Malleable Metadata Models: Representation and Inclusion of Diverse Musical Content
Music Library Association Annual Meeting 2019, St. Louis, MO
February 22, 2019

Summary written by Kristi Bergland, University of Minnesota
Session presenters: Caitlin Hunter, Library of Congress; Nancy Lorimer, Stanford University

In this session, Caitlin Hunter discussed the development of the BIBFRAME A/V Modeling Study and the issues and needs considered in its preparation. Nancy Lorimer described the development of the Performed Music Ontology, an extension of BIBFRAME for performed music.

Caitlin Hunter is Head of the Recorded Sound Section of the Library of Congress. She began her presentation by describing the genesis of the BIBFRAME A/V Modeling Study, addressing how the study happened to take place when it did and why it focused on A/V materials. The Library of Congress National Recording Preservation Plan, published in 2012, highlighted problems scholars using sound recordings in their research have with locating relevant materials. The report indicated that discovery issues affected all types of sound recordings and described the problem as "structural and stems in part from cataloging practices that do not sufficiently address issues of access and preservation that institutions might resolve through a collaborative effort." Searches for some entities and concepts in A/V materials are much better supported by existing data structures than others. The existing structure works well for searching for specific publications, but searches for specific tracks may result in "false hits", which can only be identified when the user has sifted through most of the record, because while the data points may be present, the information defining the relationships between the data points may not exist.

The other major factor in the timing of the BIBFRAME A/V Modeling Study was the message sent out in early 2013 about bibframe.org and a new service transforming existing MARC data into BIBFRAME RDF. Staff at the Packard Campus for Audio-Visual Conservation tested some audio and moving image MARC records and observed that the organization of data in BIBFRAME did not improve the discovery issues for A/V resources, but that linked data had great potential to improve A/V discovery needs if those specific needs could be identified and addressed. Meetings were held on the Packard Campus to bring the BIBFRAME A/V discovery challenges to a broader audience, and there was a consensus that there was more work to be done and that the broader library community should know these conversations were taking place. AVPreserve was hired to propose a content model and recommend potential changes to BIBFRAME. The report was released just prior to ALA in summer of 2014. Due to time constraints, not all specs for data elements and attributes for audio and moving image were included in the report, but some was later shared as part of the Linked Data for Production Performed Music Ontology project.

Among the issues considered in the development of the modeling study were the characteristics of A/V materials, the variety of needs of different communities, differences between published

and unpublished content, and varying workflows and needs at different contributing institutions. Acknowledging that many of these characteristics are not unique to A/V materials, but very important to A/V materials, the study identified the time-based nature of audio and moving image content, the multiple creators and contributors involved, the perception of recordings as events, the uniqueness of many unpublished recordings, the aggregations, collections and multiple generations of reformatted content as issues common to A/V and moving image materials. Types of content include Western art/classical music, popular music, traditional/ethnic music and events, mash-ups/remixes, oral histories, interviews, live events, natural and environmental sounds, audiobooks, training and educational materials, radio, television, feature films, production elements, and sounds effects, which draw from different communities with different needs and cultural conventions. Appendix A of the report explores some of the issues in recorded sound related to music, including methods of transmission (oral vs. text-based), performing conventions, definition and degree of emphasis on concepts such as "work" and data elements, and the degree of familiarity any given cataloger has with a particular type of music or culture. Appendix B examines descriptive practices used in discographies, common elements and differences needing to be addressed, including the inconsistency in describing "take numbers". The study also considered the processes for describing published and unpublished materials, the degree to which it is influenced by workflows and institutional needs, and whether one data structure could support the range of options. Institutional workflows are also influenced by limitations of various systems used, the ability to map data between systems, digitization needs, and expertise of staff.

The BIBFRAME A/V Modeling Study data model, referred to as the "Content Creation Domain Model", shares the "Instance" concept with the BIBFRAME data model, but where BIBFRAME uses the "Work" concept, the A/V model expands that to include the "Event" concept. Work and Event together are described as "Content". Events can represent the realization of a Work or can exist without an underlying intellectual work. Three examples from the study illustrate the model applied to various types of recorded sound. The first example, a recording of bird song, has no underlying intellectual work, but could contain extensive detail about the event of the recording. The second example, a recording of a Beethoven symphony, represents both an underlying intellectual work and an event. The final example, a performance by a cultural group in Indonesia, shows the cultural group as performers as well as members of a cultural group associated with a particular type of event.

Nancy Lorimer is Head of the Metadata Department at Stanford University. As participants in the Linked Data for Production grant, Stanford worked on a sub-project to develop an extension to BIBFRAME to use as a standard to accommodate the needs of different kinds of music and music users. The product of this effort is the Performed Music Ontology. MARC, FRBR, and BIBFRAME 1.0 lack the ability to describe recorded music accurately and effectively. Particularly problematic are the concepts that everything is a "work" and that works have a "creator" and (secondarily) "contributors". A review of MARC format bibliographic records for performed works demonstrate that MARC doesn't handle sound and video recordings particularly well, preferring title/creator as access points over performers, when users may want to use performer information to distinguish a specific recorded event. Even within Western

music, one approach to recording details of performed music does not fit all. Recording producers and engineers, production details, sources of sung text, and improvisors are examples of recording related data that is not well served by the existing approach. Another question raised was what to make of the idea that everything has to be a work. Is a ceremony considered a work, and if it isn't, what does that mean?

Lorimer introduced BIBFRAME 2.0 and describes the Work-Instance-Item structure. The Performed Music Ontology (PMO) has been developed to extend vocabularies and relationships to more fully describe data elements of performed music in a way that will be more useful, accurate, and flexible than current cataloging practice. The PMO adds Performer, Event, and Medium of Performance to Work as central concepts in describing performed music. Subclasses of Work related to music include Audio, Notated Music, and Moving Image. Through adding a Work layer, different recordings of the same opera (for example) can be linked. That model can be extended to group recordings in the original language and separate out performances in translation. Using subclasses and added properties of Event to describe performer and production details can be particularly useful for recordings without a preconceived work, Lorimer went on to describe how a recorded performance can be a recording of a performed Work, or without a pre-existing work can be simply a recording of a performance. Events can be extended, allowing them to be linked to information beyond the recording itself, including cultural and historical links. Contributors can be related to Works, Events, Instances, etc. and PMO allows for them to be described with more specific relationships. Lorimer wrapped up with a demonstration of the wireframe version of the soon-to-be-launched BIBFRAME editor Sinopia.

Questions? Contact chun@loc.gov or nlorimer@stanford.edu.