Abstract

I propose that the proper procedure for positing hierarchical levels between the level of the Work and the level of the Manifestation (Edition), in the FRBR (or WEMI) hierarchy, is to make a study of all the editions of selected works as recorded in the OCLC WorldCat database, and to determine inductively what are the most useful hierarchical levels between the top level of the Work and the lower and traditional level of the Manifestation or Edition. For this purpose, a search was made in OCLC WorldCat for all works by children's author Eileen Christelow, and 470 of the bibliographic records retrieved were assigned to works and then subclassified by key properties. The intermediate levels that popped out from this exploratory study were associated with the properties of language and format. Although the overall hierarchicalness of the WEMI hierarchy is defended, it is unclear if language and format are hierarchically related to each other.

But, there is more. I undertook a literature review, imperfect though that has been, and I have incorporated some musings on the literature review into this report. In particular, I will introduce the idea of the inequality of library users into our model of the bibliographic universe of which FRBR is part.

Introduction: What Is the FRBR or WEMI Hierarchy?

The FRBR or WEMI hierarchy -- we will suggest below a reason for preferring the term FRBR hierarchy -- was introduced in the 1998 book Functional Requirements for Bibliographic Records for the purpose of organizing bibliographic editions into larger categories of editions which share things in common. The hierarchy was articulated with four levels. From largest to smallest, they are: (1) work, (2) expression, (3) manifestation, (4) item. (Functional Requirements for Bibliographic Records and Work, Expression, Manifestation, Item -- the source of the two acronyms.) It is diagrammed in figure 3.1, page 14 of FRBR, reproduced here:
Group 1 Entities and Primary Relationships

- **Work**
  - is realized through **Expression**
  - is embodied in **Manifestation**
  - is exemplified by **Item**

Of the four hierarchical levels, the lowest two are easiest to be grasped by librarians, so we will introduce them first.

**Level 3. Manifestation** is the term used in the FRBR hierarchy for an edition. When we use "edition" to clarify "manifestation," edition refers not just to books, but to all types of creative and intellectual publications and unpublished materials, including sound recordings -- visual recordings such as videodiscs (including DVD's and Blu-Rays) and videocassettes (including VHS's and PAL's) -- computer files -- as well as books (which in turn include books you read by seeing, visual books, and books you read by touching, Braille books).

**Level 4. Item** refers to a copy of an edition. If the edition is unpublished, it might exist in only a single copy or item.

**Level 1.** But the highest level of the hierarchy, the work, is the level most often referred to by people who are not librarians or bibliographers. Three examples of works are *Pride and Prejudice* by Jane Austen and *The Origin of the Species by Means of Natural Selection* by Charles Darwin and *King Kong* (which we may or may not come back to).

**Level 2.** Finally we mention the expression, the level intermediate between the work and the edition, and the most nebulous of the four levels in my view. Coyle (*FRBR, Before and After*, p. 89) also makes this point, stating "The expression is generally considered the most conceptually elusive of the Group 1 entities." This paper is an attempt to pin down what a level between the work and the edition might look like.

The claim that the level of the expression is less clear intuitively than the other three levels is supported by the fact that *RDA: Resource Description and Access* proposes to make use of the other three levels, work, manifestation, and item, but not the level of expression. The authors of RDA, in my opinion, judged the level of expression as "not ready for prime time."
Level 3.5. In addition to these four levels, when discussing the FRBR hierarchy in the library world, it is helpful to introduce a fifth level, the level of the bibliographic record. This level is intermediate between the manifestation or edition and the item. In the library world as it currently is, it is pragmatically useful to picture the bottom levels like this:

![Diagram of the hierarchy showing Manifestation at the top, Bibliographic Record in the middle, and Item at the bottom.]

The Bibliographic Record is a level of convenience. As librarians or catalogers, we have available to us an immense store or treasure of bibliographic records, which are our inheritance from our predecessors in the cataloging business. There is an especially large inheritance of bibliographic records in the OCLC WorldCat database. Instead of setting out to create records for works, our predecessors in the cataloging business set out to create records for a level close to editions (a WorldCat design decision which we will attempt to justify in the literature review section following). But the actual result, as embodied in OCLC WorldCat, has not been by any means one record for each edition. Bibliographic records are noticeably more numerous than editions. Hence the extra level portrayed in the hierarchy just above.

**Literature Review**

I began my literature review with a study of Karen Coyle's book *FRBR, Before and After* which she has made freely available as an open-access publication on the World Wide Web. It has informed my thinking on the subject of FRBR, and serves as the jumping-off point for several of the musings which will appear near the end of this talk.

As I studied Coyle's *FRBR, Before and After* I found many points to agree with and a few where I will take a different position in the musings below. But in the most important area where I will stake out a new position, I think Coyle expresses the consensus of most thinkers and writers on the library catalog, and I think the issue is too important to be postponed and will grapple with it here and now.

**The Inequality of Library Users in the Academic Library**

It is important to pose the question, why do we have separate bibliographic records for different editions? The answer, I propose, is in an idea I will call the inequality of library users in the academic library. This can be approached through a quotation from Thomas Pynchon's novel *The Crying of Lot 49*. First the quotation (but replacing some of Pynchon's pronouns by nouns as indicated by the context), and then a little explanation.

"Oedipa showed [Professor Bortz] the paperback with the line in it. . . .
[Bortz] announced, 'I've been pirated, me and Wharfinger, we've been
Bowdlerized in reverse . . .' He flipped to the front, to see who'd re-edited his edition of Wharfinger. . . . He looked at the sun through a page or two. 'Offset.' Brought his nose close to the text. 'Misprints. Gah. Corrupt.' He dropped the book on the grass and looked at it with loathing."


In this passage, Oedipa Maas is a student. Prof. Bortz is a teacher. Prof. Bortz is intimately familiar with the edition of Wharfinger's plays which he (Bortz) edited, and can see by applying several quick tests that the edition his student has gotten hold of is inferior.

Notice that Professor Bortz uses five technical terms of bibliography and one flat-out value judgment in a single "paragraph" of speech. The five technical terms are:

-- pirated,
-- Bowdlerized in reverse (the reader might think of this as a term coined by Prof. Bortz on the spot from the existing term Bowdlerized),
-- re-edited,
-- offset,
-- misprints,

and the flat-out value judgment: corrupt.

This passage illustrates a general principle. Teachers can tell the difference between good editions and bad editions of works they are familiar with, by applying tests or criteria. Because of that, in the academic setting, teachers can be recognized as a group of library users with authority -- like power users, or even users with administrative privileges, if you will.

It is useless for academic librarians to conduct user studies, if we omit from our theory of the bibliographic environment the fact that some users have more authority than other users. They are power users of books -- of movies -- of all kinds of works collected by the library, according to the various specialties of the various teachers -- and they are power users of library catalogs, as well. That is what I propose to call the inequality of library users in the academic library. As library users, teachers have more authority than students, because teachers have more experience and more knowledge.

Furthermore, teachers are engaged in an extensive program of behavior modification of the students, which we call education. This program of behavior modification includes getting the students to read longer texts, to watch longer films and to listen to longer musical compositions -- lengthening their attention span so they can interact with and appreciate longer works. It also includes cultivating students' ability to compare works -- to isolate and appreciate the contributions of individual human creators -- and getting students to compose their own practice works (term papers, student films, musical compositions, paintings, etc.). It also includes a bibliographic component, in which the teachers' ability to cite the sources they have used rubs off on their students.
This process of behavior modification is notoriously slow. Notoriously. At the university where I'm a librarian, we have a once-a-year event called the Life of the Scholar Multidisciplinary Conference, which is an opportunity for selected mentored students to showcase their research or occasionally other forms of creativity before an audience of faculty and fellow-students. A professor recently remarked to me that he likes the Multidisciplinary Conference because it is a reminder that the behavior modification program sometimes works. (I added the "behavior modification" wording. I believe his words were "what we are doing.")

To return to my point. The Pynchon quotation and discussion illustrates why our library catalogs need to accommodate information about editions, in addition to information about works. But existing library catalogs do not merely accommodate information about editions; in catalogs built following the standards of AACR1, AACR2, and MARC, it is accurate to say that the catalog consists of records for editions.

OCLC WorldCat is a database built by thousands of users (librarians and library staff) with varying degrees of authority to make changes to the database, jointly making a very large investment of time towards creating metadata at a level near the FRBR level of the manifestation or edition. That is because real people like the fictional Professor Bortz constitute an authoritative group of users of bibliographic records who care strongly about the differences between one edition and another edition of the same work. A lot of work has been invested in metadata for editions, because an authoritative group of users of bibliographic records care about editions. The rediscovery of the importance of the work, which has been prompted by the reflections of various researchers on *Functional Requirements for Bibliographic Records*, ought not to blind us to the importance of editions.

The concept of the inequality of users pays dividends elsewhere in the analysis. For example, Coyle writes (*FRBR, Before and After*, p. 100) "Some of the attributes are only of low importance, which brings into question why they are even included as necessary attributes. For example, the Work/Work relationships summarization, adaptation, transformation, and imitation are listed as low value for identify and select. Had the focus of the report truly been user needs, it is doubtful that those attributes would have been included." -- But once we add to our model the concept that some users of bibliographic records have more authority than others, then this critique loses weight. Teachers care whether you read the whole original work or an adaptation or summarization (summary) of the work. That is why indications that a work summarizes, or is adapted from, another work are included in bibliographic records -- because some authoritative library users care whether a work is a summary, abridgment, parody, etc. of another work. Teacher needs are user needs. That's so important I'd better say it twice: Teacher needs are user needs.

**Method**

For this exploratory study, we chose to focus on works of one author with fewer than 1500 bibliographic records (since 1500 is the limit of bibliographic records which can be retrieved in one search in OCLC WorldCat).
Eileen Christelow is a United States author and illustrator of children's books. In addition to illustrating books she herself writes, she has also illustrated books by other authors for children. In this study, however, we will focus on works in which she is identified as the author.

The idea of studying the OCLC bibliographic records for her works was suggested by the example created by Fritz and Fritz, and displayed on their RIMMF website (RDA In Many Metadata Formats), of what records for one of Christelow's books might look like at the work level, at the expression level, and at the manifestation level (see bibliography). This visualization by the Fritzes is extremely helpful, and I hope you'll visit those web pages of theirs if you haven't already.

A total of 652 bibliographic records were retrieved on June 5th, 2017 from the WorldCat database using the OCLC Connexion Client interface, with the search `pn = "christelow, eileen"`. The search can be paraphrased as "bibliographic records in which the personal name Christelow, Eileen appears as an access point." Of these 652 records, 470 listed Christelow as the main entry, which means that the authorship, not only the illustration, is attributed to Christelow. So those 470 records were the focus of this study.

**Results**

The screens of the retrieval set in Connexion Client, with one line (conveniently numbered by OCLC Connexion) for each of the 652 records, were copied into a spreadsheet. A second sheet was then started, and on the second sheet the one-line entries for the bibliographic records were copied and sorted by work, for the cases where Christelow was listed as the author (main entry). This procedure isolated 29 works, distinguished by their title. It involved sublisting translated titles which were matched with their original English title (in accordance with FRBR, the figure of 29 does not count translations separately).

As noted in the introduction, the intermediate levels, in the area of "expression," that were surfaced by this procedure were associated with the properties of language and format.

**Mapping the Expression Territory: Three Views of the Levels Between Work and Manifestation**

When we speak of the "Expression territory," we use "territory" in the sense of "an area which is in the process of being invaded by a tribe, but which has not yet been fully explored or mapped by the invader." The metaphor is of librarians (or catalogers) invading the knowledge domain of the structure or taxonomy of bibliographic works. The specific reference is to the patch or stretch of the FRBR hierarchy between Work and Manifestation. Below, we present three views (or diagrams) of this territory.
(a) Fully hierarchical, with Format above Language

```
Work
  Format Format Format
  Language Language Language
  Manifestation Manifestation Manifestation
  Bib. Record Bib. Record Bib. Record
    Item Item Item
```

In this view or analysis, a work can branch into multiple format-expressions. Each format-expression can branch into multiple language-expressions. Each language-expression can branch into multiple manifestations. Each manifestation can branch into multiple bibliographic records. And each bibliographic record can branch into many items (copies, and volumes of multi-volume work-editions).

(b) Fully hierarchical, with Language above Format

```
Work
  Language Language Language
  Format Format Format
  Manifestation Manifestation Manifestation
  Bib. Record Bib. Record Bib. Record
    Item Item Item
```
In this view or analysis, a work can branch into multiple language-expressions. Each language-expression can branch into multiple format-expressions. Each format-expression can branch into multiple manifestations, and from there the branching is as before.

(c) Partly hierarchical, with Format and Language operating at the same level

In this view or analysis, a work can branch into multiple formats, and a work can simultaneously branch into multiple languages, and each combination of language and format constitutes an expression. These expressions then can branch into multiple manifestations, and the branching continues from there as before.

Between other levels, the mapping is one-to-many, and so it makes sense to speak of a hierarchy. But between Format and Language, it is unclear to me at this point whether the mapping is one-to-many or many-to-many. (For another diagram of many-to-many mapping, cf. Coyle’s diagram in *FRBR, Before and After*, p. 128.) Thus, the hierarchy could be the WFLMI or the WLFMI or the WF/LMI hierarchy. Since the acronym is for the time being indeterminate, we suggest instead calling it the Revised FRBR Hierarchy.

**Selections from the List of Works by Christelow, with Our Outlook on Duplicate Records**

In our list of works by Christelow, the grouping of two or more bib. records under a single manifestation is tantamount to suggesting the records are duplicates, aside from records which describe the same edition using different languages of cataloging. However, our criteria for grouping two or more bib. records under a single manifestation are much less stringent than OCLC’s criteria, which it uses for its Duplicate Detection and Resolution program (very briefly surveyed in Weitz 2016). One of the main differences is that OCLC's Duplicate Detection and
Resolution program refrains from merging records with missing data, whereas our criteria assume a record with missing data and a record with fuller data represent the same edition, if the available data match. (Christen 2012 provides a computer science introduction to the steps involved in deduplication within a database, and matching records across databases. He mainly deals with databases of personal information, but he touches on bibliographic databases.)

A second difference is that we examine fewer fields in the bib. record than OCLC's Duplicate Detection and Resolution program examines. We have examined the following MARC fields (some mainly to detect differences in manifestation, and others mainly to detect samenesses of the work)

- 100 (author Christelow)
- 245 (title and subtitle)
- 250 (edition statement)
- 260 $b (publisher) and $c (year of publication and/or copyright)
- 264 (year of copyright)
- 300 $a (pagination for books, and indication of non-book format) and $c (dimensions, i.e. height and width of books in centimeters) and $e (accompanying materials, for discrimination of kits)
- 336-338 (for discrimination of electronic books and Braille books from printed books)
- 240, 500, and 700 (evidence of original title for translated expressions)
- 500, 650, and 655 (for evidence of board books)

Among the many fields and subfields we have not considered are:

- 020 (ISBN) -- with a couple of exceptions noted below
- 260 $a (place of publication) (exception: 260 $a was considered as evidence for the two Chinese translations of *Five Little Monkeys Sitting in a Tree*)
- 300 $b (for a book, presence of illustrations and presence of color in the illustrations).

Since Christelow writes and illustrates recent American children's books, instead of examining the 300 field subfield $b, we have assumed that there will be no editions lacking illustrations, and that the illustrations will be in color.

Since our search for differences and samenesses in this research was not automated but conducted by the human eye, we judged taking the ISBN into account to be too time-consuming, although a couple of exceptions were made.

In the excerpts below from my list of Christelow's works, long unlabeled numbers are OCLC bibliographic record numbers (not ISBN's). Where the Encoding Level of a record is not noted, it is either Level I or Level Blank (full cataloging).

*First Selection*

<table>
<thead>
<tr>
<th>W1</th>
<th>The Desperate Dog Writes Again.</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>The Desperate Dog Writes Again</td>
</tr>
<tr>
<td>B1</td>
<td>444871628 [Clarion Books] -- 32 pp.; 29 cm</td>
</tr>
</tbody>
</table>
    [same ISBN]

W2  Don't Wake Up Mama! [Later title: Five Little Monkeys Bake a Birthday Cake]
E1a  Don't Wake Up Mama!
B1   25165520 -- 29 pp. ; 21 x 26 cm
B2   608925887 -- Encoding Level M: "Less-than-full cataloging added from tape." No 338 field. -- [No pagination, no dimensions]
    [no ISBN]

Second Selection (translated expressions of Don't Wake Up Mama!)

W2  Don't Wake Up Mama! [Later title: Five Little Monkeys Bake a Birthday Cake]
E2   Shwit! ômma kkaeují ma! [Korean translation of Don't Wake Up Mama!]
E3   Wu zhi xiao hou zi kao sheng ri dan gao [Chinese translation of Five Little Monkeys Bake a Birthday Cake]
E4   Cinco Monitos Hacen un Pastel de Cumpleaños [Spanish translation of Five Little Monkeys Bake a Birthday Cake]
    B1   173748056 -- 1 vol. (unpaged) ; 13 x 16 cm
    B2   973170213 -- Encoding level M -- 28 unnumb pp. ; 13 x 16 cm
    [same ISBN]
    B1   867610686 -- 32 unnumb. pp. ; 17 x 19 cm
    B2   902698496 -- Encoding Level M -- 32 unnumb. pp. ; 17 x 19 cm
    [same ISBN]

A Prima Facie Case for Duplicate Records

We have not demonstrated to OCLC's satisfaction that particular records in OCLC were duplicates (as of the extraction date of June 5th, 2017), but we have made a prima facie case to that effect. Here is the first case, from the first selection above:

W1  The Desperate Dog Writes Again.
E1   The Desperate Dog Writes Again
    B1   444871628 [Clarion Books] -- 32 pp. ; 29 cm
    [same ISBN]
OCLC #865334937 does not record the height (or height and width) of the book. We speculate that is why Duplicate Detection and Resolution refrained from merging it with OCLC #444871628.

Here is the second case, also from the first selection above:

W2  Don't Wake Up Mama! [Later title: Five Little Monkeys Bake a Birthday Cake]
E1a Don't Wake Up Mama!
  B1  25165520 -- 29 pp.; 21 x 26 cm
  B2  608925887 -- Encoding Level M: "Less-than-full cataloging added from tape." No 338 field. -- [No pagination, no dimensions)]
     [no ISBN]

In this case, OCLC #608925887 does not record either the pagination or the dimensions of the book being cataloged (or the ISBN, either). Again, we speculate that is why Duplicate Detection and Resolution refrained from merging it with OCLC #25165520.

Here is the third case, from the second selection above:

W2  Don't Wake Up Mama! [Later title: Five Little Monkeys Bake a Birthday Cake]
E4  Cinco Monitos Hacen un Pastel de Cumpleaños [Spanish translation of Five Little Monkeys Bake a Birthday Cake]
  B1  173748056 -- 1 vol. (unpaged); 13 x 16 cm
  B2  973170213 -- Encoding level M -- 28 unnumb pp.; 13 x 16 cm
     [same ISBN]

In this case, OCLC #173748056 does not record the number of unnumbered pages in the book being cataloged, and we speculate that is why Duplicate Detection and Resolution refrained from merging it with OCLC #973170213. In this case, the record with the lower-ranking Encoding Level (M) provides at least one piece of information not recorded in the record with the higher-ranking Encoding Level.

Here is the fourth case, also from the second selection above:

W2  Don't Wake Up Mama! [Later title: Five Little Monkeys Bake a Birthday Cake]
E4  Cinco Monitos Hacen un Pastel de Cumpleaños [Spanish translation of Five Little Monkeys Bake a Birthday Cake]
  B1  867610686 -- 32 unnumb. pp.; 17 x 19 cm
  B2  902698496 -- Encoding Level M -- 32 unnumb. pp.; 17 x 19 cm
     [same ISBN]
In this case, we have not detected a plausible reason for the non-merger of the two records before our extraction date of June 5th, 2017.

These and similar examples provide the justification for positing the bibliographic record as an additional FRBR hierarchical level, below the level of the manifestation or edition. These examples also suggest to us that the most common reason for two or more bibliographic records for the same manifestation may be non-merger of probable duplicate records, rather than different languages of cataloging (which are considered by OCLC to be non-duplicate records).

**An Additional Result and Refinement: Two Chinese Translations of Five Little Monkeys Sitting in a Tree**

One detailed result that emerged from this study was that there are two Chinese translations of *Five Little Monkeys Sitting in a Tree*. Here is some bibliographic information extracted from the relevant OCLC records:

<table>
<thead>
<tr>
<th>OCLC #</th>
<th>Title in Chinese characters</th>
<th>Title in Romanization</th>
<th>Publisher's Location</th>
<th>Publisher Capitalized</th>
<th>Year of Pub.</th>
<th>Translator</th>
</tr>
</thead>
<tbody>
<tr>
<td>437430856</td>
<td>五隻猴子在樹上</td>
<td>Wu zhi hou zi zai shu shang</td>
<td>T'aipei, Taiwan</td>
<td>Weijin Guoji Guifen Youxian Gongsì</td>
<td>2005</td>
<td>Fāng Suzhēn</td>
</tr>
<tr>
<td>769802739</td>
<td>五只小猴子坐在树上面</td>
<td>Wu zhi xiao hou zi zuo zai shu shang mian</td>
<td>Changsha, (mainland) China</td>
<td>Hunan Shaonian Ertong Chuban She</td>
<td>2010</td>
<td>Rèn Róngróng</td>
</tr>
</tbody>
</table>

The 2005 edition was published in Taiwan and written in traditional Chinese characters; the 2010 edition was published in (mainland) China and written in simplified characters. Moreover, the wording of the title as translated into Chinese is not the same. The titles can be lined up as follows (hóuzi "monkey" is two characters and two syllables; we have added tone marks):

<table>
<thead>
<tr>
<th>OCLC #</th>
<th>Title in Chinese characters</th>
<th>Title in Romanization</th>
<th>Publisher's Location</th>
<th>Publisher Capitalized</th>
<th>Year of Pub.</th>
<th>Translator</th>
</tr>
</thead>
<tbody>
<tr>
<td>437430856</td>
<td>五隻猴子在樹上</td>
<td>Wu zhi hou zi zai shu shang</td>
<td>T'aipei, Taiwan</td>
<td>Weijin Guoji Guifen Youxian Gongsì</td>
<td>2005</td>
<td>Fāng Suzhēn</td>
</tr>
<tr>
<td>769802739</td>
<td>五只小猴子坐在树上面</td>
<td>Wu zhi xiao hou zi zuo zai shu shang mian</td>
<td>Changsha, (mainland) China</td>
<td>Hunan Shaonian Ertong Chuban She</td>
<td>2010</td>
<td>Rèn Róngróng</td>
</tr>
</tbody>
</table>

The difference between traditional and simplified characters can be seen in the column for "zhī" (a word required by Chinese grammar which has no English equivalent; literally "one of a pair" or "animal") and in the column for shù (tree). But there are additional differences in wording: the Taiwan translation of the title did not bother to translate "little" or "sitting" (literally "Five
Monkeys on a Tree" -- Chinese idiom apparently demands "on a tree" instead of "in a tree", since the two translations agree on that). Also, the mainland China translation uses a two-syllable (two-character) word for "on" at the end.

These details point to the conclusion that what we have here are two different translations of the same work into the same target language (Chinese). This conclusion was confirmed by looking at the two translations, which I was able to borrow on interlibrary loan. This in turn suggests the need for a further refinement of the hierarchy we are working with. For the hierarchy used for the classification of plants and animals -- or biological taxonomy -- additional levels may be specified by the use of "sub-" and "super-" prefixed to the terms for the ranks of the hierarchy such as Class, Order, and Family. Here we propose the term "Sub-expression" for two differing translations of the same work into the same language. This phenomenon is well-known from Bible translations, where the various English translations have each their own name, as Revised Standard Version, New King James Version, Today's English version, and so on. But the phenomenon of multiple translations of a work into a single language is not limited to the Bible, as multiple people have tried their hands at translating classics such as Homer's Iliad or Shakespeare's plays into the same language. The same is true for Chinese translations of Christelow's *Five Little Monkeys Sitting in a Tree*. A detailed diagram of the hierarchy thus expanded is given below, under the third musing.

Musings on FRBR

"I was well aware my musings were far from exhaustive" -- Smiraglia 2002 p. 2.

First Musing: I Come Not to Bury FRBR But to Praise It

The publication *Functional Requirements for Bibliographic Records* belongs to an elite category of publications which I will call "publications which propose a research program which is subsequently carried out." Before coming to library science, I studied linguistics in the era when Noam Chomsky's influence was at its peak. I would suggest to you that Chomsky is a mediocre political scientist, but a brilliant linguist. His two early linguistic books *Syntactic Structures* and *Aspects of the Theory of Syntax* set forth a research program for English syntax (syntax being a subfield of linguistics) which was ably and beautifully carried out by his students. Chomsky's program of generative syntax has been given as an example of what Thomas Kuhn called a scientific revolution (Kuhn 1970:92).

My point here is that classics age slowly, and *Functional Requirements for Bibliographic Records* is a classic. It is twenty years old and it is no longer a hot topic in the library world (as it was in 2005; cf. LeBoeuf in the bibliography); but its usefulness, as I see it, is by no means at an end. Its possibilities have not been exhausted. It still deserves to be treated as it authors intended, as a framework for further research. (The first musing should be taken in the context of Coyle's 2015 talk at the Semantic Web in Libraries conference. At that talk, available on YouTube (see bibliography), she was, in my view, unfortunately too harsh on FRBR.)
Second Musing: Living Theories and Dead Theories

FRBR should be treated as a living theory, subject to modification. Living theories evolve.

This answers a conundrum posed by Coyle (FRBR, Before and After, p. 130): "This rigidity, as well as the fact that FRBR is considered “true” in its current form, means that any application must be either FRBR or not-FRBR, thus splitting the bibliographic world into noncompatible factions."

In response to Coyle's comment, we would remark that FRBR is a theory, and thus an approximation. There is no reason to consider it "true in its current form" -- that is, there is no reason to consider it a finalized or perfected version of its theory. To take a concrete example: RDA is neither completely FRBR nor completely non-FRBR. RDA is partially FRBRized.

(Cf. Reitz's definition of FRBRization: "The attempt to model in bibliographic systems the entity structure described in Functional Requirements for Bibliographic Records (FRBR), based on the concepts of work, expression, manifestation, and item. Pronounced 'furburization.' ")

Third Musing: Let the Relationship Between the FRBR Levels Be a Set-Subset Relationship

Coyle remarks (FRBR, Before and After, p. 127) "The structure of FRBR does not represent a taxonomy and the entities are not in a class/sub-class relationship to each other."

That is true as the FRBR document stands, but it is a defect which is easily cured by redefining the relationships between the terms of the hierarchy. We propose to replace "is realized through" and "is embodied in" and "is exemplified by" by the uniform relationship descriptor "is a set of." The preliminary diagram which we propose to replace FRBR figure 3.1 is this:

\[
\text{Work} \quad \rightarrow \quad \text{Expression} \quad \rightarrow \quad \text{Manifestation} \quad \rightarrow \quad \text{Item}
\]

A Work is a set of Expressions

each Expression is a set of Manifestations

each Manifestation is a set of Items

But in view of the results we obtained, this model is elaborated by the addition of the levels of Sub-expression and Bibliographic Record. This results in the following model, which we call the Revised FRBR Hierarchy:
One of the problems which restating the hierarchy in set-subset terms solves is the problem of defining a work as an idea in an author's head (cf. Smiraglia 2002:4). The actual functioning of the art of writing makes that an implausible definition for a work. Works don't exist in complete form in an author's head. One day, the author gets part of his or her idea for the work onto paper. Another day, the author gets another part of his or her idea for the work onto paper. The work is never in the author's or composer's head all at the same time. As a set of ideas, the work is diffuse, and comes out of the head onto the writing surface (or expression canvas) in installments, to which improvements are then made. There has never been any other way in which works of creative persons have been composed. Furthermore, the work and its first expression are composed simultaneously.

Redefining the relations between the levels of the FRBR hierarchy as set-subset relations allows us to say that editions inherit properties and associated information from the higher levels of the hierarchy (works and expressions), properties such as author and plot summary. That is a convenient and now permissible way of describing the facts.
Fourth Musing: Good Stewardship of the Patrimony of Bibliographic Records

The FRBR hierarchy can be thought of as an immensely powerful new tool. When combined with BibFrame, the oncoming replacement for MARC, the FRBR hierarchy (like other immensely powerful tools) has great potential -- for both good and evil. (The best envisioning I have seen of how FRBR and BibFrame might work together is the Fritzes' example of records at the work, expression, and manifestation levels for Christelow's *Five Little Monkeys Wash the Car*, mentioned previously and listed in the bibliography.) This brings us to the topic of good stewardship.

Stewardship of the patrimony of bibliographic records can be compared to stewardship of the earth. As we develop and operationalize BibFrame, we need to think of being good stewards of the MARC bibliographic records we already have, and also good stewards of the framework of new bibliographic records which will blossom in BibFrame. We have inherited the earth from our human and animal ancestors. We are trying not to poison it. In the library profession, we have inherited MARC records, a lot of them, and we are developing new frameworks for the bibliographic records of the future. Our attitude to these inherited bibliographic records and this framework must not be "sicklied o'er with the pale cast of thought." Our attitude must be practical and conservationist-minded.

This talk may strike you as far from practical -- but I am striving to contribute to the development of a *workable* version of a FRBR hierarchy theory, which can ultimately be applied to the restructuring of an enormous database (OCLC WorldCat) without disastrous results. I'll say that again: Without disastrous results!

Coyle's blog entry "The Work" (July 9th, 2017) includes the important idea of a "work decision" which is the cataloger's or bibliographer's decision that a given publication does or does not constitute a separate work from some previous publication. Making this decision very carefully -- and providing guidance to catalogers on how to make this decision -- is an important part of this good stewardship of the bibliographic patrimony. The "work decision" is cognate to, but at a higher level than, the "edition decision" which OCLC has written about extensively in the section of *Bibliographic Formats and Standards* entitled "When To Input a New Record."

Fifth Musing: Who Will Define the Levels of the FRBR Hierarchy?

Who will define the bibliographic work? Imagine a large professional organization with many committees, a large and complex website which is frequently updated, and paid employees as well as unpaid elected leaders who enable the organization to keep up with and respond to changes in the world at a respectable pace. If we focus for a minute on the committees of this large professional organization, we see that its committees allow the organization to keep its positions attuned to the needs of a complex constituency.

The defining body for the bibliographic work will need to resemble this large professional organization. Something like the organization's many committees will be needed in order to keep the definition of the bibliographic work attuned to the needs of a large complex community. The same can be said for the other levels of the FRBR hierarchy.
So -- what are we thinking about? I started out thinking about the American Library Association. Certainly we can do better than that. Let us think of the International Federation of Library Associations. The International Federation of Library Associations is a suitable body to weigh and codify the most useful and user-friendly definition of each level in the FRBR hierarchy.

**Sixth Musing: Reintegration of FRBR Levels: What Catalog Consulters Want**

We will call the people who use library catalogs *consulters* rather than *users* of the catalog, to help us remember that they trust the catalog will be only a brief preliminary step on the way to their use of books and articles, etc. If the catalog is not a brief preliminary step -- if it bogs catalog consulters down -- they will be disappointed, and they will let library staff know about their disappointment.

At the end of her blog post on the work, Karen Coyle writes,

> The "work" is very complex; I would feel less concerned if we had a strong and shared concept of what services we want the work to provide in the future, which should help us decide what to do with the messy legacy that we have today.

In this musing, I want to contribute some nuts and bolts towards "a strong and shared concept of what services we want the work to provide."

What catalog consulters want is information integrated from different levels in the FRBR hierarchy. If BibFrame -- a capital B and a capital F is the greatest plenty of capitalization for it -- If BibFrame results in information about works and editions (manifestations) being stored in separate records, then catalog consulters will want us to put that information back together before we display it to them. That is why we have the kind of legacy of bibliographic records which we do have, namely records for editions, with information from the work level integrated into the records for editions.

That is what our catalog consulters want, because that is what they are familiar with from the world of citations. All style guides known to me are in unanimous agreement on this point. They call for presenting information about the work and information about the edition in a single entry. We are free to tease the FRBR levels apart -- as part of a massive effort to organize our huge OCLC database better -- as long as we reintegrate the levels before showing them to our catalog consulters.

As evidence of this, I submit Smiraglia's citation of one of Panizzi's works. Here it is, from Smiraglia 2001 page 139:


Why do we cite this example? Because it is taken from the bibliography of a book entitled *The Nature of “a Work”: Implications for the Organization of Knowledge*. At the end of an entire
book devoted to the importance of the work (in the FRBR sense of a work), the author chooses to cite his sources in a format which mingles information about editions with information about works! -- He does so, because that is the citation format which knowledge workers in the humanities, the social sciences, the natural sciences, and library science are all equally familiar with.

In the instruction of United States college students, a lot is made of the differences between one style guide and another style guide. One of the missing ingredients in this piece of bibliographic instruction is the vast amount of agreement between style guides! Now is a good juncture for us to make an initial reckoning of what they have in common.

-- They all begin with the name of one or more human beings, if such are readily available in the source being cited, who are stated or implied to have contributed in a pivotal way to the creation of the work. This is a good design feature of citations, and it is not an accident. In our generation, and in the generations that have preceded ours, works are commonly understood as emanating from human beings. And the audience for citations is also commonly understood to be human beings. By naming one or more human beings at the beginning of the citation, the researcher hopes to pique the interest of their audience.

-- They all attempt to transcribe (that is, to copy faithfully from the source) the title of the manifestation of the work which was consulted. This is reflected in our example, which contains one example of "catalogue" with a U and one example of "cataloging" without a U because these spellings of these words were faithfully copied from the source edition!

-- They all provide additional information about the manifestation partly varying depending on the format of the item cited, but often including, as for books, the place of publication, the publisher, and the year of publication. Finally, and to return to our point:

-- They all report mingled properties of the work and of the manifestation which taken together, the citer hopes, will correctly identify both!

Smiraglia, unusually, has added one element not from the edition he consulted, which is the [1841] which we interpret as the year of the work's first publication. This is a useful piece of information for situating works in their original context, and in my humble opinion, it would not be a bad thing if this element was added to future standards of citation.

To conclude the sixth musing: OCLC is currently using a work identifier to link editions of the same work, and this decision seems like a very statesmanlike way to obtain the advantage of grouping editions into works, without forfeiting catalog consulters' ability to view information from work and edition levels reintegrated, or still integrated, in the present state of our art -- in the present state, to be really blunt, of our efforts which might not deserve the name of science, even though we strive to make them deserve it.

**Final Musing: Bibliographic Taxonomy**

In describing the FRBR hierarchy, Coyle and earlier authors have rejected the term taxonomy, but we embrace taxonomy. The levels of the FRBR hierarchy can be compared to the hierarchy of plant and animal species. The *International Code of Nomenclature for Algae, Fungi, and Plants* (2011) says in chapter 1 article 3 that the (taxonomic) ranks in descending order are kingdom, phylum or division, class, order, family, genus, and species. These seven basic ranks
for biological taxonomy are to be compared to the levels of the FRBR hierarchy for bibliography. The FRBR hierarchy is a step forward in understanding the taxonomy of bibliographic works. Biological taxonomy is a hierarchy which involves many levels. We can predict in advance that bibliographic taxonomy will also involve many levels.

Clearly, taxonomy is not all good news. We will use the example of dogs and wolves to reflect on that.

The Wikipedia article "Subspecies of Canis lupus" is very instructive. In a nutshell, it reports that the domesticated dog has been downgraded, scientifically and taxonomically speaking, to a subspecies of the wolf: Canis lupus lupus is the Eurasian or common or gray wolf, while Canis lupus familiaris is the dog (the domestic dog). Besides wolves and dogs (packed into that one species, Canis lupus), the genus Canis also includes several species for jackals, and Wikipedia describes coyote as a North American word for jackal, or perhaps a word for the North American kind of jackal.

Another subspecies of Canis lupus is Canis lupus dingo (the dingo, native to Australia and possibly New Guinea) (Wikipedia article "List of Canis species and subspecies").

The Wikipedia article cites as a source Mammal Species of the World, 3rd edition (2005), which is manifested online at the website Wilson & Reeder's Mammal Species of the World, Third Edition (https://www.departments.bucknell.edu/biology/resources/msw3/) (referred to here as the MSW3 website). I visited that website. I can report that visiting that website is like visiting a marble temple, whereas visiting the Wikipedia article is like reading an article by a group of anonymous but articulate people who have visited a number of marble temples. The MSW3 website is more authoritative, but the Wikipedia article gives a better overview of the range of opinions expressed by experts in recent times.

A clear and specific taxonomic claim is presented in the Animal Diversity Web, ADW: Canis: Dogs, Jackals, and Wolves (at http://animaldiversity.org/accounts/Canis/classification/).

The wolverine or Gulo gulo (Wikipedia article "Wolverine") is in a separate line of carnivores:

- Class: Mammalia
- Order: Carnivora
- Family: Mustelidae
- Genus: Gulo
- Species: Gulo gulo

Below is a sketch I have made -- a very small selection from the available information about the scientific relationships of dogs, wolves, jackals, dingoes, and wolverines. Wolverines have been added to this sketch because their English name is, linguistically speaking, a derivative of the English word wolf. Hence, wolverines illustrate the divergence between the less knowledgeable classifications of plants and animals implied by natural language, and the more knowledgeable classifications of plants and animals established by science. We all know that the ongoing replacement of the long-common word starfish by the new term sea star was initiated by scientists who want the man and the woman in the street to have a more accurate understanding
of the classification of animals. So far, that type of language reform has not deprecated or replaced the English word *wolverine*.

*A Taxonomic Sketch of Dogs and a Few of Their Relatives*

```
Order: Carnivora

Family: Mustelidae

Genus: Gulo
Species: Gulo gulo (wolverine)

Subspecies: Canis lupus lupus (Eurasian wolf or gray wolf)

Family: Canidae

Genus: Canis
Species: Canis lupus (wolves & dogs)
Subspecies: Canis lupus dingo (dingo)

Genus: Vulpes (many fox species)
Species: Canis aureus (golden jackal)
Subspecies: Canis familiaris (domestic dog)
Species: Canis latrans (common coyote)
```
Format and Language -- two levels whose hierarchical relationship to each other is not yet clear. We further propose two additional levels for the hierarchy not previously postulated: the level of the Sub-expression, and the level of the Bibliographic Record. In addition, we have made a number of suggestions about the proper framework or frame of mind for realizing the Revised FRBR Hierarchy where the rubber hits the road, that is to say, in the future structure of the OCLC WorldCat database.

ACKNOWLEDGMENTS

I would like to thank two audiences for their useful comments on this work in progress: first, the audience of my fellow librarians at Gardner-Webb Library, who heard a practice version or "dry run" of this research, and second, the audience at the Ohio Valley Group of Technical Services Librarians Conference on May 14th, 2018, where I presented this research as a talk. I am grateful for the suggestions and encouragement from both groups. I'm also grateful to Karen Coyle, whose writings and recorded talks on the subject have helped me get "up to speed" on FRBR.
BIBLIOGRAPHY


Coyle, Karen. *FRBR, Before and After: a Look at Our Bibliographic Models*. ALA Editions, and open access with a Creative Commons attribution license (CC-BY license), the latter downloadable from [http://kcoyle.net/BeforeAndAfter/index.html](http://kcoyle.net/BeforeAndAfter/index.html).


This is the first article in the following book/journal issue.

Information Retrieval has been co-published simultaneously as Cataloging & Classification Quarterly, Volume 33, Numbers 3/4 2002" -- title page.