

Music Discovery Requirements: A Guide to Optimizing Interfaces

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

This is an English language version of the German language article cited below. The published German language article is freely available at:

<https://oa.slub-dresden.de/ejournals/fmb/article/view/169/150>

Newcomer, Nara L. mit Rebecca Belford, Deb Kulczak, Kimmy Szeto, Jennifer Matthews, und Misti Shaw. "Anforderungen an Musik-Discovery-Systeme: Handreichungen zur Optimierung von Benutzeroberflächen." trans. Torsten Senkbeil and Cordula Werbelow. *Forum Musikbibliothek* 34, no. 2 (July 2013): 14-23. <https://oa.slub-dresden.de/ejournals/fmb/article/view/169/150>

I. Introduction

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

In the midst of this changing environment, music materials pose unique demands that must be considered for successful discovery. New discovery interfaces have exciting potential to ease the difficulties users face when searching for music materials. However, in reality, the specialized discovery needs arising from music materials are often overlooked. This reality led to creation of the Music Discovery Requirements as a guide for vendors, librarians, and anyone developing or implementing discovery interfaces.

The Emerging Technologies and Services Committee of the Music Library Association (MLA) in the United States initiated work on the Music Discovery Requirements in 2011. Following two public comment periods, the MLA Board of Directors approved the document in 2012. Given the rapidly changing bibliographic landscape and the immediate need for a document to aid in discovery interface implementations, the Music Discovery Requirements is not a standard. Instead, it gives recommendations and possible best practices, in conjunction with discussion of the factors and discovery needs that precipitated the recommendations, in three sections: musical works; expressions and manifestations; and other aspects of music discovery. The present article is an abbreviated version of the Music Discovery Requirements, translated into German. For technical details, including the most recent version of best practices for each attribute/relationship regarding indexing, display, and use in facets/limits, consult the MLA web site for the original Music Discovery Requirements and three appendixes compiling technical details in spreadsheet form: beginning with the indexes and giving the fields to include; beginning with the field and giving indexing for each field; and detailing MARC 21 bibliographic record mapping for content and carrier. To facilitate cross-referencing, the original section numbering is retained in the present article.¹

Two elements define the scope of the Music Discovery Requirements: music and discovery. First, the document focuses on musical works (scores and recordings), rather than secondary literature about music (books and articles), because musical works present more unique discovery needs. Second, the focus on discovery mostly excludes back end functions such as circulation, cataloging, and acquisitions which, like secondary literature, present fewer music-specific requirements. However, powerful searching is important for staff work with music materials, and implementing this document’s recommendations will benefit both users and library staff.

The focus on discovery also means recommendations are focused on indexing and presentation of data for discovery, rather than rules for recording data (such as AACR2, RDA, and “Musical Attributes, Refinements, and Recommendations for their Use”²) or metadata formats and encoding standards (such as MARC, Dublin Core, METS, MODS, EAD, CDWA Lite, and VRA Core³). Within this focus on indexing and display, there is a particular need to address: legacy data created according to AACR or AACR2 and encoded in MARC 21, particularly the data in OCLC’s WorldCat database; changes associated with RDA; and the particular relevance of FRBR concepts to music.

Historically, library data recording and encoding standards have been biased towards Western art music. Other musics are often not ideally served by these standards; however, insofar as these other musics have been cataloged following such biased standards, discovery interfaces will not vary much depending on the type of music. Interfaces facilitating discovery for large amounts of music beyond the Western art music tradition may benefit from further customizations, if necessary data has been encoded.

II. Musical Works

A. Introduction

Clear identification and display of information regarding musical works⁴ is crucial. This section addresses important attributes of and relationships to musical works.

B. Titles

Users' discovery and identification of musical works frequently draws on titles. Particular musical works are often referred to by many different titles in various languages, for example: Symphony no. 5, Fifth Symphony, Sinfonie C-Dur, Symphonie op. 67, etc. This reality spurred librarians’ extensive development and application of standardized titles (AACR2’s “uniform titles”) for music materials. To facilitate identification of musical works, it is crucial to display the entire standardized title, and to use authority records or other methods to lead users to the work sought, even when they begin searching with an alternate title for the work. See discussion of authority records in IV.B. Composer name is often essential to identify a work, so creator name must be displayed in conjunction with the standardized title.

When a manifestation contains expressions of multiple musical works (for example, song anthologies), standardized titles are not always assigned for every work expressed; instead, titles are merely transcribed from the item. For most comprehensive coverage, display transcribed titles and include them in title keyword indexes. See IV.C regarding further challenges of compilations.

C. Identifying Numbers

Opus, thematic index, and serial numbers are frequently used in Western art music to identify musical works. Musical works often lack a distinctive title and are instead titled only with a type of musical composition, such as “symphony,” in conjunction with a serial number. Serial number and type must display and function together. 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. General keyword searches on opus, thematic index, and serial numbers often generate many false hits, so consider a separate index for these numbers. See III.C regarding numbers associated with expressions and manifestations.

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 music for a particular medium of performance, without a particular work/expression in mind at the beginning of their search. However, data recording and encoding standards and practices have historically not made medium of performance cleanly and consistently separable from other attributes. Going forward, recording medium in a way that permits machine actionable, granular description of expressions is essential. The prevalence of manifestations with multiple expressions creates further complications, discussed in IV.C. Users may also seek an expression with a particular medium of performance (i.e. an arrangement), discussed in III.G.

E. Musical Key/Range

Original key is useful for identifying tonal works of Western art music, particularly when the work has a generic title and the identifying number is lacking. Edition statements may contain information on the key or range, but 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. Key and range as expression level attributes are discussed in III.H.

F. Dates

Creation date is an important, though neglected, access point for musical works. Users may seek works composed in a particular date range, perhaps a specific individual year or decade, or a 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 seek. 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. Dates associated with expressions and manifestations are discussed in III.I.

G. Persons and Corporate Bodies

The “created by” relationship is very important for users to find, identify, and select musical works by a particular composer. Index and display all personal and corporate names, both transcribed forms and standardized vocabulary. Index and display names associated with a work as subjects (such as music about a person), but as “subjects,” not “authors.”⁵ Once a user has identified a work, the interface should link to more materials related to that person and fully exploit existing standardized vocabulary to locate only materials related to that particular person, perhaps utilizing the bound text string for the full authorized form of name or a unique identifier functioning behind the scenes. For further discussion, see authority records in IV.B.

Users may seek works associated with persons/corporate bodies possessing particular attributes, such as date, nationality, language, or field of activity. If this data was regularly encoded, it could be exploited to answer questions like “What music do you have by French women composers born before 1950?” through use of facets or limits. It might be possible to connect to already-mined data sources such as WorldCat Identities (<http://www.worldcat.org/identities/>) or the Virtual International Authority File (<http://viaf.org/>).

Other persons, especially librettists and lyricists, can be related to the musical work. The relationship is sometimes collaborative (composer and lyricist work together on words and music) and sometimes involves an independent work (composer sets a literary text to music). The many persons and corporate bodies contributing to expressions and manifestations are discussed in III.J.

H. Topical Subjects

True topical headings describe what works are *about*, but many musical works are not objectively *about* anything. Therefore, true topical headings (in FRBR terminology, those expressing a “has as subject” relationship) are both less important and less common for music materials than for books.

In the United States, topically coded subject headings have historically been assigned to identify non-topical aspects including: music of national, religious, and ethnic groups (see II.J and III.K); the source of text set; temporal coverage (see II.F and III.I); medium of performance (see II.D and III.G); and genre/form (see II.I). Given the many attributes historically coded as subjects and the changes underway (particularly with genre/form and medium of performance), discovery systems need to allow maximum flexibility and be able to accommodate the full range of subject-related fields.

I. Genre/Form

Users regularly seek works and expressions with particular genres and forms. To the extent existing data makes the distinction, discovery services should be able to distinguish between genre/form headings and topical subjects. The Library of Congress and MLA’s Music

Genre/Form Task Force are developing the thesaurus *Library of Congress Genre/Form Terms for Library and Archival Materials (LCGFT)*, which will separate form and genre headings from topical subjects. Discovery systems should allow maximum flexibility and decisions must be regularly reviewed and revised to accommodate changes as they occur

J. Geographic Area

Geographic area of musical works (the geographic area from which they emanate, not the geographic area they are about) is particularly important for music other than Western art music but is complicated by the intellectual task of determining and prioritizing associated geographic areas, since multiple geographic areas may be associated with a piece of music including: birth, death, and residence of the composer (“art” music) or performer (“folk” music), area(s) it was intended for, area(s) it was popular, and area(s) “folk” music is associated with. In indexing and display, carefully distinguish between work, expression, and manifestation-level attributes. See III.K regarding geographic area of expressions and manifestations.

III. Expressions and Manifestations

A. Introduction

Musical works frequently exist in many expressions and manifestations,⁶ and a discovery interface allowing navigation through a FRBR Work-Expression-Manifestation-Item (WEMI) tree would be very helpful. However, users also seek versions not neatly aligned 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. Further faceting or limiting is necessary in addition to WEMI levels, and, in the absence of a WEMI-formatted discovery interface, can serve many of music users’ needs. This section addresses important attributes of and relationships to musical expressions and manifestations.

B. Format: Content and Carriers

Music as a discipline depends on 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 (CD, LP, audiocassette, videocassette, videodisc, mp3, etc.) is a manifestation-level attribute. Users need facets and limits to select content types as well as specific carriers. 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.” The display should facilitate easy identification of an item’s content type/carrier. Icons are frequently useful for this task.

C. Identifying Numbers

Whereas ISBN and ISSN serve as fairly uniform standard numbers for books and serials, many numbers are associated with musical expressions and especially manifestations, including ISMN, ISBN, ISRC, UPC, EAN, sound and video recording issue numbers, matrix numbers, and plate and publisher numbers.⁷ Index and display all associated numbers. Numbers associated with musical works are discussed in II.C.

D. Musical Presentation

Users need to distinguish between formats of notated music, such as full score, parts, vocal score, etc. This attribute does not apply to recorded performances of music. Commonly used cataloging standards and metadata formats (including AACR2, RDA, MARC 21 coding, and Library of Congress Subject Headings) do not provide a clear and concise way to support searching and limiting/faceting by musical presentation. Ensure necessary fields for identification are indexed and displayed, explore ways to exploit legacy data, and improve future data capture.

E. Edition

Users seek specific “editions” of notated music, meaning usually a search for a specific editor, publisher, or type of edition (critical edition, facsimile edition, Urtext edition, etc.), not (as is common with books), a search for a numbered edition statement (“2nd edition”). Users may also scrutinize edition information, particularly the editor of notated music, to select an expression/manifestation even if they do not begin with a particular edition in mind. See III.J regarding editors and other associated people and corporate bodies.

F. Language

Three factors contribute to a heightened need to find, identify, and select music materials based on language. First, while general users frequently focus on one or a few languages (i.e. languages they read/speak), seekers of musical works regularly desire a variety of languages, especially with vocal music, where users seek both original languages and specific translations. Second, individual music materials regularly incorporate many languages in various distinct capacities, including program notes, critical commentaries, libretti, dubbing and subtitles. Consider which capacities are most valuable, particularly for facets and limits. Third, 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, and it is unclear what users think of regarding the “language” of instrumental music.

G. Medium of Performance

Users may seek expressions of works with a particular medium of performance other than the original, i.e. arrangements. Users may also seek music for a particular medium of performance regardless of whether it is the original medium. See II.D for further discussion of medium.

H. Musical Key/Range

Musical key and range (complete span of pitches used in the piece of music) are particularly important for vocal music, but instrumental music can also exist in expressions distinguished by the key or range. While searching by key and range would be useful, necessary data is rarely present. Original key and range of musical works are addressed in II.E.

I. Dates

Like musical works, musical expressions and manifestations have many associated dates. Discovery interfaces commonly allow searching, limiting, and faceting by date, 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. For example, performance date is an important expression-level attribute for recorded music. Unfortunately, these dates may be buried in notes or completely absent from data. Ideally, take into account the full spectrum of associated dates. Minimally, clearly identify the type of dates being exploited, for example, using the label “Publication Date” rather than just “Date.”

J. Persons and Corporate Bodies

Many persons and corporate bodies contribute to music materials and hold responsibility at the expression and manifestation levels, especially the expression level, including editor, arranger, performer (individuals and groups), producer, and director. Therefore, index and display all personal and corporate names, both transcribed forms and controlled vocabulary. Index and display the person/corporate body’s function if present in the data, converting codes to the vernacular.

Western art music generally makes a clear distinction between composer (related to the work) and performers (related to the expression). Many other musics do not make this sharp distinction. Following this document’s recommendations will make related persons and corporate bodies useful for finding, identifying, and selecting music, regardless of the relationship’s exact nature.

K. Geographic Area

Geographic area is arguably a less important attribute for expressions and manifestations than for musical works. For recordings, expression geographic area is often included in a free-text note not designed for machine manipulation. Manifestation geographic area is regularly recorded and coded, providing an easily machine-actionable field for facets or limits. For books published in their original language, manifestation geographic area may be a somewhat successful approximation of work geographic area. For musical scores and recordings, this is not true. Clearly label manifestation geographic area as referring to place of publication.

IV. Other Aspects of Music Discovery

A. Introduction

This section addresses music discovery issues not focused on particular attributes or relationships.

B. Authority Records

Varying terminology is used to represent music-related access points, especially musical works because, outside of cataloging codes, a title's language, elements, and grammatical construction are inconsistent. Therefore, authority records are essential not only for back end functions like cataloging, but also in public interfaces to provide cross-references and other information and to lead users from alternate terms to the materials they seek. Historically, "see" and "see also" references in browse lists provided this guidance. Unfortunately, keyword searching has mostly ignored "see" and "see also" data. Mere spell check features, though potentially helpful, will not solve this problem. Auto-complete or "did you mean" features are useful, particularly when based on database contents or rich, authoritative sources such as the Library of Congress/NACO Authority File (LC/NAF) or the Virtual International Authority File (VIAF).⁸

The use of linked data, which relies on an identifier rather than a text string, has made initiatives like the VIAF possible. 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.⁹ Future systems could allow the library to choose which heading string to display, even if it is a "see" reference, for example, to display all titles in the local language. Future systems might further manipulate work-level attributes, such as producing a list of a composer's oeuvre organized by opus number or powering limits, facets, and other displays.

Current authority records are very similar to work records in a FRBRized environment. In the future, many important attributes will ideally be placed in work records rather than bibliographic records, making indexing and display of work or "authority" data absolutely crucial.

C. Compilations

Multiple expressions are frequently issued together in a single manifestation, sometimes called an aggregate.¹⁰ This is particularly common for recordings, but 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 between a) works; and b) works/expressions and their performers, medium of performance, or format of notated music are often recorded in notes designed for human interpretation but not linked in a manner friendly to machine processing. In future data creation, this problem could be solved by creating WEMI record trees and/or utilizing linking fields.

D. Searching: Alphabetical and Keyword

Historically, alphabetically filed “browse” listings have been indispensable in library information retrieval, beginning with the card catalog. OPACs added keyword search functionality. “Next-generation” catalogs added faceted browsing, but often eliminated alphabetical browse searching.

Alphabetical listings have been particularly important to music librarianship. “Known-item” queries are more common in music searching (both “classical” and “popular”¹¹) than in general searching, but users frequently “know” little about the entities sought. Title is a notoriously unreliable access point, especially for Western art music, and “subject” is a problematic concept. Perhaps because of these difficulties, personal names are frequently used as an access point in music searching.¹² However, many composers are very prolific, and their works exist in many versions, making personal name alone an insufficiently precise access point.

These deficiencies led to extensive development of uniform titles (especially generic uniform titles) and subject headings (which reflect true “aboutness” only occasionally) to provide access to music materials based on attributes beyond author, title, and subject. When organized into alphabetical lists, they made it possible for informed users to reliably find and browse musical works. This system does not transfer well to a keyword environment because it is extremely difficult for untrained users to predict the vocabulary used in uniform titles and subject headings. See IV.B regarding authority records and keyword searching. However, the needs historically met through alphabetical listings might be met through other means, particularly faceted browsing of attributes important to music. Retain alphabetical searching in back-end interfaces even if eliminated from public interfaces.

E. Enhancements: Third-Party Content

The ability to incorporate or link to third-party content is an important enhancement to discovery systems. Existing data (such as cover art, reviews, summaries, previews of initial/selected pages) from vendors, publishers, and other content providers can be leveraged to increase the information available in and through a discovery system. Audio or video clips would greatly increase music users’ capability to determine if an item suits their needs. For scores, initial pages

of music are much more useful than the cover to help users identify the work and evaluate style, difficulty level, and similar considerations. Using ISBN as the only match point for third party content will not function for items lacking an ISBN, as is often the case with scores and sound recordings. Create match points on additional standard numbers such as publisher numbers, recording issue numbers, or ISRCs. Most importantly, third-party content must enhance, not replace, existing data. Sound recordings are particularly at risk with the various versions of genres and “contents” that may exist across data sources.

F. Music-Specific Interface/View

As discussed throughout this document, music materials possess many unique and uniquely important attributes, and multiple expressions and manifestations are the norm. Therefore, consider creating a music-specific interface/view, especially for tools designed to discover collections containing large amounts of music. However, consider other factors also, particularly the resources needed to maintain a separate interface/view. If an alternate interface/view is created, the general interface/view should lead users seeking music materials intuitively to the alternate interface/view. This document’s recommendations can be implemented in both general and music-specific interfaces/views.

Endnotes

¹¹ Music Discovery Resources. <http://committees.musiclibraryassoc.org/ETSC/MDR> (accessed 29 January 2013). A version of the Music Discovery Requirements was also published as Newcomer, Nara L., Rebecca Belford, Deb Kulczak, and Kimmy Szeto with Jennifer Matthews, Misti Shaw, "Music Discovery Requirements: A Guide to Optimizing Resources," *Notes* 69, no. 3 (March 2013): 494-524.

² Music Library Association Bibliographic Control Committee Metadata Working Group, "Musical Attributes, Refinements, and Recommendations for Their Use," 2008, <http://bcc.musiclibraryassoc.org/BCC-Historical/BCC2008/BCC2008MSWG2.html> (accessed 29 January 2013).

³ MARC: MACHine Readable Cataloging, <http://www.loc.gov/marc/> (recommendations given are for MARC 21); Dublin Core: Metadata Initiative, <http://dublincore.org/>; METS: Metadata Encoding and Transmission Standard, <http://www.loc.gov/standards/mets/>; MODS: Metadata Object Description Schema, <http://www.loc.gov/standards/mods/>; EAD: Encoded Archival Description <http://www.loc.gov/ead/> (recommendations given are for EAD 2002); CDWA Lite: Categories for the Description of Works of Art, http://www.getty.edu/research/publications/electronic_publications/cdwa/cdwalite.html; VRA Core: data standard from the Visual Resources Association, <http://www.vraweb.org/projects/vracore4/>. CDWA Lite and VRA Core are rarely ideal for music, but they are included in this document for the benefit of situations where they may be used. (all sites accessed 29 January 2013).

⁴ Here defined according to FRBR: a distinct intellectual or artistic creation.

⁵ Mostly, subject headings for persons and corporate bodies reflect subject relationships. Authors of texts set to music are an exception. SHM H 1110 instructs addition of a subject heading for the name of the person whose writings or words are set to music, with the form subdivision "Musical settings" appended. In this case, the person has a creator, not a subject relationship to the work. However, AACR2 also instructs for added entries for the person whose words are set, making manipulating the subject headings to bring out the creator relationship less vital.

⁶ Here defined according to FRBR: expression: the intellectual or artistic realization of a *work*; manifestation: the physical embodiment of an *expression* of a *work*.

⁷ ISMN: International Standard Music Number; ISRC: International Standard Recording Code; EAN: International Article Number; matrix number: used to indicate sound recording masters, often used to identify 78s.

⁸ See, for example: Demian Katz, Ralph LeVan, and Ya'aqov Ziso, "Using Authority Data in VuFind," *Code4Lib Journal* 14 (July 25, 2011), <http://journal.code4lib.org/articles/5354> (accessed 29 January 2013).

⁹ 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 29 January 2013).

¹⁰ "Aggregate" may be defined as "a manifestation embodying multiple distinct expressions" and may represent a collection based on genres or performers, or may be an apparently "random collection of expressions." Working Group on Aggregates, "Final Report of the Working Group on Aggregates" (2011): 3. <http://www.ifla.org/files/cataloguing/frbrg/AggregatesFinalReport.pdf> (accessed 29 January 2013).

¹¹ Beth Christensen, Mary Dumont, and Alan Green, "Taking Note: Assessing the Performance of Reference Service in Academic Music Libraries: A Progress Report," *Notes* 58, no. 1 (September 2001): 52. Jin Ha Lee, "Analysis of User Needs and Information Features in Natural Language Queries Seeking Music Information," *Journal of the American Society for Information Science and Technology*, 61, no. 5 (2010): 1037.

¹² David M. King, "Catalog User Search Strategies in Finding Music Materials," *Music Reference Services Quarterly* 9, no.4 (2007): 17.