

Music Discovery Requirements: A Guide to Optimizing Interfaces

Nara L. Newcomer, Rebecca Belford, Deb Kulczak, and Kimmy Szeto
with Jennifer Matthews, Misti Shaw

Abstract

Music materials, particularly scores and recordings, pose unique demands that must be considered for their successful discovery. Some of the discoverability challenges of music materials in public search interfaces can be addressed simply by ensuring that needed bibliographic description fields are appropriately displayed and indexed in discovery interfaces. Other problems are more difficult to solve. This Music Discovery Requirements document explores the issues and gives concrete recommendations for discovery interfaces. Given that most libraries will be dealing with databases containing large bodies of legacy data recorded according to AACR2 and encoded in MARC, particular attention is paid to MARC data and to AACR2; RDA is addressed to the extent possible. The discussions and recommendations will be useful to those creating or guiding the development of discovery interfaces that will facilitate access to music materials. Furthermore, because the document identifies areas where deficient data creates particular problems for discovery, those inputting or creating standards for data can use this document to identify areas with particular needs for fuller, more consistent data.

Open Access/Copyright Notices

This publication first appeared in *Notes* 69, no. 3 (March 2013): 494-524.

<http://dx.doi.org/10.1353/not.2013.0017> This material may not be copied or reposted without explicit permission. Copyright 2013, Nara L. Newcomer, Rebecca Belford, Deb Kulczak, and Kimmy Szeto with Jennifer Matthews, Misti Shaw.

This open access posting includes three spreadsheet appendices which were not published in *Notes*.

Another version of this article (including the spreadsheet appendices, and in-text lists of fields and other metadata specifications) appeared on the Music Discovery Resources section of the Music Library Association website. Future updates will be posted to the Music Library Association website. <http://www.musiclibraryassoc.org>

Biographical Note

Nara L. Newcomer is head of music/media library at the University of Missouri-Kansas City and chair of the Music Discovery Requirements Document Group. Rebecca Belford is music cataloger/reference librarian, University at Buffalo. Deb Kulczak is database maintenance librarian/music cataloger at the University of Arkansas, Fayetteville. Kimmy Szeto is cataloging and metadata librarian at Maritime College, State University of New York. Jennifer Matthews is film, television and theatre librarian at the University of Notre Dame. Misti Shaw is music & performing arts librarian at DePauw University. All are members of the Music Discovery Requirements Document Group, which was initiated by the Music Library Association's Emerging Technologies and Services Committee. Web URLs cited herein accessed 14 November 2012.

I. Introduction

“Discovery” has become a library buzzword but it refers to a traditional concept: enabling users to find information and materials. Discovery was first facilitated by print indexes and card catalogs, then later by online library catalogs and search engines. Today, the discovery environment is changing rapidly both within libraries and externally. Within the library realm, FRBR, RDA, discovery tools, and faceted browsing are key factors related to these changes in the way our users navigate searches and encounter library data. The newest discovery interfaces aim to remove barriers between different types of data and incorporate aspects of nonlibrary online searching environments familiar to users. When discovery interfaces work well, everyone benefits: the library’s collections are more fully exposed, and the user does not have to adhere to nor struggle with the shortcomings of complicated search vocabularies and strategies. Discovery is optimized.

In the midst of this changing environment, music materials pose unique demands that must be considered for successful discovery. There is exciting potential for new discovery interfaces to ease the difficulties users face when searching for music materials. However, in reality, the specialized discovery needs arising from music materials are often overlooked. This document serves as a guide to meeting these unique demands for vendors, librarians, and anyone developing or implementing discovery interfaces of all kinds.

This work was initiated by the Emerging Technologies and Services Committee of the Music Library Association (MLA) in February 2011, with official support of the MLA Board of Directors given in June 2011. Group membership was drawn from both that committee and the broader ranks of MLA members. The group includes a liaison to MLA’s Bibliographic Control Committee, which provided many thoughtful comments and suggestions. A first draft was made available for public comment 16 November-5 December 2011. A second draft was made available for public comment 9 February-16 March 2012. The final version was approved by the MLA Board of Directors on 8 August 8 2012.

This document has roots in earlier work by Lenore Coral¹ and by the MLA Automation Subcommittee;² however, it is not a mere revision of either document but a wholly new document reflecting today’s environment. Given the rapidly changing bibliographic landscape and the immediate need for a document to aid discovery interface implementations, this document is not a standard. Instead, it gives recommendations and possible best practices, in conjunction with discussion of the factors and discovery needs that precipitated the recommendations. As the bibliographic landscape changes, this document will need revision. The group also hopes this document’s careful examination of limitations of the current landscape will inform the development and implementation of new standards for recording and encoding data.

¹ Lenore Coral, “Automation Requirements for Music Information,” *Notes* 43, no. 1 (September 1986): 14-18.

² Music Library Association Automation Subcommittee, “Automation Requirements for Music Materials, Final Subcommittee Draft,” 1997, http://library.music.indiana.edu/tech_s/mla/autoreq.txt. Note that further minor revisions were made to the document as late as 2004, but these revisions were never made publicly available.

Two elements define the scope of the present document: music and discovery. First, the document focuses on musical works (scores and recordings), rather than secondary literature about music (books and articles), because musical works present more unique discovery needs. Second, the focus on discovery mostly excludes back end functions such as circulation, cataloging, and acquisitions which, like secondary literature, present fewer music-specific requirements. However, powerful searching is important for staff work with music materials, and implementing recommendations from this document will benefit both users and library staff. The focus on discovery also means recommendations are focused on indexing and presentation of data for discovery, rather than rules and best practices for recording data (such as AACR2, RDA, and “Musical Attributes, Refinements, and Recommendations for their Use”³) or metadata formats and encoding standards (such as MARC, Dublin Core, METS, MODS, EAD, CDWA Lite, and VRA Core⁴). Within this focus on indexing and display, there is a particular need to address: legacy data created according to AACR or AACR2 and encoded in MARC, particularly the data in OCLC’s WorldCat database; coming changes associated with RDA; and the particular relevance of FRBR concepts for music materials. This document, therefore, includes recommendations for dealing with the constraints presented by this large volume of legacy data. Historically, library data recording and encoding standards have been biased towards Western art music. Other musics are often not ideally served by these standards; however, insofar as these other musics have been cataloged following such biased standards, discovery interfaces will not vary much depending on the type of music. Interfaces facilitating discovery for large amounts of music beyond the Western art music tradition may benefit from further customizations, if data have been encoded to support such customizations.

The document contains recommendations in three sections, with the first two sections aligned with FRBR concepts: musical works; expressions and manifestations; and other aspects of music discovery. Within each section, this document identifies and discusses each individual attribute or relationship, then gives a prose summary recommendation

Technical specifications for indexing, display, use in facets/limits, and related authority fields are omitted from this published document in recognition of constraints of space and formatting and the need for regular revision. The reader is referred to the “Music Discovery Resources” pages on the MLA web site,⁵ to access the most recent revision of this information, which:

- proposes best practices for the attribute/relationship regarding:
 - indexing (inclusion in the discovery tool’s searchable indexes)

³ Music Library Association Bibliographic Control Committee Metadata Working Group, “Musical Attributes, Refinements, and Recommendations for Their Use,” 2008, <http://bcc.musiclibraryassoc.org/BCC-Historical/BCC2008/BCC2008MSWG2.html>.

⁴ MARC: MACHine Readable Cataloging, <http://www.loc.gov/marc/> (recommendations given are for MARC 21); Dublin Core: Metadata Initiative, <http://dublincore.org/>; METS: Metadata Encoding and Transmission Standard, <http://www.loc.gov/standards/mets/>; MODS: Metadata Object Description Schema, <http://www.loc.gov/standards/mods/>; EAD: Encoded Archival Description <http://www.loc.gov/ead/> (recommendations given are for EAD 2002); CDWA Lite: Categories for the Description of Works of Art, http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.html; VRA Core: data standard from the Visual Resources Association, <http://www.vraweb.org/projects/vracore4/>. CDWA Lite and VRA Core are rarely ideal for music, but they are included in this document for the benefit of situations where they may be used.

⁵ Music Library Association Emerging Technologies and Service Committee, “Music Discovery Resources,” 2012, <http://committees.musiclibraryassoc.org/ETSC/MDR>.

- display (making the data viewable in patron interfaces)
- use in facets/limits (use of data to create pre- or post-search filters based on the data; traditional OPACs frequently emphasize pre-search limits, while faceted browsing discovery tools promote post-search facets for refining results)
- lists related MARC authority fields, if applicable

Also available on the MLA web site are appendixes compiling technical details of the recommendations in spreadsheet form, from two perspectives: beginning with the indexes and giving fields to include, and beginning with the field and giving indexing for each field. A third appendix details MARC bibliographic record mapping for content and carrier.

II. Musical Works

II.A. Introduction

Clear identification and display of information regarding the musical work⁶ is important to users. The following section addresses attributes of musical works which are important to users, as well as the “created by” relationships which link persons and corporate bodies (most often persons) to works and the “has as subject” relationship linking various entities to works (under the heading “topical subjects”). The following attributes and relationships are addressed:

- II.B. Titles
- II.C. Identifying Numbers
- II.D. Medium of Performance
- II.E. Musical Key/Range
- II.F. Dates
- II.G. Persons and Corporate Bodies
- II.H. Topical Subjects
- II.I. Genre/Form
- II.J. Geographic Area

II.B. Titles

Users' discovery and identification of musical works frequently draws on title, often in conjunction with creator. Particular musical works are often referred to by many different titles in various languages, for example: Symphony no. 5, Fifth Symphony, Sinfonie C-Dur, Symphonie op. 67, etc. This reality spurred librarians' extensive development and application of standardized titles (AACR2's “uniform titles”) for music materials. To facilitate identification of musical works, it is crucial to display the entire standardized title, including additions to distinguish similar works and to identify particular parts/sections of the work, as well as expression-related additions. It is also important to use authority records or other methods to lead users to the work they seek, even when they begin their search with an alternate title for the work. See discussion of authority records in IV.B.

⁶ Here defined according to FRBR: a distinct intellectual or artistic creation.

The current environment focuses on creating a standardized text string (AACR2's "uniform title" and RDA's "authorized access point representing the work") to represent the work. Cataloging codes dictate the selection of the title language, manipulation of the title, and addition of elements such as medium of performance, work numbers, and key to create this text string uniquely identifying the work. In the future, the entire authority record, or some other record, might serve as a surrogate for the work. For the present, best practice is to continue including all other elements (particularly medium of performance, work numbers, and key) in title indexes.

Musical works are nearly always associated with specific creators. For works with known composers, the composer's name is often essential to identify the work and distinguish it from works with the same or similar titles. Therefore, creator name must be displayed in conjunction with the standardized title. (Creator name should not, however, be included in title indexes.) One specific problem area with the indexing and display of standardized titles for MARC data is that some discovery interfaces create separate hyperlinks for the creator in 700 \$a\$b\$c\$d\$e\$j\$q and the work title in 700 \$k\$l\$m\$n\$o\$p\$r\$s\$t, and likewise fail to associate 100 and 240. (Standardized title combinations can also be created from 110 (corporate name) and 111 (meeting name) in combination with 240 and likewise in 710 and 711, but this is rare.) In the development of discovery systems it is important to create a robust system that will link the data from these fields together to enable users to find materials.⁷ Some musical work titles are not associated with names and are contained in MARC 130 and 730.

When a manifestation contains expressions of multiple musical works (for example, song anthologies), standardized titles are not always assigned for every work expressed. Sometimes, titles are merely transcribed from the item into contents notes. For most comprehensive coverage, these transcribed titles should also be included in title keyword indexes and displayed. See IV.C for further discussion of the challenges of compilations.

Recommendation: Index and display all title fields, except do not display coded subfields not intended for user display. Ensure that standardized title strings are associated with their creators. Pursue methods for leading users to the work regardless of the version or element of the title they use to begin a search. For data using standardized vocabulary, make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes. In linking and facet creation, and in relevance ranking, privilege standardized title fields over transcribed titles.

II.C. Identifying Numbers

Opus, thematic index, and serial numbers are frequently used in Western art music to identify musical works. Ideally, systems will exploit authority record data to easily lead users between different systems of numbering for composers (e.g. Vivaldi, Domenico Scarlatti) where multiple work-number systems exist.

⁷ Stanford University's Blacklight implementation provides an example of a combined, clickable index. <http://searchworks.stanford.edu/> The initial implementation chose to leave out subfields representing parts and expressions: 240/7xx \$o\$s\$p.

In many cases musical works lack a distinctive title and are titled only with a type of musical composition, such as “symphony.” These works are often identified by the sequential number of works of that type by a given composer. These serial numbers depend on the association of the type with the number for meaning, and must display and function together.

Successful discovery of musical works identified by serial number (e.g., symphony no. 5) is dependent on phrase searching capabilities: search terms that otherwise function as Boolean operators must be searched as *terms* when enclosed in quotes or as part of phrases within quotes, rather than as operators. With this capability, a search on “no 5” returns works with recorded serial number “5” rather than excluding “5” as a search term.

Discovery systems do not often allow for searching by opus, thematic index, or serial numbers alone, except in general keyword indexes where many false hits may be generated by numbers present in the record for other purposes. Ability to search opus, thematic index, and serial numbers in their own separate index would be very useful for music.

Other numbers are associated with expressions and manifestations, and are discussed in III.C.

Recommendation: Index and display opus, thematic index, and serial numbers. Utilize system algorithms to include in indexing all numbers in a consecutive range where only first and last numbers are recorded. Consider the interface and available screen real estate in the decision whether to include them in facets. Ensure that serial numbers are associated with the type of composition. Ensure that functionality exists for searching “no,” the English abbreviation for “number,” as a term, rather than exclusively as a Boolean operator.

II.D. Medium of Performance

Medium of performance is an attribute uniquely important to music. It is an identifying element for known musical works, and users also regularly seek works and expressions for a particular medium of performance, without a particular work/expression in mind at the beginning of their search. However, historically, data recording and encoding standards and practices have not made medium of performance cleanly and consistently separable from other attributes.

In MARC/AACR2 legacy data, the medium of performance may be stated or implied in fields for title (130, 240, 7xx) or subject/genre (650/655), encoded in 048 (medium of performance), or implied to a limited degree in certain codes in 008/18-19 and 047 (form of composition). Medium of performance may also be mentioned or implied in various descriptive fields (2xx, 5xx). As RDA records with related MARC elements are added, medium information may also be recorded in 382 (expanded to accommodate medium of performance data with MARC Proposal 2012-01, approved Jan. 21, 2012).⁸ In the future, legacy data might be massaged so that medium of performance contained in 650/655 is moved to 382.

Rules governing input of certain fields or elements within LSCH (*Library of Congress Subject Headings*) and AACR2 uniform titles add to the problem of limited or dispersed information.

⁸ MARC Proposal No. 2012-01, "New Data Elements in the MARC 21 Bibliographic and Authority Formats for Medium of Performance," <http://www.loc.gov/marc/marbi/2012/2012-01.html>.

This is particularly problematic with vocal music, folk music, recorded popular music, and compilations. Indexing, displaying, and allowing limits/facets by multiple fields will maximize users' ability to search for and identify medium.

Going forward, recording medium in a way that permits machine actionable, granular description of expressions is essential. Current LCSH patterns have many shortcomings: combination of genre/form with medium of performance in a single term, difficulty in identifying soloists and number of performers, the semantic relationships between terms, and the syntax of the terms. Current MARC 048 is also problematic, particularly for recorded popular, jazz, and folk music when some, but not all, mediums can be identified. Since the proposed *Library of Congress Genre/Form Terms for Library and Archival Materials* (LCGFT) thesaurus will not include medium of performance, the Library of Congress, in cooperation with the Music Library Association, is developing a new *Library of Congress Medium of Performance Thesaurus for Music* (LCMPT) to provide controlled vocabulary for medium of performance.⁹

Challenges relating to medium of performance are further complicated by the prevalence of manifestations with multiple expressions. (For additional discussion, see Compilations in IV.C.) Linking fields offer a potential solution in this area, but have historically not been employed.

Users may also seek an expression with a particular medium of performance (i.e. an arrangement). See III.G for further discussion.

Recommendation: Index and display fields relating to medium of performance, with codes converted to the vernacular. Explore alternative tools and interfaces to allow users to interact more easily with medium of performance. Revise interfaces as new standards for recording and encoding medium of performance are implemented.

II.E. Musical Key/Range

Original key can be useful for identifying tonal works of Western art music, particularly when the work has a generic title and the identifying number (such as an opus, work, or thematic catalog number) is unknown, absent from the cataloging data, or wholly non-existent.

Edition statements may contain information on the original key or range. The manifestation may not identify whether such statements refer to the original key/range or a transposition, so best practice is to simply include edition statements in indexing and display.

Successful discovery of musical works identified by musical key is dependent on phrase searching capabilities: terms that otherwise function as stopwords must be searched as *terms* when enclosed in quotes or as part of phrases within quotes. With this capability, a search on “a major” will return bibliographic records containing the string “a major” (a musical key) rather than “major” alone.

⁹ See also “Medium of Performance for Music: Working List of Terminology,” <http://www.loc.gov/catdir/cpsd/medprf.html>.

Another consideration for searching for works with known key is the exclusive use of the symbols for music sharps (#) and flats (b)5/12/2016 in uniform/preferred titles and the potential for symbols, rather than spelled out terms, to be used in descriptive fields. In some older data, the pound sign (#) is used as a replacement for the musical sharp, and the lowercase “b” is used as a substitute for the musical flat. A system that automatically associates search terms entered as “sharp” or “flat” with their respective symbols would not be appropriate, considering the common usage of these terms. However, this problem could be partially addressed by a feature that allows users to select from a set of frequently used symbols or diacritics. Regardless of search functionality, the system must allow entry and display of the symbols for musical sharps, flats, and naturals¹⁰ in all fields of bibliographic and item records.

An additional field that contains musical key information is MARC 031 in bibliographic and authority records. Defined for recording musical incipit¹¹ in ordinary ASCII characters, it can be used for any material containing music. While \$p contains the incipit itself, \$r is used to record the key or mode. Therefore this field could potentially be used for searching by key; equally beneficial would be the capability to convert the coded incipit to musical notation for public display. However, like many other MARC fields, 031 has not been coded regularly in bibliographic or authority records.

Key and range can also be attributes at the expression level, as further discussed in III.H.

Recommendation: Index and display musical key and range. Ensure that “a” (frequently a stopword), can be forced to be searched as a term. Ensure that the musical sharp, flat, and (if present in data) natural symbols can be entered, displayed, and searched. Evaluate data for feasibility of use as a facet. Explore the use of MARC 031 in bibliographic and/or authority records.

II.F. Dates

Creation date is an important, though neglected, access point for musical works. Users may seek works composed in a particular date range, which might be a specific individual year or decade, or a more broad era of music history (such as “baroque era” or “medieval era,” often indicative of style). Users may also use creation date to identify the work they are seeking, particularly to distinguish it from other works with similar titles and other attributes. With most books and articles (except classic literature) publication date is a fairly accurate approximation of creation date, but this is not true for musical works because a single musical work frequently exists in many expressions and manifestations. MARC 045 may be used to record period or date of composition, but this field has not been regularly used since the Library of Congress announced in March 1989 that it would cease coding 045. Some approximation of creation era might be possible by exploiting composer birth/death date information. In future RDA implementations,

¹⁰ The use of the musical natural symbol is limited in cataloging (and not currently defined in the OCLC-MARC character set), because a letter name without a sharp or flat is taken to mean “natural” in standardized titles and Library of Congress classification. However, it may be encountered in certain cases, such as call numbers created under Dickinson classification.

¹¹ Musical incipit is the opening fragment (usually the melody) of a musical work. The incipit helps to identify the work.

creation date might be recorded in musical work records. MARC 046 \$k\$l are now defined in the MARC authority format for this purpose.

Creation date, and other temporal aspects associated with musical works, may also be reflected in chronological subdivisions from topical subject headings found in \$y of MARC 650 fields. (Although \$y is defined for 655 genre/form headings, the *Library of Congress Genre/Form Terms for Library and Archival Materials* [LCGFT] thesaurus currently has no provisions for subdivision of terms.¹²) Chronological subdivisions are not consistently present due to subject cataloging rules, and they may also represent other associated temporal aspects besides creation date. Nor do existing chronological subdivisions correspond to the broad stylistic eras users sometimes seek. Therefore while index and display of chronological subdivision is useful, it is problematic for use as a creation date or style facet.

Dates are also associated with expressions and manifestations. For information and discussion on these dates, see III.I.

Recommendation: Index and display creation date and include in facets/limits if present in data. However, this access point has been neglected in much legacy data, making it difficult to implement consistently and accurately as a part of discovery interfaces.

II.G. Persons and Corporate Bodies

The “created by” relationship is very important for users to find, identify, and select musical works by a particular composer. Usually, the composer is a single person; more rarely, corporate bodies can serve as creators. It is important to index and display all personal and corporate names, both transcribed forms and standardized vocabulary.

Once a user has identified a work created by a particular composer, the interface should link to more materials related to that person. When the data utilize standardized vocabulary, the system should fully exploit the existing standardized vocabulary. For example, clicking on a link within a record for a work composed by John Adams the composer (Adams, John, 1947-) should locate only materials related to this person, not items related to either president or to other persons named John Adams. In many current systems, this will mean utilizing the bound text string for the full authorized form of name as the link. Alternately, a unique identifier could function behind the scenes, regardless of interface display. For further discussion of issues surrounding standardized vocabulary, see the discussion of authority records in IV.B.

Other names may be associated with a work as subjects (such as music about or in homage to a person), and these should be indexed and displayed, but as “subjects” rather than “authors.”¹³

¹² “Frequently Asked Questions About *Library of Congress Genre/Form Terms for Library and Archival Materials* (LCGFT),” revised June 6, 2011, http://www.loc.gov/catdir/cpsd/genre_form_faq.pdf. See Q21, Q37 and Q38.

¹³ Mostly, subject headings for persons and corporate bodies reflect subject relationships. Authors of texts set to music are an exception. SHM (*Subject Headings Manual*) H 1110 instructs addition of a subject heading for the name of the person whose writings or words are set to music, with the form subdivision “Musical settings” appended. In this case, the person has a creator, not a subject relationship to the work. However, AACR2 also

In addition to seeking specific known persons and corporate bodies, users may seek works associated with persons/corporate bodies possessing particular attributes, such as date, nationality, language, or field of activity. In existing AACR2/MARC data, the only one of these attributes recorded with any frequency is birth/death/activity dates of persons. In contrast, RDA allows for recording, for persons: date, gender, place of birth, place of death, associated country, place of residence, address, affiliation, language, field of activity, associated group, and profession/occupation of person; and for corporate bodies: associated place, associated date, associated institution, number of conference, other designation, language, address, and field of activity. MARC 046 and 370-377 have been defined in the authority format, and MARC 046 and 377 have been defined in the bibliographic format, to provide machine-actionable fields for this data. If these fields were regularly encoded, these data could be exploited to answer questions like “What music do you have by French women composers born before 1950?” through use of facets or limiters. However, MARC/ACCR2 legacy data do not contain this information; furthermore, these RDA elements, except birth and death dates for persons, and dates and locations of conferences, are not currently considered “core” RDA elements by the Library of Congress, except when needed to differentiate persons/corporate bodies. However, it might be possible to connect to other sources which have already mined this data such as WorldCat Identities (<http://www.worldcat.org/identities/>) or the Virtual International Authority File (<http://viaf.org/>).

Other persons besides the composer can be related to the musical work, such as librettists and lyricists. The relationship is sometimes collaborative (composer and lyricist work together on words and music) and sometimes involves an independent work (a literary text which the composer sets to music). AACR2 and MARC do not make the relationships clear in ways that could easily be machine-manipulated for discovery; future data-recording and -encoding might support more granularity in discovery. In addition, many more persons and corporate bodies contribute to musical expressions and manifestations; see discussion in III.J.

Recommendation: Index and display all personal and corporate names. Exploit data utilizing standardized vocabulary to allow users to differentiate between people/corporate bodies with similar names; make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes. Explore methods of allowing users to seek works based on attributes of related persons/corporate bodies.

II.H. Topical Subjects

True topical headings for music materials are confined to describing what music is *about*, but many musical works are not objectively *about* anything. Therefore, true topical headings are both less important and less common for music materials than for books. In FRBR terminology, true topical headings express a “has as subject” relationship. Subjects of musical works may include associated concepts, names of persons or bodies, and geographic areas (i.e. music about a

instructs for added entries for the person whose words are set, making manipulating the subject headings to bring out the creator relationship less vital.

geographic area, as distinguished from music emanating from a geographic area). In LCSH, topical headings for music materials are usually created by adding the form subdivisions “Songs and music” or (for dramatic musical works such as opera) “Drama” to non-musical headings. These subdivisions hold little meaning when separated from the rest of the heading, so it is important that the entire heading string be displayed as a unit and be able to function as a unit.

Currently, topically coded subject headings are also assigned to identify nontopical aspects including: music of national, religious, and ethnic groups (see also II.J and III.K), often by adding the form subdivision “Music” to an otherwise topical heading; the source of text set, by employing the title of a liturgical text or adding the subdivision “Musical settings”; temporal coverage, often via date subdivisions (addressed II.F and III.I); medium of performance (addressed in II.D and III.G); and genre/form (addressed in II.I).

Given the large number of attributes of music materials that have been historically coded as subjects and the changes underway (particularly with genre/form and medium of performance), discovery services need to allow maximum flexibility and be able to accommodate the full range of subject-related fields. Additionally, subject search algorithms should be customizable by individual libraries so they may leverage the data contained in their information systems.

Books about particular musical works are assigned subject headings which consist of the standardized title (including composer, if applicable for the work.) It could be useful to provide functionality for moving from the actual work to materials about the work, and vice versa. The same controlled vocabulary is currently recorded in both cases, but the distinction between the work and materials about it should be made clear by the display.

Recommendation: Index and display subject fields. For data using standardized vocabulary, make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes. Subdivisions are important to music materials, so make it possible for users to view and interact with the entire subject string as a unit. Because current music subject headings frequently include non-topical aspects, many interfaces will choose to combine all these aspects into one index or facet area for the near future. However, this area should be watched closely for updates as data encoding and recording standards change.

II.I. Genre/Form

Users regularly seek works and expressions with particular forms, genres, types, styles, etc. Up until now, these headings, which describe what musical works *are*, have been incorporated into the Library of Congress’ cataloging manuals for topical subjects, and coded the same as topical subjects, using MARC 650 (bibliographic) and 150 (authority). The Library of Congress is currently working with interested groups, including MLA’s Music Genre/Form Task Force, to develop the thesaurus *Library of Congress Genre/Form Terms for Library and Archival Materials* (LCGFT), which will separate form and genre headings from topical subjects and code musical genre/form headings separately from topical headings, using MARC 655 (bibliographic)

and 155 (authority). Discovery services should be able to distinguish between genre/form headings and topical subjects for purposes of searching and faceting or limiting.

While the current practice of coding genre and medium in MARC 650 topical fields will soon change as part of the LC genre projects, this does not address the issue of legacy data. Some libraries may lack the resources required for creating a genre/form index in their local systems or for updating topical 650 headings to genre 655 headings. Moreover, even libraries that do create genre/form indexes may still wish to retain 655 headings in their subject index for patrons who might not make the distinction between “is” and “about.”

Furthermore, while libraries will no doubt rely heavily on LCGFT to provide genre/form access, it remains to be seen whether this resource will adequately describe popular, ethnic, and non-Western art music. To allow maximum flexibility, discovery services should accommodate whatever genre-related thesauri are desired, including those that are locally developed.

Recommendation: Index and display genre/form terms, including displaying codes in the vernacular. Data recording and encoding standards are undergoing major change. Discovery systems need to allow maximum flexibility and decisions must be regularly reviewed and revised to accommodate changes as they occur. For data using standardized vocabulary, make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes.

II.J. Geographic Area

Geographic area of musical works (the geographic area from which they emanate, as distinguished from geographic areas they are about) is a problematic and under-coded attribute.

Geographic area is particularly important for music other than Western art music, and these musics have historically received the most geographic access, primarily through assignment of Library of Congress Subject Headings.¹⁴ However, these subject headings are problematic for facets and limits because they combine information on geographic locale and ethnic/national group, and often make use of subdivisions to create meaning. For example, the heading "Folk music -- Germany" is applied to folk music of Germans in Germany, of other national groups in Germany, and of Germans in other locales. In addition, the same subject headings used for music emanating from a geographic area are also assigned to works about the music of ethnic and national groups, with appropriate subdivisions but the same MARC field tags. For example, the heading "Music -- Uganda" is used for actual music recordings/scores. The same heading is used for a book about Ugandan music, by simply appending the subdivision "History and criticism." Furthermore, it is important that the entire subject heading be displayed and actionable as a string.

Geographic access has been considered less important for Western art music, and geographic subject headings are only assigned if geographic area is a particular focus of the manifestation.

¹⁴ See in particular *Subject Headings Manual* H 1917.

If, in the future, geographic area associated with the composer was regularly recorded (for example, in authority records, as currently allowed for in RDA; MARC 370 has been defined for this purpose), those data could be harnessed for geographic limits and facets.

A final complication of geographic access to musical works is the intellectual task of determining and prioritizing associated geographic areas, since multiple geographic areas may be associated with a piece of music including: birth, death, and residence of the composer (“art” music) or performer (“folk” music), area(s) it was intended for, area(s) it was popular, and area(s) “folk” music is associated with.

See III.K for discussion of geographic area as an expression and manifestation level attribute.

Recommendation: Geographic access to musical works is problematic. Index appropriate fields, and be careful to distinguish between work, expression, and manifestation-level attributes. For data using standardized vocabulary, make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes.

III. Expressions and Manifestations

III.A. Introduction

Musical works frequently exist in many expressions and manifestations,¹⁵ and users seek particular versions of works at the expression and manifestation levels. Users also seek versions with attributes which do not neatly align with FRBR expression or manifestation levels. For example: any vocal score, any arrangement for string quartet, any CD, or any recorded performance by soprano and piano.

A discovery interface allowing navigation through a FRBR Work-Expression-Manifestation-Item (WEMI) tree would be very helpful to music users. However, because the attributes music users seek do not always align neatly with the WEMI levels, further faceting or limiting is necessary in addition to WEMI levels, and, in the absence of a WEMI-formatted discovery interface, can serve many of music users’ needs. To date, practical implementation of full FRBR concepts in discovery interfaces has been limited. One model is Indiana University’s Variations/FRBR project, a testbed, including prototype catalog, for the FRBR conceptual model for music.¹⁶

This section will explore indexing, display, and facets/limits for important attributes of and relationships to musical expressions and manifestations. The following attributes and relationships are addressed:

¹⁵ Here defined according to FRBR: expression: the intellectual or artistic realization of a *work*; manifestation: the physical embodiment of an *expression* of a *work*.

¹⁶ Variations/FRBR: Variations as a Testbed for the FRBR Conceptual Model, <http://www.dlib.indiana.edu/projects/vfrbr/>.

- III.B. Format: Content and Carriers
- III.C. Identifying Numbers
- III.D. Musical Presentation
- III.E. Edition
- III.F. Language
- III.G. Medium of Performance
- III.H. Musical Key/Range
- III.I. Dates
- III.J. Persons and Corporate Bodies
- III.K. Geographic Area

III.B. Format: Content and Carriers

Music as a discipline depends on the use of information in different formats, including scores, sound recordings, videos, and texts. In FRBR terminology, content type (score, audio, video) is usually an expression-level attribute, while carrier type is a manifestation-level attribute. These content types, particularly recordings, exist in many different carriers, including compact disc, LP, audiocassette, videocassette, videodisc, mp3, and other formats. Users need to be able to find, identify, and select the form of information they need. They need facets and limits that enable selection of content types, such as sound recording or video recording, as well as selection of a specific carrier, such as audiocassette, compact disc, LP, streaming sound, DVD, etc. Therefore, systems must allow a single record to be assigned multiple content types/carriers, or for content types/carriers to be grouped together for patron interfaces. For example, an item must be able to be both a “sound recording” and a “CD.” Allowing multiple content types/carriers also is important for packages with accompanying material, such as score/CD combinations. The display should facilitate easy identification of an item’s content type/carrier. Icons are frequently useful for this task.

RDA content type/media type/carrier type fields could be used to create displays and search data. However, RDA carrier type "audio disc" will still be insufficient for distinguishing CDs from LPs.¹⁷ For MARC data, Leader/06, 006/00, and 007 will generally be a better source of this information; use these codes to power searching, displays, and limits/facets for formats and subformats.

RDA’s instructions for creating an authorized access point for an expression (RDA 6.27.3) indicate addition of a term indicating content type (notated music/performed music). Implementation is uncertain, but if implemented, this data might be useful for keyword searching or to create a clickable sublink of the access point string. It would be important, however, not to lose the ability to collocate all expressions of the work, regardless of content type. In addition, other coded fields and MARC 33x already provide the ability to distinguish the content type.

¹⁷ The music library community will be developing best practices which will likely address the issue of providing an eye-readable indication of compact disc in RDA records. One proposed MARC location is 300\$b. This area should be monitored for needed edits to indexes and display as best practices are implemented.

Recommendation: Distinguishing format is crucial for music materials. Users need to be able to distinguish both content type (score vs. audio recording vs. video recording) and particular carriers (CD, LP, cassette, etc.). Carefully evaluate the available data and the discovery system's capabilities in order to choose the best method(s) for distinguishing both content type and carrier. For MARC data, see Appendix C in the Music Discovery Resources pages of the MLA Web site for detailed content and carrier mapping.

III.C. Identifying Numbers

Opus, thematic index, and serial numbers are associated with musical works and are discussed in II.C.

Whereas the ISBN and ISSN serve as fairly uniform standard numbers for books and serials, music materials have a wide variety of numbers associated with expressions and especially with manifestations. These numbers include ISMN, ISBN, ISRC, UPC, EAN, sound and video recording issue numbers, matrix numbers, and plate and publisher numbers.¹⁸ Discovery systems should index and display all these associated numbers.

Plate numbers and publisher numbers are often the key to a successful search for a specific publication of a printed score. Users may seek an exact match of a score for an accompanist, teacher, conductor, or other ensemble members; users may seek a particular error or revision in a score that was printed by a specific publisher from a specific time period; catalogers may want to seek additional clues when the date of publication is not found. Much legacy data contains variations in recording plate numbers and publisher numbers.

Discovery systems often allow searches by indexed plate and publisher numbers. Systems will ideally algorithmically process letter strings, punctuation, and ranges of consecutive numbers so that these numbers can be searched, filtered, and browsed in a meaningful way.

Recommendation: Index and display all associated numbers. Utilize system algorithms to include in indexing all numbers in a consecutive range where only first and last numbers are recorded.

III.D. Musical Presentation

Users need to distinguish between various formats of notated music, such as full score, parts, vocal score, etc. This attribute does not apply to recorded performances of music.

Commonly used cataloging standards and metadata formats and the legacy data encoded via these standards and formats do not provide a clear and concise way to support searching and limiting/faceting by musical presentation. The information recorded under AACR2 and RDA is sufficient to allow informed users to identify the musical presentation when viewing a bibliographic record, but does not lend itself to machine manipulation for facets or limits. MARC coding for Format of Music (008/20, 006/03) is problematic, and mapping musical presentation

¹⁸ ISMN: International Standard Music Number; ISRC: International Standard Recording Code; EAN: International Article Number; matrix number: used to indicate sound recording masters, often used to identify 78s.

to other metadata formats is also difficult. The most promising possibility for musical presentation facets/limits comes from free-floating subdivisions in subject headings, as discussed further below.

Under AACR2, musical presentation information is scattered throughout the descriptive areas of a bibliographic record. It may be found in: physical description (1 vocal score, 1 miniature score, 1 score + 2 parts, and so forth); musical presentation statement (however, this is a transcribed statement and as such its presence, language, and formulation is dependent on the whims of publishers); and possibly in notes, title, and statement of responsibility.

RDA eliminates the distinction between musical presentation and edition statements. Musical presentation information transcribed in MARC 254 under AACR2 will be transcribed in MARC 250 under RDA.

MARC provides the opportunity to code Format of Music in 008/20 and 006/03. On the surface, this might be used to create a facet, but it is problematic for several reasons:

- Code z (music in other than score form) is an amalgamation of formats: works for solo instrument, popular music on two staves for piano with words printed between the staves, with or without guitar chords, and materials consisting solely of parts.
- Parts are ignored when issued with a score. The codes in MARC 008/21, 006/04 are for information regarding whether the entity in hand contains parts. However, OCLC and MLA's MARC Formats Subcommittee have recommended simply accepting the "blank" default so this data is not often coded, at least in WorldCat.
- Code m (several types of scores are issued together). Although the codes for any one type of score allow granularity for a single score, there is no way to encode any one score included in a manifestation with multiple types of scores issued together.
- Codes h (chorus score), i (condensed score), and j (performer-conductor part) were only approved in 2009 for MARC21 and only implemented by OCLC in May 2010; definition of code d "Voice score" was also changed at that time.

Musical presentation is not easily mapped to DC, MODS, or EAD. LC's crosswalk to MODS doesn't seem to support MARC 254. Rethinking of data recording may be necessary; perhaps musical presentation could fit in `physicalDescription` or `typeOfResource`.

The most promising source of musical presentation information comes from subject, rather than descriptive cataloging, via the list of free-floating subdivisions for Written or Performed Music; Performed Version (*Subject Headings Manual* [SHM] H 1160), which provides terms such as "Parts," "Scores," "Scores and Parts," and "Vocal scores with piano" that could be exploited to provide users with facets/limits for commonly sought formats. However, these subdivisions have historically not been comprehensively applied because of restrictions in SHM H 1160 #6, so facets based on musical presentation would exclude particular kinds of music. These subdivisions and RDA extent terms are being revised for the musical works portion of the Genre/Form Thesaurus Project.¹⁹ Preliminary recommendations call for clearly separating format

¹⁹ Genre/Form Headings for Musical Works, <http://www.loc.gov/catdir/cps0/genremusic.html>.

terms from medium of performance. For example, the current subdivision “Vocal scores with piano” would be replaced with “Vocal scores” and the accompanying instruments listed separately in a medium facet. Such changes could make the vocabulary even more promising for musical presentation facets/limits.

Recommendation: Based on the existing situation for musical presentation, best practice is to ensure that necessary fields for identification are indexed and displayed, and to explore ways to exploit legacy data as well as to improve future data capture.

III.E. Edition

Users seek specific “editions” of notated music, meaning usually a search for a specific editor, publisher, or type of edition (critical edition, facsimile edition, Urtext edition, etc.), not (as is common with books), a search for a numbered edition statement (“2nd edition”) or merely the most recent publication. Users may also scrutinize the editor, publisher, and other edition information to select a particular expression/manifestation even if they do not begin their search with a particular edition in mind. Editors frequently make substantial intellectual contributions to notated music, making the editor an important factor to users in selecting an expression. See III.J for recommendations regarding editors and other associated people and corporate bodies.

Recommendation: Many aspects of what music users mean when they talk about “edition” are addressed under other sections, including Medium of Performance (II.D, III.G), Persons/Corporate Bodies (II.G, III.J), Dates (II.F, III.I), Language (III.F), and Musical Key/Range (II.E, III.H). In addition to following recommendations in those sections, index and display edition-related fields in order to provide publisher information and edition statements transcribed from the item itself.

III.F. Language

Three factors contribute to a heightened need to find, identify, and select music materials based on language attributes: the variety of languages used in the study and performance of music, the fact that a single manifestation often incorporates many languages, and the challenges of purely instrumental music.

First, while general users frequently focus on one or perhaps a handful of languages (i.e. languages they read/speak), 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. Consider which of these capacities are most valuable, particularly for facets and limits. Including all capacities will be the most comprehensive route, but can also be confusing, particularly if the function of each language is not clearly identified. In all cases, the function of the language should be clearly identified, if this information is present in the data (as it is in MARC 041).

Finally, purely instrumental music has no language information associated with the work's primary content, yet manifestations feature an array of languages in the secondary content, making indexing and display of language problematic for instrumental music. It is unclear what users think of when considering the “language” of instrumental music. They might associate it with the composer’s language, the language used in the accompanying material, or the language used by the publisher of a particular manifestation. Or, knowing instrumental music does not have a “language” they may just ignore a language facet/limit. In the *MARC Code List for Languages*,²⁰ the code *zxx* is used to indicate “no linguistic content.”²¹ Code *zxx* could be used as a limit/facet; if so, possible display text is “no language content.” Or, it could be ignored; but it should not be replaced with language of the secondary material. Finally, *zxx* should not be confused or combined with code *und* (“undetermined”) which is used both when the language cannot be identified and when sung or spoken text is “vocalises, humming and other texts that are wordless or consist of nonsense syllables.”

Recommendation: Index and display language codes and terms, converted to vernacular when necessary. Display and allow possibility for additional limits/facets for secondary languages and indicate function of the language (e.g., subtitles, librettos/lyrics, performance instructions, original language). At local level, determine desired use and label of MARC language code *zxx* (no language content).

III.G. Medium of Performance

Users may seek expressions of works with a particular medium of performance other than the original medium, i.e. arrangements. Users may also seek music for a particular medium of performance regardless of whether it is the original medium. Both AACR2 and RDA dictate use of the original medium of performance when included in standardized titles. Medium of the expression is given in both coded and textual fields; see II.D for further detail.

Recommendation: Follow indexing and display recommendations given for Medium of Performance in II.D.

III.H. Musical Key/Range

Musical key and range (complete span of pitches used in the piece of music) are particularly important for vocal music, where works frequently exist in a variety of keys and ranges, but instrumental music can also exist in expressions distinguished by the key or range. Users would also benefit from the ability to search by more specific ranges, such as the exact pitches included, to match their ability and available resources. For example, some instruments like carillon or handbell choir do not have standard ranges, and vocalists’ ranges are a combination of physiology and training. However, key and range information is not currently or historically recorded under AACR2 nor encoded in MARC data on any regular basis. It may be included in a

²⁰ MARC Code List for Languages, <http://www.loc.gov/marc/languages/>.

²¹ Both N/A and three blanks have previously been used to designate “No linguistic content.” Although OCLC converted all instances to *zxx* in 2006, N/A and three blanks may still exist in other data sources. There was also no reliable way to convert incorrect codes for instrumental music coded for language of imprint.

note or edition statement, particularly if a statement about key or range is present on the manifestation itself.

The key in which a musical work was originally composed and its original range are work-level attributes and addressed in II.E.

Recommendation: Musical key and range are important, but because of historic data recording and encoding practices, discovery systems will not be able to do much more than provide general keyword indexing and display of the appropriate fields. More dynamic searching capabilities should be explored in the future.

III.I. Dates

Like musical works, musical expressions and manifestations have many associated dates. Discovery interfaces commonly allow searching, limiting, and faceting by date, most frequently exploiting publication date, which is regularly recorded in bibliographic data in an easily machine-actionable form. However, other dates often hold equal or greater importance for music materials. Unfortunately, these dates may be buried in notes or completely absent from bibliographic data. One particular limitation for MARC is the fact that MARC 008/07-14 only has space to record two years. Ideally, date indexing, faceting, and limiting should take into account the full spectrum of dates associated with music materials. When this is not feasible, “date” searches/limits/facets should at a minimum clearly identify the type of dates being exploited; for example, using the label “Publication Date” rather than just “Date,” if only publication dates are included.

Creation date is associated with musical works and is discussed in II.F.

Performance date is an important expression-level attribute for recorded music. Users may seek particular performances, or they may use performance date to select an expression that meets their needs. Performance date is often included in bibliographic data for recordings, though it is usually buried in notes. For popular musics, chronological subdivisions are frequently added to genre headings indicating the decade of the expression.

Publication date is primarily important as an access point for manifestations of notated music. For notated music, creation date for the expression will often have to be approximated by the date of publication of the first manifestation of that expression.

Expression date is occasionally included in standardized titles. In addition, past practice was to add the manifestation date routinely to standardized titles, and this use may still be present in some legacy data.

Recommendation: Many dates are important to music materials, and because of this variety, it is easy for discovery interfaces to be misleading regarding dates. For the present, best practice is to index and display all dates, consider the dates included in the index(es), and give an appropriate label. Also creatively explore ways to tease out the different kinds of dates and present them to the user in a useful way.

III.J. Persons and Corporate Bodies

In addition to the composer and librettists/lyricists, who are related to musical works, many additional persons and corporate bodies contribute to music materials and hold responsibility at the expression and manifestation levels, especially the expression level. The recommendations in this section build on those in II.G. and focus on the unique requirements posed by the many different people related to expressions and manifestations. These related persons and corporate bodies include: editor, arranger, performer (both individuals and groups), producer, and director.

These persons and corporate bodies are important access points for music users for finding, identifying, and selecting their desired entity. Therefore, it is important to index and display all personal and corporate names, both transcribed forms and controlled vocabulary. If codes or terms are present in the data which designate the person/corporate body's function, these should be indexed and displayed, with codes converted to the vernacular.

Western art music generally makes a clear distinction between composer (related at the work level) and performers (related at the expression level). Many other musics do not make this sharp distinction. For discovery purposes, following the recommendations in this document will make persons and corporate bodies related to the work, expression, or manifestation useful for finding, identifying, and selecting music, regardless of the exact nature of the relationship.

Recommendation: Follow recommendations in II.G. Of particular importance to expressions and manifestations, if codes or terms are present designating function, these should be indexed and displayed, with codes converted into the vernacular.

III.K. Geographic Area

While potentially useful, geographic area is arguably a less important attribute for expressions and manifestations than it is for musical works.

For recordings, geographic area of the expression is often included in a note; however, this is a free-text field not designed for machine manipulation. This note (which may contain other information about the recording, particularly the date) should be indexed in general keyword indexes and displayed. For scores, geographic area of the expression is a more abstract concept and infrequently contained in bibliographic data. MARC 033 is defined to record date/time and place of event, but is not regularly coded.

Geographic area of the manifestation is regularly recorded in the publication statement and, in MARC, coded in 008/15-17, providing an easily machine-actionable field for creation of facets or limits. For books published in their original language, geographic area of manifestation may be a somewhat successful approximation of geographic area of the work. For musical scores and recordings, this is not true. Care should be taken that any facets, limits, or displays created from geographic area of manifestation are clearly labeled as referring to place of publication.

Recommendation: Display notes relating to geographic area of expressions and manifestations, and include them in general keyword indexes. If place of publication is used to create facets, limits, or displays, ensure that it is clearly labeled as such.

IV. Other Aspects of Music Discovery

IV.A. Introduction

Several issues related to music discovery do not focus on particular attributes or relationships. The following areas are discussed in this section.

- IV.B. Authority Records
- IV.C. Compilations
- IV.D. Searching: Alphabetical and Keyword
- IV.E. Enhancements: Third-Party Content
- IV.F. Music-Specific View

IV.B. Authority Records

For music materials, authority records are essential to back end functions like cataloging. In addition, they are also extremely important for public interfaces in that they provide valuable cross-references and other information to users.

Creators—whether composers, lyricists, librettists, or even corporate bodies—are important access points for music. So too are associated contributors such as performers, arrangers, and editors. Because a single creator or contributor might be known by more than one name, a method should be provided to lead users from alternate names to the forms they seek.

Musical works also present a challenge. For instance, a work might be associated with more than one creator (as in the case of works previously attributed to another), or it might be known by different titles (possibly in different languages) or multiple work numbers. Outside of cataloging codes the choice of a title’s language, elements, and grammatical construction is far from consistent. Therefore, users may begin a search with any one of many possible titles for a known musical work, and discovery interfaces should provide a method to find the musical work even if the user starts with one of these alternate titles. The same is true for expressions and manifestations, which might be issued with new titles or differ from the original work in medium of performance or key, as is often the case with arrangements. Similarly, topical subjects or musical genres can also be represented by varying terminology.

Historically, leading the user from alternate terms to the preferred one was accomplished by “see” and “see also” references (MARC 4xx and 5xx authority record fields) in author, title, or subject browse lists. Unfortunately, keyword searching has mostly ignored “see” and “see also” data. Mere spell check features will not solve this problem, though they are potentially helpful. Auto-complete or “did you mean” features are useful, particularly when based on database

contents or rich, authoritative sources of alternate name and title data such as the Library of Congress/NACO Authority File (LC/NAF) or the Virtual International Authority File (VIAF).²²

More recently, the MARC authority format has expanded to include 7xx fields for equivalent access points found in different thesauri or authority files. The use of linked data, which relies on an identifier rather than a text string, has made possible initiatives like the VIAF; when employed in a search interface, linked data could provide a seamless experience where patrons enter search terms in their preferred language, script, or form and retrieve the desired results automatically.²³ Future systems could allow the library to choose which heading string to display to its patrons, even if it is a “see” reference. For example, a library might choose to display all titles in the local language rather than the composer’s original language. Future systems might also allow the library to manipulate various elements which identify the work in other ways to help the user identify or select the desired work, such as producing a list of a composer’s oeuvre organized by opus number.

Finally, authority records contain a wealth of other information that potentially could be used for display, pre-search limiting, or post-search facets. Such information includes the MARC 680 public “scope” note, the MARC 043-046 fields for geographic and temporal information, the MARC 336 and 37x notes for content type, associated place, occupation, gender, language, etc., and, specifically for music, the MARC 380, 382, 383, and 384 fields for form of work, medium of performance, numeric designation, and key. Current authority records are very similar to work records in a FRBRized environment; in the future, many important attributes (see the entirety of Section II. Musical Works) will ideally be placed in work records rather than bibliographic records, making indexing and display of work or “authority” data absolutely crucial to music information retrieval.

Recommendation: At a *minimum*, in back end functions discovery systems should be able to index MARC 1xx, 4xx, 5xx, and other desired fields and also to display all fields in the authority record. For public interfaces, browse indexes should display 4xx and 5xx cross-references and public notes such as the 680. Associated name and title strings must be kept together for both indexing and display.

Ideally, as discovery systems evolve, they should be able to index and display cross-references in keyword indexes and allow linking or display of alternate data. They should leverage authority information to provide autosuggestions, context-sensitive recommendations, or other functionality. Additionally, they should be able to make use of extended authority fields for display or faceting. All options should be customizable by individual libraries.

As authority records morph into work records, indexing and displaying the data they contain will become even more crucial.

²² See, for example: Demian Katz, Ralph LeVan, and Ya’aqov Ziso, “Using Authority Data in VuFind,” *Code4Lib Journal* 14 (July 25, 2011), <http://journal.code4lib.org/articles/5354>.

²³ The Library of Congress has made available as linked data many of its thesauri, including LCSH and the Genre/Form Terms, on its Authorities and Vocabularies web site, <http://id.loc.gov/>.

For data using standardized vocabulary, make it possible for users to link from within the record display from standardized vocabulary terms to other materials associated with the same attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes.

IV.C. Compilations

Multiple expressions are frequently issued together in a single manifestation, sometimes called an aggregate.²⁴ This is particularly common for recordings, which often contain expressions of more than one musical work, each of which might be by a different composer, for a different medium, of a different genre, performed by different performers, or based on different works. It also occurs with notated music, e.g. song anthologies or collections of all or some works by a single composer. Under most past and current descriptive and encoding standards, the relationships in multi-expression manifestations (a) between works; and (b) between works/expressions and their performers, medium of performance, or format of notated music are often recorded in notes designed for human interpretation but are not linked together in a manner friendly to machine processing.

In future data creation, this problem could be solved by creating WEMI record trees and/or by utilizing linking fields (such as MARC \$8, which could solve this problem but has not in reality been implemented for this purpose). However, because of the lack of linking information, this problem cannot easily be solved by discovery interfaces utilizing legacy data from MARC and AACR2. The Variations/FRBR project opted in favor of recall over precision and mapped data elements in a bibliographic record to all works/expressions on a manifestation, given the absence of machine-readable information regarding which work/expression the data element applied to.²⁵ The result is that, for example, all performers and subject headings are associated with all works/expressions contained on a recording.

Recommendation: Be aware of and explore solutions to the special challenges posed by the frequency with which compilations occur for music materials, particularly that legacy data rarely links together all the attributes of the various expressions contained in a compilation in a manner friendly to machine processing.

IV.D. Searching: Alphabetical and Keyword

Historically, alphabetically filed “browse” listings have been indispensable in library information retrieval, beginning with the card catalog, and continuing in “traditional” OPACs. OPACs introduced the additional functionality of keyword searching. “Next-generation” catalogs added faceted browsing, but many simultaneously ceased to provide alphabetical browse searching.

²⁴ “Aggregate” may be defined as “a manifestation embodying multiple distinct expressions” and may represent a collection based on genres or performers, or may be an apparently “random collection of expressions.” Working Group on Aggregates, “Final Report of the Working Group on Aggregates” (2011): 3. <http://www.ifla.org/files/cataloguing/frbrg/AggregatesFinalReport.pdf>.

²⁵ Jenn Riley, “Leveraging the FRBR Data Model for Music Discovery and Data Sharing: Autobiographical Note,” *OCLC Systems and Services* 27, no. 3 (2011): 186.

Alphabetical listings have been particularly important to music librarianship. “Known-item” queries are more common in music searching than in general searching; this is true of both “classical” and “popular” music.²⁶ However, users do not always “know” a lot about the entity they are seeking. Title is a notoriously unreliable access point, especially for Western art music. “Subject” is a problematic concept; music subject headings often reflect non-subject attributes such as form, genre, and instrumentation. Perhaps because of these difficulties, personal names are frequently used as an access point in music searching.²⁷ However, many composers are very prolific, and their works exist in many versions, making personal name alone an insufficiently precise access point.

Author, title, and subject, are, alone, insufficient access points for music. This deficiency led to extensive development of uniform titles (especially generic uniform titles) and subject headings (which in reality reflected true “aboutness” only occasionally) to provide access to music materials based on attributes beyond author, title, and subject. When organized into alphabetical lists, they made it possible for informed users to reliably find and browse musical works. Alphabetical lists facilitated browsing in this complex system because users were not required to determine and input keywords to begin a search. A significant weakness of most current keyword searching is that it does not take advantage of alternate forms found in LC/NAF authority records. These alternate forms are indispensable in music searching, given the vagueness of even “known-item” queries. It is extremely difficult for untrained users to predict the vocabulary that will be used in uniform titles and subject headings, and current systems rarely provide much guidance. See IV.B for further discussion of authority records and keyword searching.

Recommendation: Consider carefully the options for alphabetical and keyword searching. Current keyword searching has significant deficiencies for music. However, the needs which have historically been met through alphabetical listings might be met through other means, particularly faceted browsing based on attributes important to music. Recommendations have been given throughout the document. Alphabetical searching has particular value for librarians and staff, and should be retained in back-end interfaces even if it is eliminated from public interfaces.

IV.E. Enhancements: Third-Party Content

The ability to incorporate third-party content is an important enhancement to discovery systems. Existing data from vendors, publishers, and other content providers can be leveraged to increase the information available in and through a discovery system.

Content from outside a bibliographic database can be either linked or incorporated. Links to external content have been included in traditional MARC bibliographic records for some time:

²⁶ Beth Christensen, Mary Dumont, and Alan Green, "Taking Note: Assessing the Performance of Reference Service in Academic Music Libraries: A Progress Report," *Notes* 58, no. 1 (September 2001): 52. Jin Ha Lee, "Analysis of User Needs and Information Features in Natural Language Queries Seeking Music Information," *Journal of the American Society for Information Science and Technology*, 61, no. 5 (2010): 1037.

²⁷ David M. King, "Catalog User Search Strategies in Finding Music Materials," *Music Reference Services Quarterly* 9, no.4 (2007): 17.

URLs pointing to digitized versions of tangible resources, movie reviews, tables of contents, program notes, or finding aids.

Incorporated content may come from any number of sources and can include cover art, reviews, summaries, previews of initial/selected pages, or other content. In some cases content may be purchased from a vendor such as Syndetic Solutions. External content can also be incorporated from sources that gather content, provide content directly, or sell products, such as LibraryThing, Open Library, AllMusic.com, Amazon.com, iTunes, Wikipedia, IMDb, or Google Books; additional enhancements may be in the form of user-contributed tags or reviews. Looking to the future, it may be possible to incorporate audio or video clips, and systems that are able to take advantage of such a service provide music users with a greatly increased capability to determine if the item in question suits their needs.

For music formats, particularly scores, it may be difficult to incorporate content for a number of reasons: information may not be available online; in some cases the only available information may be a short description from a publisher or vendor, with no cover image available; reviews may only be published in subscription-only journals; and there may not be a way to make a reliable match to third-party content. For scores, initial pages of music will be much more useful than the cover to help users identify the work and evaluate whether it meets their needs as to style, difficulty level, and similar considerations. Even when content is available, a service that uses only ISBNs to extract content will not function for items that do not have an ISBN, as is often the case with scores and sound recordings. Creating match points on standard numbers in addition to ISBNs, such as publisher numbers, recording issue numbers, or ISRCs would allow for enrichment of records for music formats in cases where data is available.

Most importantly, third-party content must enhance, not replace, existing data. This is a risk when categories are duplicated across data sources but the data are different. Sound recordings are particularly at risk with the various versions of genres and “contents” that may exist across data sources. For a given item, cataloger-supplied contents notes and CD track listings are very different (tracks may simply be identified as “Allegro”), titles and track listings may vary in clarity greatly depending on source of the data, and specific functions of composers, performers, or other contributors with different functions may not be reliably identified (for example, there may be a single column labeled “performer/composer” containing only a single name without explanation). This information should not necessarily be blocked from catalog records--in some cases third-party data may be the only source of contents, or provide additional information of interest to users--but it must not replace cataloger-supplied notes. Furthermore, identifying the sources of data may reduce confusion where fields are duplicated.

Recommendation: Display and maintain links within bibliographic records, with ability to suppress links to inaccessible material such as subscription-only access an institution does not hold. Allow local customization and interoperability with providers of third-party content to incorporate elements not present in or linked from bibliographic records. Ensure that match points are reliable. Identify third-party content with its source, either by category (Source A genre; Source B genre) or by source (Source A: genre, contents, etc.). Do not hide or replace content in bibliographic records.

IV.F. Music-Specific Interface/View

As discussed throughout this document, music materials possess many attributes that are unique or uniquely important. In addition, while in the bibliographic universe as a whole, the norm is for each work to be realized by only one expression realized by one manifestation, for musical works, multiple expressions and manifestations are the norm. Because of these factors, it is worth considering whether a discovery tool should have a music-specific view that highlights these particular attributes and relationships, employing facets, limits, and displays that aid in the finding, identification, and selection of music materials. A separate music interface could be particularly beneficial for tools designed to discover collections containing large amounts of music. However, other factors should be considered, particularly the time and effort needed to maintain a separate interface/view. In addition, if an alternate interface/view is created, the general interface should be designed to lead users seeking music materials intuitively to the alternate interface/view. Recommendations given throughout this document can be implemented in both general and music-specific interfaces/views.

Recommendation: Consider the pros and cons of a music specific interface or view in any discovery tool implementation.