item. Location is still an expected function of the library record, but location now has an expanded function because many materials today are not sitting on a library shelf. The locate function is increasingly dynamic, linking to an OpenURL service, or a handle system that will resolve to the location of the item. In all likelihood, the location function will provide more than one route that a user or a system can take to arrive at the actual item. In particular there is the need to express alternate locations and to link location with user authentication and access permissions. This also implies that the location function is aware of the availability of different physical formats, different terms of use, and other aspects that might be part of the end user's preferences for item selection.

Purchase

The acquisition function of libraries is one that is not well-served by the records that we create today. This is partly because acquisition has generally taken place before the record is created. However, today's *purchase* function can include transmittal of the descriptive record as publishers and library book-jobbers include machine-readable records as part of the purchase. The publishing industry has developed a machine-readable record, ONIX5, that will carry both descriptive, purchase and promotional information about print and electronic books. Today's library record cannot accommodate many of the fields available through these sources.

This becomes ever more important as we develop new ways to disseminate electronic materials through libraries. No longer a simple purchase-to-own, electronic materials may have machine-readable contracts that govern the number of simultaneous copies, terms for per-copy payment, special offers for review and browsing, etc. These terms must be able to interact with the library system's acquisitions function, the circulation function, and even with display to users.

Preservation

Not all libraries have a need for detailed *preservation* information, but those that do perform an archival function know that there is no place in our current library systems record for the depth of information that is needed to perform this function well. Preservation information in the past was applied solely to older or archival materials of a certain age, often those that had already undergone some deterioration. With digital materials, preservation information must be applied at the time of acquisition and all storage of digital materials must be seen as having a preservation function.6

Preservation is an important piece of information for electronic materials that are shared widely over networks. This information is key to understanding the reliability and durability of an electronic information resource. It allows for

the cooperative sharing of the burden of digitizing in the case of materials that were not born digital, and for the sharing of the cost of diligent preservation for materials in digital form.

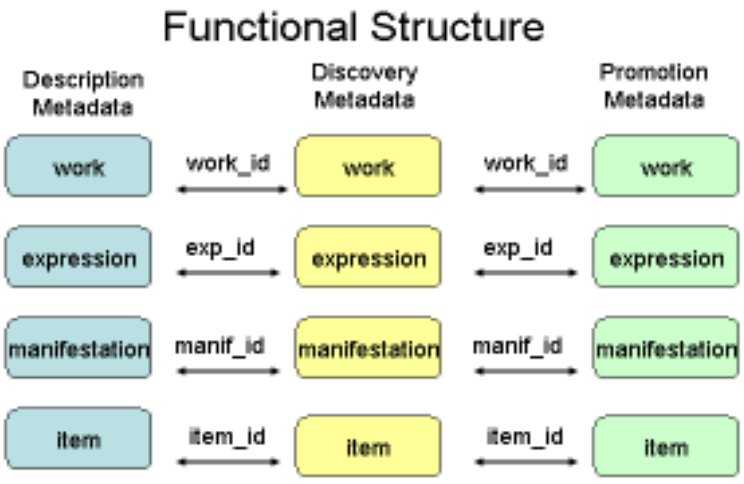
Promotion

A patron entering a library may be greeted with a colorful display promoting reading materials from the library's collection. Our "virtual" patrons, those that enter the library remotely through the online catalog, are often greeted with nothing more than plain text representing the items we hold. The plain text library catalog with its colorless entries has been shown up by the highly promotional online bookstore sites like Amazon and Barnes and Noble. Since part of the library mission is to encourage enthusiasm for reading and learning, *promotion* needs to be a function of our virtual public entryway, the library catalog. Some library systems have included the ability to display cover art with the descriptive record, but there are other ways that we can promote reading materials if we have a record structure that allows it. We could provide not only cover art but jacket blurbs and sample readings from the work. Other materials such as reviews or synopses can inspire readers. We need to have the option to store some of these promotional items locally and access others remotely, depending on their interaction with the library catalog.

Structuring the Record

Although there are many ways that one could create a machine-readable that marries FRBR levels with library systems functions, one conceptual approach is to take each of the functional areas above (and any others we design in the future) and map these to the FRBR levels.

fig 4



This simple diagram does not represent an actual data design. In particular, it does not determine whether the record will be organized primarily along the bibliographic levels:

Work
- Work: Description
- Work: Discovery
- Work: Promotion

Expression
- Expression: Description

Manifestation
- Manifestation: Description
- Manifestation: Discovery

Item:
- Item: Location

or the functional levels:

Description
- Description: Work
- Description: Expression
- Description: Manifestation

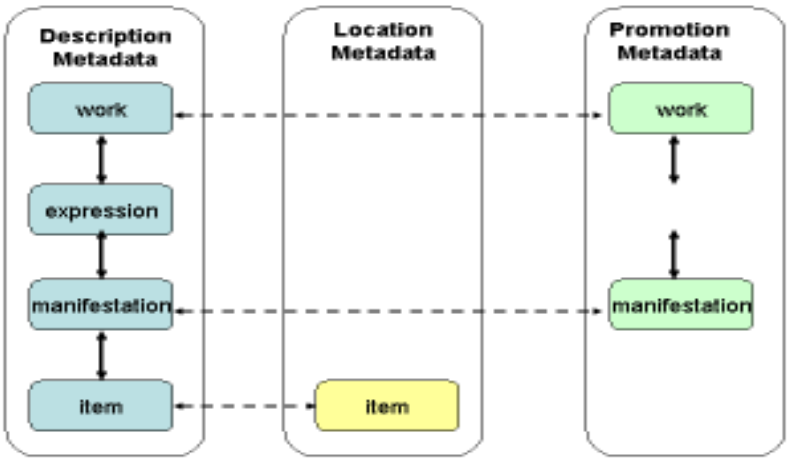
Discovery:
- Discovery: Expression
- Discovery: Manifestation

Promotion:
- Promotion: Expression
- Promotion: Manifestation

Location:
- Location: Manifestation
- Location: Item

Although each function has four theoretical levels, clearly few functions will exist at all four of them. The *location* function will be fulfilled at the manifestation level (primarily for web services for electronic materials) and the item level (for hard copy materials). The *promotion* function may have some expression and some manifestation entries. *Description* will be at the work and expression and manifestation levels. This design allows information to be stored at the appropriate level.

fig. 5



Because many functions will not have data at all levels, the identifiers allow the creation of functional records at any appropriate level as long as the rules of inheritance are obeyed such that any lower level always inherits data elements from the levels above it within its functional group. Note that between functions there is linking based on the identifiers but no inheritance takes place across functions.

As we work with this model at greater levels of detail we can explore further the issues and capabilities of the object-oriented model. For example, there may be situations where a lower level element needs to redefine an element inherited from a higher level object. This capability may help us overcome some of the areas where a data element appears to be appropriate to more than one level.

Example: Faster/Fstr

The following example is decidedly schematic and incomplete, but it will serve to show how individual views can make use of the FRBR levels and functional data structures to create coherent but different views of the same bibliographic item. The structure and linking is implemented through levels of identifiers that follow the levels of the basic data model.

Description Function

Work
work_id = wkid:123
author:
author_familyname = Gleick
author_firstnames = James
title = Faster:
subtitle = the acceleration of just about everything
statement_of_responsibility = by James Gleick

Expression
work_id = wkid:123
exp_id = 1
format = print

Manifestation
work_id = wkid:123
exp_id = 1
manif_id = isbn:0679408371
cover_title = Fstr
place_of_publication = New York
publisher = Pantheon Books
year = 1999
extent = 324 pages
isbn = 0679408371

Expression
work_id = wkid:123
exp_id = 2
edition = abridged
format = audiobook
statement_of_responsibility = read by the author

Manifestation
work_id = wkid:123
exp_id = 2
manif_id = asin:037540886X
place_of_publication = New York
publisher = Random House (Audio)
year = 2000
physical_format = audio cassette
extent = 4 cassettes
asin = 037540886X

Description Function

Work
work_id = wkid:123
sort_title = faster the acceleration of just about everything
author:
lastname = gleick
firstname = james
display_topic = Time—Psychological aspects—Popular works
sort_topic = time%psychological aspects%popular works
keywords:
time
psychological aspects
popular works

Expression
work_id = wkid:123
exp_id = 1
type = book

Expression
work_id = wkid:123
exp_id = 2
type = audio book
format = cassette tape

Manifestation
work_id = wkid:123
exp_id = 2
manif_id = asin:037540886X
creator(reader) = gleick james

Manifestation
work_id = wkid:123
exp_id = 1
manif_id = isbn:0679408371
other_title = fstr

Where we discover the greatest advantage of this record structure for library systems is when we need to make use of data that serves a function beyond the traditional catalog record. For example, if you want to add some preservation data for the audio tape, it is a matter of adding that data at the item level. That data does not need to use the same data structures as is used for the descriptive function, but it does need to clearly link to the work/expression/manifestation/item using the appropriate identifiers:

Item: Preservation
work_id = wkid:123
exp_id = 2
manif_id = asin:037540886X
item_id=2
[This can contain any set of elements from an identified preservation scheme, such as METS]

In another example, we can carry promotional elements at various levels:

Work: Promotion
work_id = wkid:123
cover_art = faster.gif
review = This latest work by James Gleick takes on our….
Expression: Promotion
work_id = wkid:123
exp_id = 2
review = The audio tape of James Gleick's "Faster" is a great way to read this book while on the go …

The key point is that required information can be added to a library system and unambiguously associated with a bibliographic record without having to follow the structure and coding of that bibliographic record beyond identification of the appropriate level. This means that library systems can make use of variable data structures in their records and that use of these data formats will not compromise the integrity of the basic description that is so vital to the full range of library functions. Each of these functions can share the descriptive record or any other functional records as needed by linking to the needed records using a standard identifier system.

Conclusion

After forty years of evolution, the time has come when we can no longer achieve our goals by incremental tweaking of the MARC bibliographic record. Libraries are passing up tremendous opportunities to serve their users by hanging on to a record structure that was brilliant in the mid-1960's but is limiting in the early 2000's.

Although library systems need the flexibility to make use of data from a variety of sources and to interact with other systems, including non-library systems, there is also a need for a highly-structured, standardized core of bibliographic description. This can be achieved by providing separate record structures for that core and for the myriad other functions that library systems may need to perform. The unitary MARC record attempts to fulfill both the core bibliographic function and a smattering of system functions, but this design threatens the integrity of the bibliographic core while it does a poor job of aiding systems design.

Although the particular design elements presented in this paper are far from fully developed, the principle of a multi-functional design and its advantages for library systems should be obvious. With such a design we can have a core bibliographic record that follows the strict rules of library cataloging, and at the same time we gain a great deal of flexibility for our library systems development.

Endnotes

1 There are other MARC record formats that might be considered functional, but the motivation behind them was not to serve library systems. The Authorities format is a digitization of the separate authorities card catalog that was maintained by the cataloging department of a library. It has been integrated into library systems, but has no structural relationship to the bibliographic record. The Classification format has so far allowed Library of Congress to automate the production of the classification schedules but it has not found its way into library systems generally.

2 IFLA Study Group on the Functional Requirements for Bibliographic Records. *Functional Requirements for Bibliographic Records.* K . G. Saur München 1998. http://www.ifla.org/VII/s13/frbr/frbr.pdf

3 Although this article uses the FRBR structure to illustrate the possibility of a multi-level, multi-functional record, it must not be construed to be in support of the FRBR 4-tiered model. Although an interesting theory of the levels of works, FRBR is untested in practice and may never be implemented as an actual cataloging record. That said, it is reasonable to assume that a future cataloging structure will embody some degree of hierarchy, especially in the need to express the relationships between multiple versions of the same work.

4 FRBR, p.9

5 See: http://www.editeur.org/onix.html

6 There are numerous national and international efforts to develop standards for digital preservation as well as preservation metadata. A lengthy, although probably not comprehensive, list is available on the IFLA web site: http://www.ifla.org/VII/s19/usefulinks.htm. The fact that there are a number of preservation metadata formats shows why a single record format will not suffice.