FRBR, Twenty Years On

Karen Coyle

Presented at FSR2014, Rome, February 2014
Published in Cataloging and Classification Quarterly, January 2014, pp. 1-22
DOI:10.1080/01639374.2014.943446

ABSTRACT: The article analyzes the conceptual model of the Functional Requirements for Bibliographic Records (FRBR) as a general model of bibliographic data and description that can be interpreted, as needed, to serve the needs of various communities. This is illustrated with descriptions of five different implementations based on the concepts in FRBR: FRBRER (entity-relation), FRBROO (object oriented), FRBRCore (FRBR entities as linked data), indecs (FRBR entities within the commerce model), and FaBiO (FRBR as a basis for academic document types). The author argues that variant models show the strength of the FRBR concepts, and should be encouraged.

Introduction

> "By far, the substance of the IFLA FRBR conceptual model specification is textual description (with tables), and only a few diagrams. While these diagrams play a very small role in model documentation and presentation, they are what is used—naturally—to describe the model to the Cultural Heritage Community and to the general public. It is difficult to appreciate the overall, emergent, characteristics of the FRBR conceptual data model—especially the more obvious interactions between model elements—from a reading of the text and then attempt to project that wealth of description into the few available diagrams."[1]

Few of us have done a close reading of the FRBR document, although undoubtedly many have glanced at the diagrams, either within the context of the document or as illustrations used in talks that they have attended. The FRBR document is 142 pages in length, including appendices, which makes it a formidable read. In the document there are three diagrams that represent the entity-relation analysis of bibliographic data. Even if each is these diagrams is worth one thousand words, they are hardly expressive of the depth of analysis of the document. Yet the structures illustrated in the diagrams dominate the discussion of FRBR. Most members of the profession can list the primary bibliographic entities: work, expression, manifestation, item (hereafter known as WEMI), and many can describe the three groups of entities that make up FRBR. Most descriptions of FRBR begin with a list of the entities in the three groups, and then illustrate these entities with one or more of the diagrams from the document.

The FRBR document clearly states that it represents a conceptual model of bibliographic data. This article analyzes the conceptual nature of the FRBR "entity-relation" (E-R) model, and shows that the conceptual model presented there has inspired a range of logical models which illustrate variable and valid interpretation of the FRBR concepts.

How the FRBR Study Group Came to Be

The FRBR study arose out of an IFLA-sponsored "Stockholm Seminar on Cataloguing" in 1990. The IFLA cataloging section had been working on international cataloging standards for decades, most prominently the creation of the vInternational Standard Bibliographic Description (ISBD). Barbara Tillett's 1994 report on the meeting and its outcomes[2] describes the factors that led to the assignment of the FRBR Study Group. In that report she refers to "the mounting costs of cataloging," the proliferation of new media, "exploding bibliographic universe,"

and the need to economize in cataloging. Regarding the concern about the costs of cataloging, Tillett states:

> "Some speakers proposed that cataloguing could be considerably simplified. One speaker stated that the number of descriptive data elements needed in a bibliographic record could be reduced without seriously affecting access."

The outcome of the Stockholm seminar was translated into a *Terms of Reference* document authored by Henriette Avram and Tom Delsey.[3] This document stated the problem thus:

> "All libraries, including national bibliographic agencies, are operating under increasing budgetary constraints and increasing pressure to reduce cataloging costs through minimal-level cataloging."

The group's charge was no less than to "delineate in clearly defined terms the functions performed by the bibliographic record with respect to various media, various applications and various user needs. The study is to cover the full range of functions for the bibliographic record in its widest sense..." including access points and organizing elements.

Significantly, included in the *Terms of Reference* was the requirement to make use of entity-relation modeling as defined in a 1984 book on database design.[4] This specific modeling requirement clearly had an effect on the outcome of the report, not the least of which was to focus the work of the Study Group on a particular moment in information technology thinking. By the time the report had been issued, fifteen years had passed, and data modeling had undergone a not-unexpected evolution that was not reflected in the Group's work.

Bibliographic records

The use of "bibliographic record" as the focus of the Study Group's task is worth noting. In information technology terms, a record is a bounded structure that holds data. Records can be standardized, semi-permanent storage for a particular community's data holdings, or they can be the opportunistic output of a process or calculation.

The *Terms of Reference* unfortunately did not define what was meant by bibliographic record, and from the group's charge it is not easy to find a demarcation between the information technology concept of a record and the library activity known as "cataloging." In essence, under the guise of a record definition the group was being asked to develop guiding principles for the library *catalog*, and perhaps for the bibliographic universe in general. In the end, the analysis of the FRBR study group was not bounded by a record concept, but instead used a modeling technique developed for and serving the functions of a relational database. Unlike the most common concept of a record, a relational database is a series of tables of data, similar to a complex spreadsheet, with relations between the elements. Any given database design can manipulate data in various ways and can often export multiple record types.

The E-R diagrams developed by the Study Group do not represent records, and they are not intended to do so. In section 1.2 of the document, captioned "Approach", the FRBR document itself says: "The study makes no a priori assumptions about the bibliographic record itself, either in terms of content or structure." However, the E-R diagrams do represent a first level database design effort.

Entity-relation Modeling

Entity-relation modeling is a technique designed to organize data elements within the tables of relational database management systems. While aimed at that particular technology, it continues to be used as a discipline for thinking about data within an information technology environment. As its name implies, it views data as entities, or things, and the relationships between those things. The E-R modeling technique was used in the FRBR study because it provided a structured approach for the group, whose task was quite broadly defined. Use of the technique was required by the *Terms of Reference* document that gave the group its charge. In the Methodology section of the FRBR document, the group explains:

> "The methodology used in this study is based on an entity analysis technique that is used in the development of conceptual models for relational database systems. Although the study is not intended to serve directly as a basis for the design of bibliographic databases, the technique was chosen as the basis for the methodology because it provides a structured approach to the analysis of data requirements..."[5]

One of the goals of an E-R model in database design is to normalize the particular data universe into atomistic units with no overlapping data elements. This model sometimes breaks logical units into artificially separated parts whose separation serves the requirements of

database applications like update and search. Again, from the Methodology section of the FRBR study:

> "The first step in the entity analysis technique is to isolate the key objects that are of interest to users of information in a particular domain. These objects of interest or entities are defined at as high a level as possible. That is to say that the analysis first focuses attention not on individual data but on the "things" the data describe. Each of the entities defined for the model, therefore, serves as the focal point for a cluster of data. An entity diagram for a personnel information system, for example, would likely identify "employee" as one entity that would be of interest to the users of such a system."[6]

E-R modeling is generally thought to have three levels of analysis, although not every project makes use of all levels:

> **Conceptual model.** A conceptual model serves to define the primary entities and relationships in the information domain at a high level. In traditional business data projects, the conceptual model is a view that can be shared by the database designers and the non-technical users of the data. It cannot be directly used as a database design or in programs as many necessary details are not included.
>
> **Logical model.** The purpose of the logical model is to add detail to the conceptual model that approximates the final database design. It completes the list of attributes, and defines the types of data values that will be stored in the database tables (text, date, currency) and the cardinality of each data element (mandatory, optional, repeatable, etc.). It then normalizes the data to remove any duplication of data within the entire database
>
> **Physical model.** The physical model is the final step in database design, and may be combined with the logical model into a single step. The physical model should reflect the actual database structure and contents.[7]

From this it should be clear that the FRBR study group developed a conceptual-level model of the bibliographic universe, as defined in E-R modeling. A conceptual model is an incomplete picture but one that is general enough to harbor both technical and non-technical discussion of the information domain. This is, in fact, what the FRBR model has indeed fostered, as we will see in the remainder of this paper.

The FRBR Conceptual Model

One characteristic of conceptual models is that they are not "actionable." Actionable means that the design is ready for use in databases and programs. This is definitely true of the FRBR model. No part of the document provides data creation rules, and the definitions of entities are not specific in a way that could be made into specific rules or programmed into algorithms. As an example, the definition of work, " a distinct intellectual or artistic creation," does not give parameters to aid a data creator in defining the work. This information would presumably be included in further documents aimed at data creators, such as input rules. FRBR also lacks the definition of values for the attributes. For example, there is no information given on how the date of publication should be entered. Also, there is no list of the many different relationships that can occur between the entities of group 2 and those of group 1, such as *author*, *composer*, *illustrator*. This makes FRBR a high level set of concepts that need further interpretation before they can be made into a functional data design. It also means, as we will see below, that the model as presented is general enough to be open to a variety of interpretations.

The three diagrams in the FRBR document provide a deceptively simple view of the conceptual model that is described in the pages of the document. There are three separate groups of entities, although the groups do not represent classes, as formal groups of entities are known. In a model with classes and sub-classes, group 2 (*person*, *corporate body*, and *family*) might be referred to collectively as "actors" or "agents," and would have some attributes that are inherited by the members of the group. Using a simple example, any member of the class could have a *name*, and any member of the class could be the *owner of a resource.* Classes gather the attributes that more specific entities have in common, not unlike the general/specific relationships that we find in taxonomies. They make it possible to communicate about the data using different levels of granularity. This can be helpful for humans who are discussing the data, but it also can provide shortcuts for programmers as they develop applications that use the data.

The lack of super- and sub-classes in FRBR is one of the hallmarks of a first-step model. Further analysis of FRBR entities would show that, for example, group 2, with its entities *person*, *corporate body*, and *family*, would surely benefit from a super-class representing the group because the entities will share many relationships with bibliographic resources. Although the Study Group did not define any classes, they did (perhaps inadvertently)

include them in the E-R diagrams.. These diagrams, using E-R modeling notation, show super-classes as boxes around the entities representing their group relationship. Dunsire[8] reports that in creating the linked data version of the FRBR model, he was informed by the IFLA Study Group that no superclasses should be impugned from the diagrams, yet because this contradicts the actual depictions using the E-R modeling technique, it may be a sign that the designers were not well-versed in technical notation they were actually using.

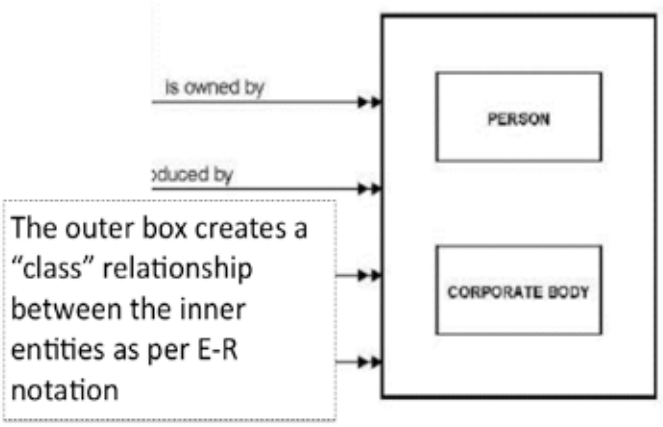

Fig. 1 Section of E-R diagram showing class/sub-class relationships

Group 1, which has entities work, expression, manifestation and item, is depicted as a kind of linear progression, either from item to work, or work to item, depending on one's starting point. The perceived linearity of the Group 1 diagram is, however, false, because the diagram includes a many-to-many relationship between manifestations and expressions. The compact notation expresses this with double arrows, which are easy to overlook. (Fig. 2)

This fact makes the group itself a network, not a hierarchy. In addition, the diagrams do not include the relationships between group 1 entities other than the primary ones that link the group members to each other. The additional relationships with groups 2 and 3, as well as the secondary relationships between group 1 entities (work/work, expression/expression, etc.), present a much more complex picture. Yet it is a complexity that should not be ignored since some of those relationships are considered essential to the success of the stated user tasks. The resulting diagram, which is limited to the group 1 relationships, shows only a tip of the iceberg of the potential complexity of the FRBR model. (Fig. 3)

If we go beyond the diagrams and look at what the text of the FRBR document says, we have the option of creating a very different picture. The four entities of group 1 are described as:

> "*work*: a distinct intellectual or artistic creation."
>
> "*expression*: the intellectual or artistic realization of a *work*..."
>
> "*manifestation*: the physical embodiment of an *expression* of a *work*."
>
> "*item*: a single exemplar of a *manifestation*."

These definitions go a long way to explain why many people have assumed that FRBR group 1 represents a hierarchical model with inheritance from the most abstract (work) to the most concrete (item). Phrases like

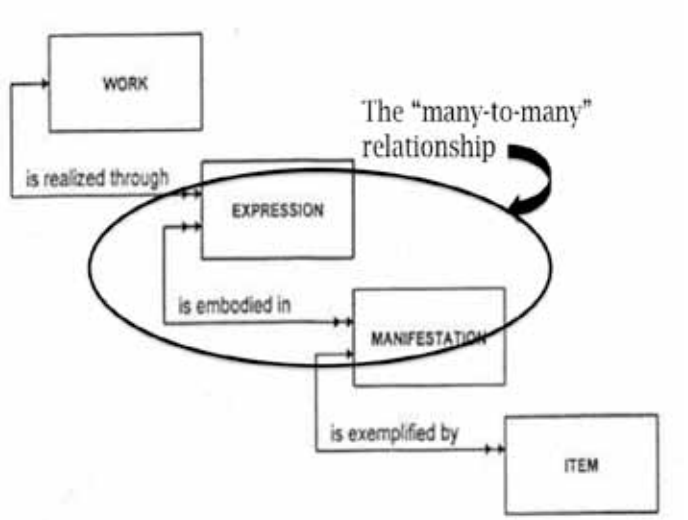

Fig. 2 Group 1 many-to-many relationship

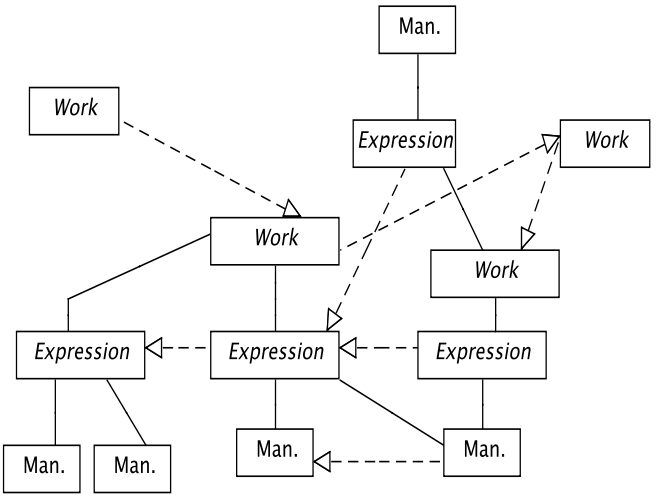

Fig. 3 Group 1 as network

"realization of a work" and "physical embodiment of an expression of a work" imply inheritance between entities, not a distinct separation. In fact, it may be difficult to see them as entities, in the data processing sense, at all, but as variable views of a complex entity that combines meaning, expression and physicality. This view is encouraged in the FRBR document, where it says:

> " The entities in the first group (as depicted in Figure 3.1) represent the different aspects of user interests in the products of intellectual or artistic endeavour."[9]

This statement, that the entities are aspects of user interests, could be interpreted to mean that the point of view of the user defines the group 1 entities. This interpretation is not borne out in the remainder of the document, however.

In the physical world, there is no way to separate the work or the expression from the manifestation that "embodies" it. Metadata, however, is an interpretation of the world as data, therefore it is not bound by physical world constraints. Decisions regarding metadata are based on use cases and expected technological constraints. On close reading, the text of the FRBR document leaves open some possibilities that are not included in the three diagrams, yet nowhere does the document explicitly describe a decision to reduce FRBR to the E-R model of the diagrams. Where the two are possibly in conflict, it isn't clear what the reader should conclude.

One possibility is that the E-R model and notation that was employed were not entirely suitable to the meaning of the text. The FRBR group made use of an early version of the E-R modeling concept and notation. There have been richer models in use since the early 1990's that allow the expression of more than entities and relationships: they can indicate inheritance, more refined cardinality, activity, and communication. The most common modeling language today is the Unified Modeling Language (UML)[10], which was developed during the 1990's, at the same time that the FRBR Study Group was active. UML, as used today, has fourteen different diagram types, modeling both structures and behaviors. Had UML been available to the Study Group the outcome of the study might have been different.

Technology and Cataloging: Uneasy Peace

The library profession, in particular that segment of the profession that works to solve information exchange problems at a global level, has a long history of accepting standards because they are standards. This is essential if libraries are to contribute their data to a central data repository, or to exchange data as part of their normal operations. Standards in that community are generally developed in a top-down manner, where, as in the case of FRBR, a small group of experts is assigned the task of developing a standard that others will use.

Unlike most other communities, the library community insists still that it can develop standards that are unaffected by technology requirements. It is, however, obvious that any data created today will be processed by computers, will probably by stored in databases, will be searched using database search capabilities, and will be accessed by its users over a computer network. FRBR was developed within the IFLA Cataloging section whereas in a traditional business case an E-R data model would be developed in a collaboration between data designers and subject specialists.

It is notable that some influential members of the FRBR Study Group were responsible for injecting technology considerations into the solution in the form of a methodology designed for relational database development, but no technology specialists were included in the process. In part this could be because the FRBR Study Group insisted throughout that the solution they had developed was technology-neutral. Actually, it was more than technology-neutral, it was at times presented as being free of any technological constraints, even though it used a known technology method to guide the work of the group.

> "FRBR is not a data model. FRBR is not a metadata scheme. FRBR is not a system design structure. It is a conceptual model of the bibliographic universe."[11]

Although the FRBR Study Group, which produced the FRBR document, insisted on technology neutrality, the IFLA FRBR Review Group, which has taken on responsibility for the model since its publication, has taken the conceptual model and translated it directly into a semantic web ontology. However, the differences between the semantic web standards and the original E-R model are such that the two, while superficially similar, have significant differences in functionality.

There was also a direct line between FRBR and the development of the newest cataloging rules, Resource Description and Access (RDA). Not only did RDA choose to adopt the FRBR model in shaping the cataloging rules, but at least two of the influential members of the FRBR Study Group, Tom Delsey and Barbara Tillett, were also deeply involved in the development of RDA. Although RDA makes use of the FRBR entities, it too claims to be technology-neutral. There also was not, and has not yet been, a similar effort to develop technology that would allow an implementation of RDA in an appropriate technology. There has been the development of an RDF ontology based on identified data elements from the RDA text, but as yet there has been no technology development that would allow the creation of RDA data.

Here one might wonder about the Library of Congress's Bibliographic Framework (BIBFRAME) effort.[12] Contrary to the development of FRBR and RDA, BIBFRAME is a technology project that is using incremental development of working code to achieve an operational linked data environment for library data. It is understandable that there is a great deal of concern about the transformation of current bibliographic data to a new technology platform, and the project has so far looked more to legacy data than to new capabilities that more modern technology can facilitate, which could make the end result less forward-looking than it could have been. Note that BIBFRAME does not claim to be an implementation of FRBR, but does state that the model of entities and relationships that have been proposed can accommodate both FRBR and RDA. This is indeed compatible with the interpretation of FRBR as a conceptual model, not a record or data format.

Variations on the FRBR Model

Since its publication, the discussion of FRBR in the library community has resulted in so many articles, books, and reports that in 2008 the FRBR review group ceased updating its bibliography.[13] While many of the early works sought to explain FRBR to fellow librarians, such as Maxwell's "Guide for the Perplexed,"[14] some recent writings, like the 2012 special issue of Cataloging and Classification Quarterly,[15] are emphasizing uses of FRBR in library and archives projects. As should be expected, there is also some criticism of the FRBR model based on experience.[16]

The above mentioned sources address the use of FRBR in a traditional cultural heritage context. However, the FRBR conceptual model has also been taken up by data developers in communities not directly involved in libraries. The great variety of solutions that have been developed gives testimony to the universality of concepts that the FRBR study group articulated, in particular those of FRBR Group 1 (WEMI). Each operational model makes its own interpretation of FRBR for its specific needs. The variability of the non-library solutions shows both a more sophisticated approach to technology in those environments as well as an indication of a cultural difference from standard library practice.

The models that follow in this section are a sample of ones that make their own use of FRBR concepts, yet provide a their own interpretation of the model. Some are formal models in themselves, but not all. Few have been tested on a significant amount of instance data, but this is also true of FRBR itself. With the exception of FRBRer, none arise out of the library cataloging community.

These variations neither disprove nor negate the work of the FRBR Study Group. A high level model not only survives variant interpretations, but those interpretations can provide a proof of the legitimacy of the core concepts of the model. In general, conceptual models do undergo revision as they move into the logical phase of data analysis.

Because we are in 2014, and because technology changes over time, the models presented here do not use the E-R methodology of the FRBR study. Had some systems developed actual relational databases extending the FRBR E-R conceptual model to a physical model in 1998 when the first version of FRBR was issued, those implementers would probably be looking today to upgrade their databases using current technology. The relational database of normalized tables responding to the SQL query language are being replaced by key/value or triple-based databases that are referred to as "noSQL" databases because they do not require the table structure of the relational database management system in order to perform efficient queries. Many of these models use semantic web standards, and thus they define classes and properties where E-R modeling uses entities and relations. FRBR's attributes, which are listed for each entity, are treated as properties in the semantic web model, as these also model the relationship between two things, in this case an entity and a value.

FRBRer

The IFLA FRBR Review Group formed a FRBR Namespace Project in 2007, led by Gordon Dunsire, to "define appropriate namespaces for FRBR in RDF and other appropriate syntaxes." RDF is the primary metadata standard of the Semantic Web, and means *Resource Description Framework*.[17] Once namespaces were developed for the FRBR family of models (FRBR, FRAD, FRSAD) it became possible to define the FRBR vocabulary as a formal RDF ontology. A version of FRBRer (which stands for the FRBR entity-relation model) in RDF is in the Open Metadata Registry[18] with a status of "published," presumably indicating this is a final and approved version.

FRBRer includes the FRBR entities, all of their relations, and the FRBR-defined attributes. It also includes the group 2 entity family which became part of the FR family in the develop of the Functional Requirements for Authority Records (FRAD). FRBRer treats all of the entities as disjoint from one another, which means that the entities cannot share attributes (called *properties* in RDF). The exception to this is that all FRBRer entities can have a "hasSubject" relationship with work. Where a relationship was listed for more than one entity in the FRBR study, FRBRer creates those as separate, entity-specific relationships, such as *is created by (corporate body)* and *is created by (person)*. In keeping with the wishes of the FRBR Review Group, FRBRer does not define super-classes for the three groups.

FRBRer appears to be intended as a strict reading of the FRBR E-R diagrams. It declares all entities and attributes as being disjoint from each other, meaning that no alternative combination of entities and attributes is allowed. It defines this aspect as well as the cardinality of the relationships between entities (one-to-one, one-to-many) using the W3C's Web Ontology Language (OWL). OWL does not support rules for creation or use of data, but is a language that describes potential interpretations of data. Use of these interpretations is generally considered to be optional in the open Web, and requires software, called a "reasoner," that is able to make the interpretations. FRBRer therefore demonstrates some conflict between the desire for a strict interpretation of FRBR and the intention of linked data in general to allow linking in an open Web environment that supports the points of view of any number of communities.

FRBRCore

FRBRCore was the first FRBR ontology created using RDF class and property definitions. It was developed by Ian Davis and Richard Newman in 2005, and was last updated in 2009. It varies in some ways from the model described in the IFLA FRBR document and was never accepted by the FRBR Review Group. It has, however, been frequently used in linked data projects. The Linked Open Vocabularies site,[19] which gathers statistics on the use of linked data on the web, reports that FRBRCore is used in twelve vocabularies, and appears in nearly 30 million instances. The primary users are the union catalogs of Bavaria, Berlin and Brandenber, and the union catalogs of Hessen and parts of Rhineland. These catalogs account for about 24 million instances of FRBRCore in triples. By far the most heavily used elements are:

- Manifestation (class)
- Item (class)
- exemplar (property)
- owner(property)

Classes *expression* and *work* are used only a few times, and *endeavor* is not used at all.

The "core" aspect of FRBRCore is that it does not include as properties the descriptive elements that in FRBR are listed as attributes of the entities. Instead, it focuses on the ten entities (*family* had not yet been added during FRBRCore's development) and the relationships between entities. The list of relationships includes those primary relationships between the entities that are shown on the diagrams in the FRBR document. These are the Group 1 relationships like *embodimentOf* (between *manifestation* and *expression*) and *realizationOf* (between *expression* and *work*), and the group 2 relationships *creator*, *owner*, *producer*, *realizer*. It also includes all of the entity/entity relationships listed in section 5 of the FRBR report, such as *adaptation* of and *part of.*

As it has been defined, FRBRCore represents what in the semantic web standard are called the object properties of the FRBR model. These are the properties that link entities to other entities. The data properties, equivalent to the FRBR attributes and used for description, are not included in FRBRCore. Anyone implementing a vocabulary that uses some or all FRBRCore classes can model their own data properties in relation to FRBRCore.

FRBRCore defines super classes for each of the FRBR groups. These are *Endeavor* (group 1), *ResponsibleEntity* (group 2), and *Subject* (group 3). This provides the capability of addressing the group as a whole, both in vocabulary development and in instance data. The relationships between group 1 entities are sub-properties of a property called *relatedEndeavor*. Some of these relationships are specific to more than one entity, such as *summarizationOf*, which can be applied to any combination of *expression* and *work*. This is one possible interpretation of section 5 of the FRBR document, which sometimes lists the same relationship between different entities.

There are a small number of sub-classes of work and expression. For example, *legal work* and *academic work* are subclasses of work, and moving picture and text are subclasses of expression. IFLA FRBR treats these as attributes of the entities, (form of work and form of expression) but in RDF these could be logically viewed as sub-classes. As we'll see below, the FaBiO ontology uses subclassing of FRBR work and expression extensively in its implementation.

FRBRCore does designate some classes to be disjoint from each other. It includes a declaration of disjointness between the super-classes that represent the groups, declares the group 1, group 2 and group 3 entities to be disjoint from each other, but allows all FRBR classes to be members of the class *Subject*. The logical result of this is that vocabularies using FRBRCore with their own defined set of data properties should not define a property with a domain of two disjoint classes if they wish to remain consistent with FRBRCore definitions. This could provide a reason to define properties common to more than one class at the super-class level. It is also possible that a "highest class," similar to the RDF definition of *Thing*, the highest level of any hierarchy of concepts, may be needed to appropriately locate any data properties that can be generalized to all of the FRBR model, such as names and identifiers.

There is conceptual overlap between the definitions of entities in the E-R model and the use of classes in the semantic web, but with some significant differences. Without going into great detail, let us just say that classes in the semantic web add meaning but do not restrict use. Because the semantic web model is intended to be applied in the open space of the web, with no limit on the numbers or types of contributions, the use of classes provides context for data not limitations. Any element can belong to any number of classes, just as you could belong to classes "librarian," "gardener," "parent," and "Tweeter." In current semantic web technology there does not exist a method to limit a property to a single class. This is because within the open context of the web some other person or community may be legitimately seeing the same information from a different point of view. The FRBR E-R model arose out of the closed environment of the database, and therefore interprets the relationship between entities (which would be classes in a semantic web analysis) and attributes (properties) as being restrictive rather than informative.

FRBR-Aligned Bibliographic Ontology (FaBiO)

FaBiO[20] is one module within the Semantic Publishing and Referencing (SPAR).[21] Other SPAR modules define the scholarly publication work-flow, citations, and more. Fabio uses the conceptual model of FRBR and the RDF ontology of FRBRCore as its starting point to create a categorization of publication types. Fabio work is a sub-class of FRBRCore *work* that is limited to published or potentially publishable items. Sub-classed to this are work types that are relevant to the SPAR effort:

> announcement, artistic work, biography, case for support, correction, critical edition, dataset, essay, examination paper, grant application, image, instructional work, metadata, model, opinion, policy, proposition, questionnaire, reference work, reply, report, research paper, review, sound recording, specification, vocabulary, work collection, work package, working paper

Similarly, FaBiO extends the FRBRCore expression to include many sub-classes.

> Gantt chart, abstract, addendum, article, audio document, book, brief report, call for applications, case for support document, chapter, comment, computer program, conference paper, conference poster, cover, data file, data management plan, database, dust jacket, e-mail, editorial, excerpt, expression collection, figure, grant application document, index, instruction manual, lecture notes, letter, manuscript, metadata document, movie, news item, oration, patent application document, patent document, periodical issue, periodical volume, personal communication, policy document, presentation, project plan, quotation, rapid communication, report document, repository, spreadsheet, structured summary, supplement, supplementary information file, table, vocabulary document, vocabulary mapping document, web content, workshop paper

Unlike library data, the FaBiO manifestation is limited to a physical manifestation, and does not include the descriptive information that is the norm in library data. Thus the FaBiO structure looks something like figure 4.

FaBiO also defines a select set elements that libraries would consider to be "descriptive" – that is, elements like *title*, *publisher*, *date*. These are defined as properties within the scope of FRBRCore *endeavor*, which means that their use is not restricted to describing only one of the FRBR group 1 entities. This is quite different from the model as described in the FRBR document, which assigns a set of descriptive attributes to each entity.

It seems that a primary goal of FaBiO is to bring out meaningful qualities of resources within the academic publishing sphere. FRBRCore facilitates this view by defining the ten FRBR entities as semantic web classes, to which FaBiO adds numerous sub-classes, and leaving the descriptive attributes relatively open. The type of detailed description that is the focus of library cataloging is of less interest to the project; the descriptive goal in FaBiO is identification of resources at the level of detail of citations.

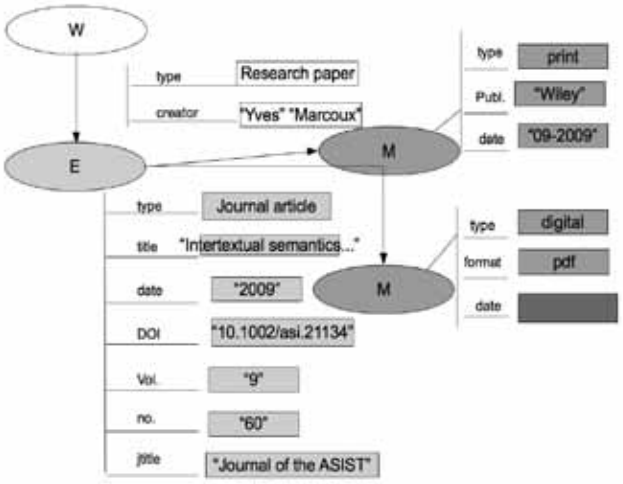

Fig. 4 A FaBiO bibliographic description (large)

Object-oriented FRBRoo

Although some people see a logical or temporal flow from *work* to *expression*, then *manifestation* and *item*, the fact is that FRBR provides a decidedly static view of the bibliographic object. This is indeed the traditional library catalog view, and it originates from the cataloger's encounter with the item in hand, which is a finished product embracing all of the FRBR entities. The entities of FRBR group 1 (WEMI) have relationships like *realize* and *embody* and *exemplify* without the intervention of sentient beings, such that it appears that the WEMI entities are actors themselves who perform these actions. This static view mimics that of database design, where tables store data and relationships that are the result of exterior workflows. While a database continues to be updated with new data, its design, and thus the types of data and relationships between those types of data, remains the same.

The object-oriented approach to data modeling taken by the International Council for Museums in the development of the CIDOC Conceptual Reference Model (CIDOC CRM)[22] lends itself to an action-based approach to the data. An effort is underway to align the CIDOC CRM with the concepts of FRBR.[23] Note that both are conceptual models, and yet the difference in the technologies underlying the models (E-R vs. object-oriented) means that different conceptual possibilities are available to the modelers.

> " FRBRer envisions bibliographic entities as static, ever-existing things that come from nowhere, and overlooks the complicated path from the initial idea for a new work in a creator's mind to the physical item in a user's hands through the dramatically important decision-making on behalf of publishers. As a matter of fact, bibliographic records do contain implicit information about that complicated path and the relationships it implies between and among bibliographic objects; FRBRoo digs that implicit information out of bibliographic structures, e.g. the precise meaning of "date of publication".[24]

In CIDOC CRM, temporal entities play a significant role. For example, there are entities and properties for *Creation*, *Production*, and *Attribute Assignment*.

The FRBRoo analysis found that the FRBR work covered a number of different defined entities in the CIDOC CRM model. To address this, the FRBR work was elevated to a new position as a super-class over more specific *work*-related classes. In FRBRoo, there are individual works (those that are expressed as a single expression), there are complex works (those that combine two or more individual works), and there are also publication works (works created with publication as an expected outcome), aggregation works (similar to the aggregating work outcome of the FRBR aggregates study)[25], and others. These sub-classes then interact with classes representing events and actors.

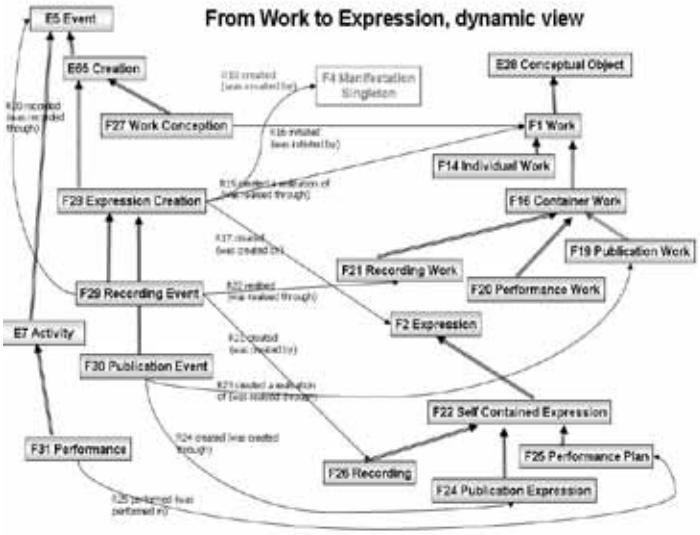

Fig. 5 Dynamic view of work and expression (large)

FRBRoo is highly complex, but it does extend FRBR into areas that look less like the traditional published materials that library cataloging emphasizes. FRBRoo is better suited to the fine arts (which can consist of unitary intellectual resources that are not easily divided into the four WEMI entities) and for the performing arts, where the interaction between works and expressions is particularly difficult to define with a static model.

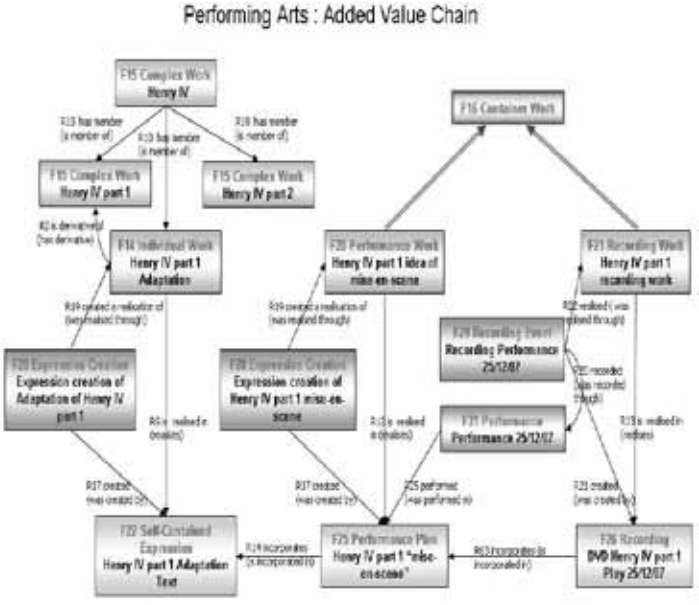

Fig. 6 FRBRoo treatment of performances (large)

<indecs> event-oriented model

The <indecs> metadata framework[26] is a model of commerce developed in the publishing environment. The original presentation of the indecs model used a simple triangle of entities and relationships to define the commercial view of the publishing activity. (Fig. 7)

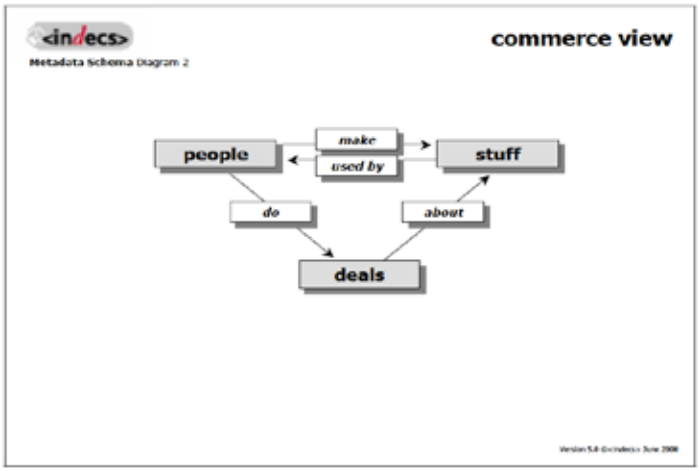

Fig. 7 "People create stuff..." (large)

<indecs> concerns itself with the many steps in the publishing flow, each of which have parties, actions, and outcomes. Important for this model is identification of the responsibilities and related rights for each of the outcomes. Of course, bibliographic data is part of the overall metadata set for publishers. <indecs> interprets FRBR group 1 entities as creation types. These are similar to the IFLA FRBR but <indecs> interprets them as events that take place in time and in a location, with responsible parties and related rights. (Fig. 8)

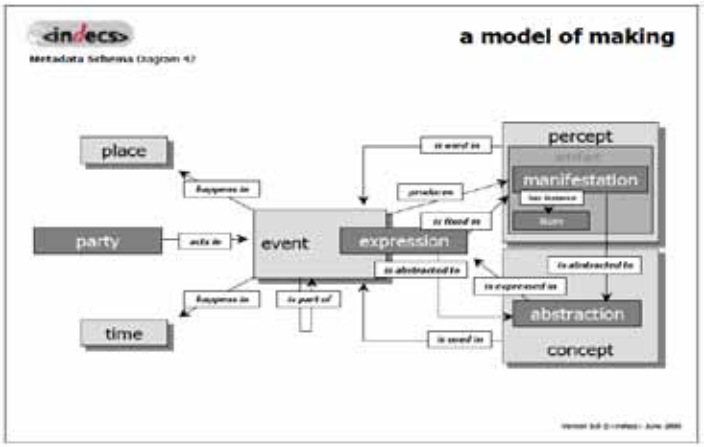

Fig. 8 Event-oriented bibliographic model (large)

The definitions of expression and of work (which <indecs> calls *abstraction*) differ from the FRBR definitions:

> expression is a performance or event that creates a reproducible record. Expression is important because there are often rights associated with the expression that are important to the publishing workflow.
>
> abstraction (FRBR work) in <indecs> follows the definition in the Berne Convention of intellectual property. Rights define an abstraction, such that two resources with different rights cannot be the same abstraction. This separates original works and their translations into separate works because the translation has separate rights.[27]

As acknowledged as a possibility in the FRBR document,[28] this is a community whose definition of work differs from that of standard library cataloging. Yet, unlike the IFLA definition of *work*, which does not provide specific rules for determining the nature of a work, the <indecs> community has an explicit rule for delimiting the boundaries between works based on the assignment of intellectual property rights. The <indecs> definition may require the use of attributes that are not the same as those assigned to the FRBR entities as defined by IFLA, because an <indecs> work description may need to include both the creator of the original work as well as the creator of the adapted work. More analysis is needed to clarify how such a view might interact with the traditional library view.

Conclusion

There is a wide range of users and creators of bibliographic data, from casual readers to booksellers to academic researchers. It would be unreasonable to expect that one set of data can respond to the great variety of needs of these users. There is growing evidence that FRBR, as a conceptual model, can be used across communities. The differences in interpretation between these communities means that the data from them will not create a single bibliographic standard. Yet at the same time, the use of the FRBR concepts provides a point of departure for the sharing of each community's standardization of the bibliographic description.

The existence of a high-level, generalized model of the bibliographic universe can create points of connection between diverse points of view. As bibliographic data develops within the context of these diverse interpretations of the FRBR conceptual model, and as that data enters the global data cloud that is the Web, we will face the challenge of working with these variations within a single data universe. The examples in this article are evidence that key aspects of a bibliographic standard need to be flexibility and extensibility.

References

1. Murray, Ronald J. "The FRBR-Theoretic Library: The Role of Conceptual Data Modeling in Cultural Heritage Information System Design." (paper presented at iPRES 2008. London, September 29-30, 2008) Accessed March 26, 2014. http://www.bl.uk/ipres2008/presentations_day1/26_Murray.pdf
2. Barbara Tillett. IFLA Study on the Functional Requirements of Bibliographic Records : Theoretical and Practical Foundations, (April, 1994). Access March 26, 2014. http://www.ifla.org/IV/ifla60/60-tilb.htm.
3. IFLA Study Group on the Functional Requirements of Bibliographic Records. "Terms of Reference for a Study of the Functional Requirements for Bibliographic Records". In: FRBR: Hype or Cure-All?. (New York: Haworth Information Press, 2005)
4. Richard Perkinson. Data analysis : the key to data base design. (Wellesley, Mass: QED Information Sciences, 1984).
5. International Federation of Library Associations, Study Group on the Functional Requirements for Bibliographic Records "Functional Requirements for Bibliographic Records." 2009. Accessed March 26, 2014. http://archive.ifla.org/VII/s13/frbr/frbr_2008.pdf p. 9.
6. IFLA Study Group on FRBR, p. 9
7. Based on: Simsion, Graeme C, and Graham C. Witt. Data Modeling Essentials. Amsterdam [etc.: Morgan Kaufmann, 2005.
8. Gordon Dunsire. "Resource and Work, Expression, Manifestation, Item." (October, 2013). Accessed March 26, 2014. http://www.ifla.org/files/assets/cataloguing/isbd/OtherDocumentation/resource-wemi.pdf
9. IFLA Study Group on FRBR, p.13
10. Object Management Group. "Unified Modeling Language." Accessed March 26, 2014. http://uml.org/
11. Barbara Tillett. "FRBR and Cataloging for the Future." In: FRBR: Hype or Cure-All?. (New York: Haworth Information Press, 2005) p. 197
12. Library of Congress. "Bibliographic Framework Initiative." Accessed March 26, 2014. http://www.loc.gov/bibframe/
13. "About the FRBR Review Group." Accessed July 3, 2014. http://www.ifla.org/about-the-frbr-review-group
14. Maxwell, Robert L. FRBR: A Guide for the Perplexed. Chicago: American Library Association, 2008.
15. " The FRBR Family of Conceptual Models: Toward a Linked Future." Cataloguing and Classification Quarterly, v. 50, n. 5-7, 2012.
16. Le Boeuf, Patrick. "FRBR: Hype or Cure-All? Introdcution." In Functional Requirements for Bibliographic Records: Hype or Cure-All?, pp. 1–13. New York: Haworth Information Press, 2005.
17. World Wide Web Consortium. "Resource Description Framework (RDF)." Accessed March 26, 2014. http://www.w3.org/RDF/
18. International Federation of Library Associations, FRBR Study Group. "The FRBRer Model." Accessed July 2, 2014. http://iflastandards.info/ns/fr/frbr/frbrer/
19. "Linked Open Vocabularies (LOV)." Accessed March 26, 2014. http://lov.okfn.org/dataset/lov/
20. David Shotton and Silvio Peroni. "FaBiO: FRBR Aligned Bibliographic Ontology." Accessed March 26, 2014. http://purl.org/spar/Fabio
21. "SPAR – Semantic Publishing and Referencing." Accessed March 26, 2014. http://sempublishing.sourceforge.net/
22. International Council for Museums. "The CIDOC CRM." Accessed March 27, 2014. http://cidoc-crm.org/
23. "Working Group on FRBR/CRM Dialog." Accessed March 27, 2014. http://www.ifla.org/node/928
24. International Working Group on FRBR and CIDOC CRM Harmonisation. "FRBR, object-oriented definition and mapping from FRBRer, FRAD and FRSAD." Version 2.0. May, 2013. p. 13
25. O'Neill, Ed, Anders Cato, Paula Goossens, Judy Kuhagen, Barbara Tillett, Carol Van Nuys, and Maja Žumer. "Final Report of the Working Group on Aggregates. Collections" (2011). Accessed March 27, 2014. http://www.ifla.org/files/assets/cataloguing/frbrrg/AggregatesFinalReport.pdf

26. Godfrey Rust and Mark Bide. "The <indecs> metadata framework; Principles , model and data dictionary." (June, 2000) Accessed March 27, 2014. http://ww.doi.org/topics/indecs/indecs_framework_2000.pdf
27. Rust and Bide, pp. 27-28
28. "The concept of what constitutes a work and where the line of demarcation lies between one work and another may in fact be viewed differently from one culture to another. Consequently the bibliographic conventions established by various cultures or national groups may differ in terms of the criteria they use for determining the boundaries between one work and another." IFLA Study Group on FRBR, p. 17