Implementing the FRBR Conceptual Model in the Variations Music Discovery System

Jenn Riley and Alex Berry
...with thanks to Paul McElwain and the rest of the V/FRBR Project Team
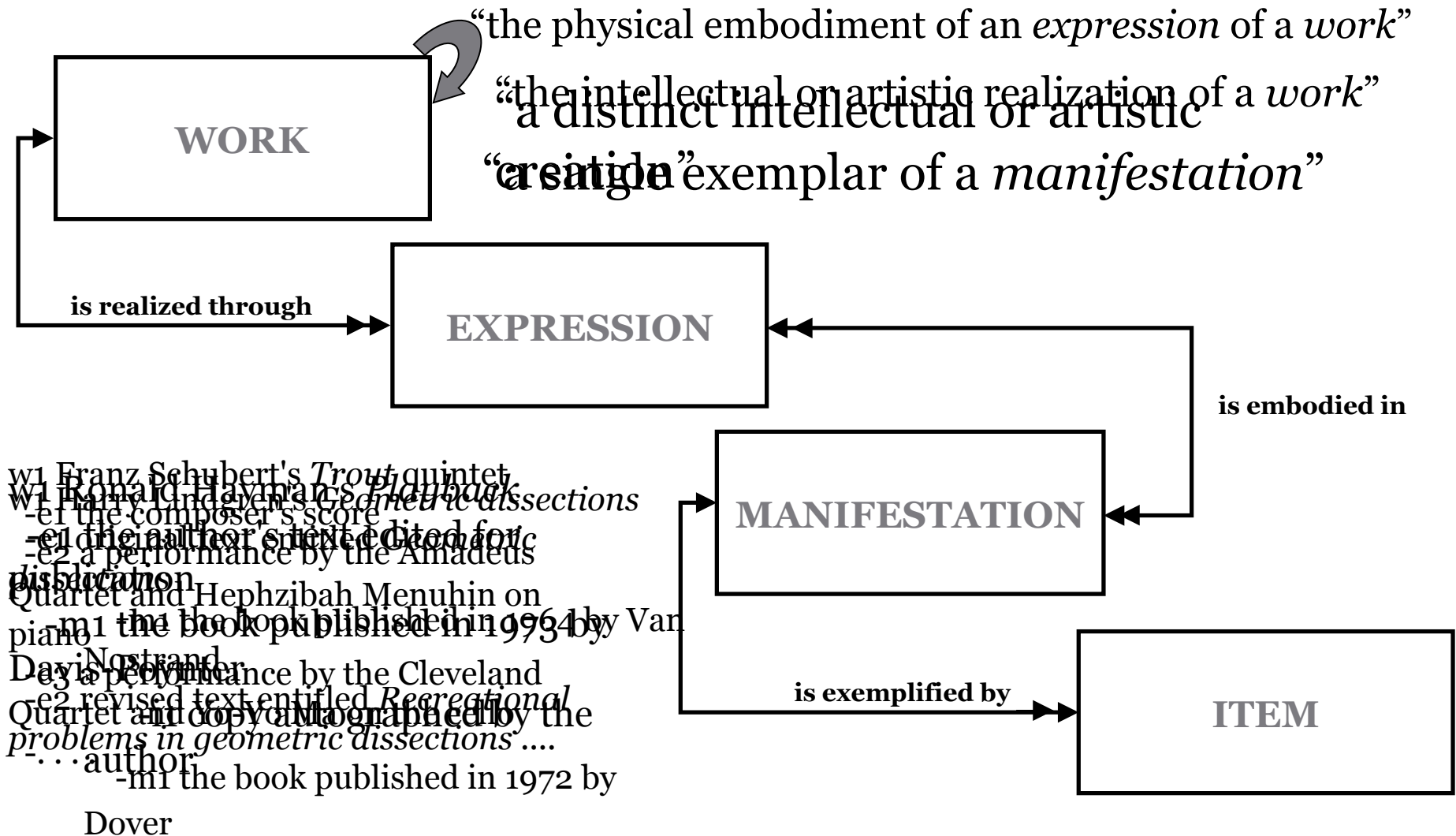
DLP Brown Bag Series
October 28, 2009

What if a catalog could answer questions like…

- Which recordings of Beethoven's early piano sonatas can I listen to from home?
- Which recordings can I listen to with Vladimir Ashkenazy as a conductor, rather than a pianist?
- How many editions of the score of [some interesting work] can I compare?
- What pieces can I as a cellist perform with my friend the soprano? Or my other friend the violinist?
- What do you have with Kirsten Flagstad singing?
- And even…
  - What works were composed by women in Vienna in the 18th century?
  - Which 19th-century composers wrote sonatas for oboe?

- Can the catalog become a *research* tool rather than just a *finding* tool?

The core of FRBR: Group 1 Entities

"the physical embodiment of an *expression* of a *work*"

"the intellectual or artistic realization of a *work*"

a distinct intellectual or artistic creation "a single exemplar of a *manifestation*"

WORK

EXPRESSION

is realized through

MANIFESTATION

is embodied in

ITEM

is exemplified by

w1 Franz Schubert's *Trout* quintet

w1 Franz Haydn's *Gsc fourth work*

w1 Harry Lindgren's *Geometric dissections*

-e1 the composer's score

-e1 the author's text edited for

-e2 a performance by the Amadeus Quartet and Hephzibah Menuhin on piano

-m1 the book published in 1964 by Van Nostrand

-e3 a performance by the Cleveland Quartet and copyrighted by the author

-e2 revised text entitled *Recreational problems in geometric dissections* ....

-m1 the book published in 1972 by Dover

(Semi)FRBRization in WorldCat Local

Basics of the V/FRBR project

- Funded by an Institute of Museum and Library Services National Leadership Grant
  - October 1, 2008 - September 30, 2011
  - $481,987 from IMLS; $482,572 contributed by IU
  - Staff from DLP, Music Library, and Tech Services are contributing some of their time to the project
- In "Demonstration" category
  - Primary mission: provide a model for other FRBRized catalogs
  - Secondary mission: provide a useful and sustainable discovery system for music at IU

Project staff

- **Jenn Riley**, Metadata Librarian, Digital Library Program (Principal Investigator)
- **Jon Dunn**, Associate Director for Technology, Digital Library Program
- **Mark Notess**, Development Manager, Digital Library Program
- **Paul McElwain**, Programmer/Analyst, Variations/FRBR Project
- **Alex Berry**, Programmer/Analyst, Variations/FRBR Project
- **Phil Ponella**, Director, Cook Music Library
- **Michelle Dalmau**, Digital Projects and Usability Librarian, Digital Library Program
- **Julie Hardesty**, Usability and Interface Specialist, Digital Library Program
- **Brian Wheeler**, System Administrator, Digital Library Program
- **Ralph Papakhian**, Head, Technical Services, Cook Music Library
- **Keith Cochran**, Associate Director/Music Collection Development Librarian, Cook Music Library
- **Spencer Anspach**, Library Systems Analyst/Programmer, Technical Services

Most visible goal: *FRBRize* Variations

- Provide a public and concrete testbed for FRBR
  - with real data
  - in a production environment
  - as a response to the call in the LC WG on the Future of Bibliographic Control report
- Follow through on long-standing vision for improved Variations searching
- Provide a sustainable path for Variations metadata outside of IUCAT

Other project goals

- To provide an openly-accessible web search interface to FRBRized data in Variations *for all scores and recordings in the music library*, for community analysis;
- To make supporting data, including data model documentation and FRBRized data, available to the community for analysis; and
- To apply innovative, evidence-based interface design techniques to Variations cataloging and search interfaces to make the most of the FRBR-compliant data model.

Planned project work products (1)

- A published FRBRization algorithm that operates on multi-Work Manifestations, and evaluation of its effectiveness
- A formal data model for FRBR
- FRBRized data made available to the community for further testing and analysis
- An openly-accessible system for searching FRBRized music data for community testing and analysis

Planned project work products (2)

- Usability evaluations of FRBR-based end-user discovery and cataloging systems
- Figures on the costs of creating FRBRized bibliographic data by both automated and manual means
- Source code for the Variations FRBR-based discovery system.

FRBR in XML

- Locally developing a suite of FRBR Schemas
- To provide a model for others encoding and sharing FRBRized data
- 3-level approach:
  - frbr – strict interpretation of FRBR report(s)
  - efrbr (extended FRBR) – make FRBR useful
  - vfrbr (Variations/FRBR) – add/remove data elements to optimize model for music
- Cover Group 1, 2, and 3 Entities, plus Relationships
- Create record packaging structure
- Current drafts use 33 different namespaces!
  - And this may grow before we're done. (Ugh.)

User testing so far

- 3 user studies performed by Julie Hardesty
  - Current Variations search observations/interviews
  - Current Variations logs analysis
  - Variations2 cataloging interface observation
- Reports available at <http://www.dlib.indiana.edu/projects/vfrbr/projectDoc/index.shtml>
- Informing design of developing search interface
- Later, will inform design of cataloging interface

Implementation Details

- Flow of Data Into Our System
- Search UI Design
- Technologies Used

Flow of Data Into Our System

- Primary Source: MARC Bib and Authority Files
- Processed by FRBRization logic creating Java objects
- Objects persisted and indexed for use

Bib Records

Total = 191804

By Type

Notated music = 104402
Musical sound recording = 84627
Manuscript notated music = 2480
Nonmusical sound recording = 251
Language material = 36
Mixed materials = 6
Projected medium = 1
Three-dimensional artifact or naturally occurring object = 1

Authority Records

Fetched from Z39.50 based on rules

For every 100 field search Personal Name using contents in |a |q |b |c |d
Import if one exact match

Potential Secondary Source of Data

MusicBrainz

- Useful for instances where we can't create any works from a record
- Helps us get around parsing contents notes for popular music

FRBRization

- Continued effort from previous Variations grants
- Attempts to create FRBR records from MARC records
- Contains work identification algorithm and mapping rules
- Geared specifically for music
- Only uses a small subset of FRBR

Work Identification Algorithm

Uses clues in MARC bib records to pull out works

- Presence of fields, subfields, and indicators
- Values of subfields compared to Collective Title and Forms lists

If the value in 240 |a equals the phrase "Chamber Music" do not identify 240 as a work

Example mapping rules

Work from Authority record

- Uniform Title 100,110,111 |t |m |n |r
- Instrumentation 100,110,111,130 |m -- make separate entries from each string delimited by comma; do not include (x); map value inside () to number

Results of FRBRization

- Java objects representing FRBR entities with attributes and relations
- Stored in a relational database and indexed for retrieval
- Used for VFRBR search

Search UI Design

The following slides were developed by Julie Hardesty during a series a meetings involving the VFRBR UI design team.

IU MU*sic* SE*arch*

Search

Creator/Composer: [   ]

Work Title: [   ]

Performer/Conductor: [   ]

Recording/Score Title: [   ]

Media Format: ● Recording   ○ Score

☐ Available online

[Search]   [Clear]

Information

Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

URL: http://www.namegoeshere.indiana.edu
Last Updated: Day, Month Day, Year, HH:MM:SS
Cook Music Library & Digital Library Program

Variations/FRBR Logo

beethoven piano sonata   **SEARCH**

35 results for "beethoven piano sonata" as keyword.   < >

Available Online [28]
Recordings [34]
  CD [27]
  Analog Disc [5]
  Tape [2]
Scores [1]
  Full Score [1]
Date of Publication
Date of Composition
Composer
Performer/Conductor
Instrumentation
Publisher
Language
Genre/Form

53 works matching "beethoven piano sonata."

1. Sonatas, piano, no. 1, op. 2, no. 1, F minor, Beethoven, Ludwig van (1770-1827). more..

2. Sonatas, piano, no. 2, op. 2, no. 2, A major, Beethoven, Ludwig van (1770-1827). more...

3. Sonatas, piano, no. 3, op. 2, no. 3, C major, Beethoven, Ludwig van (1770-1827). more...

more results>

Sort by: relevance   **1** 2 3 4 5 ... >> >|

☐ Select/Unselect Current Page   Print/Email/Save

☐ 1. **Sonate** per **piano**forte. **Beethoven**, Ludwig van. Milano: Ricordi, c1919.

Contents: **Sonatas, piano**, no. 1, op. 2, no. 1, F minor; more...
People: **Beethoven, Ludwig van**
Media: 2 v. of music; 30 cm.
Access: Variations I Copies: See Catalog

☐ 2. The complete **piano sonatas** on period instruments. **Beethoven**, Ludwig van (1770-1827). Thun, Switzerland: Claves, p1997.

Variations/FRBR Logo

beethoven piano sonata   **SEARCH**

< Back to search results

| < | > |

The complete piano sonatas on period instruments.

Access: Variations I Copies: See Catalog

CONTENTS | -- | + |

Sonatas, piano, no. 1, op. 2, no. 1, F minor

Composed by
  Beethoven, Ludwig van
Contributors
  Bilson, Malcolm (Performer)
Performance Dates
  1996-03
  1996-06
  1996-07
  1996-11
Place of Performance
  Maria Minor Church in Utrecht (Holland) and Masterview Sound Studio in Ithaca (New York)
Instrumentation
  Piano
Subject Headings
  Sonatas (Piano)
Copyright Declarations
  ©1997 Claves Records, Recording

  hide

Sonatas, piano, no. 2, op. 2, no. 2, A major more...
Sonatas, piano, no. 3, op. 2, no. 3, C major more...
Sonatas, piano, no. 4, op. 7, E♭ major more...

Technologies Used in Search UI

- Hibernate
- Solr
- Wicket
- jQuery

Hibernate

- Object-relational mapping framework for Java
- Well established in industry
- Works with a wide variety of databases

Solr

- A search server built on top of the popular Lucene library
- Use HTTP requests to index data and query index
- Uses simple XML documents throughout
- Enables search interfaces instead of just acting as a full text index

Wicket

- Component based web application framework
- Supports reuse at various granularities
- Uses XHTML templates that are manipulated via Java code (no JSP)

jQuery

- Cross browser JavaScript library
- Makes interaction between JS, CSS, and XHTML easier

Current status of work

- Project programmers hired April 2009
- XML schemas will be released "soon"
- Software development taking "agile" approach
- Starting to process extract of 191,804 IUCAT music library score and recording records
- Will have a discovery interface for people to look at "soon"
- Will start work on cataloging tool later this fall

To stay informed

- These presentation slides: http://www.dlib.indiana.edu/~jenlrile/ presentations/bbfall09/vfrbr.pptx
- Project home page: http://www.dlib.indiana.edu/ project/vfrbr
- Subscribe to project RSS feed from the home page
- Talk to a member of our project team!

<div align="center">

jenlrile@indiana.edu

berry3@indiana.edu

pbmcelwa@indiana.edu

Others listed on project home page

</div>