Report of the ALCTS Cataloging and Metadata Management Section (CaMMS) Catalog Management Interest Group Meeting, ALCTS Virtual Interest Group Week, June 2020

Marina Morgan
Florida Southern College, Lakeland, FL, USA
mmorgan@flsouthern.edu
http://orcid.org/0000-0003-1862-7226

Dan Tam Do
University of Pittsburgh, Pittsburgh, PA, USA
dad193@pitt.edu
https://orcid.org/0000-0003-1862-7226

The 2020 ALA Annual Conference & Exhibition scheduled for June 25-30, 2020, in Chicago, IL, was canceled. In lieu of face-to-face meetings at the Annual Conference, ALCTS held a Virtual Interest Group Week, June 8-12, 2020. The ALCTS CaMMS Catalog Management Interest Group meeting during the ALCTS Virtual Interest Group Week consisted of two presentations, entitled "Breaking Rules and Preserving Legacies for Digital Access" and "Improving Local Holdings Records." A general discussion and questions from the audience followed the presentations. 513 registered for the virtual event and approximately 265 attended the live presentations.

Breaking Rules and Preserving Legacies for Digital Access

The first presentation, entitled "Breaking Rules and Preserving Legacies for Digital Access," was given by Michelle Hahn, Assistant Librarian, Sound Recordings Cataloger at the William and Gayle Cook Music Library, Indiana University Bloomington. She is also past Chair of the Music OCLC Users Group and the Assistant Web Editor for the International Association of Music Libraries, Archives and Documentation Centres. Hahn's presentation focused on the Indiana University School of Music's re-cataloging practices for 30,000 unique and rare sound recordings, to allow for the world-renowned School of Music's legacies to be more discoverable and accessible to a wider audience.

Hahn started the presentation with a quote from George M. Logan,[1] who wrote about the Indiana University School of Music as being recognized as the best music school in North America since the 1960s and one of the five largest music schools in the world in the same era. 2020 is an important year for Indiana University as it celebrates its bicentennial, while the School of Music will celebrate its centennial next year, in 2021. Hahn continued the institutional history, highlighting notable individuals and performance legacies, and noted that the school now presents more than a thousand performances each year, out of which more than four

hundred are preserved as digital recordings, which become part of the music library's collection. This vast collection of about 30,000 performance recordings includes performance recordings dating from the 1940s, most of which are now preserved digitally.

Hahn set the stage by looking back at the origins of the Variations project at Indiana University, started 25 years prior. She highlighted chronologically all the stages of building a digital library focused on sound recordings using the emerging technologies available starting in the early 1990s. In 1996, the first iteration was called Variations. Variations3, an open source version, preceded the Avalon Media System, introduced in 2013.

A leader in digital access and preservation of audio content, Indiana University holds over 560,000 audio and audiovisual materials. Most of the rare, unique, and volatile items have now been digitized through the Media Digitization and Preservation Initiative (MDPI), which originally intended to digitize 280,000 recordings; that number, for just the first stage, is now up to 320,000. Within this initiative, the Music Library is responsible for around 50,000 items, with another 150,000 showing potential, but this presentation is focused on 30,000 of them.

Hahn continued by mapping the path of the library's collection from the card catalog era to the present day. While the retrospective conversion of the print collection was completed, the AV collection still has tens of thousands of cards remaining. Moreover, with the music cataloging changes that occurred over the years, especially in the 26X (Publication, Distribution, etc.) and 30X (Physical Description) MARC fields, records with old and new data coexisted in the ILS.

The local cataloging practice was the next focus of the presentation. As MARC and cataloging standards for performance recordings evolved through the decades, so did local cataloging decisions. For the 30,000 digitized recorded performances collected in the Music Library, much descriptive data was generic or repetitive, including the publication place and entity, performance dates and locations, and the names of performers, who give multiple performances. While the student and faculty performances were unique, the music performed generally was not. Consequently, certain rules and guidelines were adapted to include local practices. One example is the changing approach over the years to the title, which is expected to be transcribed as it appears on the primary source of information. Titles were standardized to include the academic year and a numbering system used by the School of Music; some titles are in square brackets, to indicate that they were supplied by the cataloger, while others are not. Subsequently, to collocate like performances, the title was the performance type. The performer, their instrument and the date of performance were then added to distinguish between performances by the same person during the same degree program. As Hahn argued, the name should be the title proper, the type should be the other title information, and the date should not be included. However, breaking the rules allowed for continued collocation by performance type, descriptive information in the title and thus meaningful distinction between search results, and an added prominence to the performer, a legacy, now or in the future, of the School of Music.

Another example Hahn discusses is publishing ownership; the publisher name had been applied inconsistently or left out altogether. Currently, it matches the appropriate corporate body for the

time in which the recording was made. Main entries were also an issue: some records had no main entry, while others had the performer or the School of Music as the creator. Now, the principal performer is listed in the 511 MARC field, in the 505 field in some situations, and generally as an access point in a 700 field as well.

Hahn noted that 75% of Indiana University performance recordings have data about them only on 3" x 5" cards -- which leave out some information due to space limitations -- or are missing data.

With the addition of digital access to 30,000 performance recordings, she undertook a project to re-catalog them all. The goals of the project are standardization and uniformity, with consideration for future automation. Records for new recordings have benefited from no longer being subject to the prior data space restrictions and from guidelines and current standards that allow for more descriptive and identifying data. For information that is repetitive, a template in OCLC Connexion is used to add constant data, and text strings are also available. The template includes instructions for the standard data to be applied to the old data of existing records in WorldCat.

Hahn continued the presentation by detailing the responsibilities of the digital access lab and the MDPI, as well as her own responsibilities, for ensuring a smooth project workflow. In her role, Hahn verifies that steps are completed within the streaming platform and then re-catalogs the content, ensuring accuracy and completeness. She applies appropriate constant data and text strings for the physical and digital objects and makes updates and corrections to the records as necessary. In doing so, she bends more rules, adding descriptive data for the digital surrogate to the record for the physical object -- taking advantage of the MARC subfield 3 -- in order to facilitate user discovery and access. Provider neutral guidelines are finessed when persistent URLs for the stream are added to the WorldCat records, since Indiana University is the only provider there is or ever will be. With every URL, Hahn has added a statement on availability to authorized users. Authorizations expand organically as content moves to the public domain, so the statement will not need to be revised. Additionally, Hahn imports the descriptive data into the streaming platform through an automated MARC to MODS crosswalk.

The last part of the presentation focused on the importance of detailed documentation, from providing complete directions and using visual cues, to laying out a clear and hierarchical structure. The goal of any proper documentation is to allow someone to replicate successfully and without any errors all the steps of the project. At the end of the presentation, Hahn shared two examples of WorldCat and IUCAT records, workflow documentation for the project, as well as the publications that detail the MDPI (Media Digitization and Preservation Initiative) timeline.

Slides from this presentation are available on ALA Connect.

Improving Local Holdings Records

The second presentation, entitled "Improving Local Holdings Records," was given by Amy Wood, Head of Technical Services, Center for Research Libraries (CRL). Wood described a project that will support resource sharing for CRL and its members over the long term, that has helped strengthen relationships between departments and develop staff skills, and that has provided remote work for employees operating under restrictions related to Covid-19.

The presentation opened with an overview of CRL, a Chicago-based international consortium of over 200 academic and research libraries that supports research and teaching in the humanities, science, and social sciences and preserves and provides access to rare and uncommon primary source material. Founded in 1949 by 10 libraries in the Midwest, CRL's initial deposits included trade journals, medical journals, and government documents. As the serials collection grew through acquisitions programs, rarely held material continued to be an area of focus.

Access to the collection has always been an important part of the CRL mission. Early on, every member received copies of catalog cards. Later, CRL joined OCLC and began contributing to the National Union Catalog. In 1988, CRL started providing tape loads of its records to members. Accurate records indicating what CRL held became increasingly important as its membership and collection grew.

Since it began contributing serial holdings to OCLC, CRL has maintained two types of holdings records: one followed national standards in WorldCat, while the other followed local policies in CRL's catalog. Holdings in WorldCat used the 866 field, whereas in CRL's catalog, holdings information was entered in 590. While the WorldCat records were more consistently formatted, local policies meant that CRL catalog records often had more detailed holdings information. However, inconsistent formatting was an obstacle to making the data structured and sharing it with other systems for use in resource sharing or collection analysis. CRL has been involved in past decades with national shared print efforts to improve serial holdings data that support collective decision making about collections, which made clearer the need to improve its data and inspired the local holdings improvement project.

The goal of the project was to create detailed and structured 863 fields for all CRL serial holdings. The scope of the project covered print serial holdings that have 866 fields, which numbered around 24,000 records, or somewhat under half of CRL's approximately 50,000 total print holdings records.

The original plan called for cataloging staff to compare, side by side, holdings in a catalog record with the local holdings record (LHR) in WorldCat and to make enhancements to the LHR directly in Connexion. If there was not enough information available for the 866 fields to be expanded into 863 fields, colleagues in stacks would check the physical material. However, it became clear that the side by side approach was neither efficient nor error-proof.

The second plan began with a cataloger pulling LHRs from WorldCat by creating a query collection in WorldShare Collection Manager. The records were sent to Backstage Library

Works, which provides numerous services to libraries, museums, and archives, to process the 866 fields into 863 fields. Cataloging staff would check the records and mark for updating in WorldCat those that they found acceptable. Whenever records did not contain enough information for 866 fields to be translated into 863 fields, colleagues in stacks would check the physical material and write the updates on printouts, which catalogers would then use to correct the LHRs in WorldCat. It was Backstage's idea to insert "TBD" anywhere in a record that information was not available.

Covid-19 and the need to transition quickly to remote work made it impossible to continue with this plan, so a third plan was developed to accommodate the need of the stacks staff for a project that would not require in-building handling of physical material. With the pull of LHRs from WorldCat and processing by Backstage already complete, what needed to be planned included distribution of the work, staff training, controlling the flow of records, quality assurance, and how to work as a team while physically separate.

The team came to include seven staff from the stacks unit and four from technical services: the expert cataloger, the records manager, and two additional staff who review records. The presenter became the project coordinator, and her initially large role grew smaller as the team's expertise increased and leadership emerged among the stacks supervisor, expert cataloger, and records manager. At the start of the project, there was much anxiety, frustration, and confusion. Most stacks staff had not received much, if any, cataloging training, and none had received LHR training. However, team members did possess an understanding of serials as well as deep knowledge of CRL's serials collection, and the stacks supervisor has been instrumental in keeping up staff morale.

Tools and resources supporting the project include Microsoft Teams, used for meetings and calls and as a place to place documentation and other files and to post questions, the Library of Congress' manual on MARC 21 holdings data, and documentation from a member library for their local policies.

Formal LHR training, which might have overwhelmed the staff, was eschewed in favor of a "learning by doing" approach. An overview of the project and process was provided, and staff were encouraged to read the documentation and try to apply what they learned to records. For the first two months of the project, question-and-answer sessions up to an hour in length were held daily in Microsoft Teams. Team members raised questions and used screen sharing to examine specific records as a group. The frequency of the Q&A sessions fell to three times a week as skills improved and questions became fewer in number.

The presenter noted that not all LHRs could be updated without examination of the physical material, so she often had to make decisions about when enough time and effort had been expended on attempts to process a record without access to the physical resource.
The workflow is as follows. The Backstage file with 24,000 records was split into files of 50 records, which were loaded to Microsoft Teams. Each member of the CRL team can claim a file by modifying the file name, and upon completion of work on a file, it is saved in a specified

folder. Once records are checked for errors, they are sorted into categories so that it is clear what action needs to be taken with them later. Records are ready for WorldCat, with either a new or reconciled LHR, or they need additional review for cataloging or physical validation reasons.

Using snippets of MARC records, the presenter showed examples of a CRL LHR in WorldCat, a record that has gone through processing by Backstage, and a record after the CRL team has reviewed and updated it. She also showed a checklist used by the team when reviewing records, noting how much team members have encouraged each other's learning and high-quality work.

As of presentation time, between about 2500 and 3000 records had been processed, with around 1200 of them fully checked and ready either for validation against the physical material, whenever work at CRL can resume, or for contribution back to WorldCat. Only around a quarter have required physical validation; although that proportion may change as the project progresses further, awareness of the figure has helped with evaluating the project's costs and benefits.

One final positive effect noted by the presenter is on catalog maintenance: examining records individually has revealed problems that would not have been found using other methods such as general search or batch processes.

Slides from this presentation are available on ALA Connect.

Q&A

After the presentations, there was time for the presenters to answer questions from the virtual attendees. The first question, addressed to Wood, was with regard to pulling LHRs from WorldCat and whether she sent an API request. Wood explained that she did not send an API request but rather used the query collection functionality in WorldShare Collection Manager; her last slide includes a link to an OCLC video that explains the steps.

The second question, also addressed to Wood, concerned the number of staff and hours per week spent working on the project. Wood explained that including herself, there were a total of 12 individuals working a range of hours on the project.

An attendee then asked Wood about the decision to add 863 fields and whether it was related to the OPAC display, noting that some discovery systems display item records, making an 863 for each item unnecessary. Wood replied that the decision to translate the 866 fields to 863 was based first on considerations related to future catalog migration and to the demands of record sharing (e.g., work with Rapid) and, second, because the 863 presents fewest problems for the scale of record manipulation that is done to compare collections for shared print programs. The attendee asked whether Wood would recommend that member institutions add 863 fields in

their LHRs, and Wood replied that she thinks it is an investment in the future and that the more detailed and exact holdings data are, the better they can be shared with other institutions.

The next question was for Hahn and concerned what she has been using as a primary source of information if there is no printed program available. Hahn responded that they have printed programs in multiple places, including a folded copy in the container; recordings without one tend to be donor collections with information written on the container (often in handwriting that is difficult to read). If there is no printed program, the chief source is a label on the medium itself.

The question that followed concerned institutional support to finish the rest of the retrospective conversion, to which Hahn replied that it would take money to continue the project.

Hahn was then asked about the statement she had made about access increasing as things come into the public domain: was she referring to the entire performance or the individual works? Hahn replied that it depends on factors such as the university's decision, the performer's policies, which can change over time, and copyright licenses. She gave the example of the famous violinist Joshua Bell, an Indiana University alumnus, who worked with the School of Music to provide worldwide access to some of the performances he has given at Indiana University.

Reference

[1]Logan, G. M. (2000). The Indiana University School of Music: A History. Bloomington, Indiana: Indiana University Press.