Richard P. Smiraglia — University of Wisconsin - Milwaukee

Is FRBR a Domain? Domain Analysis Applied to the Literature of *The FRBR Family of Conceptual Models*

Abstract

Domain analysis helps visualize the semantic intellectual content of a coherent group, or domain. A domain is a group with an ontological base, an underlying teleology, common hypotheses and epistemology, and social semantics. *FRBR* has spawned a family of conceptual models, and much writing. A recent second anthology about the *FRBR* models raises the question of whether a coherent domain has formed around the *FRBR* family. Domain analysis is used here to visualize the semantic content of the *FRBR* family domain, and to compare its two main component groups, scholar authors and practitioner authors. Results show a common teleology with some subtle differences surrounding implementation of the *FRBR* family of models.

FRBR family and domain analysis

Domain analysis provides a set of techniques for extracting and analyzing the semantic intellectual content of a coherent group. A major tool for knowledge organization, domain analysis has been successfully applied to such diverse domains as "musicianship" (Lam 2011), "gourmet cooking" (Hartel 2010), and "scientific computing" (Tanaka 2010). Domain analysis has also been applied successfully to less deliberately coherent groups, such as those whose research incorporates a classic work (Patrick Wilson's *Two Kinds of Power*; cf. Smiraglia 2007). Recently, Smiraglia defined a domain (2012, 114) in these terms: "a group with an ontological base that reveals an underlying teleology, a set of common hypotheses, epistemological consensus on methodological approaches, and social semantics." Domain analytical techniques draw out the concepts that form these components of domain coherence. But it remains a question yet to be informed by empirical analysis whether these components rely on overall coherence or require equal representation in a domain.

FRBR (*Functional Requirements for Bibliographic Records*) is the acronym standing for a group of conceptual models promulgated by IFLA (1998), and serving as the basis of a re-engineering of library bibliographic services. So powerful has the set of *FRBR*-based conceptual models become, that 2012 saw the publication of a special anthology of papers in the journal *Cataloging & Classification Quarterly* (vol. 50 nos. 5-7). The volume (edited by Smiraglia, together with Pat Riva and Maja Žumer, which also appeared in May 2013 as a Taylor & Francis monograph) contained 24 papers contributed by authors worldwide concerning implementation, expansion and research about the *FRBR* "family" of conceptual models. Table 1 shows the table of contents from the anthology.

Table 1. Table of contents of *The FRBR Family of Conceptual Models: Toward a Linked Bibliographic Future* (*Cataloging & classification quarterly* v. 50 nos. 5-7 2012).

Authors	Title
Patrick Le Boeuf	Foreword
Richard P. Smiraglia	Introduction: Be Careful What You Wish For: Lacunae in the FRBR Family of Models
Implementations	

2

John Espley and Robert Pillow	The VTLS Implementation of FRBR
Michaela Putz, Verena Schaffner, Wolfram Seidler	FRBR: The MAB2 Perspective
Corinne Bitoun, Aurelie Signoles and Asuncion Valderrama	Implementing FRBR to Improve Retrieval of In-House Information in a Medium-Sized International Institute
Extensions	
Patrick LeBoeuf	A Strange Model Named FRBRoo
Norberto Manzanos	Item, Document, Carrier: An Object Oriented Approach
Maja Zumer and Edward T. O'Neill	Modeling Aggregates in FRBR
Carlo Bianchini	FRBR Without FRBR?
Jonathan Furner	FRSAD and the Ontology of Subjects of Works
Martin Doerr, Pat Riva, Maja Zumer	FRBR Entities: Identity and Identification
FRBR and Cataloging Rules	
Mirna Willer and Ana Barbaric	FRBR/FRAD and Eva Verona's Cataloguing Code: Toward the Future Development of the Croatian Cataloguing Code
Pat Riva and Chris Oliver	Evaluation of RDA as an Implementation of FRBR and FRAD
Manolis Peponakis	Conceptualizations of Cataloguing Object: A Critique on Current Perceptions on FRBR Group 1 Entities
Alberto Petrucciani	From the FRBR Model to the Italian Cataloguing Code (and Vice Versa?)
Research Using FRBR	
Virginia Ortiz-Repiso and Paola Picco	The Contribution of FRBR to the Identification of Bibliographical Relationships: The New RDA-based Ways of Representing the Relationships in Catalogs
Clement Arsenault and Alireza Noruzi	Work-to-Work Bibliographic Relationships from FRBR Point of View: A Canadian Perspective
Ray Schmidt	Composing in Real Time: Jazz Performances as "Works" in the FRBR Model
Takuya Tokita, Maiko Koto, Yosuke Miyata, Yukio Yokoyama, Shoichi Taniguchi and Shuichi Ueda	Identifying Works for Japanese Classics toward Construction of FRBRized OPACs
Hyewon Lee and Ziyoung Park	FRBRizing Bibliographical Records without Main Entry Headings and Uniform Titles
Yin Zhang and Athena Salaba	What do Users Tell us About FRBR-Based Catalogs?
FRBR and The Semantic Web	
Gordon Dunsire	Representing the FR Family in the Semantic Web
Jane Greenberg, Ketan Mayer-Patel and Shaun Trujillo	YouTube: Applying FRBR and Exploring the Multiple Description Coding Compression Model
Lynne C. Howarth	FRBR and Linked Data: Connecting FRBR and Linked Data

An interesting question that arose during the compilation of this volume was whether these authors represented in any way a coherent domain. The authors present an interesting mix of researchers and librarians, and papers range from empirical studies to conceptual analyses to descriptions of implementations. Domain analysis based on the citations in these papers is one approach to answering that question.

In fact, preliminary observations suggested there might be divergent citation patterns in the different papers. Although a small core of papers and monographs are cited in most of the papers, the rest is quite surprisingly diverse. A large proportion of the references are not to published materials, but rather to web-based services. Although a domain-like core seems to be shared among these authors, there are quite different citation practices between, for example, scholarly research papers and implementation descriptions. There also appear to be geographical or geopolitical differences present in the divergent citation practices across the group. Thus this domain-like group might have a common ontological base and share social semantics, but at the same time incorporate diverse epistemological stances due to divergent teleological imperatives. In other words, *FRBR* is a domain united by a conceptual model governing bibliographic information retrieval. But it is possible that quite divergent subgroups make up the domain.

Methodology

During the preparation of the *FRBR Family* volume, a bibliography of works about FRBR was compiled. The list was examined carefully to remove news bulletins. The final list was combined with the papers from the *FRBR Family* volume to yield 91 papers from practitioners and scholars, all devoted to some aspect of *FRBR* or its implementation. The citations in these 91 papers were gathered into an Excel spreadsheet. A total of 1511 citations were recorded. The citations required manual "cleaning" because they were not in author-date format, nor were names inverted, and some appeared not in reference lists but rather, in endnotes, so the process of delimiting the data for analysis was time-consuming. After cleaning, 1499 citations remained. Then the papers were divided according to first-author affiliation into two groups "scholar" and "practitioner"—632 citations appeared in the papers by scholars, and 867 citations were in the papers by practitioners. A much larger number of papers fell into the practitioner category: 67 papers versus 24 attributed to scholars. Thus, the number of works cited per paper differed for the two groups. The mean number of works cited per paper in the scholars group was 21.75; in the practitioners group it was 15.16. Overall then the scholars cited more heavily than the practitioners.

Year of cited work

Analyzing the age of works cited in a domain tells us something about obsolescence and therefore also about the rate of absorption of new knowledge. In scientific domains, for instance, most cited works are relatively recent, because science is cumulative. That is, data reported in journals are incorporated in successive studies, which in turn are cited as the most recent research in a productive domain.

The citations were delimited to separate the dates of publication. Interestingly, 135 of the citations had no date. Those with dates were arrayed chronologically and the age

4

of cited work was calculated. The mean age of cited work was 15.47 years; the median was 11 years and the mode was 8 years. The range of dates stretched from 1722 to 2012. The citation dates were clustered in groups that seemed meaningful when compared to the distribution of years in the data. These clusters and the proportional frequencies are shown in Figure 1.

Figure 1. Clustered dates of cited works

FRBR was first published in 1998 so it is no surprise most of the publications cited fall after that date. The clear majority were published between 2000-2007 and interest seemed to wane after that. Of course the 19th century citations are primarily references to Cutter's *Rules* (1876); the large cluster dated 1950-1999 includes citations to the various editions of *Anglo-American Cataloging Rules* as well as Patrick Wilson's *Two Kinds of Power* and research by Smiraglia, Tillett, Yee and others that pre-saged *FRBR* by emphasizing bibliographic relationships and the importance of works in the library catalog. Overall this distribution is comparable to the results in most domain analyses of information studies or its subdomains (such as knowledge organization), and the results resemble those of a social scientific domain, with a moderate rate of absorption but a continued reliance on classic texts.

The comparison of the two groups "scholar" and "practitioner" is shown in Table 2.

Table 2. Dates of works cited by practitioners and scholars

Practitioners			Scholars		
Date range	Number cited	%	Date range	Number cited	%

177

5

0	104	11	0	31	4
18th c.	3	.03	18th c.	0	0
19th c.	10	1	19th c.	1	.01
1900-1949	4	.04	1900-1949	6	.09
1950-1999	247	28	1950-1999	213	33
2000-2007	431	49	2000-2007	334	52
2008	49	5	2008	35	5
2009	13	1	2009	10	1
2010	3	.03	2010	2	.03
2011	2	.02	2011	0	0
2012	1	.01	2012	0	0
	867			632	

The mean age of works cited by practitioners was 16.4 years, the median was 11 years and the mode was 8 years. The mean age of works cited by scholars was 14.3 years, the median was 12 years and the mode was 11 years. So it appears that practitioners cited slightly older works than scholars. In fact, the practitioners had more citations to pre-twentieth century works, including several 1876 citations to classical texts. The scholars had slightly more citations to recent works dated 2000 or later (58% vs. 54%). The year of publication of *FRBR* was 1998; both groups had a large number of citations dated in that year, 60 among the scholars (or 9%) and 61 among the practitioners (or 7%). Both groups had large numbers of undated citations, although proportionally the scholars had fewer—4% vs. 11%. This apparently minor result reflects a real difference between the two groups. The practitioners cite web resources such as OCLC WorldCat frequently, whereas the scholars do not. The difference is subtle—it means that the scholars are referencing works they cite, but the practitioners are referencing uncited but relevant professional resources.

Countries of Author Affiliation
FRBR is a product of international cooperation through the International Federation of Library Associations and Information Institutions (IFLA) so it is no surprise that interest in the *FRBR* model is also international. Authors of the papers in this study listed twenty countries of affiliation. These are given in Table 3.
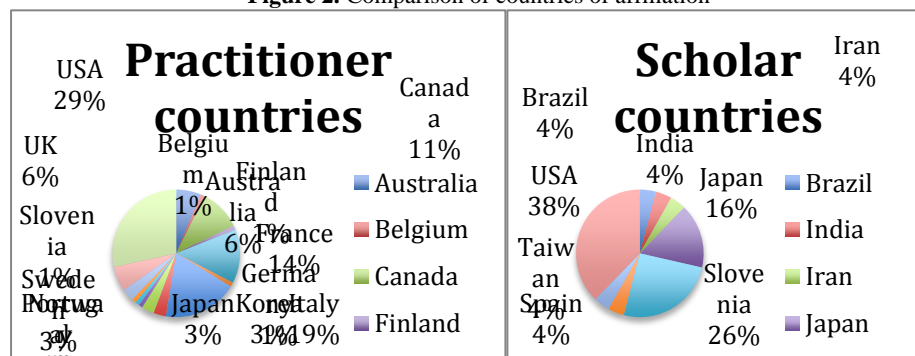
Table 3. Countries of affiliation

Australia
Belgium
Brazil
Canada
Finland
France
Germany
India
Iran
Italy

6

Japan
Korea
Norway
Portugal
Slovenia
Spain
Sweden
Taiwan
UK
USA

Figure 2 shows the comparative proportions of country affiliations for the two groups of authors, scholars and practitioners.

Figure 2. Comparison of countries of affiliation

Obviously there are more countries of affiliation among the practitioner authors than among the scholar authors. Although the USA dominates both distributions, the relative prominence of European authors among the practitioner authors is noticeable, as is the Asian influence among the scholar authors. Whether these distinct differences constitute any intellectual difference is arguable. It is more likely that this is a representation of strengths in the bibliographic control community; strong and innovative libraries in the practitioner grouping, iSchools and research institutes in the scholar grouping. A more appropriate question, then, is whether there is knowledge sharing between or among the various components of the *FRBR Family* domain.

Author co-citation analysis
If there is intellectual coherence within the domain, one way to visualize it is with author co-citation analysis. When authors are co-cited it means they are perceived to be writing on similar topical threads in a research front. The larger the proportion of author co-citation the better the evidence that members of the domain have common points of view, or at least common theoretical mile-posts. To begin, the 1491 citations were sorted by cited author. 1397 recognizable author names were retrieved from the citations, and of those, 137 were cited more than once. Thirty-two authors were cited six times or more and these are shown in Table 4.

7

Table 4. Most cited authors in the *FRBR Family* domain

Tillett, Barbara Ann Barnett	44	Hegna, Knut	10
Yee, Martha M.	32	Madison, Olivia M. A.	10
Le Boeuf, Patrick	30	Bowen, Jennifer	8
Hickey, Thomas B	29	Heaney, Michael	8
Delsey, Tom	27	Albertsen, Ketil	7
Smiraglia, Richard P.	27	Antelman, Kristin	7
Carlyle, Allyson	21	Cutter, Charles A.	7
O'Neill, Edward T.	19	Guerini, Mauro	7
Svenonius, Elaine	16	Jonsson, Gunilla	7
Bennett, R.	15	Kilner, Kerry	7
Taniguchi, Shoichi	13	Leazar, Gregory H.	7
Zumer, Maja	13	Lubetzky, Seymour	7
Vellucci, Sherry L.	12	Howarth, Lynne C.	6
Ayres, Marie-Louise	11	Lagoze, Carl	6
Wilson, Patrick	11	Riva, Pat	6
Aalberg, Trond	10	Verona, Eva	6

As usual it is an interesting list. A few names—Charles Cutter, Eva Verona, Seymour Lubetzky—are frequently cited classic authors from the late nineteenth and early to mid twentieth century. Otherwise the list looks like the list of the most prominent members of the *FRBR* community in general. After removing the classical authors, the remaining twenty most cited authors were searched in *Web of Science*™ for co-citation totals. These were entered into SPSS™ and a multi-dimensionally-scaled plot was generated. This plot is shown in Figure 3, and is a visualization of the entire domain.

Figure 3. Author co-citation in the *FRBR Family* domain
(stress = .08060 R-squared = .98024)

8

Bearing in mind that the visualization represents the manner in which the citing community perceives the domain, we can see that there are several points of coherence. The accompanying dendrogram (not shown here) tells us there are essentially four clusters, which ultimately are loosely associated in two groups. At the left we see "Ayres Heaney and O'Neill" who form one distinct cluster, associated with the OCLC research division and its work on theoretical issues underlying implementation. At the far right we see a sort of classical *FRBR* cluster including most of the authors who wrote about works and super-works, and oddly including Delsey who was editor of *RDA*. Associated with them are those who worked with IFLA to create *FRBR* as well as some who worked on the object-oriented *FRBRoo*. There is no circle around Le Boeuf, Aalberg, Hegna or Hickey, because they are individually associated with the rest of this large group but not otherwise clustered. A broad interpretation of this map, again remembering we are looking at perceptions of the citing community, is that there are two groups of contributors to the *FRBR Family* at the conceptual level, those who wrote about works and those who wrote *FRBR* itself, and there is a small research front working on theoretical issues surrounding implementation.

To see whether there are differences in author co-citation between the two groups "scholar authors" and "practitioner authors" all cited authors were sorted by group. Table 5 shows the authors most cited by the two groups using a cut-off point of 5 citations or more.

Table 5. Most-cited authors by group

Most cited by practitioner authors		Most cited by scholar authors	
IFLA Study Group on the Functional Requirement for Bibliographic Records	36	IFLA Study Group on the Functional Requirements for Bibliographic Records	26

9

Tillet, Barbara	28	Delsey, Tom	21
Smiraglia, Richard P.	21	Tillet, Barbara B.	16
Yee, Martha M.	17	Le Boeuf, Patrick	15
Hickey, Thomas B	16	Yee, Martha M.	14
Joint Steering Committee for Development of RDA	16	Hickey, Thomas B. et al.	13
Library of Congress	14	Svenonius, Elaine	11
International Federation of Library Associations and Institutions	13	O'Neill, Edward T	10
Carlyle, Allyson	12	Carlyle, Allyson	8
Le Boeuf, Patrick	12	Joint Steering Committee for Development of RDA	8
Dublin Core Metadata Initiative	11	Library of Congress	8
Vellucci, Sherry L	10	Zumer, Maja	8
O'Neill, Edward T.	9	Heaney, Michael	7
OCLC	9	Ayres, Marie-Louise	6
Wilson, Patrick	9	Library of Congress, Network Development and MARC Standards Office	6
Bennett, Rick	8	Madison, Olivia M. A.	6
Delsey, Tom	7	Smiraglia, Richard P.	6
Taniguchi, Shoichi	7	Taniguchi, Shoichi	6
Albertsen, Ketil	6	Aalberg, Trond	5
Bowen, Jennifer	6	ALCTS CCS CC:DA	5
JISC	6	Bennett, Rick	5
Jonsson, Gunilla	6	IFLA	5
Library of Congress, Network Development and MARC Standard Office	6	Lagoze, Carl	5
Aalberg, Trond	5	Riva, Pat	5
Antelman, Kristin	5	Weinstein, Peter C	5
Ayres, Marie-Louise	5		
Cutter, Charles Ammi	5		
Guerrini, Mauro	5		
Hegna, Knut	5		
Jones, Edgar A.	5		
Lubetzky, Seymour	5		
Ranganathan, S.R.	5		

Now we see that the two groups are distinctly different in a number of ways. For one thing, there are many more citations to institutional documents in the practitioner group. Even removing those leaves two quite different lists, in particular with regard to the order of citedness. Still, if we remove the institutions and make the cut-off point six citations or more, we will be left with the same list of core authors whose names appear in Figure 3. (A methodological note: names with low co-citation counts removed from the plot in Figure 3 are those at the bottom of the practitioner distribution, including Bowen, Albertsen and Jonsson). Thus it appears that despite the differences in approach represented in Table 5, there remains an intellectual core common to both groups. This is a sign of domain coherence overall.

**Co-word analysis of the 91 titles**
Co-word analysis can be used to visualize themes within a domain, by using software such as WordStat™ that calculates term frequencies and using co-occurrence data makes three-dimensional plots. Often in domain analytical research, co-word analysis can provide a methodological triangulation that adds to the interpretation of author co-citation analyses. For this study three plots of term frequency were created utilizing all of the titles of the 91 papers under study, and then making separate plots of the titles in papers by scholars and practitioners. The overall plot appears in Figure 4.

Figure 4. Co-word plot of all title keywords (stress = .24487 R-squared = .8164)

This picture is quite similar to the picture we saw in Figure 3. We have here three distinct groups loosely affiliated but nested near each other. The "universe, models, implementation" cluster is a parallel for the co-citation cluster anchored by O'Neill. Now we have also a small cluster including "cataloging and resources," which perhaps helps us understand the central position of Carlyle and Delsey in Figure 3, representing resource description and catalog display as priorities. The rest of the domain is anchored by FRBR and the words that constitute the acronym, but also incorporates "metadata, conceptual, entity, and expression" among others. Figure 5 is a side-by-side plot of title co-words from the practitioner and scholar authors.

Figure 5. Practitioner title keywords (stress = .23328 R-squared = .8453) and scholar title keywords (stress = .17797 R-squared = .9100)
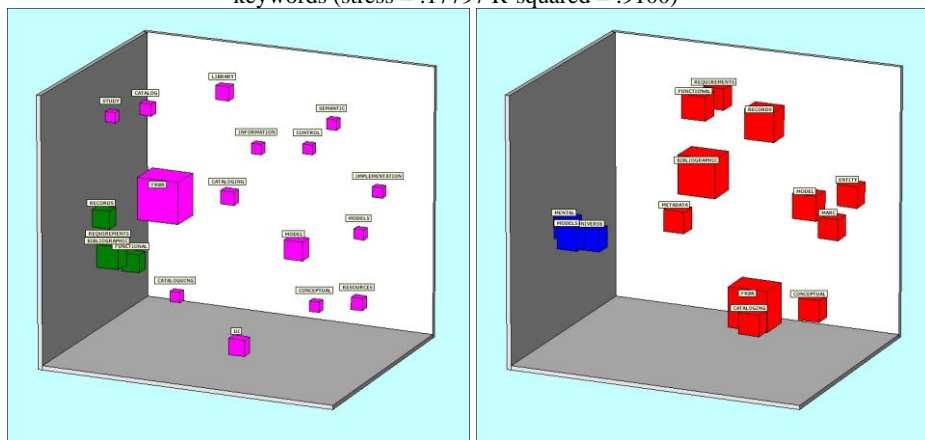
In each of these visualizations there obviously are only two clusters so there is no need to outline them. And here the subtle difference between the two clusters is apparent and reflects what we saw above in the co-citation analysis. That is, there is a common core of FRBR conceptual material, but the practitioners have an extra concern for functional requirements, where the scholars show extra interest in universal mental models. The subtle differences now appear to reflect divergent approaches to implementation.

Concluding observations

We began with the question of whether the FRBR family of conceptual models had spawned a domain and the answer appears to be affirmative. There is a common ontological base incorporated in the FRBR models, an underlying teleology in the goal of reinventing the catalog to separate formerly disregarded entities, and there is a shared epistemological base in the shared hypotheses, particularly concerning FRBR entities. There is also ample evidence of social semantics, particularly visible in common citation patterns. Overall the domain has characteristics that mirror those of information studies in general, or knowledge organization (sometimes described as a sub-domain of information studies), but it also has its own FRBR-like character.

There were internal differences as well. Scholars cited more heavily than practitioners. Both scholar and practitioner authors had large numbers of undated citations, reflecting a real but subtle difference between the two groups. The same

subtle difference emerged in co-word and author co-citation analyses. That is, practitioner authors cite institutional documentation, particularly in the form of web resources more frequently than their scholar counterparts. There are some interesting geopolitical implications in the observation that European authors predominate among the practitioners while Asian authors have influence among the scholars, although the domain is dominated by authors in the US.

An attempt to discover divergence between the author and scholar groups highlighted some differences in approach toward implementation. Practitioner authors had greater interest in resource description practice, and in particular the definition of *FRBR* entities, while the scholar authors had definitive bent toward universal applications of the *FRBR* models.

A secondary research question was whether the components of a domain rely on overall coherence or require equal representation in the domain. In this case, overall coherence is demonstrated, but without requiring equal representation throughout the domain. In other words, the domain can be coherent in its extension and still tolerate divergence in its intension.

Acknowledgements

References

Hartel, Jenna. 2010. Managing documents at home for serious leisure: a case study of the hobby of gourmet cooking. Journal of documentation 66: 847-74,

IFLA Study Group on the Functional Requirements for Bibliographic Records. 1998. Functional requirements for bibliographic records: Final report. München: K. G. Saur.

Lam, Margaret. 2011. Towards a "musicianship model" for music knowledge organization. OCLC systems & services 27: 190-99.

Smiraglia, Richard P. 2007. Two kinds of power, insight into the legacy of Patrick Wilson. In Dalkir, Kimiz and Clément Arsenault eds., Information sharing in a fragmented world: crossing boundaries, Proceedings of the Canadian Association for Information Science Annual Conference, May 12-15, 2007, Montréal. Available http://www.cais-acsi.ca/search.asp?year=2007.

Tanaka, Michiko. 2010. Domain analysis of computational science: fifty years of a scientific computing group. In Gnoli, Claudio and Fulvio Mazzocchi eds., Paradigms and conceptual systems in knowledge organization, Proceedings of the 11th International ISKO Conference, 23-26 February 2010, Rome, Italy. Advances in knowledge organization v. 12. Würzburg: Ergon-Verlag, pp. 248-52.