APPLICATION OF THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR) TO MUSIC

Jenn Riley

Indiana University
Digital Library Program
jenlrile@indiana.edu

ABSTRACT

This paper describes work applying the Functional Requirements for Bibliographic Records (FRBR) model to music, as the basis for implementing a fully FRBR-compliant music digital library system. A detailed analysis of the FRBR and Functional Requirements for Authority Data (FRAD) entities and attributes is presented. The paper closes with a discussion of the ways in which FRBR is gaining adoption outside of the library environment in which it was born. This work benefits the MIR community by demonstrating a model that can be used in MIR systems for the storage of descriptive information in support of metadata-based searching, and by positioning the Variations system to be a source of robust descriptive information for use by third-party MIR systems.

1. THE FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS (FRBR)

Functional Requirements for Bibliographic Records (FRBR) is a conceptual model released in 1997 by the International Federation of Library Associations and Institutions (IFLA). FRBR provides a "framework that would provide a clear, precisely stated, and commonly shared understanding of what it is that the bibliographic record aims to provide information about, and what it is that we expect the record to achieve in terms of answering user needs" [9] (p. 2). In short, FRBR provides a model for the "bibliographic universe" by describing the entities and attributes that make up library metadata.

FRBR entities are divided into three groups. Group 1 entities are the "products of intellectual or artistic endeavour," and include Work ("a distinct intellectual or artistic creation"), Expression ("the intellectual or artistic realization of a Work"), Manifestation ("the physical embodiment of an Expression of a Work") and Item ("a single exemplar of a manifestation"), as seen in Figure 1 [9] (p. 13). The FRBR Group 2 entities are those that create or act upon Group 1 entities: Person and Corporate Body. FRBR Group 3 entities are the subjects of Works: Concept, Object, Event, and Place.

The application of FRBR to music has been described (though mostly in passing) at past ISMIR conferences

[4][8][18][19][22][23]. More in-depth studies of FRBR for musical materials appear in the library and information science literature. LeBoeuf performs a semiotic analysis of FRBR for musical works, raising crucial questions regarding the boundaries of Works and proposing revised definitions for the FRBR Group 1 entities [12]. Miller and LeBoeuf [14] analyze the application of the FRBR Group 1 entities to live events of performing arts, focusing on mixed-media works including drama or choreography. Vellucci, in the most systematic published analysis to date, considers what the Work and Expression entities and their various attributes mean for music [25].

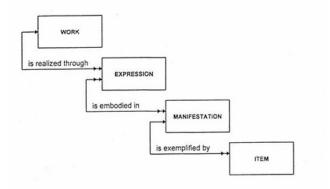


Figure 1. The FRBR group 1 entities [9] (p. 14)

FRBR has also served as the underlying conceptual model for a small number of music digital libraries, most notably MusicAustralia [1] and the evolving Probado system at the Bavarian State Library [4]. While research reports from some of these initiatives are available, specific details needed as a guide to local FRBR adoption efforts are generally more difficult to obtain.

Despite the work that has been done to study and test implementation of FRBR in music digital libraries, it is still not clear how the model should be applied to a newly-developed system. Even taking into account the move of some FRBR implementations for music towards production-level systems, the music library and MIR communities' understanding of FRBR exists largely at a theoretical rather than a practical level. The work presented in this paper represents a significant step forward in demonstrating how this theory can be turned into practice.

2. THE VARIATIONS SYSTEM AND FRBR

2.1. Variations music metadata

Like FRBR, the Variations digital music library system at Indiana University has been discussed at previous ISMIR conferences [5][6][8][19][22], demonstrating the value of contributions from the academic music library and digital library communities to MIR research. Variations2 was a research and development project investigating such wide areas as system architecture, metadata standards, component-based application architecture, and network services. A system that provides access to streaming audio and scanned scores developed as part of the Variations2 project has been put into production at the Cook Music Library in the Indiana University Jacobs School of Music. A follow-on project known as Variations3 is currently underway, which aims to provide a version of the software that can be deployed at other institutions, and continues some of the music metadata research started in the earlier project. The source code for the Variations Audio Timeliner¹ and the Variations metadata model, described below, both show promise for use within MIR systems.

The Variations system uses a work-based metadata model which has been described in more detail at previous ISMIR conferences [8][22] and elsewhere [16]. While the Variations metadata model is similar in many ways to FRBR, it was not envisioned as a FRBR implementation at the time of its creation, relying heavily instead on the music modeling work of Richard Smiraglia [24]. Table 1 shows a partial summary of the differences between the Variations model and FRBR.

Variations2/3 Entity	FRBR Group 1
	Entity
Work	Work
(more concrete than FRBR Work)	
Instantiation	Expression
(can only appear on one	
Container)	
Container	Manifestation
(includes some copy-specific data)	
Media Object	Item
(defined as a digital file)	

Table 1. Comparison of Variations and FRBR models

The FRBR model is increasingly taking hold in the library environment, and shows promise for improving interoperability of data within libraries and beyond, including with the MIR community. The increasing uptake of FRBR led the Variations team to embark upon an in-depth study of how the FRBR model applies to

musical materials. In order to make our work as useful as possible to others outside of our home institution, we structured our investigations as direct analysis of FRBR, rather than presenting our analysis as incremental changes to our existing data model.

We also limited our work to applying the FRBR model to scores and recordings, rather than to other types of material held by music libraries or managed by MIR systems. With this approach we did not model musical works as separate from texts included in them, in contrast to the method that some have proposed as ideal [1]. This decision is largely a practical one, limiting the scope of study to materials in the Variations system.

To function as a production service, Variations also requires a variety of structural and technical metadata. While some attributes described in the FRBR report could be used to support these functions, FRBR is primarily focused on descriptive metadata. Our team took the approach of using FRBR for the descriptive metadata in Variations only, with plans to model structural and technical metadata separately.

2.2. FRBR Group 1 Entities

The Variations team's initial efforts to study the FRBR model focused on the FRBR Group 1 entities: Work, Expression, Manifestation, and Item. We created operational definitions of each of the Group 1 entities as they applied to music, evaluated each attribute and relationship for its utility for the description of musical materials, and noted areas of description needed for music that were not covered in the FRBR report. Detailed results of our analysis are documented in a white paper freely available on the project web site [20].

2.2.1. Defining Work and Expression

One area of particular interest for music is the boundary between FRBR Works and Expressions. Phrases in the FRBR report suggest a Work for music should be interpreted broadly rather than strictly, including passages such as "addition of parts or an accompaniment to a musical composition" and "musical transcriptions and arrangements" as referring to new Expressions rather than new Works [9] (p. 17). Based on this approach, our operational definition for a Work is liberal and abstract, following the FRBR characterization of "adaptation" as a relationship between two Works, and "arrangement" as a relationship between two Expressions of the same Work.

This basic definition functions well for music in the canon of Western art music, where the composition takes precedence over any given performance of it. For other types of music, the notion of an abstract work is less natural. The Variations team proceeded nevertheless to create definitions of "Work" for music from these other traditions, believing that compromises for the modeling of

¹ http://sourceforge.net/projects/variations/

specific cases are worth the benefits of using a single model throughout the system. For jazz, our operational definition of a FRBR Work for music is the "tune." Performances of that tune, even widely diverging in nature, would be Expressions of that same Work. A fundamental transformation of the work would be considered a new Work. For pop music (itself a difficultto-define category), we defined the "song" as a Work. Covers and different performances by the same artist or group are therefore considered separate Expressions of that same Work. An album, when it represents a cohesive artistic whole, can also be considered a Work, with a whole-part relationship to the individual songs on the album. World and non-Western music, including systems that operate in an oral rather than a written tradition, pose a particular challenge to a work-based model. The Variations team's definition of a Work for music from these traditions is necessarily vague, and specifies that the cultural or ethnic group responsible for the music and the location and context in which the music is generally performed should be major factors in the decision as to whether two performances reflect two Works or different Expressions of the same Work.

A specific definition of Expression for music requires even more interpretation of the FRBR report than Works. It is clear from the FRBR report that Expressions have a defined form; that is, music represented in visual notation is a different Expression than music represented as an audio recording. Yet since different arrangements are considered different Expressions of the same Work, for recording Manifestations, each Expression represents both an arrangement and a specific interpretation of that arrangement as a musical performance. The combining of the abstract notion of arrangement together with the slightly more concrete idea of realizing a Work in a particular notation or performance seems to reflect a weakness of the FRBR model for the performing arts. Vellucci in response to this issue proposes that FRBR entities have "subentities." In this approach, notation and performance are separate Expressions, each with "sub-Expressions" for particular notations or performances [25] (p. 142-4). The Variations team did not take this approach, as it is possible to connect performances of the same arrangement Expression with mechanisms described in the FRBR report through the use of Expression-Expression relationships. While our solution might not give the notion of arrangement the intellectual primacy it holds in typical discussions of Western art music, it appears to be sufficient to support the requirements of Variations.

2.2.2. Balancing rigor with practicality

In our work applying FRBR to music, the Variations team faced difficult compromises between rigorous data

modeling and practical considerations. In some cases our team determined additional levels of complexity were not required, and in others we believed there was benefit to following the FRBR model to its fullest. We took the former approach in two notable cases. First, due to our early decision that our model would only create entities representing musical content, we decided to track language of liner notes and other material appearing on the Manifestation as an attribute of that Manifestation, although a different FRBR interpretation might model the liner notes as a separate Work with an Expression that appears together on the recording Manifestation. Second, although multiple volumes in a set can be modeled as separate Manifestations with whole-part relationships to a set-level Manifestation according to the FRBR report, the Variations team determined that this would not be necessary in most cases for our purposes.

We took the approach of following more complex features of FRBR more often, allowing for more robust services to be built on our data. The current Variations system allows us to track the date of first performance of a musical Work (i.e., its first performance Expression), as this data can be used to support a wider range of research questions than most current library catalogs. In order to support this type of discovery in a FRBR-based system, an Expression should be modeled for the original performance, even if the Variations system did not include a recording of that performance. We also determined that the recording held by the library should be modeled as a separate Manifestation than the WAV file generated from it, and that the MP3 file for streaming to end-users was a third Manifestation.

FRBR Items presented a challenge in our analysis. The FRBR report describes Items as physical entities, and does not explicitly cover the case where a Manifestation exists only as a digital file rather than in physical form. While there is some disagreement in the community as to how digital files should be handled [7], the Variations team concluded that the intent of the FRBR authors was that digital files should be FRBR Items. The Item-Item "reproduction" relationship will be used to record which copy of a physical recording was digitized to create a particular WAV file. For digitized scores, each individual page might be modeled as a separate Item², although our team is reserving a final decision on this issue until we complete our work on specifying technical metadata needed for effective system operation. These decisions represent a fairly complex model for Items, beyond what is implied by the text of the FRBR report. We expect, however, that most Item attributes and relationships will be handled by Variations at the system level and not require manual intervention by a cataloger.

² The Probado project's implementation of FRBR introduces a "file" entity to avoid this situation. [4]

2.2.3. Additional attributes needed

The Variations team excluded a number of attributes and relationships defined in the FRBR report from local interpretation of the model, as they were not appropriate to the description of musical material. Our analysis also uncovered a few attributes that were not present in the FRBR model, but nevertheless would be necessary for the Variations system. The "missing" attributes generally fell three categories: those necessary simplifications to FRBR we instituted for practical reasons, those needed for efficient system operation, and those that appear to have simply been overlooked in the development of the FRBR report.

The first case, attributes necessary due to our decision to preference musical content over other types, can be seen in the need to record a language for a Work, since text settings are an integral part of musical works and in our interpretation are not considered to be separate Works with their own Expressions. The second case, attributes to assist with system operation, represent both new attributes and refinements of FRBR-defined attributes. An identifier attribute for Work and location, call number, and copy number for Item are examples of this second case.

The third case, attributes missing entirely from FRBR, represent our most significant departure from the FRBR model. Based on initial reactions to our findings from members of the FRBR Review Group³, we are optimistic that our findings will influence the growth of the FRBR model into the future. The attributes we noted as necessary for our local implementation but not obviously present in the FRBR report are place of composition and genre/form/style for Works, and place of performance, key, and genre/form/style for Expressions. We also noted that data essential for the understanding of non-Western music, such as geography, culture, event, and function could be recorded within FRBR-defined Work attributes, but that FRBR could be adjusted to model this type of information in a more prominent way.

2.3. Next Steps and Implementation Plans

The Variations team's initial study of the FRBR model extended only to the FRBR Group 1 Entities, assuming that if they matched the needs of the Variations system, then the rest of the FRBR model would as well. Our analysis demonstrated that moving the Variations system to a fully FRBRized model would be both possible and desirable. This decision was only the first step in what will be a much longer process.

Our next area of study was to extend our detailed analysis of the applicability of FRBR to musical material to the Group 2 (Person and Corporate Body) and Group 3 (Concept, Object, Event, and Place) entities and their attributes. We also considered in this second phase of our

analysis the model presented by the FRBR companion report Functional Requirements for Authority Data (FRAD), released in draft in April 2007 [10]. This second phase of our FRBR analysis proved more straightforward than the first, with relatively little interpretation of FRBR/FRAD entities and attributes needed for musical materials. A report describing this phase two analysis is available from the Variations3 project web site [21].

The Variations team concluded that modeling Names/Identifiers and Controlled Access Points as separate entities, as outlined in FRAD, was not necessary for a single-language system of the scale of Variations. We adjusted our phase one interpretations in a few cases, modeling features such as form/genre and instrumentation as FRBR/FRAD Group 3 Concept entities rather than as attributes on Work or Expression, but stopped short of large-scale conversion of attributes to relationships to Group 3 entities, in keeping with the spirit of the FRBR report limiting Group 3 entities to "subjects" of Works.

We expect to embark upon a third planning stage in the late summer and fall of 2008, where we will adopt a FRBR encoding syntax, or develop our own if necessary. FRBR as a conceptual model does not define a formal data structure for recording FRBR data, and one has not emerged from the library community from a trusted maintenance agency that would represent an obvious choice for adoption by Variations. A few FRBR encodings have emerged from efforts outside of the core library community, however, and the Variations team plans to study each of these to determine if they meet the functional requirements for our system, and to analyze their likely sustainability and supportability over time.

3. ADOPTION OF FRBR AND IMPLICATIONS FOR MUSIC INFORMATION RETRIEVAL

While FRBR originated within libraries as an effort to better understand the bibliographic universe, it has also attracted attention from outside the library community. An expression of FRBR in RDF, allowing FRBR data to interact in Semantic Web environments, has arisen from a small but diverse group of information professionals [2]. The museum community recognized similarities between FRBR and its CIDOC Conceptual Reference Model⁴, and organized an effort to create an object-oriented version of FRBR that aims to harmonize the two models [11]. The enthusiast- and researcher-driven Music Ontology, expressed in the OWL Web Ontology Language⁵, uses FRBR together with other existing frameworks such as Friend of a Friend (FOAF)⁶ to provide a "formal framework for dealing with music-related information on

³ http://www.ifla.org/VII/s13/wgfrbr/

⁴ http://cidoc.ics.forth.gr/index.html

⁵ http://www.w3.org/2004/OWL/

⁶ http://www.foaf-project.org/

the Semantic Web, including editorial, cultural and acoustic information" [18] (p. 417). The Music Ontology especially falls within the area of interest for the MIR community, and these developments can be interpreted as a sign that the FRBR approach has proven useful beyond libraries, including in MIR research.

Within the library environment, the FRBR model is gaining significant traction as well. While a number of systems have implemented FRBR or FRBR-like models in test environments, more production-scale implementations are starting to emerge, such as OCLC's WorldCat.org. A recent report made to the Library of Congress concerning the future of library cataloging recognized the potential of the FRBR model for enhancing library services, but called for more testing of it in production situations [13]. A new content standard for library cataloging records called Resource Description and Access (RDA)⁸ is in development, which will be based on FRBR principles. These and other changes are driving investigations into the use of record structures other than the current MARC standards in ubiquitous use in libraries [3][15]. It seems clear that FRBR has taken hold as a fundamental guiding model for ongoing metadata development in libraries. We expect that Variations, as a FRBR-based system, will serve as a model for these developments.

The library community has recognized that this time of change in our system architectures and standards presents us with an opportunity to participate more fully in the global information environment. The activity in this area most relevant to the ISMIR community is the work of the Dublin Core Metadata Initiative (DCMI)/RDA Task Group. This group has three primary goals: the definition of the "elements" presented by RDA as Resource Description Framework (RDF)⁹ properties and classes, the exposure of in-line vocabularies defined in RDA for use by Semantic Web tools, possibly in RDF Schema (RDFS)¹⁰ or Simple Knowledge Organization System (SKOS)¹¹, and the development of a DCMI application profile for FRBR and FRAD.¹² When library-generated metadata is openly exposed in formats such as these, communities such as MIR can benefit.

Indiana University's plans to move Variations to a FRBR-based model, if a pending funding application is awarded to support the implementation phase of our work, would be of particular advantage to the MIR community in two ways. First, structured music metadata in standards-based and commonly-understood forms would be available to MIR researchers long before the vast

10 http://www.w3.org/TR/rdf-schema/

amounts of library metadata as a whole could undergo a similar transformation. Such wide exposure of librarygenerated music metadata would allow it to be used in new environments, such as the emerging linked data movement [17]. This metadata could be used as a source of ground truth, as information for display to users in systems demonstrating MIR research, in support of production music discovery and use systems, or even as primary data worthy itself of analysis and study. Second, the interpretation of the FRBR model for musical materials that the Variations team has performed could serve as a model for MIR systems that wish to store descriptive metadata as the basis of metadata-based searching. The potential for advancing the state of MIR research as a result of the Variations/FRBR harmonization and other related initiatives in the library sector is indeed

4. ACKNOWLEDGMENTS

This paper is based in part upon work supported by the Institute of Museum and Library Services. Any views, findings, conclusions, or recommendations expressed in this paper do not necessarily represent those of the Institute of Museum and Library Services.

5. REFERENCES

- [1] Ayres, Marie-Louise. "Case Studies in Implementing Functional Requirements for Bibliographic Records (FRBR): AustLit and MusicAustralia." *ALJ: The Australian Library Journal* 54(1), 2005, pp. 43-54.
- [2] Davis, Ian, Richard Newman, and Bruce D'Arcus. "Expression of Core FRBR Concepts in RDF." August 10, 2005. http://vocab.org/frbr/core
- [3] Delsey, Tom. "RDA Database Implementation Scenarios." January 14, 2007. http://www. collectionscanada.gc.ca/jsc/docs/5editor2.pdf
- [4] Diet, Jürgen and Frank Kurth. "The Probado Music Repository at the Bavarian State Library." Proceedings of the International Conference on Music Information Retrieval, Vienna, Austria, 2007. http://ismir2007.ismir.net/proceedings/ISMIR2007_p 501_diet.pdf
- [5] Dunn, Jon W. "Beyond VARIATIONS: Creating a Digital Music Library." Proceedings of the International Symposium on Music Information Retrieval, Plymouth, Massachusetts, 2000. http:// ismir2000.ismir.net/papers/invites/dunn_invite.pdf
- [6] Dunn, Jon W., Mary Wallace Davidson, and Eric J. Isaacson. "Indiana University Digital Music Library Project." Proceedings of the International Symposium on Music Information Retrieval,

⁷ http://www.worldcat.org. Note that this system is not a full FRBR implementation; it lacks a robust Expression entity.

⁸ http://www.collectionscanada.ca/jsc/rda.html

⁹ http://www.w3.org/RDF/

¹¹ http://www.w3.org/2004/02/skos/

¹² http://dublincore.org/dcmirdataskgroup/

- Bloomington, Indiana, USA, 2001. http://variations2.indiana.edu/pdf/dml-ismir2001.pdf
- [7] Floyd, Ingbert and Allen Renear. "What Exactly is a FRBR Item in the Digital World?" Grove, Andrew, Ed. *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, Milwaukee, Wisconsin, 2007. http://hdl.handle.net/2142/5254
- [8] Hemmasi, Harriette. "Why Not MARC?" Proceedings of the International Conference on Music Information Retrieval, Paris, France, 2002. http://ismir2002.ismir.net/Proceedings/02-FP08-3.pdf
- [9] IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional Requirements for Bibliographic Records. Munich: K.G. Saur, 1997, with amendments in 2008. http://www.ifla.org/VII/s13/frbr/frbr_2008.pdf
- [10] IFLA Working Group on Functional Requirements and Numbering of Authority Records. Functional Requirements for Authority Data: A Conceptual Model. Draft, April 1, 2007. http://www.ifla.org/VII/d4/FRANAR-ConceptualModel-2ndReview.pdf
- [11] International Working Group on FRBR and CIDOC CRM Harmonisation. "FRBR Object-oriented Definition and Mapping to FRBR-ER." Version 0.9 draft, January 2008. http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR_oo_V0.9.pdf
- [12] LeBoeuf, Patrick. "Musical Works in the FRBR Model or 'Quasi la Stessa Cosa': Variations on a Theme by Umberto Eco." *Cataloging and Classification Quarterly* 39(3/4), 2005, pp. 103-124.
- [13] Library of Congress Working Group on the Future of Bibliographic Control. "On the Record." January 9, 2008. http://www.loc.gov/bibliographicfuture/news/lcwg-ontherecord-jan08-final.pdf
- [14] Miller, David and Patrick Le Boeuf. "Such Stuff as Dreams are Made On': How Does FRBR Fit Performing Arts?" *Cataloging & Classification Quarterly* 39(3/4), 2005, pp. 151-178.
- [15] Mimno, David, Gregory Crane, and Alison Jones. "Hierarchical Catalog Records: Implementing a FRBR Catalog." *D-Lib Magazine* 11(10), October 2005. http://www.dlib.org/dlib/october05/crane/10crane.html
- [16] Notess, Mark, Jenn Riley, and Harriette Hemmasi. "From Abstract to Virtual Entities: Implementation of Work-Based Searching in a Multimedia Digital Library." *Proceedings of the European Conference*

- on Digital Libraries, Bath, UK, 2004. http://mypage.iu.edu/~mnotess/ECDL/ecdl-04-reprint.pdf
- [17] Raimond, Yves, Christopher Sutton, and Mark Sandler. "Automatic Interlinking of Music Datasets on the Semantic Web." *Linked Data on the Web Conference*, Beijing, China, 2008. http://events.linkeddata.org/ldow2008/papers/18-raimond-sutton-automatic-interlinking.pdf
- [18] Raimond, Yves, Samer Abdallah, Mark Sandler, and Frederick Giasson. "The Music Ontology." Proceedings of the International Conference on Music Information Retrieval, Vienna, Austria, 2007. http://ismir2007.ismir.net/proceedings/ISMIR2007_p 417_raimond.pdf
- [19] Riley, Jenn. "Exploiting Musical Connections: A Proposal for Support of Work Relationships in a Digital Music Library." Proceedings of the International Conference on Music Information Retrieval, London, England, 2005. http://ismir2005.ismir.net/proceedings/1108.pdf
- [20] Riley, Jenn, Caitlin Hunter, Chris Colvard, and Alex Berry. "Definition of a FRBR-Based Metadata Model for the Indiana University Variations3 Project." September 10, 2007. http://www.dlib.indiana.edu/ projects/variations3/docs/v3FRBRreport.pdf
- [21] Riley, Jenn, Casey Mullin, Chris Colvard, and Alex Berry. "Definition of a FRBR-Based Metadata Model for the Indiana University Variations3 Project. Phase 2: FRBR Group 2&3 Entities and FRAD." July 23, 2008. http://www.dlib.indiana.edu/projects/ variations3/docs/v3FRBRreportPhase2.pdf
- [22] Scherle, Ryan and Don Byrd. "The Anatomy of a Bibliographic Search System for Music." Proceedings of the International Conference on Music Information Retrieval, Barcelona, Spain, 2004. http://ismir2004.ismir.net/proceedings/p089-page-489-paper241.pdf
- [23] Smiraglia, Richard. "Musical Works as Information Retrieval Entities: Epistemological Perspectives." Proceedings of the International Symposium on Music Information Retrieval, Bloomington, Indiana, 2001. http://ismir2001.ismir.net/pdf/smiraglia.pdf
- [24] Smiraglia, Richard. *The Nature of 'The Work.'* Lanham, Md: Scarecrow Press, 2001.
- [25] Vellucci, Sherry L. "FRBR and Music." In Understanding FRBR: What it Is and How it Will Affect Our Retrieval Tools, ed. Arlene G. Taylor; pp. 131-151. Westport, CT: Libraries Unlimited, 2007.