Towards User-Centered Displays of Resources in Global Digital Libraries

Zorana Ercegovac, zercegov@ucla.edu

University of California Los Angeles (USA)

Abstract

We report findings from experiments with IFLA's *Functional Requirements for Bibliographic Records* (FRBR) as applied to the domain of science fiction (Abbott's *Flatland*) in the OCLC's *WorldCat*. The objective is to gauge the characteristics of bibliographic entities under study, to examine the types of relationships these entities exhibit, and to collocate bibliographic entities according to the FRBR group 1 hierarchy of entities. The study's findings may shed some light on a navigational capability in global digital libraries by assembling items into interrelated clusters and displaying them according to the FRBR model in order to make it easier for the user to find desired bibliographic entities.

Keywords

user-centered displays, bibliographic relations, bibliographic entities, online catalogues, digital libraries, *Functional requirements of bibliographic records*, usability

1 Introduction

Digital libraries are viewed as systems providing a community of users with coherent access to a large, organized repository of information and knowledge (Lynch 1995). Coherent access to DL collections hinges on the organization of resources in a collection. Organization is influenced by the nature of resources and patterns of users' tasks as well as by technology, protocols, and policy. The content of resources has increasingly become *dynamic*, ranging from minor updates to multiple releases. Contents may be represented in different *forms of expression* (e.g., textual, numeric, graphical, video); new *genres* or *classes of materials*; used in various *formats* (e.g., HTML, postscript, gif, pdf); and encoded with mixed *metadata* (e.g., MARC; XMLMARC;[1] XML;[2] Dublin Core;[3] RDF;[4] automatically generated). DLs, as defined above, use disparate architectures and metadata to describe resources, ranging from loosely federated resources on the web that are searchable by

1. <http://www.oasis-open.org/cover/marc.html>
2. <http://www.w3.org/TR/REC-xml>
3. <http://dublincore.org/>
4. <http://www.w3.org/TR/rdf-primer/>

any of the available search engines, to well structured library catalogs. While the latter produces a list of more consistently described entities than the former, there is room for improvement, especially in the ways to link entities together and display them in order to facilitate users in their decisions to select and obtain desired expressions and manifestations from collections. The FRBR study offers entity-relationship analysis that may well be the blueprint for knowledge organization in Global Digital Libraries (GDL).

The quality of digital libraries depends on the ability to *link* entities through bibliographic relationships; to *access* entities (through names, words, or phrases in a record that may be used to find that record in a retrieval system); and to *describe* entities (through attributes of a «creator» and title of a work, edition information, publication and distribution data, series, notes, subjects, availability, and intellectual rights management). Conventional libraries have emphasized the description part; GDL will most likely focus their attention to the ways these resources may be linked, displayed, and accessed. Since collections in GDL are to be shared and used internationally, GDL will ultimately become more interoperable, standardized, and integrated than ever before, which is the goal of Universal Bibliographical Control (UBC). A critical aspect of UBC is the work on navigational pathways or *links* between and among entities (ERCEGOVAC 2000; LIBRARY 2003; TILLETT 1992a, 1992b, 1996, 2001). We started off by considering the group 1 entities (work, expression, manifestation, and item) as defined in *Functional Requirements for Bibliographic Records* (IFLA 1998). The requirements were defined in relation to the four user tasks: to find, identify, select, and obtain resources from a given collection of resources.

The following studies have contributed in major ways to our understanding of the interrelated concepts of multiple versions, defined as manifestations in which a single work can appear, and of bibliographic relationships, defined as an association between at least two entities, or components of entities. A bibliographic family has been defined as «a set of related bibliographic works that are somehow derived from a common progenitor» (SMIRAGLIA 1999, p. 494).

2 MULTIPLE VERSION ENTITIES: AN OVERVIEW

2.1 Taxonomy of bibliographic relationships

The problem of multiple versions goes back to the cataloging practice during the pre-electronic era (VERONA 1959; WILSON 1968; LUBETZKY 1969). However, the advent of desktop publishing produced an abundance of documents each with multiple alterations as they get packaged for different purposes and equally varied audiences.

In the electronic environment, the UNIMARC Format (IFLA 1980) recognized vertical (hierarchical), horizontal relationships between different versions of an item in different languages, formats, media, and

chronological relationships. Graham (GRAHAM 1990) reviewed proposals to deal with the crisis of multiple versions in cataloging in view of the overall goals to improve access, streamline cataloging process, and reduce costs. Tillett (2001) identified bibliographic relationships as follows: equivalence, derivative, descriptive, whole-part, accompanying, sequential, and shared relationships. Tillett defined *equivalence* for exact copies that can be used interchangeably, such as reproductions from the same typesetting, photocopies such as reprints, microfilm, scanned images, or other formats including bitmap, gif, pdf, html, and XML. *Derivative* relationships are expressions, representing different editions, translations, adaptations, illustrations, and renditions. *Descriptive* relationships include critical and evaluative reviews, criticism and interpretation, annotated editions, (these are all new works). *Whole-Part* relationships are hierarchical relationships between component parts, such as individual parts in a series (e.g., individual books in the Dover Thrift Editions). *Shared* relationships include different works that share an attribute, such as ID, title, subject, and author. *Accompanying* relationships are supplements, indexes, and individual maps within magazines.

Smiraglia (1992, p. 28) subdivided Tillett's derivative bibliographic relationships into seven types as follows:

— Simultaneous derivations (works that are published in two editions simultaneously, or nearly so);
— Successive derivations (revisions of the same work one or more times, and labeled as numbered editions);
— Translations;
— Amplifications (with illustrations, commentaries);
— Extractions (abridgements, condensations);
— Adaptations (simplifications; screenplays);
— Performances (sound or visual recordings).

2.2 Multiple-version entities in Global Digital Libraries

Researchers pointed out that «the information that was sufficient to identify particular items and bibliographic relationships within a local collection is inadequate in this new environment» (O'NEILL 1989, p. 172). Recently, the Library of Congress (LIBRARY 2002) has acknowledged that the «Issue has become even more pressing as a result of the increasing use of digital technologies both to create original materials and to convert existing materials to digital form, and the resulting proliferation of alternative *manifestations* of the same content».

Currently, researchers have experimented with algorithms that group existing bibliographic records of general nature into works and expressions (HICKEY 2000). Others have developed models that facilitate interoperability between metadata ontologies from different domains (LAGOZE 2001). The emergence of the web has had the effect of producing an abundance of multi-versioned XML resources across disciplinary re-

search. This has also been discussed in computer science literature (ACM SIGMOD).[5]

2.3 How prevalent are bibliographic families?

Tillett (1992) found that between 11.2 and 19.4 percent of LC catalog records exhibit derivative bibliographic relationships. Drawing on a sample of library cards from the Georgetown University library, Smiraglia (1992) determined that nearly 50 percent of all works in that library possessed derivative bibliographic relationships. Vellucci (1995) found that 97 percent of the scores in her sample of musical bibliographic entities exhibited at least one relationship. Smiraglia and Leazer (SMIRAGLIA 1999) determined the proportion of families exhibiting derivative relationships to be 30.2 percent in a sample of 477 progenitor works in the *WorldCat*. While reviewed results vary in distributions of established derivative bibliographic relationships, researchers agree that bibliographic families are prevalent in large numbers and should be accounted for in designing online library catalogs.

3 FRBR: THE BASIS FOR THE NEXT GENERATION OF KNOWLEDGE ORGANIZATION

According to the FRBR Study, there are three groups of entities: group 1 includes works, expressions, manifestations, and items. group 2 includes persons and corporate bodies that are responsible for intellectual and/or artistic content, while Group 3 pertains to subjects of works such as concepts, objects, events, and places.

Accordingly, *work* is a distinct intellectual or artistic creation (e.g., Shakespeare's *Romeo and Juliet* or Abbott's *Flatland*). Works that are substantially modified through the medium of expression become new works (per AACR2R Rule 21.9; also Rule 21.10A.).

FRBR considers *expressions* as the intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of forms. A new expression is any change in language (AACR2R, Rule 21.14A.), revision and/or edition (Rule 21.12A1.), no matter how minor the modification may be.

Manifestation is defined as the physical embodiment of an expression of a work, encompassing a wide range of materials, such as books, periodicals, maps, posters, sound recordings, films, video recordings, multimedia kits. Manifestation represents all the physical objects that bear the same intellectual and artistic characteristics, both in regard to content and physical form. The physical embodiment of a given expression of a work

---

5. (ACM SIGMOD) *ASSOCIATION FOR COMPUTING MACHINERY (ACM) Special Interest Group on Management of Data (SIGMOD)* contains state of the art and research directions and projects that relate to clustering, historical and temporal XML documents, time series algorithms, indexing and mining streams, conceptual modeling, and others.

may be paper or computer file as well as text, sound recording, visual material such as canvas, plaster, clay, or other material. In some cases, there may be a single physical exemplar (e.g., the Van Gogh's *Irises*; Abbott's first edition of 1884). In other cases there are multiple copies reproduced to facilitate commercial and educational uses (*Irises* on note cards, on posters). Each manifestation necessarily has the publisher's layout and typesetting, color, size, packaging, and other details. When catalogers identify and describe an object in hand, they do so at the level of manifestation and generalize to all exemplars or copies that might be located in different library shelves. So an *item* is an exemplar of the manifestation, such as an autographed copy, or a copy having notes in margins.

The E-R model enables us to draw relationships between a work and related works, between a work and its expressions, and so on. So, we can relate the work of Abbott's *Flatland* to the work of criticism and interpretations, to various translations, and to new works such as a film or a sound-recording based on the progenitor.

4 Research questions

We asked the following questions:

1. What categories of *bibliographic relationships* are found in a sample of science fiction entities in a given collection?
2. What is the capability of expressing bibliographic entities and relationships in current cataloging standards?

One of the objectives of information retrieval systems, library catalogs included, is to help the user understand the options and content of the database the user is searching. User studies have pointed out that users have serious problems when they search online library catalogs with regard to scanning long displays (Wiberley 1995). However, research suggests that display of bibliographic records in current OPACs if reorganized differently, might be more helpful to end-users. For example, authors who studied bibliographic families have demonstrated that derivative bibliographic relationships are highly prevalent in library collections. If the high prevalence of bibliographic families is accounted for in the design of library catalogs, assembling entities into smaller groups could be helpful to users in their decision making process. Other studies have suggested ways to improve bibliographic displays by means of clustering entries together (Carlyle 2001; Carlyle 2002; Buckland 1993). Yet another set of studies on human cognitive information processing (Card 1983; Miller 1957) reminds us that human capacity to process and remember information is limited to about seven units of information.

Therefore, we need to take advantage of distributions that seem to be operating among derivative bibliographic relationships. When applied, the entity-relationship model would have the effect of collocating like entities together and displaying them in a way that is native to human in-

formation processing capacity, which in turn would better facilitate searchers' decision-making process. Specifically, the model would collocate different manifestations and expressions of the same work, and by extension, show new works that are based on the progenitor. Expressive linkages between entities would enhance the navigational capability in terms of enabling us to group related entities together and show how these entities are collocated. By showing explicit relationships among groups of records, the model would help the user navigate search results and facilitate tasks of resource discovery and data management. Design of different views should support various communities of users for their specific tasks and purposes.

4.1 Toward user-centered display of bibliographic entities

The proposed model in Table 1 is derived empirically; it shows various search options that the user might explore. Accordingly, a searcher who wishes to get an overview of Abbott's *Flatland* uses an author/title query. After having identified Abbott's *Flatland* as the desired entity, s/he observes two options: Work #1 with 79 records subdivides printed material manifestations into English language expressions and translations, all based on the English publication by Seeley in 1884. The user also notices a path with four new works (W2-5).

TABLE 1: Distribution of Abbott's Flatland under study

Work #1: Textual entities	
—English language books	n=70
—Non-English language books	n= 9
Works #2-5: Non-textual entities	n= 7
Total # entities in collection under study:	**N=86**

If one chose the non-textual entities such as animated films, three releases or manifestations are presented as follows (Table 2): a release produced by Media Magic in 1994, a manifestation in color produced by Great Media, and the third one produced in Toronto in 1995.

Records in Table 2 are from the OCLC's *WorldCat*. New works are represented with seven scattered records throughout the list (3, 6, 15, 86, 83, 84, 89).

Anecdotal evidence suggests that users do not know about various expressions and manifestations if they do not find them in the catalog. Current library catalogs bury valuable machine-readable information behind user interfaces. For example, if one searches for Abbott's *Flatland*, library catalogs return a list of over 80 records that are displayed in a way that does not support aforementioned objectives of IRS. Partitioning of larger retrieval results into smaller units, and then displaying these into meaningful clusters would make sense both from the perspective of high prevalence of bibliographic families in collections and from the human informa-

TABLE 2: Non textual entities (new Works 2-5)

WORK2 [motion picture]

#3 motion picture
Contemporary Films/McGraw-Hill, 1965

WORK3 [videorecording] animation workshop

#6 Cambridge, Mass.: Harvard Film Archive, 1996, 1990

1 videocassette (12 min.) col.; 1/2 in.

#15 Cambridge, Mass.: Film Study Center, Harvard Univ., 2000. 1 videocassette (11 min.) col.; 1/2 in.

WORK4 [videorecording] animated films

#86 Nicasio, Calif.: Media Magic, 1994

1 videocassette (ca. 25 min.), b&w, 1/2 in.

Film adaptation. «Previously released in 1980».

#83 Nicasio, Calif.: Great Media Co., 1994?

1 videocassette (VHS) (23 min.), col. 1/2 in. «Originally produced in Italy in 1980».

#84 Toronto, Ont.: Distributed by International Tele-Film, 1995. 1 videocassette (ca. 25 min.), col., 1/2 in.

WORK5 [videorecording] descriptive relationship

#89 Speaker: Thomas Banchoff, recorded by the Fairfield Univ. Media Center, 1998. From the series on «Humanities lecture series 1998».

tion processing research. Current displays lack capability to group manifestations into expressions, display those under a given work, and alert users of related new works. We need to exploit available information from MARC records, much of which is hidden from the user, yet potentially useful in decision making during the processes of browsing and searching.

5 METHODOLOGY

Descriptive survey methodology was used to examine cataloging entries of records from UCLA's online library catalog *Orion*, University of California *Melvyl*, and the OCLC's *WorldCat*. The *Orion* library catalog was selected to allow for physical examination of local resources.

TABLE 3: Abbott *Flatland* entities in three OPACs

Name/title	WorldCat	Orion	Melvyl
Abbott *Flatland*	86 items	10 items	19 items

All 86 items for Abbott's *Flatland* were searched exhaustively (April 2004). An author title search ((au: abbott and au: edwin) and ti: flatland) for Abbott's *Flatland* produced a total of 86 bibliographic records in OCLC's *WorldCat*. Next, the FRBR Display Tool (http://www.loc.gov/marc/marc-functional-analysis/tool.html) was applied on MARC21

Flatland records that were retrieved from the Library of Congress catalog (fewer number than from the set of previously obtained *WorldCat*). The most recent search on the newly installed library catalog at UCLA (http://catalog.library.ucla.edu/webvoy.htm) yielded again ten entries (as of July 15, 2004). Since the Library of Congress search set was larger, it was used with the FRBR Display Tool. It «*sorts* and *arranges* bibliographic record sets using the FRBR model. It then generates useful hierarchical displays of these record sets containing works that consist of multiple expressions and manifestations» (LIBRARY 2002).

There are serious issues with the newly created display of «frbr-ized» records. Mainly, the display depends on current cataloging rules (AACR2R) and the existing MARC format, neither of which is based on the FRBR model. How to collocate expressions and display together all such realizations of a work that is in a given database is not simply an interface question.

A data set of 86 records was examined manually, and for future experimentation, exported into *FileMaker®* program. Records were coded with MARC tags. New fields were added to facilitate this project's analysis, matching, sorting and reporting. These include the field titled «parent», separately tagged subfields in the statement of responsibility (245 c_1, $c_2, \dots, n_3$) to recognize various authorial responsibilities (e.g., illustrator, writer of prefaces or introductory notes, a bibliographer who writes annotations and prepares references), a progenitor, and the expression_type field. The field *progenitor* is the first expression, which in the case of Abbott's *Flatland*, was the manifestation published in London by Seeley in 1884 as English text. It has no parent. These newly added fields have facilitated the analysis in this project because the FRBR Display Tool, for expression level, matched Leader/06 and field 008/35-37 for the language of expression. In addition to different language expressions, we wanted to create clusters based on derivative bibliographic relationships, including amplified, numbered editions, and special editions.

The field *parent* is used to code those entities that produced respective children. The newly created *parent* field in this study made it possible to study clustering of entities within the bibliographic genealogy of *Flatland*. Findings may shed light on ways to dynamically change distances between nodes.

The field *expression_type* is defined as a holder for specific types of expression that were found in the dataset under study (e.g., amplified edition, amplified and numbered edition, translated edition, illustrated).

6 FINDINGS

6.1 Big picture

By means of including the parent field, we could see the *Flatland* genealogy. For example, the progenitor parented four prolific editions represented in records #22, #28, #41, and #44 (Tables 4-7). Record #22 pro-

duced two siblings, one of which (record #24) parented 4 editions of which record #30 produced 8 new manifestations, all by Little, Brown between 1927 and 1939. As shown in Tables below, four children populated 57 percent of English language texts. This sort of distribution falls within Bradford's distribution of various information units.

TABLE 4: Expressions and manifestations produced by #22

#22 (n=15) Boston: Little, Brown

#23 Boston: Roberts Brothers, 1891 155 p. ill. 18 cm.

#24 Boston: Little, Brown and Co, 1899. Manifestations:

#25-#27 (1907, 1912, 1915)

#30 is new expression (1926 «with introd. by William Garnett.») It gives 8 manifestations, #31-4; 36-9, 1927-1941.

TABLE 5: Nine expressions produced by #28

#28 (n=10) Oxford: Blackwell

(3rd ed., rev. with intro by William Garnett, 1926)

#29 3rd ed., rev. 1926, intro by W. Garnett. (incomp. cataloging)

#30 4th ed., rev. 1932, with intro by W. Garnett.

#40 5th ed., rev. 1944, with intro by W. Garnett.

#42 5th ed. rev. 1950, with intro by W. Garnett.

#43 6th ed. rev. 1950, intro…

#48 1962, intro…

#47 1962, intro… (#47 and #48 may be the same entity)

#54 2nd rev. ed., 1874, with intro by W. Garnett.

#56 1978, with intro…

TABLE 6: Expressions produced by #41

#41 (n=8) New York: Barnes & Noble

(6th ed., rev. with intro by William Garnett, 1950)

#50 5th ed., rev. with intro by William Garnett, 1963, series University paperbacks 45.

#49 5th ed., rev. with intro by William Garnett, [1963], series Barnes & Noble everyday handbooks BN210.

#52 … with intro by William Garnett, 1964, «Followed by Dionys Burger's Sphereland.» (shared relationship by topic)

#51 6th ed., rev., with intro by William Garnett, 1964?

#53 5th ed., rev. with intro by William Garnett, 1966. «Followed by Dionys Burger's Sphereland». (shared relationship, as in #52)

#4 5th ed., rev., 1969 (incomplete cataloging)

#63 (reprint of: 5th ed., rev., 1963, with foreword by Isaac Asimov); with intro by William Garnett; series Everyday handbook; EH/573.

TABLE 7: Expressions produced by #44

#44 distribution (n=5) New York: Dover

(6th ed., rev. with introduction by B. Hoffmann, 1952)

> #2, 6th ed., rev. with intro by B.H., 1952
> #45 6th ed. re., with intro by B.H., 1953
> #46 7th ed., rev. with intro by B.H., 1957
> #71 with intro by B.H., series Dover thrift eds, 1991.

6.2 The *Flatland* family displayed in current OPAC*s*

Each of the three databases searched in this project arranges entries differently. For example, the first entry in *WorldCat* was the one that got the highest library holding score, the piece of information that end users most likely do not regard as significant in their searching. The next entry represents the Dover edition published in 1952, and owned by 760 libraries, etc. Moving through the UCLA's library catalog, the first entry (of the total of ten) was published in Pasadena by Grant Dahlstrom in 1978.

It was unclear how these records were arranged in the respective three lists. Informal probing of students who searched the *Orion* library catalog revealed their confusion with regard to sequencing of entries that was presented to them. In short displays of up to two screens, this may not be as bothersome for users to briefly scan lists of records as it might be in long search results. Carlyle writes, «displays that organize retrieved record sets into intelligible categories may communicate search results more quickly and effectively to users than current catalog displays…» (Carlyle 2002, p. 3). For users who do not have the patience to scroll through multiple screens, the first displayed entry might be the one they would select, ignoring all other editions that are available to them in a given collection.

6.3 Derivative bibliographic relationships in *Flatland*

— *Amplified editions*
Of 86 bibliographic records of Abbott's *Flatland*, 79 records were clustered under the same work, and subdivided into the English language material (n=70) and translations to six non-English language materials (n=9). Of the English language materials, twenty-six editions contained introductory texts by various writers (e.g., W. Garnett, B. Hoffmann, I. Asimov, K. Feiden, R. Bradbury, D. V. Davies, A.K. Dewdney, I. Stewart, A. Lightman, and T. Banchoff). In the subset of these editions, sixteen are both amplified and numbered. There are ten amplified editions alone (containing additional introductory texts); nine of the ten editions transcribed personal names in a subfield c of the statement of responsibility, 245$c, and once in the notes field (500) with an added entry in the 700 field. The general note reads: «With an introd. by Ray Bradbury.» However, the writer's name was not given in the statement of respon-

sibility field. In the amplified expressions that are also coded as numbered or consecutive editions (n=16), personal names were given five times in the edition statement field 250, but not in 245$c.

— *Numbered or successive editions*
   In the subset of 70 English language materials, numbered editions were indicated with wording such as «nth ed. rev.», «1st Shambhala ed.», «New nth ed.», and «New and rev. ed.» All numbered editions gave edition information in the edition statement field 250.

— *Translated editions*
   Indicative of translated editions (n=9) is presence of fixed field 008/35-37 with a three letter language code excluding «eng.» Other fields are 041 1_, 240 $l, 245 $c, 260, 5xx, the language note field (546), series statement 490, and added entry for names of translators, 700 (55 percent, n=5). Abbott's *Flatland* is translated to Italian, French, Greek, Hebrew, Persian, and Russian.

To automatically cluster translated editions as expressions, the only reliable field would be 008/35-37. Physical format in the 006 MARC field informs us if an item is printed language material. The 008 field contains a mandatory three-letter code for language materials. In addition, all records but one coded language information in the 041 field. Six of nine records coded language information in the uniform title field 240.

6.4 Other bibliographic relationships

All major derivative bibliographic relationships were *illustrated* and coded in the 245 $c, the 300 $b, and oftentimes in the notes field. In addition, *equivalence* was present in the form of reprints, microforms, and electronic resources. Indicative of «form of item» is given in the fixed field 008. Microforms and microfiche are coded with a single letter «a» and «b». Large print material is indicated with a «d», regular print reproductions with a «r», while electronic items with a «s». Physical description is coded in the fixed field 007, and the extent of carrier is transcribed in the companion physical description field 300. General material designation is given in 245 $h and in the 5xx fields. Other important fields that may be used to link different editions are 775 (other edition entry) and 776 (additional physical form entry). Mapping of MARC21 linking entry fields 760-78X to FRBR and Tillett's bibliographic relationships has been discussed in Riva (RIVA 2004).
   Some amplified editions are part of series (4xx), such as Princeton Science Library, Dover Thrift Editions, Penguin Classics, and Project Gutenberg. In nine translations, five are part of series. These are all considered as whole-part relationships. Information for shared relationships is given in the general notes field 550. One record, using full cataloging, describes a shared bibliographic relationship in fields 245 $c (Sphereland: a fantasy about curved spaces and an expanding universe / by

Dionys Burger; translated from the Dutch by Cornelie J. Reinboldt; foreword by Isaac Asimov). General note mentions about «two separate books, each with forwards by Isaac Asimov. This edition with the texts bound back to back and upside down». Again, it is noted (in 500) that «Flatland was originally published in 1880. Shereland was originally published, in Dutch, in 1965». Personal names for the writer, Burger, and the translator from Dutch, Reinboldt, are given added entries.

6.5 New works

Of the new works, indicative coding includes letters «m» for motion picture, and «v» for videorecording, in the fixed field 008. Fixed field 007 uses «v» for videorecording and motion picture, as well as specific characters to code color, format, sound on medium, medium for sound, dimensions, and playback channels. For motion pictures, there is additional coding for production elements, generation, base of film, kind of color stock or print, deterioration stage, and completeness. AACR2R identifies physical format in the general material designation, GMD, (MARC field 245 $h [form of carrier]); however, this is optional and may be treated differently in different Anglo-American cataloging traditions. Other fields that contain information of non-book materials are 260, 300, 5xx, 650, and name added entries (700, 710).

7 DISCUSSION

Inconsistent treatment of important contributors to a given work by transcribing them into various fields (245 $c, 700, 500, and 250) makes it difficult to collocate expressions automatically. Furthermore, there are presently no separate tags to code various authorial responsibilities in a 245$c. Computer algorithm would have to match the presence of truncated text strings for authorial responsibilities. Examples in the set of 78 records under study are: «with introd# by», «with ill# by», «forward# by», «with notes by», and «includes bibl# reference#». Numbered or consecutive editions collocate with the edition statement field 250. Translated editions to non-English languages use other than «eng» three-letter code in the fixed field 008/35-37.

The main types of expressions found in the 86 records under study are formalized below as follows:

EXP: = A | B | C | D | E | F | EXP AND EXP
where
A is amplified edition
B is numbered edition
C is translated edition
D is illustrated edition
E is with reference notes
F is simultaneous or near simultaneous edition

Given ABD, the grammar rule above could be used to parse the expression to inform us if the expression ABD is valid:

$$
\left.\begin{array}{l}
A \rightarrow EXP \\
B \rightarrow EXP \\
D \rightarrow EXP
\end{array}\right\} \; EXP \left.\begin{array}{l} \\ \\ \end{array}\right\} \; EXP
$$

According to FRBR, expressions are embodied in manifestations, and these are exemplified in items. Manifestations are typically coded in MARC fields 260, 300, and 5xx. Looking at the display of records for multiple manifestations of the same work, and for single record for manifestations of an expression (from http://www.loc.gov/marc/marc-functional-analysis/functional-analysis.html), it is already possible to display translations of Abbott's *Flatland* automatically. Since FRBR has not been implemented into cataloging rules or extended into MARC format, collocation of other types of derivative relationships has not been implemented in commercial library catalogs. Expressions and manifestations are typically not distinguished in current cataloging practice. They are all treated as different manifestations of a single work.

8 CONCLUSIONS AND FURTHER WORK

We need design mechanisms that would guide users in browsing and searching of resources that are linked together by explicitly expressed relationships. This paper demonstrated a need for implementing the FRBR E-R model and shortcomings of conventional library catalogs to provide meaningful navigational capability. The metadata schema must be able to link multiple versions of the same resource, to show how these resources are inter-related expressively, and how these are related to other similar resources in a digital collection. For example, links could be constructed to show that:

Entries should be arranged in a helpful way to facilitate users in their choice of a desired entity of a given work. Current and future catalogs should build this sort of capability in order to make participation between systems and end-users more collaborative and symmetrical than is currently possible.

In particular, linking of entities on the basis of derivative bibliographic relationships would be particularly useful for technical, environmental, earth sciences, pharmaceutical, and biomedical literatures where resources are likely to evolve over a long period of time within large projects, interdisciplinary teams, and among distributed laboratories. The reviewed literature revealed no published studies that examined the nature of bibliographic relationships between and among such entities in digital repositories. This author is currently examining relative importance of weights between parts of entities: text files, source codes, interim solutions, draft case studies, experimental data, drawings, and anno-

tated versions. Whether it concerns assembling all expressions and manifestations of Abbott's *Flatland*, or XML Manuals, searchers in Global Digital Libraries ought to see a pathway representing a work along with all associative expressions.

REFERENCES

(AACR2r 2002) *ANGLO-AMERICAN CATALOGUING RULES*. Prepared under the direction of the Joint Steering Committee for Revision of AACR, a committee of the American Library Association, et al. 2nd ed., 2002 revision. Chicago: ALA, 2002.

(BUCKLAND 1993) BUCKLAND, M. K.; NORGARD, B. A.; PLAUNT, C. «Filing, filtering, and the first few found». *Information Technology and Libraries*, 12 (1993), p. 311-319.

(CARD 1983) CARD, S. K.; NEWELL, A.; MORAN, T. P. *The Psychology of Human-Computer Interface*. Mahwah, N.J.: Lawrence Erlbaum, 1983.

(CARLYLE 2001) CARLYLE, A. «Developing organized information displays for voluminous works: a study of user clustering behavior». *Information Processing & Management*, v. 37 (2001), p. 677-699.

(CARLYLE 2002) CARLYLE, A.; SUMMERLIN, J. «Transforming catalog displays: record clustering for works of fiction». *Cataloging & Classification Quarterly*, v. 33 (2002), p. 13-25.

(ERCEGOVAC 1998) ERCEGOVAC, Z. «Minimal level cataloging: what does it mean for maps in the contexts of card catalogs, online catalogs, and digital libraries?». *Journal of the American Society for Information Science*, v. 49 (1998), p. 706-719.

(ERCEGOVAC 1999) ERCEGOVAC, Z. (guest ed.) «Introduction». *Special Topic Issue: Integrating Multiple Overlapping Metadata Standards*. v. 50 (1999), p. 1165-1168.

(ERCEGOVAC 2000) ERCEGOVAC, Z. «Toward a global access to bibliographic information [electronic resource]: converging patterns, new paradigms». In: COBISS/SICRIS USER CONFERENCE (2000: Maribor, Slovenia). Invitational keynote presentation. *Obvestila*, v. 5 (2000), p. 4-28. <http://www.cs.ucla.edu/Leap/zer/maribor.htm>. [Cited 10 April 2005].

(GRAHAM 1990) GRAHAM, C. «Definition and scope of multiple versions». *Cataloging & Classification Quarterly*, v. 11 (1990), p. 5-32.

(HICKEY 2002) HICKEY, T. B.; O'NEILL, E. T.; TOVES, J. «Experiments with the IFLA Functional Requirements for Bibliographic Records (FRBR)». *D-Lib Magazine*, v. 8, n. 9 (2002).

(IFLA 1980) IFLA-Working Group on Content Designators. *UNIMARC: Universal MARC Format*. 2nd ed. London: IFLA International Office for UBC, 1980.

(IFLA 1998) INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. *Functional Requirements for Bibliographic Records: Final Report. IFLA Study Group on the Functional Requirements for Bibliographic Records*. IFLA Universal Bibliographic Control and International MARC Programme, Deutsche Bibliothek, Frankfurt am Main. Munich: K. G. Saur, 1998. (UBCIM Publications, v. 19).

(LAGOZE 2001) LAGOZE, C.; HUNTER, J. «The ABC ontology and model» [electronic resource]. *Journal of Digital Information*, v. 2, n. 2 (2001). <http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Lagoze/>. [Cited 10 April 2005].

(LIBRARY 2002) LIBRARY OF CONGRESS. Network Development and MARC Standards Office. *Displays for Multiple Versions from MARC 21 and FRBR* [electron-

ic resource], 2002. <http://www.loc.gov/marc/marc-functional-analysis/multiple-versions.html>. [Cited 10 April 2005].

(LIBRARY 2003) LIBRARY OF CONGRESS. Cataloging Directorate. Bicentennial Conference on Bibliographic Control for the new Millennium (December 12, 2003). Available online: <http://www.loc.gov/catdir/bibcontrol/actionplan.html>. [Cited 10 April 2005].

(LUBETZKY 1969) LUBETZKY, S. *Principles of Cataloging. Phase 1, Descriptive Cataloging.* «Final Report». Los Angeles: Institute of Library Research, University of California, 1969.

(LYNCH 1995) LYNCH, C.; GARCIA-MOLINA, H. *Interoperability, scaling, and the digital libraries research agenda* [electronic resource]*: a report on the May 18-19, 1995.* IITA Digital Libraries Workshop, 1995. <http://www-diglib.standford.edu/diglib/pub/reports/iita-dlw/main.html>. [Cited 10 April 2005].

(MILLER 1956) MILLER, G. A. «The magical number seven, plus or minus two: some limits on our capacity for processing information.» *The Psychological Review*, v. 63 (1956), p. 81-97.

(O'NEILL 1989) O'NEILL, E. T.; VIZINE-GOETZ, D. «Bibliographic relationships: implications for the function of the catalog.» In: *The Conceptual Foundations of Descriptive Cataloging*. (San Diego: Academic Press), p. 167-179.

(RIVA 2004) RIVA, P. «Mapping MARC 21 linking entry fields to FRBR and Tillett's taxonomy of bibliographic relationships». *Library Resources & Technical Services*, v. 48 (2004), p. 130-143.

(SMIRAGLIA 1992) SMIRAGLIA, R. P. *Authority Control and the Extent of Derivative Bibliographic Relationships*. Urbana, Il.: Graduate Library School, University of Chicago, 1992.

(SMIRAGLIA 1999) SMIRAGLIA, R. P.; LEAZER, G. H. «Derivative bibliographic relationships: the work relationship in a global bibliographic database». *Journal of the American Society and Information Science*, v. 50 (1999), p. 493-504.

(TILLETT 1992a) TILLETT, B. A. B. «Bibliographic relationships: an empirical study of the LC machine-readable records». *Library Resources & Technical Services*, v. 36 (1992), p. 162-188.

(TILLETT 1992b) TILLETT, B. A. B. «The history of linking devices». *Library Resources & Technical services*, v. 36 (1992), p. 23-36.

(TILLETT 1996) TILLETT, B. A. B. «International shared resource file [electronic resource]: from authority control to access control». In: OCLC. *Authority Control in the Twenty-First Century: An Invitational Conference, March 31-April 1, 1996: proceedings*. (Dublin, Ohio: OCLC, 1996). <http://oclc.org/oclc/man/authconf/confhome.htm>. [Cited 10 April 2005].

(TILLETT 2001) TILLETT, B. A. B. «Bibliographic relationships». In: *RELATIONSHIPS in the Organization of Knowledge*. C.A, Bean and R. Green, eds. (Boston: Kluwer Academic Pub., 2001), p. 19-35.

(VELLUCCI 1995) VELLUCCI, S. L. «Bibliographic relationships among musical bibliographic entities: a conceptual analysis of music represented in a library catalog with a taxonomy of the relationships discovered». Unpublished doctoral dissertation, [Columbia University, 1995].

(VERONA 1959) VERONA, E. «Literary unit versus bibliographic unit». *Libri*, v. 9 (1995), p. 79-104.

(WIBERLEY 1995) WIBERLEY, S. E., Jr.; DAUGHERTY, R. A.; DANOWSKI, J. A. «User persistence in displaying online catalog postings: LUIS». *Library Resources & Technical Services*, v. 39 (1995), p. 247-264.

(WILSON 1968) WILSON, P. *Two Kinds of Power: An Essay on Bibliographic Control*. Berkeley; Los Angeles: University of California Press, 1968.