Exploring a Faceted World: Discovering Music Resources Using Medium of Performance and Genre Terms
MLA Annual Meeting 2015, Denver, Colorado
Summary written by Charles Peters, Indiana University
Session presenters: Beth Iseminger, Harvard University; Casey Mullin, Stanford University; Hermine Vermeij, UCLA; Kevin Kishimoto, University of Chicago

This session was sponsored by the MLA-BCC Subject Access Subcommittee and Genre/Form Task Force. The presentation included a discussion of ongoing developments in music vocabularies which could make finding and selecting music resources with online library catalogs easier and more efficient for users and public service staff. A summary follows; the full session can be viewed here:
http://www.musiclibraryassoc.org/page/mla_2015_media

Catalog searching / Kevin Kishimoto
In library catalogs there are four general techniques used in searching for resources: browse, keyword single box, advanced keyword, and facets. The techniques complement one another in that they allow the user to approach the search from a variety of perspectives, using a single one or combining them, depending on what information is already known and what resources are sought.

Browse: The browse search has been around the longest and it was used in card catalogs. There are separate lists for different aspects: author, title, subject, or call number. Browsing is very efficient for known item searches, or when the user knows the exact phrase found in the catalog, the exact title of the resource, the exact form of the author or composer's name access point, the exact subject access point. While the browse search is still very effective and not obsolete, it seems to be less and less common in library catalogs and discovery tools. The University of Chicago catalog includes a browse search, and is labeled "Begins with." The user includes a search term and selects an index to browse ("Subject"). The catalog sends the user to the place within the alphabetical list where the term is found.

Keyword single box search: This is a popular search with patrons because of the single box. The search returns bibliographic records that include most or all of the search terms. This search is simple for users to understand and is a good one to use for exploratory searching. The keyword search also allows command line searches constructed with Boolean operators (AND, OR, NOT, etc.). Google is probably the most well-known user of the single search box. Google searches are not exactly keyword searches; however, when library users see a single search box they sometimes expect the search to behave in the same manner as when they use Google, drawing on vast worldwide sources of information.

Advanced keyword search: Using this search, it is possible to conduct multiple types of keyword searches simultaneously. Each search box can perform a different type of search, and Boolean searches can be constructed. Pre-search limits can be assigned. However, this type of search is found to be intimidating to some users.

Facets: Each facet describes a distinct aspect, such as format, author, subject or language. Facets are applied post-search in order to narrow the search results. They are common on commercial websites, and are included in newer discovery layers used by libraries.

Introduction to Music Genre and Medium Vocabularies / Beth Iseminger
Work on the music Library of Congress Genre/Form Thesaurus (LCGFT) was begun in 2009. The BCC Subject Access Subcommittee began work in 2011 on the medium of performance part of the project.

These are the definitions used by those working on the projects:
- Genre: a class, type, or category, sanctioned by convention;
- Form: the constructive or organizing element in music;
- Medium of performance: the voices, instruments, and other entities necessary to perform a piece.

Faceting terms is part of the process of establishing the vocabularies. Benefits of facets include higher indexing precision achieved when using the medium as a unique facet, rather than in a combined medium/genre string. Facets will also be necessary in the current and future linked data environment.

LCGFT includes terms for several subject areas including music, moving images, law, cartography, literature and religion. The process of coming up with the music terms began with an examination of existing LCSH music headings, followed by adding missing terms to the list. New terms were added from reference sources, including New Grove, 2nd ed., and the Garland Encyclopedia of World Music. Also consulted were the IAML code list, the RILM thesaurus, and the All Music Guide, among others. The list of terms was large, resulting in the decision to limit terms primarily to those in LCSH. Literary warrant (the idea that terms should apply to actual resources) was also a factor.

The vocabularies for both LCGFT and LCMPT are available at id.loc.gov. Music genre terms are also available in Connexion. New terms may be submitted through the SACO music funnel. Iseminger stressed that using music genre terms in cataloging is now recommended.

New Music Vocabularies / Hermine Vermeij
Vermeij explained the hierarchies of the two new music vocabularies. In the first release, 567 genre/form music terms were added. The structure is a poly-hierarchy and is a true thesaurus. This means that each term has at least one broader term, sometimes more. One example would be Folk songs, with broader terms of Songs and Folk music.

At the top of the hierarchy is the broadest term, "Music." Art music, Folk music and Popular music are three broad terms that are found one step down in the hierarchy, and many other terms are found under these three. World music is another term being defined for inclusion as a second-tier term. There's still an effort to "figure out what World music is exactly." Functional music is a new category and includes music such as ritual music,

music for holidays, and dance music.

LCMPT was first available in February of 2014, when over 800 terms were released. The top terms are Ensemble, Performer, and Visuals. Probably the largest set of terms in the thesaurus is under Instrument.

An example of all the steps in a hierarchy of terms, narrowest to broadest, is: alto saxophone; saxophone; single reed instrument; reed instrument; woodwind instrument; wind instrument; instrument; performer.

Vermeij explained that in bibliographic records, old and new practices will exist side-by-side for a while: traditional subject headings along with the newer medium of performance terms and genre/form terms will be found in the same records.

Music Vocabularies in Library Catalogs / Kevin Kishimoto
Kishimoto pointed out that many library catalogs are not yet set up to use LCGFT and LCMPT to their full potential. The development of the library system follows after the development of the data, simply because programmers need actual data to use during development. It is necessary to develop the library system in four different areas: display, indexing, facets, and search.

Genre/form display could ideally be a separate category from subject headings, and listed apart from the subject headings. It might be useful to locate the two groups near each other, yet label them clearly as "Genre" and "Subject", as one example.

It might be more difficult to arrange a display of medium of performance terms. First, the MARC coding for these terms is more complex than that for genre/form terms. Then there is the question of where to locate the terms, identified to the user by an easily understandable label. Some examples of user-friendly labels were Medium of performance; Instrumentation; Performance medium; and Performing forces.

For examples of bibliographic displays of genre/form and/or medium of performance terms, see the Stanford University catalog (http://searchworks.stanford.edu/), the University of Chicago catalog (https://catalog.lib.uchicago.edu/vufind/), or the UCLA catalog (http://catalog.library.ucla.edu/vwebv/searchBasic).

For the genre/form terms to be functional in the catalog, the MARC 655 field needs to be indexed. Likewise, for medium of performance terms the MARC 382 field needs to be indexed.

For facets to function as desired, the catalog would ideally have separate facets for each of the new vocabularies rather than just one "subject" facet, including: genre/form; medium of performance; and demographic groups including "creator" and "audience." The facets would generally be found in a sidebar, used to narrow a search. MARC fields and subfields would need to be indexed.

At this time there are very few library catalogs that have a separate genre/form search, probably fewer that can search medium of performance terms, although some of these terms are included in general keyword searches. The University of North Texas and Ball State University both have medium of performance searches available that have been constructed from older MARC data (the MARC 048 field).

A future search might allow users to navigate through the hierarchies of genre/form or medium of performance terms, narrowing the terms until the desired term is reached. This type of search capability might arrive as an add-on, rather than a built-in function of the catalog.

Technical Issues / Casey Mullin
Mullin noted that it is desirable to pull data from "old" LCSH headings with some type of machine conversion, because the process would be too lengthy if done by hand. The idea is that each LCSH form/genre heading is likely to produce at least one LCMPT and or one LCGFT field. It's necessary to develop a program that will work equally well with score and sound recording records. The generated terms should be real terms in the thesauri, and should conform to MLA best practices. Duplicate fields should be removed. Since any such program will be imperfect, it's important to keep the current data, so existing LCSH will be retained for the immediate future. Also, the subject headings contain many other facets belonging to vocabularies that haven't been finalized yet, such as demographic groups, dates, and geographic areas.

There are challenges to such a conversion process. For example, an implied medium of performance, such as "symphonies," is immediately understood by a person, but may not be by a machine. Differences in vocabulary could cause problems, such as "mixed voices" in LCSH, but "mixed chorus" in LCMPT. Some LCSH terms have been broken out into multiple new terms: "canons, fugues, etc." has become "canons" or "fugues." Some terms are completely new, such as "art music." The identification of solo performers is challenging since the designation would be different in the new data. Some existing subject headings are truly topical, and so shouldn't be converted. Video recordings present a special challenge, since they could be the music itself, or a documentary about the music, for example.

When the medium of performance terms have been broken out from existing LCSH headings, presumably only topical headings will remain. Discovery systems will need time to catch up with the function of new fields and vocabularies.

Conclusion / Kevin Kishimoto
Kishimoto reiterated that new music vocabularies have already been created, for genre/form and medium of performance terms. A thesaurus for demographic terms is expected soon. Music catalogers have already begun to use these new vocabularies. But discovery tools lag behind in functionality because programmers need the data to exist first. Conversations need to begin between catalogers, user services, ILS departments, vendors and any other decision makers in libraries.