Metadata for Music: Understanding the Landscape

Jenn Riley

Indiana University Digital Library Program

One metadata standard for all music is a pipe dream

- "Music" is diverse
  - Many styles
  - Many cultures
  - Many formats
  - Many uses
- Institutions that need to describe music are diverse
- Users are diverse

Guiding principles

All metadata standards represent compromise between treating one type of material extremely well and being more generally applicable.

Good metadata is *fit for a purpose.*

Some types of standards

- Descriptive metadata structure standards
  - General
  - Music-specific
- Music markup languages
- Content standards
- Controlled vocabularies
- Conceptual frameworks
- (Will leave out technical, structural, rights, etc., for today)

Descriptive metadata structure standards - general

- Often include elements that can be used for features specific to music (instrumentation, key, etc.)
- Profiles and application guidelines allow music-specific description
- Using these allows easier interoperability with other types of collections
- Best practice to examine these options first, and move on to other solutions only if none meet the specific needs of the project

MAchine Readable Cataloging (MARC)

- Used for library metadata since 1960s
  - Adopted as national standard in 1971
  - Adopted as international standard in 1973
- Actually a family of MARC standards throughout the world
  - U.S. & Canada use MARC21
- ANSI/NISO Z39.2, ISO 2709
- Sample implementation: any library OPAC

Good times to use MARC

- Integration with other records in OPAC
- Resources are like those traditionally found in library catalogs
- Maximum compatibility with other libraries is needed
- Have expert catalogers for metadata creation

Metadata Object Description Schema (MODS)

- Developed and managed by the Library of Congress Network Development and MARC Standards Office
- Emerged with the recent explosion of metadata standards
  - First released for trial use June 2002
  - MODS 3.2 released late 2006
- Influenced by MARC, but intended to be simpler
- XML with textual tag names
- Sample implementation: IN Harmony from Indiana University

Good times to use MODS

- Materials lend themselves to library-type description
- Want to reach both library and non-library audiences
- Need a reasonably robust format, but not the formalities of MARC
- Want XML representation to store within complete digital object metadata
- Have or can build a MODS-aware search system

Dublin Core (DC)

- Base version known as “simple” or “unqualified” DC
- 15-element set
- ANSI/NISO Z39.85/ISO 15836
- DC principles
  - “Core” across all knowledge domains
  - No element required
  - All elements repeatable
  - 1:1 principle
- Sample implementation: Sheet Music Consortium

Good times to use Unqualified DC

- Cross-collection searching
- Cross-domain discovery
- Metadata sharing
- Describing some types of simple resources
- Metadata creation by novices

Qualified Dublin Core (QDC)

- Adds some increased specificity to Unqualified Dublin Core
- Types of DC qualifiers
  - Additional elements
  - Element refinements
  - Encoding schemes (vocabulary & syntax)
- Sample implementation: Pacific Northwest Sheet Music Collection

Best times to use QDC

- More specificity needed than simple DC, but not a fundamentally different approach to description
- Are using a content management system with QDC as the default metadata format
- Want to share DC with others, but need a few extensions for your local environment
- Describing some types of simple resources
- Metadata creation by novices

Encoded Archival Description (EAD)

- Markup language for archival finding aids
- Maintained by the Society for American Archivists EAD Working Group
- Designed to accommodate multi-level description
- Developed along with the first of the metadata standards of the SGML/XML era
  - EAD 1.0 released in 1998
  - EAD2002 finalized in December 2002
- Sample implementation: Duke Historic American Sheet Music

Good times to use EAD

- When working in a special collections environment
- When you don't need/want item-level description
- When provenance of the collection significantly adds to its understanding
- Have or can build an EAD-aware search system

Descriptive metadata structure standards – music-specific

- Require specialized systems to implement
- Allow for more robust (and more expensive) description
- Some examples
  - Variations2 at Indiana University
  - Probado at the Bavarian State Library
  - Music Ontology Specification
  - ID3 tags for MP3 files
  - MANY local databases!

MPEG-7 in the middle

- For “multimedia” (still images, moving images, audio)
- Includes many “low-level” features that could be provided by recording devices
- Covers more than just descriptive metadata
- Not generally a good choice for supporting bibliographic searching in libraries

Music markup languages

- Encode musical notation itself
- Usually have some basic descriptive metadata capabilities
- Some examples
  - MusicXML
  - Music Encoding Initiative (MEI)
  - Standard Music Description Language (SMDL)

Content standards

- Theoretically designed to be used with any metadata structure standard
- Can adopt in whole or in part
- Some examples
  - AACR2
  - Describing Archives: A Content Standard (DACS)
  - Cataloging Cultural Objects (CCO)
  - Resource Description and Access (RDA)
  - Cataloging Sheet Music: Guidelines for Use with AACR2 and the MARC Format

Controlled vocabularies

- Many options exist, both large and small
- For each field, determine if vocabulary control is needed, and select an appropriate vocabulary
- Different vocabularies can be used for various classes of metadata elements
  - Names
  - Places
  - Genres
  - Topical subjects
- Increasingly moving towards mixing and matching

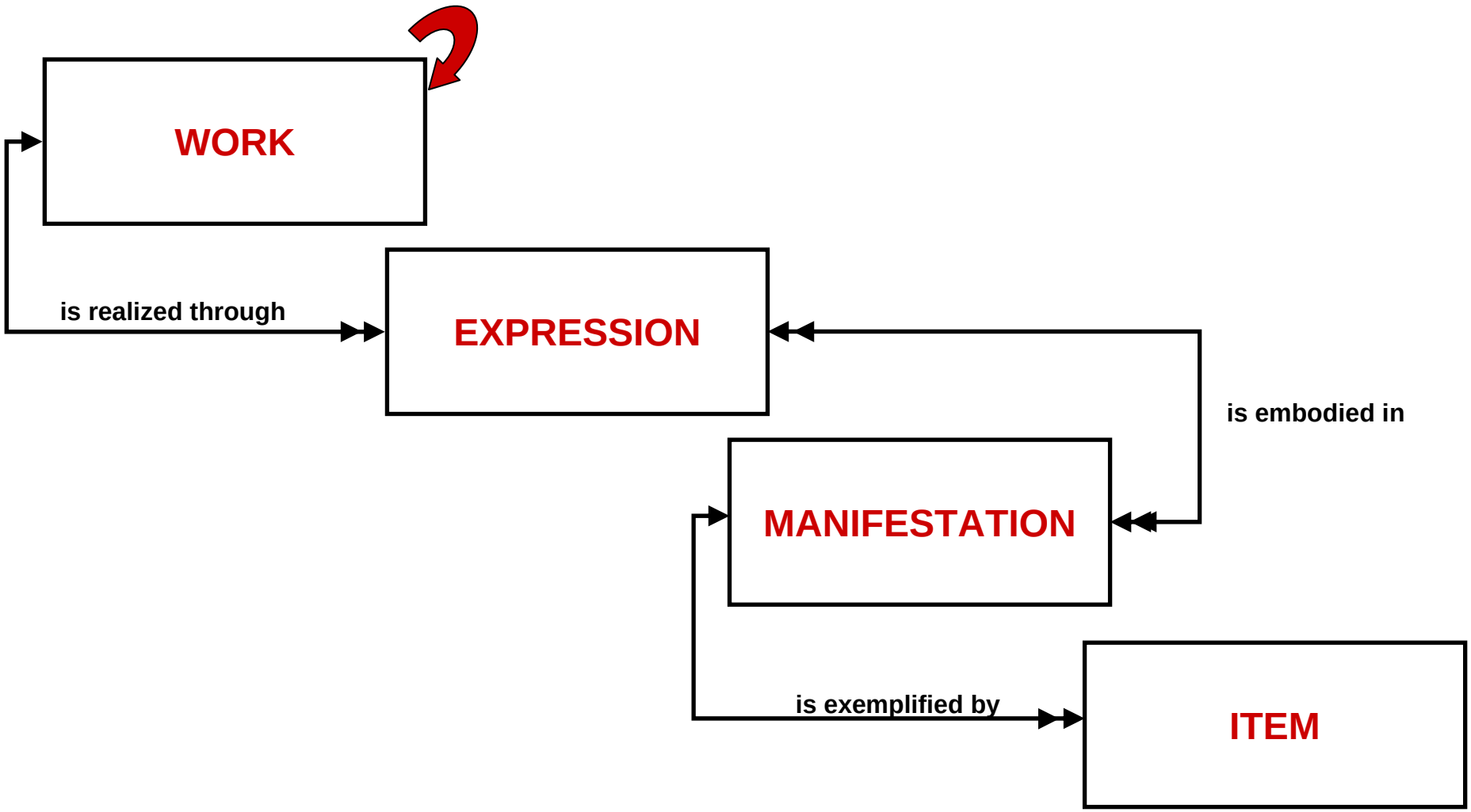
Conceptual frameworks

- Provide theoretical description of *entities* operating in a specific *domain*
- Aim to provide conceptual clarity in complex environments
- Operate independently of specific metadata syntaxes
- Usually intended to be invisible to users
- Can underlie specific metadata structure standards and content standards
- Useful more as a way of thinking about metadata than as a set of prescriptive guidelines to follow

Functional Requirements for Bibliographic Records (FRBR)

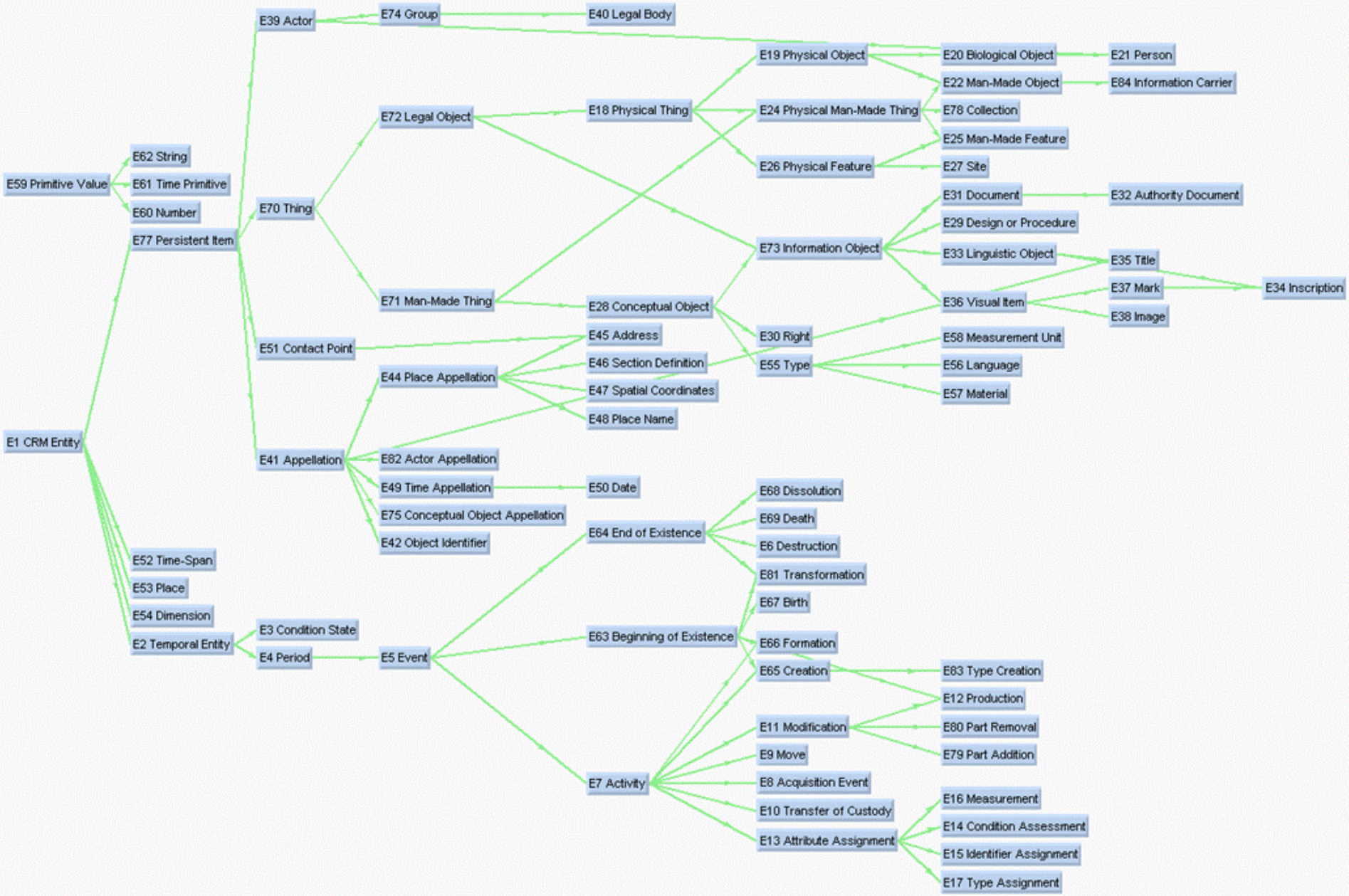
- 1998 IFLA report
- Defines the following for the bibliographic universe:
  - Entities & their relationships
  - User tasks
  - Basic requirements for national bibliographic records
- Work-based nature a natural fit for music
- Will underlie new RDA content standard
- No explicit moves yet to make MARC FRBR-based, therefore;
- Slow adoption by ILS vendors

FRBR Group 1 Entities



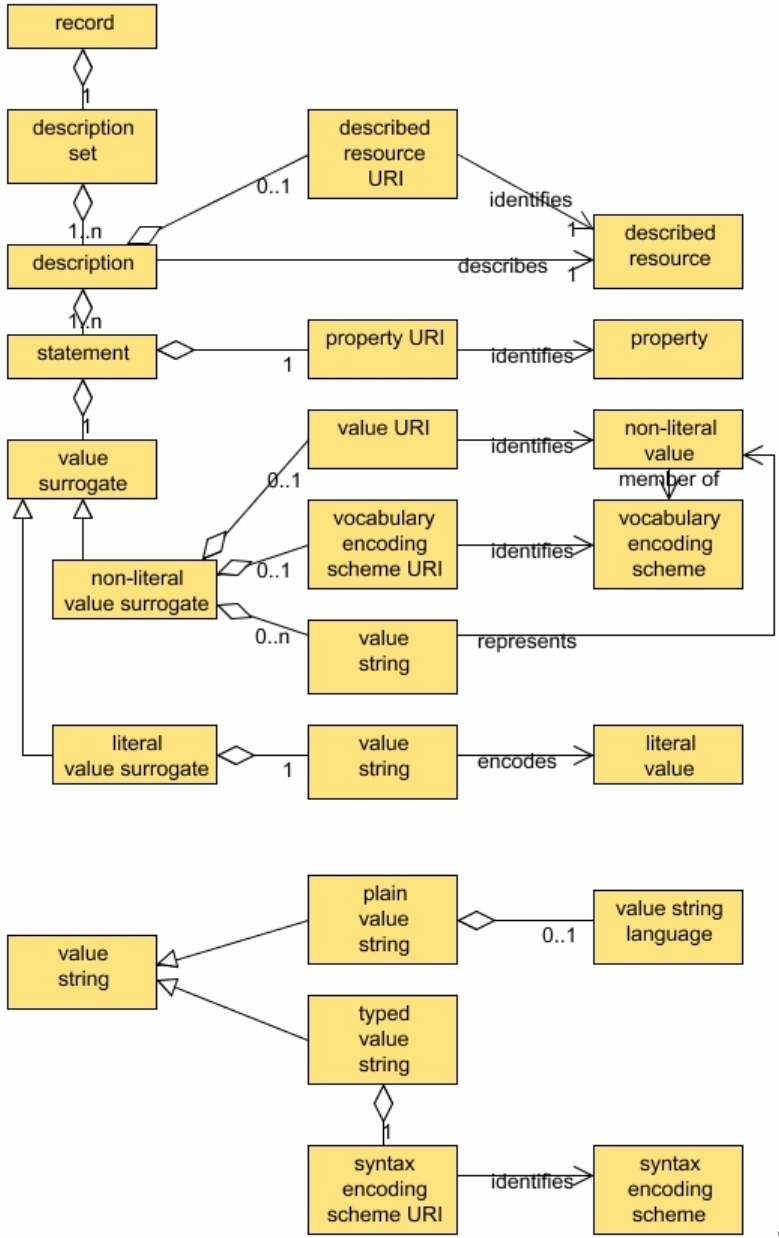
CIDOC Conceptual Reference Model

- ISO 21127:2006
- "… intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework that any cultural heritage information can be mapped to."
- More natural fit for museums than libraries
- Work is ongoing on CIDOC/FRBR harmonization


E1 CRM Entity
E59 Primitive Value
E62 String
E61 Time Primitive
E60 Number
E77 Persistent Item
E39 Actor
E74 Group
E40 Legal Body
E70 Thing
E72 Legal Object
E18 Physical Thing
E19 Physical Object
E20 Biological Object
E21 Person
E22 Man-Made Object
E84 Information Carrier
E24 Physical Man-Made Thing
E78 Collection
E25 Man-Made Feature
E26 Physical Feature
E27 Site
E73 Information Object
E31 Document
E32 Authority Document
E29 Design or Procedure
E33 Linguistic Object
E35 Title
E34 Inscription
E36 Visual Item
E37 Mark
E38 Image
E71 Man-Made Thing
E28 Conceptual Object
E30 Right
E55 Type
E58 Measurement Unit
E56 Language
E57 Material
E51 Contact Point
E45 Address
E44 Place Appellation
E46 Section Definition
E47 Spatial Coordinates
E48 Place Name
E41 Appellation
E82 Actor Appellation
E49 Time Appellation
E50 Date
E75 Conceptual Object Appellation
E42 Object Identifier
E52 Time-Span
E53 Place
E54 Dimension
E2 Temporal Entity
E3 Condition State
E4 Period
E5 Event
E64 End of Existence
E68 Dissolution
E69 Death
E6 Destruction
E81 Transformation
E67 Birth
E63 Beginning of Existence
E66 Formation
E65 Creation
E83 Type Creation
E12 Production
E11 Modification
E80 Part Removal
E79 Part Addition
E9 Move
E7 Activity
E8 Acquisition Event
E10 Transfer of Custody
E13 Attribute Assignment
E16 Measurement
E14 Condition Assessment
E15 Identifier Assignment
E17 Type Assignment

Dublin Core Abstract Model

- Still evolving
- Major shift in thinking for DC community
- Aims to provide a framework to underlie all metadata standards, not just DC


record
1
description set
1..n
description
1..n
statement
1
value surrogate
0..1
0..1
0..n
non-literal value surrogate
literal value surrogate
1
described resource URI
0..1
identifies
1
described resource
describes
1
property URI
1
identifies
property
value URI
identifies
non-literal value
member of
vocabulary encoding scheme URI
identifies
vocabulary encoding scheme
value string
represents
value string
encodes
literal value
plain value string
0..1
value string language
value string
typed value string
1
syntax encoding scheme URI
identifies
syntax encoding scheme

So what now?

- There are no easy answers
- Choose the best tool for the job
- Systems supporting these standards becoming more readily available
- Regardless of what is chosen as the native format, will need to map to other formats as well
- Share your experiences!

Thank you!

- For more information
  - jenlrile@indiana.edu
  - These presentation slides: <http://www.dlib.indiana.edu/~jenlrile/presentations/mlamw2007/mlamw2007.ppt>
  - Presentation handout: <http://www.dlib.indiana.edu/~jenlrile/presentations/mlamw2007/handout.doc>
  - Riley, Jenn and Michelle Dalmau. "The IN Harmony Project: Developing a Flexible Metadata Model for the Description and Discovery of Sheet Music." *The Electronic Library* 25, no. 2 (2007): 132-147.