Music Discovery Requirements

Version 2 (2017)

Created by the Music Discovery Requirements Task Force
Music Library Association

Endorsed by the Music Library Association Board of Directors xx/xx/yyyy

DRAFT - for public comment
July 7, 2017
Comments due August 7, 2017

Comments welcome in any form, but especially encouraged on Google doc at:
https://docs.google.com/document/d/1e1Mvo5SP7ywZFvZsv0csdbRZHLvyCMKKFzo2Jxq
FRo/edit?usp=sharing
or, email Task Force Chair Nara Newcomer newcomern@umkc.edu
Appointed by the Music Library Association Board of Directors
Formed jointly by MLA’s Emerging Technologies and Services Committee and the Cataloging and Metadata Committee.
Nara Newcomer, University of Missouri-Kansas City, Chair
Anna Alfeld LoPrete, Indiana University
Rebecca Belford, Brown University
Jessica Harvey, University of South Carolina
Christopher Holden, Library of Congress
Keith Knop, University of Georgia
Nancy Lorimer, Stanford University
Karen Peters, Bates College
Patricia Sasser, Furman University, webmaster

Music Discovery Requirements Version 1 Group (2011-2012)
Created by the Music Library Association’s Emerging Technologies and Services Committee.
Endorsed by the Music Library Association Board of Directors, Aug. 8, 2012
Nara Newcomer (leader), East Carolina University
Rebecca Belford, University at Buffalo
Deb Kulczak, University of Arkansas
Jennifer Matthews, University of Notre Dame
Misti Shaw, DePauw University
Kimmy Szeto, Maritime College, State University of New York
Executive Summary

This document explores discovery needs specific to and especially important for music materials, particularly scores and recordings. Music materials pose unique demands that must be considered for successful discovery. For example, most books are published only once; for musical works, however, multiple versions (for example, score, parts, and recordings of different performances, arrangements, and transcriptions) are the norm. These different versions also often have different titles (different languages or grammatical formulations), making titles transcribed from items much less useful for finding and identifying musical works. Music materials also possess unique attributes not found in books, such as instruments/voices used to perform the music. Some of the unique needs posed by music materials can be solved simply by ensuring that needed fields are appropriately displayed and indexed in discovery interfaces. Other problems are more difficult to solve. This document discusses the issues outlined in the table of contents and when possible gives concrete recommendations for display and indexing. Two appendices compile technical details of the specific indexing recommendations in spreadsheets.

This document employs technical language, and the target audience is those creating or guiding the development of discovery interfaces that will include music materials. Developers of future recording and encoding standards will also benefit from reading this document and creating solutions to facilitate better discovery interfaces. 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 where there is particular need for fuller, more consistent data. Given that most libraries will be dealing with large bodies of data encoded in MARC and recorded in RDA or AACR2, particular attention is paid to these standards, as well as looking to the future of BIBFRAME and linked data. The document also includes indexing recommendations for non-MARC metadata standards. As data recording and encoding changes, specific recommendations in this document will need revision in order to achieve and maintain this document’s underlying principles in an evolving environment.
Table of Contents

Executive Summary
Table of Contents
I. Introduction
   A. Introduction to MDR2 (2017)
II. Musical Works
   A. Introduction
   B. Titles
   C. Identifying Numbers
   D. Medium of Performance
   E. Musical Key/Range
   F. Dates
   G. Persons and Corporate Bodies
   H. Genre/Form
   I. Topical Subjects
   J. Geographic Area (and Ethnic/National Group)
III. Expressions and Manifestations
   A. Introduction
   B. Format: Content and Carriers
   C. Identifying Numbers
   D. Format of Notated Music
   E. Edition
   F. Language
   G. Medium of Performance
   H. Musical Key/Range
   I. Dates
   J. Persons and Corporate Bodies
   K. Geographic Area (and Ethnic/National Group)
IV. Other Aspects of Music Discovery
   A. Introduction
   B. Authority Records
   C. Compilations
   D. Searching: Alphabetical and Keyword
   E. Phrase searching
   F. Music-Specific Interface/View
   G. Enhancements: Third-Party Content
V. Metadata Schemas
   A. Introduction
   B. MARC
C. Dublin Core
D. EAD
E. EBUCore
F. MODS
G. PBCore
H. BIBFRAME 2.0
VI. Schema Versions Used
Appendices
Works Cited
I. Introduction

A. Introduction to MDR2 (2017)

The discovery landscape shifted rapidly in the years after the Music Discovery Requirements’ 2012 release.¹ RDA replaced AACR2 as the major library cataloging standard. The Library of Congress launched the Library of Congress Medium of Performance Thesaurus (LCMPT) and Library of Congress Genre/Form Terms for Library and Archival Materials (LCGFT), providing stand-alone vocabularies for these attributes. BIBFRAME development is solidly underway, and linked data is an emerging library reality, not a future idea. Though music users’ needs are relatively unchanged, expectations continue to rise and technological shifts have opened new discovery possibilities.

In this environment, the Music Library Association (MLA), at the joint request of the Emerging Technologies and Services Committee and Cataloging and Metadata Committee, appointed a task force to revise the Music Discovery Requirements with the following charge:

Update the Music Discovery Requirements, with particular focus on: RDA updates; MARC updates; incorporating LCGFT and LCMPT, and updating appendices B & C (appendix A will be dropped.) Strongly consider creating an HTML version, and evaluate whether a pdf version needs to be retained. Set a mechanism in place for ongoing revisions to the MDR.

The task force began work in summer 2016. [Brief additional details of timeline once they have occurred: reviews, comment periods, publication.]

This 2017 revision of the Music Discovery Requirements (MDR2) reflects changes in encoding and recording standards since 2012. LCMPT and LCGFT are now incorporated throughout. MDR2 focuses on RDA with considerations for AACR2 legacy data. MARC and all metadata schemas were carefully reviewed for updates. MDR2 drops three metadata schemas (VRA Core and CDWA Lite, due to low use for music materials; EAD due to considerations detailed in section V), adds two schemas (PBCore and EBUCore)² and introduces discussion of

---


² A future version of the MDR will incorporate any changes based on RDA adoption of IFLA-LRM as a conceptual model. The first changes are expected in April 2018, per the RDA Steering Committee: RDA Steering Committee, “Implementation of the LRM in RDA,” http://rda-rsc.org/ImplementationLRMinRDA (accessed 23 March 2017).
BIBFRAME, including some basic mappings.\(^3\) The chart below details the metadata schemas and versions in MDR1 and MDR2.

MRD2 retains MDR1’s scope of music and discovery, focusing on the unique discovery needs of musical works (scores and recordings) rather than secondary literature about music (books and articles), and on discovery rather than back end functions such as circulation, cataloging, and acquisitions. This continued focus on discovery means recommendations are focused on indexing and presentation of data for discovery, rather than rules and best practices for recording data or metadata formats and encoding standards. The organization is likewise unchanged, consisting of three FRBR-aligned sections and subsections consisting of: prose discussion; prose summary recommendation; best practices for indexing, display, and facets/limits; and related MARC authority fields. MARC field labels are given as currently defined, even when they do not align with RDA terminology. Library data standards continue to privilege Western art music, but strides have been made towards serving other musics, and those changes are incorporated as applicable to discovery. Interfaces facilitating discovery for large amounts of music beyond the Western art music tradition may benefit from further customizations, if data has been encoded to support such customizations.

MDR2 contains three appendices: Appendix A compiles technical details of the recommendations in spreadsheet form. Appendix B details MARC bibliographic record mapping for content and carrier. Note that MDR2 drops MDR1’s Appendix A: Compiled Details of Indexing and Display Requirements (Index Focused). The tag focused spreadsheet (Appendix B in MDR1) can meet the needs served in Appendix A. Therefore, in MDR2, the tag focused version is now Appendix A and MARC Bibliographic Record Mapping for Content and Carrier is now Appendix B. The new Appendix C is not a mapping document but a spreadsheet outlining history of the coding and definition of Format of Notated Music, to be used in conjunction with III.D Format of Notated Music.

The MDR is now available in both pdf and an easily navigable HTML version. [The HTML version will be put in place after review of drafts.]

MLA’s Emerging Technologies and Services Committee (ETSC) will initiate (by forming a working group) future revisions to the MDR, informed by a running list relevant changes to content standards, encoding standards, and vocabularies which MLA’s Cataloging and Metadata Committee (CMC) will maintain and share with the ETSC annually in September. Full maintenance plan details are posted to the Music Discovery Resources section of the MLA web site.\(^4\)

\(^3\) BIBFRAME mapping for discovery depends a great deal on context, vocabularies used, and local implementation decisions. The mappings provided here are not complete, and are intended only as guidance about where to start. See Section V.H. for links to current information.

\(^4\) http://www.musiclibraryassoc.org/?page=mdr

Music Discovery Requirements 7
II. Musical Works

A. Introduction

Clear identification and display of information regarding the musical work 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:

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

B. Titles

Users' discovery and identification of musical works frequently draws on titles, often in conjunction with creator. Particular musical works are often referred to by many different titles in various languages. For example, Beethoven’s Fifth Symphony might also be called Symphony no. 5, Sinfonie C-Moll, or Symphonie op. 67. This reality spurred librarians’ extensive development and application of standardized titles (RDA’s “authorized access point representing the musical work”; AACR2’s “name/uniform titles”) for music materials. To facilitate identification of musical works, it is crucial to display the entire standardized title. 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.

The current environment focuses on creating a standardized text string (RDA’s “authorized access point representing the musical work”; AACR2’s “name/uniform title”) 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 linked data functionality might be used to represent the work and its attributes. 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 (composers), and the composer’s name is often essential to identify the work and distinguish it from works with the same or similar titles. Therefore, display the creator name in conjunction with the standardized title. However, do not include creator name in title indexes. One specific problem 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 for musical works. 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. Some musical work titles are not associated with names and are contained in MARC 130 and 730.

While having a separable link to creators is desirable (see II.G), creator and title must be considered together in any literal-string-based search tool to function properly. Remember that music is frequently encountered in compilations containing numerous other works, often by multiple composers, and a match on creator and title separately in the same record is no guarantee of a correct match. Search tools that harness URIs to identify works unambiguously do not have this limitation.

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, display these these transcribed titles and include them in title keyword indexes. 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 ($x$0$1$5$8). A special case is $i Relationship information: RDA encourages the use of relationship designators such as “Opera adaptation of (work)” and “Container of (expression)” to make relationships between works explicit in a way that is human readable while providing anchor terms for future machine manipulation. Subfield $i should display but should not index (or be hyperlinked) as part of the creator name or title. Additionally, the parenthetical WEMI terms are meaningless jargon to most patrons; consider options to display only the main body of the relationship designator and suppress the parenthetical terms.

Ensure that standardized title strings are associated with their creators; when $e is present, it should display but not function as part of the linked text string. 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 standardized vocabulary terms within the record display to other materials associated with the same attribute/entity. This could be accomplished through use of bound text strings for fully authorized terms or via

---

5 This is true for music in the Western Art music tradition, but less true for other musics such as “popular” and “folk” musics.
identifiers functioning behind the scenes. In linking and facet creation, and in relevance ranking, privilege standardized title or name/title fields over transcribed titles.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 130 Main Entry--Uniform Title
MARC: 240 Uniform Title (display all subfields; in particular, associate with names in 100, 110, 111)
MARC: 245 Title Statement
MARC: 246 Varying Form of Title
MARC: 380; 382; 383; 384 Form of Work, Medium of Performance, Numeric Designation of Musical Work, [Musical] Key (keyword)
MARC: 505 $t Formatted Contents Note Title ($a could also be added to include titles found in unenhanced contents notes. However, this would also add many creators/contributors/authors to the title index. Consider local data to make a decision balancing precision and recall.)
MARC: 700; 710; 711 Added Entry--Personal Name, Corporate Name, Meeting Name (only include title subfields, $k$l$m$n$o$p$r$s$t, in title keyword indexes but display all subfields intended for public display) (in particular, associate name and title subfields)
MARC: 730 Added Entry--Uniform Title
MARC: 740 Added Entry--Uncontrolled Related/Analytical Title
MARC: 800, 810, 811, Series Added Entry--Personal Name, Corporate Name, Meeting Name (display all subfields intended for public display) (only include title subfields, $k$l$m$n$o$p$r$s$t, in title keyword indexes but display all subfields intended for public display) (in particular, associate name and title subfields.)
MARC: 830 Series Added Entry--Uniform Title
Dublin Core: title
EBUCore: Title, abridgedTitle, alternativeTitle, identifier, mainTitle, originalTitle, publishedTitle, translationTitle, workingTitle
MODS: titleInfo + type, authority (for name+uniform title access points such as those defined in the NAF, use matching nametitlegroup attributes in titleInfo and name to link creators and titles; this also ensures the uniform title is associated with the correct name if there are multiple names in the record)
PBCore: pbcoreTitle, pbcoreIdentifier (evaluate data for whether musical work numbers have been recorded in pbcoreIdentifier), pbcoreDescription (for contents of compilations or collections); pbcoreAssetType provides added context for determining if a title pertains to a recording or its contents

**Facets/Limits (Bibliographic/Descriptive Metadata):**
Musical work titles do not lend themselves to pre-search limits.
Post-search facets based on musical work title could be useful, but their long strings make them difficult to employ in the limited real estate usually available for facets. Any such facets should use standardized title fields, not transcribed titles. Facets by medium of performance (which
Currently is frequently included in authorized access points/uniform titles) could be useful; see II.D and III.G for further discussion.

**Related MARC Authority Fields:**
MARC: 100; 110; 111; 130; 380; 382; 383; 384

**Related MADS/RDF Authority Fields:**
MADS/RDF: Classes: madsrdf:Authority, madsrdf:NameTitle, madsrdf:TitleElement, madsrdf:PartNumberElement, madsrdf:MainTitleElement, madsrdf:SubtitleElement, madsrdf:LanguageElement; Properties: madsrdf:hasCharacteristic, madsrdf:elementValue, madsrdf:authoritativeLabel, madsrdf:isIdentifiedByAuthority, madsrdf:identifiesRWO

### 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, D. Scarlatti) where multiple work number systems exist.

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. See IV.E for special phrase searching considerations surrounding the English “no.” (as in “Symphony no. 5”) in systems which treat it as a Boolean “NOT” operator.

Keyword searching by opus, thematic, and serial numbers can be particularly problematic because of abbreviations used frequently in library cataloging and to varying degrees in the broader world. For example:

- no., number, Nr., Nummer
- op., opus
- BWV, S, Schmieder (for J.S. Bach)
- K, KV, Köchel (for W.A. Mozart)

It is difficult for uninformed library users to predict which term will appear in the data. Various solutions are possible:

- Automatically include alternate possibilities; however, care is needed when some of the possibilities are actual words (like “no” which additionally may be a stop word) or single letters (like “S”), because this could greatly decrease precision. Make the adjustment transparent, to match expectations created by common search engines (“showing results for Y, search instead for X”).
- Program the interface to use “did you mean” when the user’s search contains a non-preferred term. (A user types “opus 18” and the interface delivers results for “opus 18” but with a message “did you mean op. 18?”.)
Integrate search tips instructing users to include only the numeral in keyword searches for work numbers. (“Numeral only; omit op., BWV, etc.)

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. Evaluate best way to address the numerous abbreviations often associated with opus, thematic index, and serial numbers.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 130 $n; 240 $n; 700 $n; 730 $n Number of part/section of a work, including work numbers for music materials.
MARC: 383 $a$b$c$d$e Numeric Designation of Musical Work
Dublin Core: identifier
EBUCore: partNumber
MODS: identifier
PBCore: no field defined for musical work identifying numbers; pbcoreIdentifier is more likely to refer to a show, segment, or recorded performance than the musical work itself
BIBFRAME 2.0: Classes: bf:Iswc; Properties: bf:musicOpusNumber, bf:musicSerialNumber, bf:musicThematicNumber, bf:partNumber

**Facets/Limits (Bibliographic/Descriptive Metadata):**
Work numbers do not function well for pre-search limits. They could be used for post-search facets, but in situations where real estate is limited for post-search facets, they can be difficult to employ and may be best omitted as facets.

**Related MARC Authority Fields:**
MARC: 100 $n; 130 $n; 383 $a$b$c$d$e

**Related MADS/RDF Authority Fields:**
MADS/RDF: madsrdf:PartNumberElement, madsrdf:elementValue

**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. The Library of Congress Medium of Performance Thesaurus (LCMPT) was only launched in 2014.\(^6\)

In current practice, medium of performance information is/may be recorded in MARC 382 using terms from LCMPT.\(^7\) 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 (245, 246, 250, 500, 505, 511). Descriptive fields are particularly important for recordings, especially for jazz and popular music, where the only explicit medium of performance information may be in a listing of performers and their instruments in field 511. In the future, legacy data might be massaged so that medium of performance contained in 650 is moved to 382.

Rules governing input of certain fields or elements within LCSH and RDA authorized access points representing the musical work/AACR2 uniform titles add to the problem of limited or dispersed information. 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 of performance.

Going forward, configuring discovery interfaces to utilize LCMPT vocabulary encoded in MARC 382 is essential, as is massaging legacy data to ensure uniform presence of the field. 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. MARC 048 is also problematic, particularly for recorded popular, jazz, and folk music when some, but not all, mediums can be identified. Now that 382 is defined and in use, strongly consider suppressing 048 when 382 is present, and focus development efforts on 382 data.

Challenges relating to medium of performance are further complicated by the prevalence of manifestations containing 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.

---


7 MARC 382 was defined in MARC Update 14, April 2012. [https://www.loc.gov/marc/up14bibliographic/bdapndxg.html](https://www.loc.gov/marc/up14bibliographic/bdapndxg.html) Further changes to 382 were made in MARC Update 20 (added $e Number of ensembles (R) and MARC Update 22 (added $r Total number of individuals performing alongside ensembles (NR); added $t Total number of ensembles (NR); renamed $e to Number of ensembles of the same type). MLA’s Best Practices documents do not yet reflect the most recent additions, but will be updated to do so in the future.
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. For detailed display recommendations for MARC 382 Medium of Performance information, see the WorldCat Discovery Preferences for Medium of Performance from the Music OCLC Users Group.

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: 048 Number of Musical Instruments or Voices Code (display code in vernacular; use only if 382 is not present)
MARC: 130 $m; 240 $m; 7xx $m Medium of performance for music
MARC: 382 Medium of Performance. See WorldCat Discovery Preferences for Medium of Performance from the Music OCLC Users Group for detailed recommendations.
MARC: 505 Formatted Contents Note
MARC: 511 Participant or Performer Note
MARC: 650 Subject Added Entry - Topical Term
Dublin Core: type
EBUCore: hasGenre, hasSubject (if using LCSH)
MODS: subject (if using LCSH)
PBCore: pbcoreSubject (if using LCSH)

**Facets/Limits (Bibliographic/Descriptive Metadata):**
MARC: 048 Number of Musical Instruments or Voices (display code in vernacular)
MARC: 382 Medium of Performance
MARC: 650 Subject Added Entry--Topical Term
EBUCore: hasGenre, hasSubject
MODS: subject
PBCore: pbcoreSubject (if using LCSH; LCSH is not a valid vocabulary for pbcoreGenre)

Medium of performance information may be too complex to use as a limit/facet within a general search interface. Before LCMPT and MARC 382, several institutions developed separate, complementary search tools utilizing MARC 048 data. These tools are provided as examples of functionality, though these tools may undergo revisions in the near future, such as to employ LCMPT and MARC 382 data.

  Kent State: [http://libguides.library.kent.edu/c.php?g=278264&p=1854323](http://libguides.library.kent.edu/c.php?g=278264&p=1854323)
  University of North Texas: [http://iii.library.unt.edu/search~S7/X](http://iii.library.unt.edu/search~S7/X)
Related MARC Authority Fields:
MARC: 100 $m; 130 $m; 150; 155; 382.

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 key or voice 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.

The musical note "a" presents special considerations in systems which consider it a stop word.
Phrase searching must be configured so terms that otherwise function as stop words 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. See IV.E for more details.

Another consideration for searching for works with known key is the exclusive use of the
symbols for music sharps and flats 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. One possible solution is to associate the search terms “sharp” and “flat” with
their respective symbols and substitutes, and vice-versa. This may result in some false hits,
especially for searches unrelated to music, but this could be mitigated by prompts to include or
exclude musical symbols from searches. For example, a message along the lines of “including
results for # (musical sharp sign). Search only for the word sharp" (linked to a search excluding
the sharp sign) would provide a solution for searches containing numerous false hits. Another
option is a feature that allows users to select from a set of frequently used symbols or diacritics.

---

8 For details on the development of Ball State’s Media Finders, see Kelley McGrath, “Media Finders: Improving the
Browsability of Media Collections via the OPAC,” Internet Reference Services Quarterly 11, no. 3 (2006): 19-34. Though
written before the Chamber Music finder was developed, the underlying processes are the same.
9 See “UNIMARC field 146 - Medium of Performance” http://www.iaml.info/unimarc-field-146-medium-performance for further information on these codes, from the
International Association of Music Libraries, Archives and Documentation Centres (IAML). See also search
interface at http://www.urfm.braidense.it/risorse/searchmedium_en.php. Retrieve a complete list by selecting
"Display all terms": http://www.urfm.braidense.it/risorse/medium_query_en.php

Ball State: http://cms.bsu.edu/Academics/Libraries/ResearchTools/MediaFinders.aspx
(the Chamber Music finder)

Note: The Ball State search relies on detailed encoded information drawn from a non-MARC
thesaurus. The addition of IAML medium of performance codes to MARC codes provides a high
level of specificity, but the IAML thesaurus is not widely used.
Regardless of search functionality, the system must allow entry and display of the symbols for musical sharps and flats 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 incipit\(^{10}\) in ordinary ASCII characters, it can be used for any material containing music. While $p$ contains the incipit itself, $sr$ 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.\(^{11}\)

MARC 384 is defined to record key information, including a first indicator to designate original or transposed key. Though defined in both bibliographic and authority formats, it is more commonly applied to authority records.\(^{12}\)

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 stop word), can be forced to be searched as a term. Ensure that the musical sharp and flat 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.

**Index and Display (Bibliographic/Descriptive Metadata):**
- MARC: 130 $r; 240 $r; 7xx $r; 031 $r Key for music
- MARC: 250 Edition Statement
- MARC: 384 [Musical] Key (generate display label based on first indicator)
- Dublin Core: No field defined to include musical key/range information. Examine data set for whether "description" or possibly "coverage" may contain this data.
- EBUCore: No field defined to include musical key/range information.
- MODS: No field defined to include musical key/range information.
- PBCore: No field defined to include musical key/range information.
- BIBFRAME 2.0: Properties: bf:musicKey

**Facets/Limits (Bibliographic/Descriptive Metadata):**
In a music-specific interface, a facet or limit for musical key could be useful. However, in most legacy data, key is rarely recorded for distinctively-titled works, so such a facet would be far from comprehensive. MARC 384 can be coded for either work (original) key (in bibliographic or

\(^{10}\) Musical incipit is the opening fragment (usually the melody) of a musical work. The incipit helps to identify the work.

\(^{11}\) The Moravian Music Foundation is one institution which regularly codes 031. See OCLC numbers: #923024892, #905230517, and #905237411 for examples.

\(^{12}\) MARC 384 is not repeatable in either bibliographic or authority format. As a result, common practice is to not include 384 in bibliographic records for manifestations containing multiple works. MARC 384 allows key to be recorded for expressions that are transpositions (key other than the original work); currently RDA provides instructions for recording key at the work level (original key) only.
authority records) or expression key (in bibliographic records). Key in MARC $r$ in MARC 130, 240, or 7xx may be misleading in cases of arrangements. Therefore, distinguishing between work key and expression key could be very problematic in facets/limits.

MODS: note type="key"

**Related MARC Authority Fields:**
MARC: 031; 100 $r$; 130 $r$; 384

**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. In current practice, creation date(s) can be recorded in MARC 046 $k$l in both bibliographic and authority formats. Additionally, MARC 388 (Time Period of Creation) has been available in both bibliographic and authority formats since October 2014, but is not yet widely used. Some approximation of creation era might be possible by exploiting composer birth/death date information. 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 it would cease coding 045, in March 1989.

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 $y$ and other subfields are defined for 655 genre/form headings, the LCGFT manual prohibits subdivision of genre/form terms. Chronological subdivisions are not consistently present due to subject cataloging rules (some headings are allowed to subdivide, and others are not) 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, especially for jazz and popular music, 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.1.

---

13 For example, as of January 2016, only 159 records in OCLC’s WorldCat database contained field 388. [http://experimental.worldcat.org/marcusage/388.html](http://experimental.worldcat.org/marcusage/388.html)

**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.

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: 045 Time Period of Content. (Was not used consistently after LC announced in March 1989 that it would cease coding 045. Legacy data should be evaluated for consistent presence of this field before using as a facet. Display code in vernacular.)
MARC: 046 $k$l Special Coded Dates. (Use codes in other subfields to generate vernacular display labels and to properly interpret dates coded in $k$l for single works and $o$p for aggregates. Remember that concluding dates are not always clear, particularly for works with complicated histories.)
MARC: 388 Time Period of Creation
MARC: 650 $y Chronological subdivision
Dublin Core: date.created
EBUCore: dateCreated, eventStartDate, eventEndDate, eventPeriod
MODS: originInfo.dateCreated, subject.temporal
PBCore: pbcoreCoverage (coverage, coverageType = Temporal); pbcoreInstantiation (instantiationDate + dateType)
BIBFRAME 2.0: Classes: bf:Temporal, Properties: bf:originDate, bf:temporalCoverage

**Facets/Limits (Bibliographic/Descriptive Metadata):**
MARC: 045 Time Period of Content. Was not used consistently after LC announced in March 1989 that it would cease coding. Legacy data should be evaluated for consistent presence of this field before using as a facet. (display code in vernacular)
MARC: 046 $k$l Special Coded Dates--Beginning or single date created, Ending date created. Evaluate data for consistent presence before using as facet. (use data and subfield labels to create eye-readable form)
MARC: 388 Time Period of Creation
MARC: 650 $y Chronological subdivision. Problematic due to subject cataloging rules, but commonly used for popular music and jazz headings.
Dublin Core: date.created
EBUCore: dateCreated; eventStartDate; eventEndDate; eventPeriod
MODS: originInfo.dateCreated
PBCore: pbcoreCoverage (coverage, coverageType = Temporal); pbcoreInstantiation (instantiationDate + dateType)
BIBFRAME 2.0: Classes: bf:Temporal, Properties: bf:originDate, bf:temporalCoverage

**Related MARC Authority Fields:**
MARC: 150 $y, 046 $k$l (display code in vernacular), 388

---

15 Some libraries continued to use MARC 045 after this time.
16 Recent changes to the DCM Z1 for name and series authority records instruct the cataloger to “use the Extended Date Time Format (EDTF) schema in all cases except for centuries.” With the new prominence of the EDTF schema, it may be necessary to manipulate 046 content to make it suitable for public display.
G. Persons and Corporate Bodies

The “created by” relationship is very important for users to find, identify, and select musical works associated with a particular composer or other creator(s). Usually, the composer is a single person; more rarely, corporate bodies or multiple persons can serve as creators. It is important to index and display all personal and corporate names, both transcribed forms and standardized vocabulary.\textsuperscript{17}

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 utilizes 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.\textsuperscript{18} 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 \textit{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 “creators/contributors/authors.”\textsuperscript{19}

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 legacy 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

\textsuperscript{17} Outside of art music traditions, the lines between creators and contributors such as performers or arrangers (see \textit{III.J}) may be blurred, and contributors are often presented and publicly perceived as having creative responsibility while creators may be identified nebulously or not at all, share responsibility in unclear ways, or be conflated (correctly or not) with copyright owners. In such cases contributors may be far more important than creators for locating and identifying a resource, thus underscoring the need to make all names available in records.

\textsuperscript{18} Including another composer: Adams, John Luther, 1953-.

\textsuperscript{19} Mostly, subject headings for persons and corporate bodies reflect subject relationships. Authors of texts set to music are an exception. SHM H 1110 (\url{https://www.loc.gov/aba/publications/FreeSHM/H1110.pdf}) 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 relationship, not a subject relationship to the work. However, RDA and MLA Best Practices instruct creating an authorized access point for creators and related persons, families, and corporate bodies (RDA 19.2.1.3, RDA 18.4.1.1, MLA BP 18.4.1.1), and AACR2 instructed to create added entries for the person whose words are set, making manipulating the subject headings to bring out the creator relationship less vital.
activity. MARC 046, 370-377, and 386 have been defined in the authority format, and MARC 046, 370, 377, and 386 have been defined in the bibliographic format, to provide machine-actionable fields for this data. If these fields were regularly encoded, this 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/AACR2 legacy data does 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/), ORCID (https://orcid.org/), ISNI (http://www.isni.org/) 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.\(^{20}\) 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). RDA/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.\(^{21}\) 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 or identifiers to allow users to differentiate between people/corporate bodies with similar names; make it possible for users to link from standardized vocabulary terms within the record display 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.

**Index and Display (Bibliographic/Descriptive Metadata):**  
*As “creators/contributors/authors”:\(^{22}\)*

**Standardized vocabulary:**
MARC: 100; 110; 111 Main Entry--Personal Name, Corporate Name, Meeting Name  
MARC: 700; 710; 711 Added Entry--Personal Name, Corporate Name, Meeting Name  
MARC: 800, 810; 811 Series Added Entry--Personal Name, Corporate Name, Meeting Name

---

\(^{20}\) Audiences are also related, but do not contribute to the work. MARC 385 was defined in 2013 to contain information on audience characteristics. If this field is regularly coded in the future, evaluate best ways to leverage it in discovery.

\(^{21}\) 1) MARC proposal 2017-01, passed January 21, 2017, will allow subfield $4 to contain URIs. Once this change is implemented, $4 will need a readable label and machine manipulation possibilities may expand. 2) RDA’s relationship designators, recorded in MARC $i, are readable text strings but the terminology is not straightforward for patrons. $4 codes (translated to vernacular) and $e text strings could be displayed or manipulated to facets, but first assess data for field presence and consistency. Per PCC practice, RDA records may contain both a $4 code and an $e text string, so displaying all could result in displays like: Doe, Jane (Actor). singer, singer. See II.B for additional discussion of these issues.

\(^{22}\) The specific label for the creator/contributor/author index will vary among discovery interfaces.
MARC: $4 Relator code; $e Relator term (assess data for presence and potential duplication) (display codes in vernacular) (index for use as facets, but do not include in the creator/contributor/author index)

Non-standardized vocabulary:
MARC: 508, 505 $r Creation/Production Credits Note, Formatted Contents Note--Statement of responsibility
Persons and corporate bodies referred to in notes will often, but not always, be included in controlled vocabulary fields. Including these notes in the “author” keyword index will give fullest coverage of persons and corporate bodies. 245 $c might also be included, but this subfield may also contain title and other non-creator/contributor/author data. 505 $a could also be included to include creators/contributors/authors found in unenhanced contents notes. However, this would also add many titles to the creator/contributor/author index. Consider local data to make a decision balancing precision and recall.

Analyze data for standardized/non-standardized vocabulary (best practices call for adopting an authority source, but this may not always be done)
Dublin Core: creator ; contributor
EBUCore: hasContributor; hasCrewMember; groupName
MODS: name / namePart + type; role + roleTerm; nameIdentifier contains standard identifying numbers or codes, such as an ORCID identifier, which should be indexed but many not necessarily be appropriate for unmodified display
PBCore: pbcoreCreator (creator, creatorRole)

As “subjects”:
Standardized vocabulary:
MARC: 600; 610; 611 Subject Added Entry--Personal Name, Corporate Name, Meeting Name
Dublin Core: subject, relation
EBUCore: hasSubject
MODS: relatedItem
PBCore: pbcoreSubject + subjectType
BIBFRAME 2.0: Classes: bf:Topic; Properties: bf:subject + Agent classes & properties

Facets/Limits (Bibliographic/Descriptive Metadata):
Include persons and corporate bodies contained in standardized vocabulary “creator/contributor/author” fields in a “creator/contributor/author” facet.23
If evaluation of data suggests roles are regularly and uniformly encoded, consider incorporating this information into faceting (for example, to allow easy filtering of Leonard Bernstein as composer vs. Leonard Bernstein as conductor).

Related MARC Authority Fields:

23 The specific label of this facet will vary among discovery interfaces. Balance brevity with clarity and consideration of non-text formats.
H. Genre/Form

Users regularly seek works and expressions with particular forms, genres, types, styles, etc. These headings, which describe what musical works are, were 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). More recently, however, the Library of Congress and other interested groups, including MLA’s Music Genre/Form Task Force, developed the thesaurus Library of Congress Genre/Form Terms for Library and Archival Materials (LCGFT), which separates form and genre headings from topical subjects and codes musical genre/form headings separately from topical headings, using MARC 655 (bibliographic) and 155 (authority). Discovery services need the capability 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 eventually change as part of the LC genre projects (at present, the practice is to code genres in both 650 and 655), 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.”. 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 standardized vocabulary terms within the record display to other materials associated with the same attribute/entity. This could be accomplished through use of bound text strings for full authorized terms or via identifiers functioning behind the scenes.

Because current music genre information is often encoded in topical subject fields, 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. To avoid frustrating or misleading results, faceting systems should recognize


Music Discovery Requirements
and respect the distinctions between different subject and genre thesauri (LCSH, LCGFT, FAST, etc.) even if they share identical strings as headings.

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: 006/01-02 Form of composition code (display code in vernacular)
MARC: 008/18-19 Form of composition code (display code in vernacular)
MARC: 047 Form of Musical Composition Code (display code in vernacular)
MARC: 380 Form of Work (note: 380 is defined in both bibliographic and authority formats, but is primarily used in authority records. Therefore, display and index occurrences in bibliographic records, but rarely will data be present in bibliographic records with enough regularity to use it as a facet.)
MARC: 650 Subject Added Entry--Topical Term (until such time as genre/form terms are moved to 655, including those in legacy data)
MARC: 655 Index Term—Genre/Form
Dublin Core: subject, format, description
EBUCore: hasGenre
MODS: genre, subject, note
PBCore: pbccoreGenre
BIBFRAME 2.0: **Classes:** bf:GenreForm; **Properties:** bf:genreForm

**Facets/Limits (Bibliographic/Descriptive Metadata):**
MARC: 006/01-02 Form of composition code (display code in vernacular)
MARC: 008/18-19 Form of composition code (display code in vernacular)
MARC: 047 Form of Musical Composition Code (display code in vernacular)
MARC: 655 Index Term--Genre/Form
*Note:* MARC 650 Subject Added Entry--Topical Term is not recommended for inclusion in facets because it will visibly include many subjects, essentially making the “genre” and “subject” facets identical.
Dublin Core: subject, format
EBUCore: hasGenre
MODS: genre, subject
PBCore: pbccoreGenre
BIBFRAME 2.0: **Classes:** bf:GenreForm; **Properties:** bf:genreForm

**Related MARC Authority Fields:**
MARC: 150; 155; 380

---

**I. 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 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 non-topical 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”; and temporal coverage, often via date subdivisions (addressed in II.F and III.I).

Medium of performance (addressed in II.D and III.G); and genre/form (addressed in II.H) are special cases because LCMPT and LCGFT vocabularies are now complete, but cleanup of LCSH and legacy data is still in progress, and subject terms are still assigned in addition to Medium of Performance and Genre/Form terms. Therefore, 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.

Musical works may be the subject of topical works. These topical works can be assigned subject headings consisting of the standardized title of the work (including composer, if applicable). The same controlled vocabulary is currently recorded for musical works as both titles and subjects, but the indexing and display should make the distinction clear. It could be useful to provide functionality for moving from the actual work to materials about the work and vice versa.

**Recommendation**: Index and display subject fields. For data using standardized vocabulary, make it possible for users to link from standardized vocabulary terms within the record display 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. To avoid frustrating or misleading results, faceting systems should recognize

---

25 MLA Best Practices for Using LCMPT (Feb. 14, 2017) and MLA Best Practices For Using LCGFT for Music Resources version 1.0 (June 8, 2015) both instruct to continue using LCSH until LCGFT and LCMPT “terms are fully implemented, and a method for generating genre and medium of performance terms… from headings currently coded as LCSH have been developed and deployed.” See most current versions on MLA Best Practices website: [http://www.musiclibraryassoc.org/BlankCustom.asp?page=cmc_mlabestpractices](http://www.musiclibraryassoc.org/BlankCustom.asp?page=cmc_mlabestpractices)
and respect the distinctions between subject and genre thesauri (LCSH, LCGFT, FAST, etc.)
even if they share identical strings as headings.

Index and Display (Bibliographic/Descriptive Metadata):
Also note: II.F, III.I (dates); II.D, III.G (medium of performance); II.H (genre/form); II.G
(persons/corporate bodies); II.J, III.K (geographic area)
MARC: 043 Geographic Area Code (display code in vernacular)
MARC: 6XX Subject Access Fields
Dublin Core: subject
EBUCore: hasSubject
MODS: subject
PBCore: pbcCoreSubject + subjectType
BIBFRAME 2.0: Classes: bf:Topic; Properties: bf:subject

Facets/Limits (Bibliographic/Descriptive Metadata):
MARC: 043 Geographic Area Code (display code in vernacular)
MARC: 6XX Subject Access Fields
Dublin Core: subject
EBUCore: hasSubject
MODS: subject
PBCore: pbcCoreSubject + subjectType
BIBFRAME 2.0: Classes: bf:Topic; Properties: bf:subject

Related MARC Authority Fields:
MARC: 100; 110; 111; 130; 150; 151; 180

J. Geographic Area (and Ethnic/National Group)

Geographic area of musical works--the geographic area from which they emanate, as
distinguished from geographic areas they are about--is a challenging attribute to record and
therefore undercoded. A somewhat common tendency to conflate or confuse geographic area
with a corresponding ethnic or national group may be a contributing factor to the problematic
nature of this attribute. RDA, for example, lists Place of Origin as an attribute of a work (RDA
6.5) but makes no similar mention of ethnic or national group. While it might be assumed that
this information would inherit from a work’s creator (although, again, Associated Places are
considered attributes, while the place of ethnic/national identity is as yet somewhat obscure),
such is only possible if the creator is a specific person or persons, corporate body, or family
rather than an ethnic or national grouping of some sort.

Despite (or perhaps because of) this circumstance, geographic area is a particularly important
attribute for music other than Western art music. Descriptions of these musics (including
non-Western classical music genres and genres often referred to as “popular” or “folk” music)
have historically included more geographic access, primarily through assignment of geographic
subdivisions to Library of Congress Subject Headings. However, the resulting subject heading strings may be problematic for search faceting and limiting because of the need to combine information on geographic locale and ethnic/national group to create meaning: as suggested above, these two data points do not always overlap. For example, the heading "Folk music -- Germany" is applied to folk music: (a) Germans in Germany; (b) of other ethnic/national groups in Germany; and (c) of Germans in other locales. To clarify which of these is intended, a second subject heading string is required: for example, “Turks--Music--Germany” (for a particular non-German national group in Germany); “Germans--Music--United States” (for Germans in a particular locale outside of Germany); but “Germans--Music” (that is, without a geographic subdivision--for Germans in Germany). In cases such as these, it is important that the entire subject heading be displayed and machine-actionable as a string.

In addition, the same subject headings used for music emanating from a geographic area are also assigned--with the addition of appropriate subdivisions--to works about the music in question. For example, the heading “Music--Uganda” is used for music recordings and scores; the same heading is used for a book about music in Uganda by simply appending the subdivision “History and criticism.” Again, it is important that the entire subject heading be displayed and actionable as a string to avoid loss of meaning.

A final complication of geographic access to musical works is the intellectual task of determining and prioritizing associated geographic areas. Geographic areas associated with a composer or work are now commonly recorded in name and name/title authority records in MARC 370, though this information may not be present in legacy data. When it is present, harnessing this data for limits or facets can become problematic, as creators often change residence over the course of their lives, and none of these places may be pertinent to a specific work by that creator. Furthermore, the geographic terms in use themselves, like the boundaries of the entities that they name, often change over time and may be misleading to users who are not aware of pertinent historical circumstances. Finally, as noted above, the distinction between geography and ethnicity may not be apparent to the user (or to the cataloger providing the data, for that matter) or it may even be a matter of controversy. In the future, some clarification may be achieved through use of the new MARC 386 field (Creator/Contributor Characteristics).

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 standardized vocabulary terms within the record display to other materials associated with the same

---

26 See in particular SHM H 1917, Music of Ethnic, National, and Religious Groups. Use of geographical subdivisions for jazz and popular music is found in SHM H 1916.5 (2.); for Western art music, SHM H 1160 (3.).
27 As of Nov. 2013, the Library of Congress Descriptive Cataloging Manual instructed use of the LC/NAF form of name in MARC 370. [https://www.loc.gov/aba/publications/FreeDCM/freedcm.htm](https://www.loc.gov/aba/publications/FreeDCM/freedcm.htm)
28 As of January 2017, the Library of Congress Descriptive Cataloging Manual “encourages” LC/PCC catalogers to wait until the PCC develops guidelines before using MARC fields 385 (Audience Characteristic) and 386. Per correspondence with LC representatives, LC practice is to not use either $m or $n.
attribute/entity. This could be accomplished through use of bound texts strings for full authorized terms or via identifiers functioning behind the scenes.

**Index and display (Bibliographic/Descriptive Metadata):**

*(note that MARC 651 is not included; this field is for topical use)*

MARC: 043 Geographic Area Code (display code in vernacular) (problematic because MARC fields codes do not distinguish between “is” and “about”)

MARC: 370 Associated Place (more common in authority records)

MARC: 650 Subject Added Entry--Topical Term (problematic both “is” and “about” are contained in MARC 650; distinction could be made based on presence of other subdivisions)

MARC: 650 $z Subject Added Entry--Topical Term, Geographic subdivision (problematic because MARC field tags do not distinguish between “is” and “about”; distinction could be made based on presence of other subdivisions)

Dublin Core: coverage

EBUCore: hasCoverage

MODS: geographic

PBCore: pbcoreCoverage (coverage, coverageType = Spatial)

BIBFRAME 2.0: **Classes**: bf:Place; **Properties**: bf:geographicCoverage, bf:originPlace, bf:place

**Facets/Limits (Bibliographic/Descriptive Metadata):**

MARC: 043 Geographic Area Code (display code in vernacular) (problematic because MARC field codes do not distinguish between “is” and “about”)

MARC: 650 Subject Added Entry--Topical Term (problematic; both “is” and “about” are contained in MARC 650; distinction could be made based on presence of other subdivisions)

(also problematic because not always recorded due to subject cataloging rules)

EBUCore: hasCoverage

MODS: geographic

PBCore: pbcoreCoverage (coverage, coverageType = Spatial)

BIBFRAME 2.0: **Classes**: bf:Place; **Properties**: bf:place, bf:originPlace

**Related Authority Fields:**

MARC: 150; 151; 370 (150 is problematic because it does not distinguish between “is” and “about”)
III. Expressions and Manifestations

A. Introduction

Musical works frequently exist in many expressions and manifestations, 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. Therefore, a discovery interface should group or connect versions of a work in some way and simultaneously aid users in finding, identifying, and selecting a version which meets their needs.

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:

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

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 physical carriers such as compact disc, LP, audio cassette, videocassette and videodisc, along with digital files encoded using formats such as MP3, MPEG, and WAV. 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 audio cassette, compact disc, LP, digital audio, DVD, digital video, 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/carry. Icons are frequently useful for this task.

RDA content type/media type/carry type fields (encoded in MARC 336, 337, and 338) could be used to create displays and search data. However, RDA carry type "audio disc" is insufficient for distinguishing CDs from LPs, and additional RDA elements such as “sound characteristic” or “digital file characteristic” are necessary to distinguish the specific carry for recorded audio. 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. Content types "spoken word" and "performed music" (coded in MARC 336) do provide data to distinguish between these two kinds of sound recordings.

RDA’s instructions for creating an authorized access point for an expression of a musical work (RDA 6.28.3) direct catalogers to add a term indicating content type (notated music/performed music), but implementation is stalled and uncertain. 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.

**Recommendation:** Distinguishing format is crucial to finding 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 B for detailed content and carrier mapping.

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: Leader/06 (Type of Record); 006/00 (Form of material); 007 (Physical Description Fixed Field). Use this coded data to create icons or labels indicating material type. See Appendix B for detailed content and carrier MARC mapping.
MARC: 344 Sound Characteristics
MARC: 347 Digital File Characteristics
Dublin Core: format
EBUCore: hasMedium; hasFormat (relevant subproperties: “Audio format,” “container format,” “encoding format,” “file format,” “video format”); hasAudioEncodingFormat; hasContainerFormat
MODS: typeOfResource, typeOfResource + form or internetMediaType

---

29 The Music Library Association’s Best Practices for Music Cataloging alternate 300 field examples include a MARC 300 $b with differentiating information, e.g. “CD audio, stereo” but direct that “these alternative 300 encodings should only be used in systems where the 34x fields do not adequately display to the user.” Supplements to Best Practices for Music Cataloging, version 1.5, 4/12/2016. Supplement 1.2. Examples.
30 LC-PCC PS 6.28.3 states LC practice is to not add another characteristic, and PCC Practice is “pending outcome of report/recommendations from the PCC Access Point for Expression Task Group.” That group’s January 2013 report recommends “allowing $h for content type in access points for musical expressions, if not for other expressions” and states that “more specific guidelines should be offered.” However, no movement seems to have happened as of January 2017.
Facets/Limits (Bibliographic/Descriptive Metadata):
MARC: Leader/06, 006/00, 007 (Use this coded data to create facets/limits in the vernacular.)
(Allow selection of both a general content type (printed music, audio recording, video recording)
as well as more specific carriers (CD, LP, audio cassette, streaming audio, DVD, VHS, etc.).)
(See Appendix B for detailed content and carrier MARC mapping.)
Dublin Core: format
EBUCore: hasMedium; hasFormat (subproperties: “Audio format,” “container format,”
“encoding format,” “file format,” “video format”); hasAudioEncodingFormat;
hasContainerFormat
MODS: typeOfResource; typeOfResource+form; internetMediaType
PBCore: pbcoreInstantiation (instantiationMediaType, instantiationPhysical, instantiationDigital)
BIBFRAME 2.0: bf:Audio, bf:NotatedMusic, bf:Content, bf:Media, bf:Carrier,
bf:SoundCharacteristic (Subclasses: bf:RecordingMethod, bf:RecordingMedium,
bf:PlayingSpeed, bf:GrooveCharacteristics, bf:TrackConfig, bf:PlaybackChannel,
bf:PlaybackCharacteristic), bf:DigitalCharacteristic (Subclasses: bf:FileType,
bf:EncodingFormat, bf:EncodedBitrate), bf:AppliedMaterial, bf:BaseMaterial; Properties:
bf:carrier; bf:content, bf:appliedMaterial, bf:baseMaterial, bf:soundCharacteristic,
bf:digitalCharacterestic, bf:media

Related MARC Authority Fields:
MARC: 100 $h; 130 $h (assess data for presence before implementing)

C. Identifying Numbers

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

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, publisher, and distributor numbers.31.
Discovery systems should index and display all of these associated numbers, including ISBN and
ISSN, sometimes also found on notated music.

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,

Article Number; matrix number: used to indicate sound recording masters, often used to identify 78s.
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 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.

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: 020 $a$z ISBN ($z$ is a canceled/invalid standard number or code; make local decision on inclusion)
MARC 022 $a$z ISSN ($z$ is a cancelled/invalid standard number or code; make local decision on inclusion)
MARC: 024 $a$z Other Standard Identifier ($z$ is a canceled/invalid standard number or code; make local decision on inclusion) (use first indicator to generate display label)
MARC: 028 Publisher or Distributor Number (use first indicator to generate display label) (display but do not index $q$) 32
Dublin Core: identifier
EBUCore: identifier
MODS: identifier
PBCore: pbcoreInstantiation (instantiationIdentifier)

**Facets/Limits (Bibliographic/Descriptive Metadata):**
Facets/limits by identifying number are rarely useful for expressions and manifestations; such facets/limits would only occasionally yield more than one result.

**Related Authority Fields:**
MARC: 020; 024 (024 is defined, but rarely used)

**D. Format of Notated Music**

Users need to distinguish between various formats of notated music (AACR2 term: musical presentation), such as full score, parts, vocal score, etc. This attribute does not apply to recorded music.

32 MARC 028 renamed and redefined to include distributor numbers in 2016.
performances of music. Most musical works are published in more than one format of notated music, and the different formats are not interchangeable in actual use. Users may casually refer to these variations as “editions,” but “format of notated music” is used here in order to make the distinction between format and intellectual content. Editions are covered in III.E.

Commonly used cataloging standards and metadata formats, and the legacy data encoded via these standards and formats, do not cleanly support searching and limiting/faceting by format of notated music. The information recorded under RDA and AACR2 is sufficient to allow informed users to identify the format of notated music 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 due to changes over time that have resulted in inconsistencies among records. Mapping format of notated music to other metadata formats may present similar difficulties. MARC 348, Format of Notated Music, defined in 2015, accommodates controlled vocabularies in both textual and coded form. Form of notated music (e.g. staff notation) is coded in MARC 546 $b.

Format of notated music information may be scattered throughout the descriptive areas of a bibliographic record, including:

- extent/physical description (1 vocal score, 1 study score, 1 score + 4 parts, etc.). This vocabulary is consistently present, but multiple terms may be included in a subfield, e.g. 1 score + 4 parts.
- transcribed in edition (RDA, coded in MARC 250) or musical presentation (AACR2, coded in MARC 254) statements. This transcription relies on how, or if, information is presented on the resource. These statements, furthermore, may appear in various languages and may be abbreviated or even misleading.
- statement of responsibility in AACR2 records may include transcriptions or cataloger-supplied descriptions of musical presentation (e.g., vocal score).
- free-text notes.

The most promising source of format of notated music information comes from controlled vocabularies, in the form subdivisions currently used in LC subject headings and in format terms in LCGFT. The free-floating subdivisions in Pattern Headings: Musical Compositions (Subject Headings Manual (SHM) H 1160),33 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, the Subject Headings Manual dictates that these terms not be used with particular kinds of music (notably: works for solo instrument, voice and piano works, and popular music). LCGFT’s removal of restrictions based on medium or genre of music means that terms indicating format of notated music can be included in all records for notated music. For example, the LC subject heading coded in MARC 650 for a work for solo piano cannot contain the subdivision $v Scores, but the LCGFT term Scores can be assigned to that same work and coded in MARC 655 in the same bibliographic record.

---


Music Discovery Requirements 32
Redundancy or near-redundancy of terms in current LC subject heading form subdivisions, LCGFT, and MARC 348 pose the risk of a proliferation of identical or similar terms in a facet/limit. A system that can de-duplicate identical strings from facet terms would avoid this problem.

An alternative for extracting consistent data for facets/limits may be to use MARC coding in 008/20 or 006/03, which will include the preponderance of legacy records without MARC 348 and local legacy records without MARC 655. This alternative presents its own problems, however, especially in the way the definitions have changed over time and the lack of a way to separate out parts, an important user task. See Appendix C - Format of Notated Music History for full details. Some of the most difficult problems are highlighted below.

- Parts are not coded in 008/20 when issued with a score. When issued alone, parts are lumped into Code z (Other), as described below. While MARC (008/21, 006/04) provides for coding information regarding the presence or absence of parts, OCLC and MLA’s MARC Formats Subcommittee have recommended simply accepting the “blank” (“no attempt to code”) default. As a result, this data is not often coded, at least in WorldCat.
- Code m (Multiple score formats), while necessary, does nothing to indicate the nature of the individual scores included within a manifestation consisting of multiple types of scores issued together.
- Code z (Other) is an amalgamation of formats, including materials consisting solely of parts and manifestations where “none of the other defined codes are appropriate.”
- Over time, codes and definitions have been added, changed, and dropped from use. For example:
  - Code z (Other), defined in AACR2 as “Music in other than score form,” also included works for solo instrument or popular music on two staves for piano with words printed between the staves. Works for solo instrument are now coded l (“el”) in RDA. RDA does not explicitly address how to code popular music on two staves for piano with words printed between the staves, and to date it appears coding has been inconsistent.\(^{34}\)—both of which are now coded l (“el”) in RDA.
  - Codes h (Chorus score), i (Condensed score), and j (Performer-conductor part) were added after their approval in 2009; code p (Piano score) was added in 2016.
  - In 2012, code d was changed from “Voice score,” defined as including vocal parts with accompaniment arranged for piano, to “Voice score with accompaniment omitted”; the code for vocal scores with arranged accompaniments was then designated k (Vocal score).

Type of score is not easily mapped to DC or MODS. Type of score combines content and carrier, but DC elements distinguish between the two. LC’s crosswalk to MODS doesn’t seem to support MARC 254. Rethinking of data recording may be necessary; perhaps type of score could fit in physicalDescription or typeOfResource.

\(^{34}\) Possible codes include l (“el”) (Score); i (Condensed score); and k (Vocal score) according to cataloger’s judgement and interpretation of the definitions.
**Recommendation:** Based on the existing situation for format of notated music, 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.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 250 Edition Statement (RDA)
MARC: 254 Musical Presentation Statement (pre-RDA)
MARC: 300 $a $e Physical Description--Extent,Accompanying material
MARC: 348 Format of Notated Music
MARC: 546 $b Language Note, Information code or alphabet
MARC: 650 $v Subject Added Entry--Topical Term,Form subdivision
MARC: 655 Index Term--Genre/Form
Dublin Core: type, format
EBUCore: no applicable properties
MODS: typeOfResource, physicalDescription
PBCore: n/a; information for display can be recorded in pbcoreAnnotation

**Facets/Limits (Bibliographic/Descriptive Metadata):**
(problematic; see discussion above)
MARC: 008/20, 006/03 Format of music
MARC: 008/21, 006/04 Music parts
MARC: 650 $v Form subdivision (Free-floating subdivisions for Written or Performed Music; Performed Version)
MARC: 655 Index Term--Genre/Form
DC, MODS, EBUCore: musical presentation is not easily mapped to these formats.

**Related MARC Authority Fields:**
MARC: 150 $v; 185 $v

**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, 2-piano edition, high voice 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), Musical Key/Range (II.E, III.H), and Format of Notated Music (III.D). In addition to following recommendations in those sections, index and display fields identified below in order to provide publisher information and edition statements transcribed from the item itself.

**Index and display (Bibliographic/Descriptive Metadata):**

MARC: 250 Edition Statement (in RDA records, 250 will also contain Format of Notated Music information)

MARC: 260 $b$f Name of publisher, distributor, etc.; Manufacturer

MARC: 264 $b$ Name of producer, publisher, distributor, manufacturer

MARC: 5XX Note Fields (may include information about the editor and the edition, such as being a critical edition, etc.)

Dublin Core: description

EBUCore: “Edition” not easily mapped

MODS: edition under originInfo

PBCore: pbcoreInstantiation (instantiationIdentifier); different instantiations of the same asset can be collocated through pbcoreIdentifier

BIBFRAME 2.0: Properties: bf:editionStatement

**Facets/Limits (Bibliographic/Descriptive Metadata):**

Edition information does not readily lend itself to facets/limits. While users often seek “authoritative”, “Urtext” or “scholarly” editions, this data has, historically, been neither recorded nor encoded in any standardized way. The subjective nature of these designations would make it difficult to record and encode them in the future as well.

**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, the code zxx is used to indicate “no linguistic content.” 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).

**Index and Display (Bibliographic/Descriptive Metadata):**
MARC: 041, 008/35-37 Language Code (may be better suited to facets) (If used, ideally make clear the function of the language within the item, if this info is coded in the data. Consider carefully subfields of 041 to include. All subfields contain language codes except: $2 (Source of code), $6 (Linkage), and $8 (Field link and sequence number.)
MARC: 130 $l; 240 $l; 7XX $l Language information in standardized titles (subfield is titled “Language of a work”)
MARC: 377 Associated Language (assess data for presence before implementing)
MARC: 546 $a Language Note (older data may contain language notes in 500, where it was recorded before 546 was defined)
Dublin Core: language
EBUCore: hasLanguage; hasOriginalLanguage
MODS: languageTerm
PBCore: pbcoreInstantiation (instantiationLanguage)
BIBFRAME 2.0: Classes: bf:Language; Properties: bf:language

---

36 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.
37 The facets in the video view of University of Virginia’s Blacklight implementation provide an example of clearly labelled facets for different language functions. [http://search.lib.virginia.edu/catalog?portal=video](http://search.lib.virginia.edu/catalog?portal=video)
Facets/Limits (Bibliographic/Descriptive Metadata):
Allow facets/limits by all languages, not just those in MARC 008/35-37. Ideally, allow the possibility to facet/limit by the function of the language.
Dublin Core: language
EBUCore: hasLanguage; hasOriginalLanguage
MODS: languageTerm
PBCore: pbcoreInstantiation (instantiationLanguage)
BIBFRAME: Classes: bf:Language; Properties: bf:language

Related MARC Authority Fields:
MARC: 100 $l; 130 $l; 377

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 RDA and AACR2 dictate use of the original medium of performance when included in standardized titles. Arrangements are identified by adding a term to the end of the standardized title: “arr.” in AACR2, or “arranged” in RDA (both coded in MARC $o). LCSH reflects the expression medium, with “Arranged” added to the end of the term. Similarly, LCMPT is coded for expression medium. Expression medium is typically coded in bibliographic records in uncontrolled textual fields (e.g., notes), coded fields (MARC 048), and controlled vocabularies (e.g., LCMPT); see II.D for further detail.

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

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, expression key and range information is not currently or historically recorded under AACR2 or RDA nor encoded in MARC data on a comprehensive basis. It may be included in a 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:** Follow work-level indexing and display recommendations at II.E. Display notes and edition statements and include both in general keyword indexes so that users can minimally identify specific arrangements when looking at a record, and potentially keyword search this information. Explore more dynamic searching options in the future.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 250 Edition Statement
MARC: Edition information may also appear in other areas including Title Statement (245) and various Note Fields (5XX)
Dublin Core: description; creative use of coverage
EBUCore: No field defined to include musical key/range information.
MODS: note type="key"
PBCore: pbcoreAnnotation
BIBFRAME 2.0: Properties: bf:musicKey

**Facets/Limits (Bibliographic/Descriptive Metadata):**
Coded data for range could lend itself to some sort of coded search capability, though this data has historically not been coded in MARC.
MODS: note type="key"
BIBFRAME: Properties: bf:musicKey

**Related MARC Authority Fields:**
MARC: 031; 384

**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.E.
Performance date is an important expression-level attribute for recorded music. Users may seek particular performances, or they may use the performance date to select an expression from a particular decade. 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 (pre-1990s) practice was to add the manifestation date routinely to standardized titles, and this use may still be present in some legacy data. Adding manifestation date to bibliographic records is now allowed and some libraries do encode this 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.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 008/07-10 (Date 1); 008/11-14 (Date2) (consider using code in 008/06 to generate vernacular display label)
MARC: 033 Date/Time and Place of an Event (display code in vernacular) (has not been regularly coded for music)
MARC: 046 $b$c$d$e Special Coded Dates, (use codes in $a to generate vernacular display labels and to properly interpret dates coded in $b$c$d$e. $k$l contain work level date information and are addressed in II.F )
MARC: 111; 711; 811 $d Date of meeting
MARC: 100, 110, 111; 130; 240; 700; 710; 711; 730; 800; 810; 811 $f Date of a work (100, 110, and 111 are defined but in reality rarely used because AACR2 title main entry information is placed in 240)
MARC: 260 $c Date of publication, distribution, etc.
MARC: 264 $c Date of production, publication, distribution, manufacture, or copyright notice
MARC: 502 $d Dissertation Note, Year degree granted
MARC: 518 $d Date of event
MARC: 534 $c Original Version Note, Publication, distribution, etc. of original, (includes dates)
MARC: 650 $y Subject Added Entry--Topical Term, Chronological subdivision
Dublin Core: date
EBUCore: dateBroadcast; dateCreated; dateDigitised; dateDistributed; dateIssued
MODS: dateCreated, dateModified and copyrightDate in dateIssued, and possibly dateOther

---

While 518$a contains date information in legacy data, it also contains additional free-text information that would be difficult for machine parsing, so 518$a is not recommended for inclusion in date indexes.
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, director, publisher, and typesetter.

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. However, first assess data for presence and possible duplication; see II.G for further discussion.

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 names related to the work, expression, or manifestation useful for finding, identifying, and selecting music, regardless of the exact nature of the relationship.

Publishers, typesetters, printers, etc. are particularly important persons and corporate bodies for earlier music, and reprints of early music. Insofar as this information is coded in the data, following these recommendations will allow searching by and identification of publishers and typesetters.
**Recommendation:** Follow recommendations in II.G and IV.B. 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, after assessing the data for presence and possible duplication.

**Index and Display:**
See recommendations in II.G.

MARC: 511 Participant or Performer Note

**K. Geographic Area (and Ethnic/National Group)**

For Western art music, geographic area is arguably a less important attribute for expressions and manifestations than it is for musical works. For other musics, however, geographic area of expression or manifestation may be of interest to users, not only because geographic area may be disputable at the work level, but also due to its reflection in the changes that occur as a musical work travels from place to place. These other musics have traditionally received the most subject access, as detailed in II.J. Note that the concerns expressed in that section regarding conflation or confusion of geographic area with ethnicity apply here as well.

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. As it becomes more common to parse this information out in various MARC 518 subfields, locations recorded in $p$ may become useful in this regard.

Geographic area of the manifestation is regularly recorded in the publication statement and coded in MARC 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 area 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.

**Index and display (Bibliographic/Descriptive Metadata):**
MARC: 008/15-17 Place of publication, production, or execution (display code in vernacular)
MARC: 033 Date/Time and Place of an Event
MARC: 260 $a$e Publication, Distribution, etc. (Imprint)-Place of publication, distribution, etc.; Place of manufacture
MARC: 264 $a Place of production, publication, distribution, manufacture
MARC: 518 $p Date/Time and Place of an Event Note-Place of event (also display $a, and index $a in general keyword index, but not in a place of publication index.)
Dublin Core: coverage
EBUCore: hasCoverage
MODS: geographic; originInfo.place
PBCore: pbcoreCoverage (coverageType = Spatial)
BIBFRAME 2.0: Classes: bf:Place, bf:; Properties: bf:place, bf:geographicCoverage

Facets/Limits (Bibliographic/Descriptive Metadata):
MARC: 008/15-17 Place of publication, production, or execution (display code in vernacular)
Dublin Core: coverage
EBUCore: hasCoverage
MODS: geographic
PBCore: pbcoreCoverage (coverageType = Spatial)
BIBFRAME: Classes: bf:Place, bf:; Properties: bf:place, bf:geographicCoverage
IV. Other Aspects of Music Discovery

A. Introduction

Several issues related to music discovery do not focus on particular attributes or relationships. The following areas are discussed in this section. The topics in this section include recommendations, indexing recommendations for specific fields are not relevant and therefore are omitted.

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

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. Musicians regularly use notated music with a title page is printed in a language they do not speak, because the language of the title page (which often reflects the country where the publisher is located) is irrelevant to their use of the musical notation. Even when the musical work includes a text, musicians frequently seek the language originally used by the composer, with no regard for the title page language. Similarly, the language used on a recording label or container is irrelevant to the primary purpose of listening to the recorded (including video recorded) sound. 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).

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. 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 that 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. In a linked data environment, the possibility exists to use data from additional information sources such as VIAF, or even from sources external to the library community such as MusicBrainz or Wikidata.

---

39 See, for example, an author browse on mozart, johann in Cornell University’s Blacklight instance. “Mozart, Johann Chrysostom Wolfgang Amadeus, 1756-1791” (a non-authorized term) has a link to “See Mozart, Wolfgang Amadeus, 1756-1791” (an authorized term.)
41 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/ (accessed 23 March 2017).
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 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 auto suggestions, 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.

For data using standardized vocabulary, make it possible for users to link from standardized vocabulary terms within the record display 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.

C. Compilations

Multiple expressions are frequently issued together in a single manifestation, sometimes called an aggregate. 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 consistently 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 as yet). However, because of the lack of linking information, this problem cannot easily be solved by discovery interfaces utilizing MARC-encoded RDA and AACR2 legacy data. 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 indicating to which work/expression the data

The result is that, for example, all performers and subject headings are associated with all works/expressions contained on a compilation 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.

### D. Searching: Alphabetical and Keyword

Historically, alphabetically filed left-anchored “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. Discovery layers and “next-generation” catalogs added faceted browsing, but many simultaneously ceased to provide alphabetical browse searching.

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. However, users do not always “know” a lot about the entity they are seeking. Transcribed title is a notoriously unreliable access point, especially for Western art music. “Subject” is a problematic concept because much music is arguably not “about” anything. Perhaps because of these difficulties, personal names are frequently used as an access point in music searching. However, many composers are very prolific, and their works exist in many versions, making personal name alone an insufficiently precise access point.

Because author, (transcribed) title, and subject were not sufficient to find, identify, and select musical works, music librarians exploited left-anchored title and subject card catalogs to provide access beyond personal names. They developed complex uniform titles (especially generic uniform titles) and subject headings (which commonly covered attributes of form, genre, and medium of performance and only occasionally reflected true “aboutness”) to allow informed users to reliably find and browse musical works, as well as to quickly ascertain when pieces represented a part of a larger work or selections or arrangements of a work. These complex uniform titles and “subject” headings persisted even when catalogs moved online and keyword searching was introduced, and alphabetical searching remained vital because few users possessed knowledge to determine the complex vocabulary in advance.

---

More recently, facetable, linked data-friendly controlled vocabularies including the Library of Congress Genre/Form Thesaurus (LCGFT) and the Library of Congress Medium of Performance Thesaurus (LCMPT) have been developed to encode a piece of music’s genre and medium of performance separately from its subject. Similarly, RDA focuses on recording discrete, machine-manipulable data elements. LCMPT, LCGFT, and RDA were created at a time when keyword searching had become the default search strategy of choice, in contrast to LCSH and AACR2, which were developed in an era when alphabetical searching predominated.

However, the need to determine complex vocabulary in advance remains a major weakness of most current keyword searching, which rarely exploits alternate forms found in LC/NAF authority records. 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, especially with regard to authorized access points and controlled vocabularies, and should be retained in back-end interfaces even if it is eliminated from public interfaces.

**E. Phrase searching**

Phrase searching configured to override Boolean operators, proximity operators, and stop words is basic functionality for all materials. When Boolean operators and stop words are enclosed in quotes, they must be searched as terms, not Boolean operators, proximity operators, nor stop words. Otherwise, searches will fail for titles like "Gone with the wind" and "I am NOT going to get up today." Additionally, phrase searching should prevent a system from automatically searching for alternate forms of the word, such as a plural or possessive version of a singular term, or vice-versa.

Identifying numbers and musical key provide three music-specific situations where correctly configured phrase searching is especially important. The English "no." (as in "Symphony no. 5") presents a special challenge in systems that treat "no" as a Boolean NOT operator. With correctly configured phrase searching, a search on “no 5” returns works with recorded serial number “5” rather than excluding “5” as a search term (as would happen if “no” is treated as Boolean NOT operator). Similarly, the musical note "a" presents special considerations in systems which consider it a stop word. With correctly configured phrase searching, a search on “a major” will return bibliographic records containing the string “a major” (a musical key) rather than “major” alone. Finally, the pound sign (#), might be used in a query as a substitute for the musical sharp sign and, when enclosed in quotes, should function as a search term, not an operator such as a truncation or wildcard symbol.
See II.C, III.H, and I.I.E for additional considerations regarding musical key/range and identifying numbers.

**Recommendation:** Configure phrase searching to override Boolean operators and stop words when they are enclosed in quotes, particularly: “no”; “a”; “#”.

**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. While many attributes are relevant to music in special ways, selected attributes are unique to music: opus, thematic index, or serial numbers; medium of performance; musical key/range; musical genre/form; musical presentation.

A separate music-specific 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.

**G. 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. A related issue is that many discovery systems now integrate metadata from a variety of sources (local catalog, electronic resources, digital repositories, etc.) A discovery interface needs to support ingest and use of metadata from all these sources, as well as actual retrieval of items available digitally.

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:

---

47 The University of Virginia’s Virgo Music Search is one example of a music-specific view. Built on Blacklight, the view provides music-specific facets. [http://search.lib.virginia.edu/music](http://search.lib.virginia.edu/music)
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 ProQuest Syndetics Unbound.\textsuperscript{48} 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.\textsuperscript{49} 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 is 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 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.

Other possible pitfalls of incorporating third-party data include a lack of local control, the possibility of broken links or “link rot,” and a search giving the false impression of completeness.

\textsuperscript{48} \url{http://proquest.syndetics.com/}

\textsuperscript{49} As of early 2017, authors of this document were not aware of any vendors that make initial pages of scores available for library incorporation, except occasional incidental inclusion; for example, some Dover scores may show up because they are treated along with the Dover’s book publications. Many commercial vendors do make initial pages available on their websites for selling music.

Music Discovery Requirements
when there may be other available resources not integrated into the discovery system. Weigh the advantages and disadvantages of such issues, and consider labelling third-party data to make it obvious to the patron.

**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. When the discovery system integrates metadata from various sources, support ingest and use of all metadata.
V. Metadata Schemas

A. Introduction

Particular metadata schemas pose their own opportunities and challenges for music discovery. Individual attributes and relationships are discussed in depth in sections II-IV. Section V provides overarching comments on each schema as it relates to music discovery. For details on the versions of each schema used to create the Music Discovery Requirements, see VI Schema Versions Used.

B. MARC
C. Dublin Core
D. EAD
E. EBUCore
F. MODS
G. PBCore
H. BIBFRAME 2.0

B. MARC

In North American library communities, MARC boasts a long history, with the result that most of our library discovery systems need to accommodate large bodies of MARC-encoded data. OCLC’s WorldCat database, which many of our libraries use and contribute to, is a common source of this data. Because of this large body of data, MARC bibliographic and authority formats, following the standards published by the Library of Congress, receive extended review and attention in this document.

MARC originated as a way to create catalog cards containing data devised according to a book-centric cataloging code. While music-specific fields have since received extensive--and intricate--development, cleanly teasing out important music-specific elements like medium of performance or work numbers often eludes our best efforts. Development continues at a rapid pace, with an eye towards enabling us to take advantage of the perceived benefits of Linked Data, not to mention facilitating the coming transition to BIBFRAME. Readers should keep in mind that as they are reading this document, additional changes will be in the works, if not in fact already implemented in some cases.

MARC legacy data poses additional challenges. As MARC standards change, few libraries have the resources to completely update all existing data, especially when new data points are added. Furthermore, there are some fields/subfields that have never achieved widespread implementation. These factors may present difficulties for discovery interface creators because many features will not work well unless data is regularly and uniformly present. This document identifies fields which have been irregularly or recently implemented, alerting discovery interface creators of the need to assess their data for the presence or absence of these fields.
C. Dublin Core

The Dublin Core Metadata Element Set is a "vocabulary of fifteen properties for use in resource description." The name comes from the schema's origins in 1995 at a workshop in Dublin, Ohio. Dublin Core was designed to be intentionally simple to use and flexible, in order to promote international usage for content creators beyond trained information professionals. All elements are optional, all elements are repeatable, and there is neither a mandatory syntax for the terms used nor a prescribed order of elements.

Simple Dublin Core refers to the original fifteen elements, and is the most widely used iteration of the schema. Qualified Dublin Core refines some elements and adds additional elements, and also allows for the usage of controlled vocabularies. Application profiles can be defined for specific communities of users.

Dublin Core's simplicity can be a boon for those inputting metadata, but it also means that many of the nuances in describing musical materials may be lost. Crosswalking MARC data to Dublin Core will inevitably mean a loss of granularity, and Dublin Core's small set of metadata elements mean that it can be difficult to find an exact match for music-specific metadata. Dublin Core also poses a problem for aggregate works, which are common in music, and is not well-equipped to separate the attributes of the various FRBR entities. Because it was developed primarily as a tool for description, Dublin Core is also not well-equipped to take advantage of recent developments aimed at aiding discovery of musical materials. Those looking to use Dublin Core for musical materials should investigate specific application profiles, and be willing to accept a certain lack of granularity and specificity.

D. EAD

EAD (Encoded Archival Description) is “a non-proprietary de facto standard for the encoding of finding aids for use in a networked (online) environment.” As an encoding schema it does not prescribe descriptive practices. The official standard for description of archival collections is found in Describing Archives: A Content Standard.
The focus of archival description tends to be at the level of collections and file series, not items, making it essentially different from other schemas represented in MDR. Even when applied to items, archival description is presented in such a wide variety of methods that general guidelines for harvesting metadata at the item level are not useful. Therefore, EAD has been excluded from this edition of MDR.

**E. EBUCore**

EBUCore is a metadata schema based on Dublin Core developed by the European Broadcasting Union. The “specification addresses the creation, management and preservation of audiovisual material.” As such, it has limited applicability to describing music materials, particularly as regards musical Works and Expressions. It can be used to describe carriers for audiovisual materials as well as for describing events, such as the recording of a live performance. For archival collections with a large number of spoken word or field recordings EBUCore could be very useful. Beyond using EBUCore for structured descriptions of musical resources, EBUCore can be used to annotate the recordings themselves.

**F. MODS**

MODS is a general-purpose XML-based metadata schema developed by the Library of Congress as a response to perceived failings of MARCXML (its strict adherence to MARC conventions and lack of human-readable tags) and Dublin Core (its deliberate simplicity). It bears resemblance to MARC in terms of the types of metadata it can encode and to Dublin Core in the way its elements are named and organized, and it is generally highly compatible with both. MODS is commonly used in many digital library applications.

For musical description, MODS affords a high degree of granularity and controlled access to creators and works, with support for linked data URIs associated with names, titles, and subject or genre headings. However, not all of the more recent additions and changes made to MARC to better accommodate musical description translate seamlessly into MODS. In particular, there is no structured way to provide medium of performance information through a controlled vocabulary such as LCMPT, although it is possible to use LCSH genre/medium of performance headings in a genre element.

The Library of Congress provides mappings of MODS elements to and from MARC and Dublin Core and from pure RDA description. Also available are XSLT stylesheets for converting between MODS, Dublin Core, and MARCXML, or outputting MODS data to an HTML definition list.

**G. PBCore**

[^54]: https://tech.ebu.ch/docs/tech/tech3293.pdf
[^55]: ibid, 7.
PBCore is an XML-based schema for describing sound and video recordings. It was originally designed for use by the public broadcasting community in the United States and puts focus on intellectual property and rights information. It has also seen adoption in media archives. Current development is headed by the American Archive of Public Broadcasting and the Association of Moving Image Archivists. PBCore elements may make use of several existing thesauri, although there are small controlled vocabularies specific to PBCore as well; currently efforts are underway to make PBCore compatible with the EBUCore RDF ontology.

From a musical perspective, the structure of PBCore and the defined scope of its elements encourage the treatment of recordings (such as radio programs or program segments) as the works being described, rather than individual performances or works being performed. In essence, the schema is designed to describe programs about music or of music more thoroughly than the music itself. That means it is more difficult to provide structured, controlled access to specific musical works in PBCore than it is in some other schemas, which, depending on the purpose of the collection being described, may be a more significant drawback for some musical traditions than others. Crosswalking between PBCore and metadata schemas that provide more granular access to musical work data will almost certainly result in some loss of fidelity in either direction.

In environments that contain a significant mix of musical and nonmusical recordings (which seems likely given the impetus behind PBCore’s creation), care should be used not to bend the scope of PBCore elements used for musical description in ways that produce unpredictable or inconsistent results for searchers.

H. BIBFRAME 2.0

Music Discovery in BIBFRAME and other Linked Data ontologies

Discussing music discovery in a linked data environment is difficult. The flexibility and extensibility of linked data on the one hand permits an unprecedented ability to customize the discovery experience, depending on the type of search being done. On the other hand, these same characteristics and the potentially boundless linking require the developer of the discovery interface and indexes to make a lot of choices, the biggest being where to set data boundaries—what you actually display or link to on screen.

That being said, linked data is simply a carrier of data, and basic music discovery requirements are unchanged for the user. The benefits (one hopes!) is that the user experience will be improved through more flexible search possibilities. On the flipside, this may also raise user expectations to improve our data and become even more embedded into the greater web.

Initial linked data discovery as part of library catalogs will likely need to work in tandem with MARC data as well, since not all data may be converted immediately and other aspects of library systems, such as acquisitions, may need still to rely on a MARC record from which to hang ordering and fiscal data. Faceted indexing may well be the bridge, and still serve its purpose in refining searches in a user-friendly manner. That being said, it may end up that such a hybrid discovery layer will result in a catalog that looks much the same as it does now! At the start, this is not necessarily a bad thing.
The primary thing to recognize in developing discovery in a linked data environment is that there will not be a single ontology or framework to work in. While this document has only referenced BIBFRAME 2.0, that ontology, by design, is lightweight and does not include the granular vocabularies necessary for actual description. Additional subclassing is needed in places and terms for named individuals must be drawn from outside vocabulary lists, thesauri, or wholesale ontologies to fill out the description. An ontology may also be extended to provide modelling for areas not present in the primary ontology, either from established ontologies or original modelling. An experimental, though fairly conservative extension of BIBFRAME 2.0 to accommodate audio recordings is currently being developed through the **Performed Music Ontology subproject** of the **Linked Data for Production grant**. Although unfinished as yet, the ontology uses individual terms and subclasses drawn from RDA value vocabularies ([http://www.rdaregistry.info/termList](http://www.rdaregistry.info/termList)), RDA unconstrained properties, the Library of Congress [Linked Data Service](http://www.loc.gov/ld/), and other sources, as well as classes, subclasses, and individuals unique to the project and the performed music domain. A discovery interface will need to accommodate potential differences such as these. Allowing for this variability is a major challenge for discovery.

As stated near the beginning of this document, BIBFRAME mapping for discovery depends a great deal on context, vocabularies used, and local implementation decisions (including use of local and pre-existing extensions). The mappings provided here are not complete, and are intended only as guidance about where to start. BIBFRAME is also still evolving and it is important to check the most current, machine-readable form of the ontology from the Library of Congress. Extended notes on some parts of the ontology and the ontology in human-readable form are available via [BIBFRAME Model, Vocabulary, Guidelines, Examples, Notes, Analyses](http://www.loc.gov/bibframe/modellinks). Basic tooling and mappings for the 2.0 version are in development; what is currently available may be found at [BIBFRAME Implementation, Tools and Downloads](http://www.loc.gov/bibframe/implementation).

**Making the ILS linked-data friendly**

Even before any wholesale conversion to linked data, it is possible to make a library system more linked-data ready, and even emulate, in a highly restricted fashion, some of the benefits of linking.

1. **Allow for the enhancement of MARC records for conversion to linked data**

   Linked data links through URIs, not strings. Without existing URIs, URIs must either be created locally in the conversion process and then reconciled with existing internationally-recognized URIs, or else a look-up in the relevant authority file for every term in a bibliographic record. A better way is to add as many URIs as possible through machine-processing to either bibliographic records, associated authority records, or both. Currently, URIs are recorded in the $0 of many MARC fields in bibliographic records and in the 024 or local fields in authority records, and in the authority URI (for the authority record) and value URI (for the entity) in MODS 3.6. A MARC discussion paper

---

56 The Library of Congress. BIBFRAME Implementation, Tools, and Downloads. at [http://www.loc.gov/bibframe/implementation](http://www.loc.gov/bibframe/implementation) or [http://www.loc.gov/bibframe/docs/](http://www.loc.gov/bibframe/docs/). (Click on the “RDF view”.)
scheduled for ALA Midwinter 2017, proposes the use of $1 as well, to record URIs for entities, as opposed to authority records which are recorded in the $0.

A library system should permit the addition and indexing of URIs in the $0 and possibly the $1 in both bibliographic and authority records. Note that for bibliographic records, potential URIs are not restricted to those generally found in ILS authority files. They may be pulled from other internationally recognized authority files such as ISNI and VIAF, or from vendor-supplied URIs, or be local in origin. The system should also permit URIs already present in records to be loaded into the system along with the rest of the bibliographic or authority data.

Even before converting the MARC record to linked data, some inserted URIs may be put to use to explore other resources if they are active links. Potentially, a discovery interface may display the URI as an icon of some sort beside the string, and link out to other data sources beyond the library catalog.

2. **Actively use the $2 in MARC records to identify vocabularies and authority files.**

   Subject, genre, demographic, and many other terms recorded in MARC records come from thesauri that are specified through the $2 in the relevant 65X or 38x field. An ILS authority module should use the $2 to identify the vocabulary and be able to provide authority records for any and all vocabularies that the client wishes to load into the authority module. Currently, some systems only allow for one of these vocabularies, since the module uses the second indicator “7” rather than the $2 to separate the vocabulary.
## VI. Schema Versions Used

This chart lists the specific versions of each schema that were the basis for recommendations in the both Music Discovery Requirements 2012 (MDR1) and this second edition of the Music Discovery Requirements (MDR2). Consult official documentation for specific metadata schema regarding schema changes since the last update to the Music Discovery Requirements. For additional discussion regarding each standard, see V. Metadata Schema.

<table>
<thead>
<tr>
<th>Schema</th>
<th>MDR1 version</th>
<th>MDR2 version</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDWA Lite</td>
<td>CDWA Lite 1.1 (last updated 2006)</td>
<td>Dropped</td>
</tr>
<tr>
<td>Dublin Core</td>
<td>DC Metadata Element Set Version 1.1</td>
<td>DC Metadata Element Set Version 1.1 (unchanged) Simple DC, with notes about applicable aspects of qualified DC and that the system needs to work with qualified DC.</td>
</tr>
<tr>
<td>EAD</td>
<td>EAD 2002</td>
<td>Dropped</td>
</tr>
<tr>
<td>MARC</td>
<td>MARC21 Format Through MARC update no. 13 (Sept. 2011)</td>
<td>MARC21 Format With updates nos. 14-23 (Nov. 2016)</td>
</tr>
<tr>
<td>MODS</td>
<td>MODS 3.4</td>
<td>MODS 3.6 (released May 5, 2015)</td>
</tr>
<tr>
<td>VRA Core</td>
<td>VRA Core 4 (last updated 2007, except a new introduction in 2014)</td>
<td>Dropped</td>
</tr>
<tr>
<td>PBCore</td>
<td>(not included)</td>
<td>PBCore 2.1</td>
</tr>
<tr>
<td>EBUCore</td>
<td>(not included)</td>
<td>Tech 3923 (EBU Core)</td>
</tr>
<tr>
<td>BIBFRAME</td>
<td>(not included)</td>
<td>BIBFRAME 2.0 <a href="http://www.loc.gov/bibframe/docs/index.html">http://www.loc.gov/bibframe/docs/index.html</a></td>
</tr>
</tbody>
</table>
**Appendices**

Appendix A is a spreadsheet compiling the indexing, display, and facet/limit recommendations in the document. The Appendix B spreadsheet details recommended MARC bibliographic mapping and display terms for content and carrier for printed music, audio, and video.

The spreadsheets should be used in conjunction with the full document, particularly because in some cases multiple options are given for addressing discovery needs, and the extended discussion is contained in the document proper. The spreadsheets are not exhaustive mapping documents; their scope is the same as the document: areas which are music-specific or particularly important for music. When fields are repeatable, ensure discovery systems are set to display and/or index all occurrences.

Note that MDR2 drops MDR1’s Appendix A: Compiled Details of Indexing and Display Requirements (Index Focused). The tag focused spreadsheet (Appendix B in MDR1) can meet the needs served in Appendix A. Therefore, in MDR2, the tag focused version is now Appendix A and MARC Bibliographic Record Mapping for Content and Carrier is now Appendix B. Appendix C is not a mapping document but a spreadsheet outlining history of the coding and definition of Format of Notated Music, to be used in conjunction with III.D Format of Notated Music.

A. Compiled Details of Indexing and Display Requirements (Bibliographic/Descriptive): Tag Focused version (spreadsheet)
B. MARC Bibliographic Record Mapping for Content and Carrier (spreadsheet)
C. Format of Notated Music History (spreadsheet)
Works Cited


McGrath, Kelley. “Media Finders: Improving the Browsable Media Collections via the OPAC.” Internet Reference Services Quarterly 11, no. 3 (2006): 19-34.


“UNIMARC field 146 - Medium of Performance”

Variations/FRBR: Variations as a Testbed for the FRBR Conceptual Model.