

CREATING A DIGITAL MUSIC LIBRARY
(DLI2 PROJECT, 2000-2004)

METADATA DELIVERABLES

Mary Wallace Davidson, Faculty Investigator
Harriette Hemmasi, Faculty Investigator

Preamble: There are many definitions of metadata, and the same terminology is often used with different meanings. In the document that follows, “descriptive” means bibliographic data (e.g. a MARC record) describing the object at hand; “administrative” means other data necessary for decision-making (e.g., copyright status); and “structural” means data necessary for navigation (e.g. links).

There are also many standards for metadata available, and much literature about them. The web-site, “Metadata and Music” <http://www.musiclibraryassoc.org/BCC/metadata.html> provides links to several. But other standards are emerging that are particularly appropriate for sound, graphic, and moving image files, and we will be investigating those throughout the grant period.

Academic Year 2000-2001

1. Gather requirements from 1) music librarians, 2) application developers (faculty and computer scientists), and 3) library users (faculty and students in various disciplines – historians, theorists, performers, generalists)

Timeline: January – March, 2001

Activity: Develop one or more set of questions appropriate to these groups, focused on what they would like to be in a digital music library, and how they would use it. Use both written objective questions and interviews, in such a way as to get useful answers.

Result: A white paper summarizing information so gleaned, with an appendix distilling this information in some graphic form (chart).

Deliverable: April 1, 2001

2. Based on these discussions:

a) Analyze MARC records for diverse sample of sound recordings in light of information gathered or anticipated in (1) above; determine which fields from existing US MARC records for these artifacts will be transferred into new schema; enumerate preliminary set of data elements (descriptive, administrative, and structural). Work with technical staff to plan for transfer of data; ask for review and feedback from the same three groups.

Timeline: January 1 – April 30, 2001

Activity: Begin to work on sample collection of artifacts, selected by Collection Development Librarian Michael Fling, to collect and enumerate this data from real examples.

Result: A comprehensive list of elements for which we will enter metadata information, including the potential elements to be transferred from MARC to the selected metadata schema

Deliverable: May 1, 2001

b) Develop metadata architecture; determine types of searches (i.e., by composer, publisher, key, main theme, track number, copyright information, etc.); determine which elements should use standardized vocabulary; identify standardized vocabularies; ask for review and feedback from the same three groups.

Timeline: January 1 – April 30, 2001

Activity: Determine how metadata will be organized and how elements will interrelate; how authority control for those fields that use it (e.g., personal and corporate names, uniform titles, subjects, added authors and titles.) might be established or transferred into the database under conditions where global change will be possible. Consider mapping MARC 240 tag to a combined name/title metadata element.

Result:

Deliverable: May 1, 2001

3. Review and analyze how elements (in 2) are treated in existing and emerging non-MARC schemes for descriptive, administrative, and structural metadata.

Timeline: January 1- April 30, 2001

Activity: Assisted by guidance in L697: Metadata (Professor Elin Jacob), and further consultation with those who have concentrated on music and metadata (sessions at annual meeting of Music Library Association, February, and further consultation with member of advisory board), discover the best existing frameworks for recording data elements and expressing necessary relationships.

Result: [Working notes]

Deliverable: May 15, 2001

4. Make final selection of elements, develop schema for this project, and determine inputting standards appropriate for desired search/retrieval results.

Timeline: May 15 – June 30, 2001

Activity: Chart source and style of original schemes from which we are selecting elements, to express schema as developed.

Result: Rough metadata schema for Developers ready to test.

Deliverable: July 15, 2001

Academic year 2001-2002

5. Review use of standardized form of personal and corporate names, uniform titles, and subjects, based on Library of Congress Name Authority File and Library of Congress Subject

Headings; investigate whether and how other data elements might be standardized (e.g. form/genre, instruments, publishers).

Timeline: August 1 – September 30, 2001

Activity: Review currently available systems to determine whether the manner in which they handle these data elements may be applicable to this project.

Deliverable: October 15, 2001

6. Develop preliminary thesaurus structure for selected descriptors.

Timeline: November 1, 2001 – May 1, 2002

Activity: Define descriptors selected in (2), (4), and (5) above. Ask for review and feedback on preliminary hierarchies from music librarians, Developers, and library users.

Result: Compiled representation of thesaurus structure chosen (paper or web)

Deliverable: May 15, 2002

7. Work with software developers to establish user interface search and retrieval screens, including auto-index feature, display of LC authority records, thesaurus records, lead-in vocabulary, and syndetic structures when available.

Timeline: June 1 – November 1, 2002

Activity: Working now with selected records from the Variations project, demonstrate: 1) transfer of MARC data into new schema; 2) indexing schemes; 3) hyperlinks; 4) search and retrieval screens; 6) use of standardized names, titles, descriptors.

Result: Sample database ready to test search and retrieval, that would meet needs of other investigators, including usability testing.

Deliverable: November 15, 2002

N.B. Reviewing, testing, and refining results will continue throughout the project.