The Detail Behind Web-Scale:

Selecting and Configuring Web-Scale Discovery Tools to Meet Music Information Retrieval Needs

Nara L. Newcomer

Abstract:
Web-scale discovery tools are rapidly gaining popularity as a purported "one-stop search" for discovering library information. Music, particularly printed music and recordings, presents unique information retrieval needs. This article identifies, explores, and makes recommendations regarding key music-related aspects to consider when selecting and implementing a discovery tool, considering scope, metadata, and interface.

Introduction

Web-scale discovery tools are rapidly gaining popularity as an option for discovering library information. Products currently on the market include EBSCO Discovery Service, Encore Synergy (Innovative Interfaces), Primo Central (Ex Libris), Summon (Serials Solutions), and WorldCat Local (OCLC). These tools are often presented as the holy grail of a one-stop search for library information, particularly because they combine traditional "catalog" data with article-level indexing, employ a central index to provide superior speed and relevance ranking, and utilize a faceted interface to aid users in exploring results. With all the hype, the specialized information retrieval needs of music materials can easily be lost in the shuffle. This article will identify and examine special considerations music materials pose in discovery tool selection and implementation, providing recommendations regarding optimizing discovery tools for music information retrieval.

Literature Review

The literature on web-scale discovery and music materials is sparse due to the emerging nature of web-scale discovery. Two papers, while not scholarly studies, evaluate individual discovery tools for music use. The Music OCLC Users' Group Reference Services Subcommittee's "WorldCat Local Enhancement Recommendations for Music" calls for specific changes to WorldCat Local, highlights display issues and uniform-title related needs specific to music materials, and states that music researchers' needs regarding faceting differ from those of general searchers.[1] The author's review of Summon for music materials indirectly identifies some areas to consider in discovery tool evaluation and highlights the need for music librarians to understand the inner workings of discovery tools in order to advocate for music information retrieval needs.[2]

Web-scale discovery tools have many similarities to any tool with a faceted interface. Faceted searching, popularized by "next-generation" library catalogs like AquaBrowser, Endeca, and Blacklight, helps users navigate, refine, and sort their results via facets relevant to their results set. Facets commonly created from library data include author, subject, format, language, and publication date. Only one study exists of music needs in a faceted catalog. Snyder examined the needs of music researchers in relation to the University of Chicago's AquaBrowser implementation, particularly noting music researchers' confusion with facets for genre, geographical region, and time period, as well as highlighting problems created by poor handling of uniform titles, an access point informed music researchers wanted to use.[3]

Given the sparse nature of the literature on web-scale discovery and music materials, it is worth examining papers which consider only half of the present topic, that is, only web-scale discovery

---

[1] Music OCLC Users' Group Reference Services Subcommittee, "WorldCat Local Enhancement Recommendations for Music," (April 2010), accessed Mar. 22, 2011 http://www.musicoclcusers.org/WorldCatLocal20100412.pdf

[2] Nara L. Newcomer, "Electronic Resources Review Column: Summon," *Music Reference Services Quarterly* 14/1&2 (2011): 59-62.

[3] Tracey Snyder, "Music Materials in a Faceted Catalog: Interviews with Faculty and Graduate Students," *Music Reference Services Quarterly* 13/3&4 (July 2010): 66-95.

or only music information retrieval. Articles related to web-scale discovery in general, primarily evaluations of specific discovery tools, have begun to appear in the literature. Rowe's review of three web-scale discovery tools (Summon, EBSCO Discovery Service, and WorldCat Local) also identifies some areas to consider when selecting a discovery tool.[4] An analysis of usage statistics at Grand Valley State University found Summon was widely adopted by patrons in its first semester of use and increased access to and usage of library resources.[5]

Again broadening the scope to faceted catalogs in general, two studies conducted on next-generation catalogs (AquaBrowser and Endeca) found user behavior differed between faceted catalogs and traditional OPACS. Olson studied twelve users conducting dissertation level research in the humanities and discovered that a faceted search helped these users find materials they had not previously located.[6] Examination of usage log statistics at New York Law School revealed that discovery layer facets and tag clouds were employed fifteen times more frequently than traditional OPAC limiters.[7]

Music information retrieval needs are also represented in the literature. In 2007, King summarized and synthesized the research to that date regarding user search strategies for printed music and recordings, concluding that music searching differs in specific ways from general searching, particularly in that known-item searching is more common for music materials and personal name is used more frequently as an access point.[8] Similarly, a study of music reference service found that, in comparison to general reference, music questions "focus much more heavily on locating specific, known items" and were answered via different means, with music reference employees relying on "their local catalogs, the circulating collection and their own knowledge much more frequently than librarians in general reference situations."[9]

This literature review shows that web-scale discovery and faceted searching represent a significant change in the way users can encounter library information. Simultaneously music, particularly printed music and recordings, presents unique information retrieval needs. As institutions increasingly implement web-scale discovery tools, music librarians must guide their institutions in discovery tool selection and advocate for discovery tool optimization to meet music information retrieval needs, whether through local configuration or conversations with vendors. At the time of this article's publication, there exists no guide for music librarians seeking to aid their institutions in this way. The present article seeks to fill that gap, identifying and exploring key music-related aspects to consider when selecting and implementing a discovery tool. This article examines scope, metadata, and interface, focusing on the challenges

---

[4] Rhonda Rowe, "Web-Scale Discovery: A Review of Summon, EBSCO Discovery Service, and WorldCat Local," *The Charleston Advisor* 12/1 (July 2010): 5-10.

[5] Doug Way, "The Impact of Web-Scale Discovery on the Use of a Library Collection," *Serials Review* 36/4 (Dec. 2010): 214-20.

[6] Tod A. Olson, "Utility of a Faceted Catalog for Scholarly Research," *Library Hi Tech* 25/4 (2007): 550-561.

[7] Terry Ballard, "Comparison of User Search Behaviors with Classic Online Catalogs and Discovery Platforms," *The Charleston Advisor* 12/3 (Jan. 2011): 65-66.

[8] David M. King, "Catalog User Search Strategies in Finding Music Materials," *Music Reference Services Quarterly* 9:4 (2007): 1-24.

[9] Beth Christensen Mary Dumont, and Alan Green, "Taking Note: Assessing the Performance of Reference Service in Academic Music Libraries: A Progress Report." *Notes* 58/1 (Sept. 2001): 52.

presented by printed music and recordings without neglecting issues related to other varieties of music information.

Scope: What Do Discovery Tools Really Discover?

The first step in discovery tool analysis is determining the discovery tool's purported scope. Most vendors provide some information about scope in their promotional materials and documentation. Key content areas to consider include:

Local catalog content: Discovery tools generally have the ability to include local catalog data, exported from the Integrated Library System (ILS) and updated on a regular basis. The exception is WorldCat Local, which provides metadata for locally held materials via WorldCat records with the institution's holdings marked.

Noncatalog local repository content: Examples include music-specific digital collections, music electronic theses/dissertations, and music materials in an institutional repository.

Remote public domain repository content: The number of public domain scores and books freely available online grows daily and includes HathiTrust scores and music books, International Music Score Library Project (IMSLP) scores, and scores and music books from other digitization projects such as the Library of Congress' American Memory collections. A discovery tool can help facilitate access without bloating the local catalog.

Subscription e-content, particularly for printed music and recordings: Subscription e-content has proliferated in recent years and includes Naxos Music Library audio, Naxos Video Library, the Database of Recorded American Music, and Alexander Street Press' Music Online audio, video, and scores components.

Article-level content: Inclusion of article content is the major defining factor between web-scale discovery tools and next-generation catalogs. However, the article indexing usually comes from the discovery tool's central cloud index, which is not necessarily in any direct relationship with article databases like *International Index to Music Periodicals*, *Music Index*, *JSTOR*, and *RILM Abstracts of Music Literature*.

Metadata and Interface: Complementary Partners in Discovery

Once a clear conception of a discovery tool's intended scope is established, the next step is determining how well the discovery tool actually indexes the material it supposedly covers and its success at manipulating and displaying the data. The remainder of this article will focus on each of the above content areas in succession, examining in tandem the complementary aspects of metadata and interface. Individual institutions generally have the most direct influence over local catalog content, and so it will be discussed at greatest length. The next three content areas (non-catalog local content, remote public domain repository content, and subscription e-content) are similar in many ways and will be discussed simultaneously. Finally, the unique aspects of article-level content will be examined.

Local Catalog Content

Individual institutions generally have the power to determine what local catalog content will be included and how it will be indexed and (to a lesser extent) displayed within the discovery tool.[10] As already noted, printed music and recordings present the largest segment of music-specific discovery challenges. For most libraries, a large, important portion of printed music and recordings represented in any discovery tool will come from locally owned physical collections, so indexing and display of catalog data can make or break a discovery tool's success in locating music materials.

The specialized requirements of music information retrieval and the faceted nature of discovery tools lend themselves to an examination organized according to FRBR terminology, and a brief review of selected FRBR terminology follows. Two categories of FRBR terms are particularly relevant: "user tasks" provide a framework for analyzing users' information needs; "group 1 entities" represent aspects of the products of intellectual or artistic endeavor which users seek. The FRBR user tasks are to "find" entities that correspond to a user's stated search criteria; "identify" an entity, including distinguishing similar entities; "select" the entity appropriate to the user's needs, such as a particular format; and "obtain" access to the entity described, for example through locating an item on the shelf or accessing it electronically. The FRBR group 1 entities are the "work" (a distinct intellectual or artistic creation, such as Mozart's *Le Nozze di Figaro*); the "expression" (the intellectual or artistic realization of a work, such as the March 1968 performance with Karl Böhm conducting); the "manifestation" (the physical embodiment of an expression of a work, such as the 1996 Deutsche Grammophon CD re-release of the performance); and the "item" (a single exemplar of a manifestation, such as a particular CD you can physically hold in your hand).[11]

Music materials pose relatively few unique challenges regarding the user task "obtain" and the group 1 entity "item" and so this paper will focus on finding, identifying, and selecting works, expressions, and manifestations. The key needs of music users in relation to discovery tools may be organized into three broad groups: finding and identifying the desired musical work; finding, identifying, and selecting the desired expression or manifestation, including particular formats, versions, and languages; and finding, identifying, and selecting works, expressions and manifestations which meet subject type needs, including specific instrumentations. The first two groups can be described as primarily constituting known-entity searching. (The term "known-item" is popularly used for such searches. Here, "known-entity" is substituted to avoid confusion with the FRBR-defined "item".) The third group contains searches which frequently begin without a specific entity in mind. Discovery tools meet these needs through metadata and indexing decisions in tandem with display and interface design.

---

[10] The exception is WorldCat Local, which utilizes the WorldCat version of bibliographic records with the institution's holdings marked. Therefore, materials not represented in WorldCat will be excluded from WorldCat Local. Specific material groups many music libraries have historically chosen not to catalog in WorldCat include local concert recordings and collections of sheet music and choral octavos.

[11] For further information on FRBR, see IFLA Study Group on the Functional Requirements for Bibliographic Records' "Functional Requirements for Bibliographic Records", http://www.ifla.org/publications/functional-requirements-for-bibliographic-records. For a brief summary, see Barbara Tillet, "What is FRBR: A Conceptual Model for the Bibliographic Universe" http://www.loc.gov/cds/downloads/FRBR.PDF.

Indexing is a silent but key partner to meeting all three groups of needs. Poor indexing will cause searches to inexplicably fail while good indexing will transparently bring the user to the material sought.   It is crucial for someone with music cataloging knowledge to examine the catalog data indexing, lest key music-specific fields be improperly indexed or wholly excluded. The actual technical details of implementing the indexing are normally handled by the vendor or a systems librarian.

Two sources are particularly helpful in determining indexing. OCLC's indexing choices for MARC data recorded in "US-MARC Fields And Subfields In WorldCat FirstSearch indexes" provide a solid and thorough starting point for each institution's decisions.[12] "Musical Attributes, Refinements, and Recommendations For Their Use" (from the Music Library Association's Bibliographic Control Committee Metadata Working Group) describes fourteen attributes of music materials, makes recommendations for the appropriate capture of metadata for each attribute, and includes mappings to Dublin Core, MARC, METS, and MODs metadata formats.[13]

As noted, the first broad group of needs (finding and identifying the desired musical work) is met by known-entity searching. However, users may begin a known-entity search for a particular musical work with any one of a wide variety of titles due to music's frequent use of generic titles and the needs of music users for works originating from and published in many languages. Music librarianship's answer to the malleable nature of musical work titles has been the uniform title, including "see" references from alternate titles and "see also" references to related works. Historically, a composer's works could be browsed in alphabetical order by name-title uniform title, first via the card catalog and in recent decades via OPAC alphabetical browses. Even the International Music Score Library Project (http://www.imslp.org), a digital score repository, includes alphabetical browsing arranged by composer, then title. Discovery tools, however, generally dispense with the alphabetical browse and rely on keyword searching exclusively. How then can they meet the needs of music users, especially their needs to find and identify particular musical works?

An autocomplete or autosuggest feature in the initial search or beginning results list which takes into consideration see and see also references from records in the Library of Congress/NACO Authority File (LC/NAF) could help guide users towards the musical works they seek. The mere presence of an autocomplete or autosuggest feature is no guarantee it is utilizing data from the LC/NAF. There are numerous sources of data for such features and determining what is powering the feature can require direct questions to vendors in combination with hands-on testing of the feature.

Careful display configuration will aid users in identifying and selecting the desired musical work. Ensure that appropriate notes (including MARC 500, 505, 508, 511, 518, and 546) and

---

[12]"US-MARC Fields And Subfields In WorldCat FirstSearch Indexes," OCLC, accessed Mar. 22, 2011, http://www.oclc.org/support/documentation/firstsearch/z3950/z3950_databases/specs/worldcat_linked_file_marc_fields.htm

[13] "Musical Attributes, Refinements, and Recommendations For Their Use." Music Library Association Bibliographic Control Committee, Metadata Working Group. Accessed Mar. 22, 2011. Last revised February 11, 2008. http://bcc.musiclibraryassoc.org/BCC-Historical/BCC2008/BCC2008MSWG2.html

personal and corporate names (MARC 700, 710, and 100/240), are displayed.[14] Furthermore, uniform title access points should promote further discovery of similar materials. This is often done through hyperlinking, which will be most successful when the name and title portions of name-title uniform titles are associated with each other regardless of whether they originate from a MARC 100/240 combination or from a MARC 700 field.

The second broad group of needs regards finding, identifying, and selecting the desired expression or manifestation. The fact that musical works exist in many different expressions and manifestations complicates meeting this group of needs. Specifically consider the ability of the discovery tool to help users: identify particular physical formats; identify and select the desired expression and manifestation based on displayed information; and find, identify, and select materials with particular language attributes.

Music users frequently need a specific physical format such as printed music, an audio recording, or a video recording. Music users also seek particular sub-formats: for example, not just any audio recording, but a CD specifically. Traditional OPACs usually meet this need through use of limiters applied at the point of search construction, but discovery tools more commonly provide facets in the initial results list, with pre-search faceting sometimes available from advanced search screens. In addition to filtering and sorting results, music users also need the ability to quickly determine the format of a particular item in a results list. Discovery tools (as well as some traditional OPACs) commonly use icons to aid in this task.

In addition, formats pose metadata and indexing considerations. Discovery tools use various data sources for assigning formats, and it is important for music librarians to understand the basis for format assignment and how this compares with format assignment in other tools, such as a traditional OPAC, the institution provides.   Two common data sources for format assignment are bibliographic records (usually the fixed fields) and item records. Consider what method is used by the institution's existing OPAC. If that method works well, it may be preferable to use the same method in the discovery tool for consistency's sake. Conversely, using a different method may provide superior faceting and offer an opportunity to connect users with the desired format more easily.

Ideally, a discovery tool will allow materials to be assigned to more than one format facet, including facets for sub-formats. For example, a DVD could be considered both a video recording and a DVD; a streaming audio file could be both an e-resource and an audio recording; a score with accompanying CD could be a score, a CD, and an audio recording. WorldCat Local combines good display and good interface to exemplify intuitive use of format facets. Materials are grouped into formats as well as sub-formats, with the relationships clear from the display. (See figure 1).

---

[14] Here and following, MARC field tags are often referenced in relation to library catalog data. Though not all music information is encoded in the MARC format, most library catalog data still uses MARC and thus the tags add specificity to the discussions.
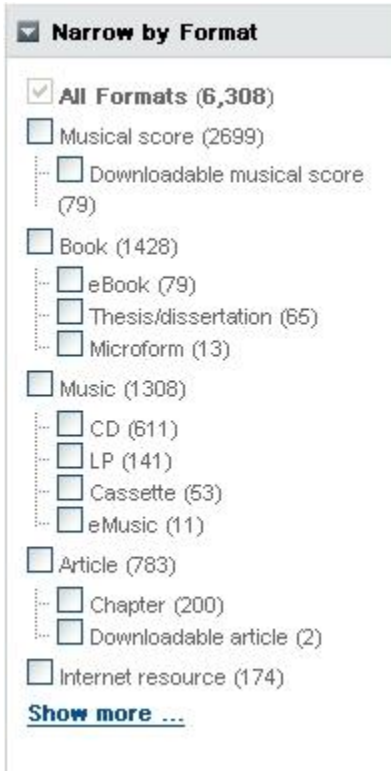
Figure 1: WorldCat Local allows items to be assigned multiple facets for formats and sub-formats, and also groups formats and sub-formats in an intuitive hierarchy for display.

Continuing with the second broad group (finding, identifying, and selecting an expression or manifestation), the full display should be configured to help users identify and select the particular expression or manifestation sought. Some discovery tools provide full displays; others do not, instead linking out to the "native" interface, such as a traditional OPAC or a database. The following recommendations serve as guidelines for optimizing full displays, regardless of the tool powering the interface. As with indexing, the technical aspects of actually configuring displays will generally be handled by the vendor or a systems librarian, but it is crucial for someone with music cataloging knowledge to examine the fields chosen for inclusion.

Both printed and recorded music are frequently the work of multiple contributors: composer, librettist, arranger, editor, performer, producer, etc. The full display will ideally make the role of each contributor clear. Translating MARC $4 relator codes into vernacular terms is one way to provide this information. In addition, printed and recorded music each pose further individual display needs.

For printed music, it is particularly important to be able to determine the physical presentation as well as the arrangement/edition, i.e. score, parts, vocal score, arrangement, etc. This can be accomplished by ensuring that the following are displayed:

- musical presentation statement (MARC 254)
- edition statement (MARC 250)
- physical description (MARC 300)

- appropriate notes (in general, all MARC 5xx, except any intended as internal or system notes)

For recorded music, it is particularly important to display performers, recording date/location, and the particular works contained within the recording. Ensure displays include:

- personal, corporate, and meeting names, including title portions (MARC 700, 710, 711)
- titles not associated with names (730, 740)
- participant/performer notes (MARC 511)
- creation/production notes (MARC 508)
- date/time and place of event notes (MARC 518)
- contents notes (MARC 505)
- general notes (MARC 500)

The final aspect of the second broad group regards the special importance of language attributes in music information retrieval. Several factors contribute to this heightened need to find, identify, and select music materials based on language attributes. First, while general users frequently focus on one or at most a handful of languages, seekers of musical works regularly desire a variety of languages. This is particularly true for vocal music, where users seek both original languages and specific translations. Second, individual music materials regularly incorporate a number of languages in various distinct capacities, including program notes, critical commentaries, libretti, dubbing and subtitles. Finally, purely instrumental music has no language information associated with the work's primary content, yet published manifestations feature an array of languages in the secondary content.

Discovery tools usually include a language facet, with vendors allowing individual institutions varying degrees of control over this facet's content and behavior. A variety of possibilities exist regarding which specific language aspects to include in the facet. For example, all languages could be included, or the facet could consider only those languages relating to the primary content, perhaps with selected secondary language aspects added in. Discovery tool language facets handle lack of linguistic content in varying ways. Possibilities include ignoring these materials, grouping them together with a label like "no linguistic content", and assigning a language based on secondary content. For MARC data, the best places to obtain language information are 008/35-37 [Lang] and 041, with the specific subfields of 041 chosen based on which language aspects the institution desires to include in the language facet. In the full display, facilitate identifying language attributes by displaying languages notes (MARC 546).

Despite the importance of known-entity searching, music users do have subject searching needs, represented by the third broad group of needs: finding, identifying, and selecting works, expressions and manifestations which meet particular subject needs, including particular instrumentations. Discovery tools frequently meet these needs through the use of subject facets.

Subject facets are often derived from subject headings in catalog data and subject-like keywords in data from other sources. Some discovery tools keep the entire subject heading string together as one facet while others split each subdivision into its own facet. Splitting apart subdivisions is an attempt to provide faceted browsing, but it can deliver odd results since Library of Congress

Subject Headings (LCSH) were not designed to be used for faceted access and the subdivisions often derive meaning from their placement in the entire string.

Classification numbers provide another possibility for creating subject facets for library catalog content. WorldCat Local is one discovery tool which uses this approach. Classification numbers group materials into broader classes than LCSH and are useful for easily disambiguating terms which have different meanings in different disciplines. For example, "mass" means one thing to musicians and another to physicists. Unfortunately for the success of the classification number approach to subject faceting, many libraries do not classify their sound recordings, so users applying this facet may unknowingly exclude large swaths of sound recordings.

Music users also seek music with specific instrumentations. Many library catalogs rely on LCSH to convey this information, though imperfectly since LCSH mix medium of performance (e.g., piano) and genre/form (e.g., sonatas). The MARC 048 field, containing coded instrumentation information, provides one alternative, but 048 has rarely been uniformly coded in bibliographic records and few discovery options of any type effectively use the 048. While not a web-scale discovery tool, the University of Virginia's implementation of Blacklight shows how the 048 field can be used in faceted catalogs. Within the tool's special "music view", an "instrument" facet allows users to easily add and remove performing forces, including instruments, voice, and ensembles, from a search. (See figures 2 and 3.)

Figure 2: "Instrument" facet in University of Virginia's Blacklight "music view".
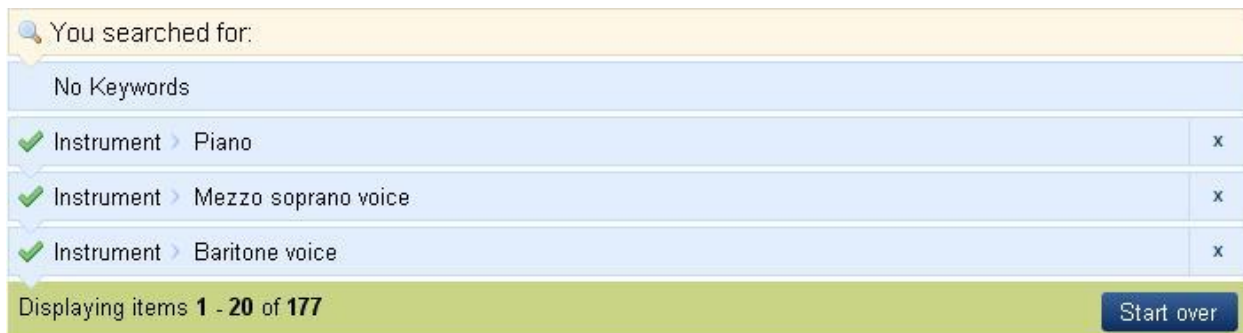

Figure 3: Once performing forces have been added using Blacklight's instrument facet, the choices made a displayed and each individual performing force can be easily be removed.

Non-catalog local repository content, remote public domain repository content, and subscription e-content

From a discovery tool standpoint, other local repositories, remote public domain repositories, and subscription e-content pose many of the same configuration problems and questions as local catalog data. Therefore, many of the same principles already discussed in relation to catalog content apply to metadata considerations and display issues for these other repositories. In all cases, attributes must be indexed appropriately, mapped to useful facets, and displayed helpfully. Individual institutions generally have more control over these areas for local repositories than for remote repositories, whether public domain or subscription. Aligning the configuration of local catalog data, local repositories, and remote repositories with each other will provide users with a uniform experience regardless of which source ultimately provides the music information sought.

Another consideration stems from the reality that in recent years, many libraries have attempted to make their OPAC function as a single search point, particularly by loading record sets for e-content into their catalog. In addition to music books and journals contained in general e-content packages and local and public domain repositories, music librarians should particularly be aware of record sets loaded for electronic audio, video, and printed music. Discovery tools may include indexing for these materials in their central index, and if the local catalog also contains records for these materials, duplicate hits may occur in the discovery tool if the entire local catalog is loaded into the discovery tool. In such situations, options include: 1) allow the duplication to exist; 2) configure or massage the catalog export so as to exclude the records; 3) reconsider the role of the catalog as a single search point and delete the records from the catalog entirely; or 4) configure the discovery tool to exclude its indexing for the materials in question. Ultimately, the option chosen will depend on the purposes envisioned for the discovery tool vis-à-vis the OPAC, the quality and comprehensiveness of the indexing in each source, and the technical feasibility of the various options.

Article-level content

Web-scale discovery tools use a variety of methods to provide article content, including a central cloud index, "traditional" federated searching, and web services connections with content providers (APIs). However, most discovery tools focus on a central cloud index created and maintained by the vendor for a major portion of article content. Individual institutions generally have limited control over the data contained in these central indexes beyond scoping them to materials available to the institution's users. It is important to contrast these central indexes with "traditional" federated searching, which broadcasts a search to multiple databases before returning the results in one interface. The metadata in central indexes is not merely pre-harvested from databases; indeed it may not be received from database vendors at all but instead come from other content providers such as publishers. Therefore, searches conducted via discovery tools may retrieve different results from the same searches conducted in individual databases.

In analyzing discovery tool article content, the ultimate question regards the breadth and indexing quality of a discovery tool's article content coverage. Specifically, to what extent can searching a discovery tool substitute for searching individual music databases like *International Index to Music Periodicals*, *Music Index*, *JSTOR*, or *RILM*? This question can be difficult to

answer, particularly if a discovery tool relies heavily on a central index. Most discovery tool central indexes are so large that defining their scope is a little like defining Google's scope. Vendors usually provide some assistance in determining scope, perhaps in the form a list of publishers or titles covered or an offer to compare the institution's electronic journal content to the titles indexed in the discovery tool. However, the lists can be long and vendors present the information in different formats and with varying degrees of completeness, making comparison tricky. For example, OCLC provides a spreadsheet of all the collections and databases in WorldCat Local, including additional details and the source of the data.[15] Serials Solutions, on the other hand, provides: 1) a list of publishers participating in Summon; and 2) a list of titles represented in Summon which includes title, ISSN, and coverage dates, but does not indicate the specific source of the data.[16]

Moreover, even equivalent coverage dates and titles do not guarantee equivalent indexing. This is particularly true for materials which come from a central index, since article metadata may come from publishers or other content providers, not databases. Music databases provide various kinds of metadata for their content. For example, searches in *JSTOR* and *International Index to Music Periodicals Full Text* can query the full text of articles. *Music Index* includes basic citation information plus various types of subject terms and abstracts. *RILM Abstracts of Music Literature* likewise adds subject terms and extensive abstracts to its indexing. A discovery tool might index the same content but provide only citation-level metadata. If a discovery tool receives its data from a source other than the database provider, it will not be able to serve as a complete substitute for individual database searching in cases where users need a truly comprehensive search. Though vendors may share some information about the metadata powering their central index, direct questions may be needed and even then the provided information may not give a complete picture of the situation.

In the quest to understand a discovery tool's central index and its usefulness for discovering article-level content, it may be necessary to move beyond analyzing information provided by the vendor and conduct a variety of strategic searches in the discovery tool and databases, noting differences and similarities. Searches on broad or common terms like "symphonies", "Mozart", or "piano" will give an idea of extent of coverage and the kinds of materials most common and most highly weighted by the ranking algorithm. Searches on very specific terms such as names of obscure people or titles of minor musical works will allow more careful comparison of indexing quality.

Even when their metadata and/or scope do not allow them to fully substitute for music databases, discovery tools can still be useful for article content. They are generally beneficial for introductory research and gaining a general sense of the information available and in situations when comprehensiveness is not required. Their broad scope also makes them particularly valuable for interdisciplinary topics, where they may uncover information not included in music-specific databases.

---

[15]"WorldCat Local Databases and Collections List," OCLC, accessed Mar. 20, 2011, last revised February 25, 2011, http://www.oclc.org/worldcatlocal/overview/content/dblist/WorldCat_Local_Content_List.xls
[16] "Summon Content and Coverage," Serials Solutions, accessed Mar. 22, 2011, http://www.serialssolutions.com/summon-content-and-coverage/

Conclusion

Web-scale discovery tools combine data from a myriad of sources in their quest to achieve one-stop searching. While current systems ultimately fall short of this goal, web-scale discovery has the potential to meet many aspects of music information retrieval needs if the tools are carefully configured and thoughtfully developed. Because of the special considerations of music information, particularly printed music and recordings, music librarians must be involved in the selection and configuration of these increasingly popular products. On the institutional level, the greatest customization is possible for content from the local catalog and other local repositories. On a larger scale, music librarians must band together and advocate to vendors for music information retrieval needs. In so doing, we will come closer to our ultimate goal: connecting users with the music information and materials they need.

Works Cited

Ballard, Terry. "Comparison of User Search Behaviors with Classic Online Catalogs and Discovery Platforms." *The Charleston Advisor* 12/3 (Jan. 2011): 65-66.

Christensen, Beth, Mary Dumont, and Alan Green. "Taking Note: Assessing the Performance of Reference Service in Academic Music Libraries: A Progress Report." *Notes* 58/1 (Sept. 2001): 39-54.

King, David M. "Catalog User Search Strategies in Finding Music Materials." *Music Reference Services Quarterly* 9:4 (2007): 1-24.

Music OCLC Users' Group Reference Services Subcommittee. "WorldCat Local Enhancement Recommendations for Music." (April 2010), accessed Mar. 22, 2011. http://www.musicoclcusers.org/WorldCatLocal20100412.pdf

"Musical Attributes, Refinements, and Recommendations For Their Use." Music Library Association Bibliographic Control Committee, Metadata Working Group. Accessed Mar. 22, 2011. Last revised February 11, 2008. http://bcc.musiclibraryassoc.org/BCC-Historical/BCC2008/BCC2008MSWG2.html

Newcomer, Nara L. "Electronic Resources Review Column: Summon." *Music Reference Services Quarterly* 14/1&2 (2011): 59-62.

Olson, Tod A. "Utility of a Faceted Catalog for Scholarly Research." *Library Hi Tech* 25/4 (2007): 550-561.

Rowe, Rhonda. "Web-Scale Discovery: A Review of Summon, EBSCO Discovery Service, and WorldCat Local." *The Charleston Advisor*, 12/1 (July 2010): 5-10.

Snyder, Tracey. "Music Materials in a Faceted Catalog: Interviews with Faculty and Graduate Students." *Music Reference Services Quarterly* 13/3&4 (July 2010): 66-95.

"Summon Content and Coverage. Serials Solutions. Accessed Mar. 22, 2011. (http://www.serialssolutions.com/summon-content-and-coverage/)

"US-MARC Fields and Subfields in WorldCat FirstSearch Indexes." OCLC. Accessed Mar. 22, 2011. http://www.oclc.org/support/documentation/firstsearch/z3950/z3950_databases/specs/worldcat_linked_file_marc_fields.htm

Way, Doug. "The Impact of Web-Scale Discovery on the Use of a Library Collection." *Serials Review* 36/4 (Dec. 2010): 214-20.

WorldCat Local Databases and Collections List." OCLC. Accessed Mar. 20, 2011. Last revised
         February 25, 2011.
         http://www.oclc.org/worldcatlocal/overview/content/dblist/WorldCat_Local_Content_Lis
         t.xls