

Mission

Editorial Committee

Process and Structure

Code4Lib

ISSN 1940-5758

Current Issue Music

Issue 38, 2017-10-18

Previous Issues

Issue 37, 2017-07-18 Issue 36, 2017-04-20 Issue 35, 2017-01-30 Issue 34, 2016-10-25 Older Issues

For Authors

Call for Submissions Article Guidelines

Look What We Got! How Inherited Data Drives Decision-Making: UNC-Chapel Hill's 19th-Century American Sheet Music Collection

Have you inherited a digital collection containing valuable, but inconsistent metadata? And wondered how to transform it into a usable, quality resource while accepting that it can't meet your idea of perfection? This article describes such an experience at the University of North Carolina at Chapel Hill University Library with its CONTENTdm-based 19th-Century American Sheet Music Collection, addressing issues such as field construction, the use of controlled vocabularies, development of a project data dictionary, and metadata clean-up.

by Renée McBride

Introduction

In December 2009 the University of North Carolina at Chapel Hill (UNC-Chapel Hill) University Library completed migrating its 19th-Century American Sheet Music Collection – a collection of approximately 3,500 popular vocal and instrumental titles from the 19th century – from a locally developed PHP application backed by a MySQL database to CONTENTdm, a software package enabling libraries to manage and provide access to their digitized collections. UNC-Chapel Hill adopted CONTENTdm as its digital asset management system in 2005. Since then we have engaged in a broad effort to eliminate digital silos (e.g., collections in Access, FileMaker Pro, MySQL) by migrating them to CONTENTdm or our Millennium-based library catalog, and making them available through our information access platform Endeca, depending on the nature of the collection. Our goals in this area are to offer our users easier access to digital collections, to create more consistent workflows and interfaces across the collections, to extend the benefits of enhancements of one collection to all collections, and to achieve interoperability.

Background

A brief history of our project is necessary to provide the context in which our migration and project decisionmaking took place. In the 1940s or early 1950s, 125 bound volumes of sheet music came to UNC-Chapel Hill's Wilson Library, followed by a few additions in subsequent years, resulting in a total of 131 volumes, now housed in the UNC-Chapel Hill Music Library. Cataloging of various sorts occurred in the early 1950s and during the 1960s, and a major scanning and cataloging effort took place in the mid-1990s. In 2004 the collection was migrated from static HTML pages to a locally developed PHP/MySQL application and provided with a custom-made public interface. Scanning and cataloging continued in fits and starts and with a variety of staffing over the next few years. In 2008 serious planning for the migration of this collection to CONTENTdm began. A student from UNC-Chapel Hill's School of Information and Library Science (SILS) exported the data from MySQL into a format that CONTENTdm could read and began a project wiki that we use to this day. Another SILS student loaded a backlog of previously scanned images - many of less than ideal quality - into CONTENTdm and helped with the process of rescanning images and getting them into CONTENTdm. The purpose of rescanning was to replace poor-quality with high-quality images. The rescanned images were created and stored in our digital archive as TIFF images, then converted to JPEG2000 in CONTENTdm, allowing for features such as loading only part of an image when zooming in on an image. In January 2009 I assumed management of the project. At that point in time we were in the process of rescanning images of music for which metadata had already been created. Once a volume was rescanned, the images were uploaded into CONTENTdm with their metadata. Rescanning was completed in April 2010. As of the writing of this article 121 of the 131 total volumes are completed, and we anticipate finishing all cataloging by June 2011.

As you may imagine, there was something of a lack of consistency in workflow, supervision, policies and procedures, and staffing over the many years this project has been in existence. By extension, there was also a lack of consistency in the quality of the metadata. Project staff ranged from librarians to SILS students to musicology graduate students. Once in a while, we struck gold with staff experienced in both cataloging and music. Students with systems expertise were invaluable in helping us learn how to scan, organize and transfer files, and work in CONTENTdm, but were rarely involved in creating metadata. The uncertainty of our staffing was a driving force behind some of the decisions we made in this project. The other primary factor was the amount of metadata of uneven quality that we inherited.

Two broad areas required decision-making and subsequent implementation before we could entertain the thought of going live with the CONTENTdm version of this collection: the technical area, responsible for the content of the collection, and the public interface area, responsible for offering an attractive and friendly product to our users. My primary area of responsibility was the technical area, although I consulted closely with our Public Services Music Librarian on both technical and public interface issues throughout the decision-making

process. Additionally, we received invaluable input from student staff and systems librarians.

This article focuses on the following technical issues and areas of decision-making:

- which fields to retain, delete, or change;
- which Dublin Core fields our CONTENTdm fields would map to;
- which fields required controlled vocabularies;
- creating and maintaining a project-specific data dictionary;
- which fields could reasonably be cleaned up, and which we would choose to live with (or, as one of my systems colleagues put it, which fields offered us an "acceptable level of embarrassment"); and,
- decisions affecting browsability.

Field Construction and Dublin Core Mapping

CONTENTdm allows complete freedom in the naming of fields. At UNC-Chapel Hill we maintain a general digital collection data dictionary that is dynamic, enabling the addition of fields as required by collections with specialized information, e.g., latitude and longitude for map collections, or artist for art-related collections. A very few of our digital collections exist in non-compliance with the data dictionary, for reasons such as the inheritance of metadata of uneven quality in an otherwise valuable collection, precisely the case with our 19th-Century American Sheet Music Collection. Once we determined that this collection would exist, at least for the time being, outside our ideal digital collection world, we began the process of making decisions about its content.

The collection's original 41 fields now number 35.



Figure 1. CONTENTdm Fields

Some decisions were fairly simple matters, such as deleting fields determined to have no use and containing highly inconsistent and confusing metadata. Examples of such fields were Series Instrumentation, Series Dates, and Series Description, none of which existed in any recent project staff's memory. Over the years the terms "series" and "volume" appeared to have become tangled up with each other, so that by the time we were looking at these fields, they simply made no sense. We also did away with Publisher's Quote, Colophon, and Advertisement, which contained no valuable metadata. We reduced five title fields to three fields, incorporated the Cover Design field into the Illustration field, and changed the names of two fields: Cover Inscription became Inscription, and Subject became Topic; these fields are discussed in more detail below. An exception to our general policy of deleting fields of questionable use and containing inconsistent and confusing metadata is the field labeled Unknown – by whom, when, and for what purpose no one knows. After determining that this field contains a variety of possibly useful information, we decided to retain the field but hide it from public view and not use it for current cataloging. We hope to be able to extract useful metadata from this field at a later time and move it to appropriate fields to enhance description and access.

We added a few fields to bring the collection partially in line with our digital collection practice:

- Type (resource type, i.e., Sheet music)
- Repository (entity responsible for publishing the digital object, i.e., University of North Carolina at Chapel Hill. Music Library)
- Digital Collection (name of the digital collection of which the digital object is a part, i.e., 19th-Century American Sheet Music)
- Host (University of North Carolina at Chapel Hill)

More challenging were fields devoted to titles, illustrations, instrumentation, and topic. Our original five title fields:

- Full Title
- Short Title
- Caption Title
- Inside Title
- Other Titles

were compressed into three, hopefully less confusing, fields:

- Title (title proper)
- Caption Title (if different from Title)
- Other Titles (all other expressions of title)

The original fields Cover Design and Illustration were reduced to a single field, Illustration, incorporating selected content from Cover Design into Illustration (more on this in the following discussion of metadata cleanup). We also eliminated the earlier practice of describing such illustrations as calligraphic swashes and decorative designs, limiting the use of this field to descriptions of scenes, portraits, and photographs.

Originally there was an Instrumentation field and a Subject field, with a majority of Subject entries expressing in fact instrumentation or genre. We kept the Instrumentation field and changed the name of the Subject field to Topic, using it only for topics of vocal works. No vocabulary control had been applied to either of these fields, resulting in many conflicting headings for the same concept, e.g., Choir music and Choral music, or Solo song and Solo voice. Our primary purposes in all our decision-making were to simplify both the user's understanding of the metadata and project staff's creation of metadata. As we made decisions about field construction, we took efforts to create self-explanatory field labels.

Concomitant with the construction of our fields was their mapping to Dublin Core fields (see the "DC map" column of Figure 1). Our CONTENTdm collections are harvested by the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), which employs Dublin Core as its common metadata format, allowing the sharing of digital collections created in varying metadata formats. Once mapped to Dublin Core, data can be harvested by anyone who has adopted OAI-PMH.

Controlled Vocabularies and Data Dictionary

CONTENTdm contains several thesauri, e.g., the Getty Research Institute's Art & Architecture Thesaurus, the Library of Congress's (LC) Thesaurus for Graphic Materials I, and MeSH (the U.S. National Library of Medicine's Medical Subject Headings). It does not contain nor can it be connected to the LC Name Authority File (LCNAF) or LC Subject Headings (LCSH). Therefore, we enabled controlled vocabularies for the fields for which authority control and consistent metadata quality were most important. Such control is important not only for consistent search results, but also for public browse capability. A crucial question we addressed was: What fields do we want our users to be able to browse?



Figure 2. Sheet music collection opening screen

The fields for which we offer browsing capability are shown in Figure 2. Effective browsability requires that metadata be consistently applied, therefore browsable fields tend to have controlled vocabularies. For example, a sheet music researcher could likely be interested in studying the history of music publishing in a particular location and would find the ability to browse this collection by publisher location highly useful.

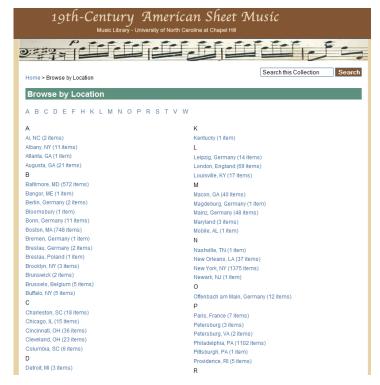


Figure 3. Browse by Location

When a field is made browsable, a hyperlinked, alphabetical listing of the contents of that field throughout the entire digital collection is made available to the user. If this particular vocabulary were not controlled, state abbreviations would appear in various forms, as would names of non-U.S. cities, e.g., Mainz and Mayence, Germany, and Brussels and Bruxelles, Belgium. The same logic extends to providing reliable browsing of, for instance, composer and lyricist names.

As we made controlled vocabulary decisions, we developed guidelines for creating entries for all fields, controlled or not, in the form of a project-specific data dictionary. [1] (See Appendix A for Data Dictionary.)

As with our decisions regarding field construction, our primary purpose in creating the data dictionary was to simplify the creation of metadata for project staff. Our fundamental principles in the construction of the data dictionary were simplicity and consistency. These principles allowed us to make steady progress dealing with a mass of inherited metadata of uneven quality with no control mechanisms, and to successfully handle our situation of uncertain staffing, which required fairly frequent training of new, and always temporary, project staff.

Clearly, names need to be controlled, so we enabled controlled vocabularies for our four name fields:

- Composer
- Lyricist
- Arranger
- Lithographer, Engraver, or Artist

After completing metadata clean-up (to be discussed later) of the name fields, we combined the originally four separate name controlled vocabularies into one large shared controlled vocabulary that contains all the names used in these four fields. Due to the sheer number of names we inherited, we couldn't consider bringing them all into conformity with LCNAF, AACR2, and local digital collection practice. But we could ensure consistency within the database, making sure that one, and only one, form of each name exists in the controlled vocabulary, and adding new metadata according to the guidelines laid out in our data dictionary. As much as possible, we drew on previous project practice as our guide to constructing names in order to avoid recreating everything from scratch, and we formalized practice in the data dictionary. The "Comments" column of our data dictionary contains instructions about how to construct names, with examples given in the "Content Examples" column and guidance about where in the musical score to locate name fields metadata in the "Vocabulary Source" column.

Other fields for which we enabled controlled vocabularies are:

- Publisher; Publisher Location: Left uncontrolled, these fields would give users unpredictable search results, and their browse views would be a mess, with geographical terms such as state abbreviations and names of foreign cities appearing in various forms. As with personal and corporate names, we did not follow LCNAF and AACR2 practice, but ensured consistency within the database.
- Topic: We fairly quickly decided against using LCSH for our Topic field due to training challenges with our frequently changing project staff. LCSH is simply too complex to use with student staff that experiences regular turnover. After considering several thesauri, we decided to rely on Wikipedia, with Webster's Third New International Dictionary, Unabridged (available via our Library's website) as our backup. Wikipedia has served our needs beautifully. Wikipedia concisely describes the usually simple topics addressed by vocal works in our collection, e.g., Christmas, Death, Loneliness, and Patriotism. More complex topics are expressed much more simply by Wikipedia than LCSH, e.g., American Civil War vs. United States—History—Civil War, 1861-1865; Battle of Prague (1757) vs. Prague, Battle of, Prague, Czech Republic, 1757; and Silver standard vs. Bimetallism, or Silver question, or Currency question, or Money, or Coinage. As in other areas of decision-making, we chose Wikipedia as our primary topical thesaurus in order to simplify the creation of metadata for project staff, as well as to easily ensure consistency in the Topic field.
- Volume: Consistency in this field is most important to staff working behind the scenes. For instance, when we uploaded rescanned images with their metadata, we had to make sure that the volume part of the image file names exactly matched the volume numbering in the metadata, which in turn had to be consistent throughout the individual records for the musical works from each volume. Additionally, a public use is served when someone wants to look at the contents of a full volume, as at least one local musicology professor assigns his students to do.
- Type; Repository; Digital Collection; Host: These fields contain constant data, so enabling controlled vocabularies for them ensures consistency, as well as enables extremely easy metadata input for these fields. The cataloger simply double clicks on the entry in the controlled vocabulary, and the field is accurately populated.

Metadata Clean-up

When metadata clean-up could be accomplished using CONTENTdm's "Find & replace" function, we were able to tackle large amounts of data. The function "Find and replace a single field" served our needs best, since we attacked clean-up on a field-by-field basis.

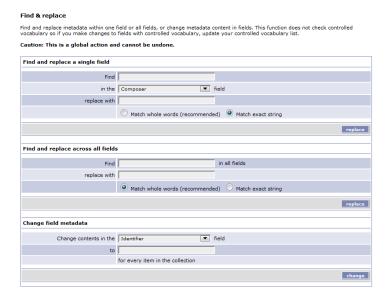


Figure 4. Find & replace function

Additionally, making changes across all fields is a high risk activity for the reason that words or phrases needing correction in some fields occasionally occur in contexts not requiring correction. When clean-up could be achieved only one record at a time, we cleaned up fields involving manageable amounts of data, but sometimes had to opt for the "acceptable level of embarrassment."

Examples of large clean-up efforts include the following fields:

- Composer; Lyricist; Arranger; Lithographer, Engraver, or Artist
- Topic
- Illustration

Once guidelines for creating names were established and controlled vocabularies for each name field were enabled, clean-up began with the controlled vocabularies. Controlled vocabularies were originally populated with the existing contents of their corresponding fields. Many names appeared in multiple forms, so duplicate forms were deleted after the authoritative form was selected. When the controlled vocabulary was clean, we proceeded to cleaning up that field's metadata, accomplished by using the "Find & replace" function mentioned above. The incorrect form of name was entered in the "find" box, and the correct form in the "replace with" box. Once this process was completed, we ran the Verify command for that field, comparing a field's metadata across all records with the field's associated controlled vocabulary.

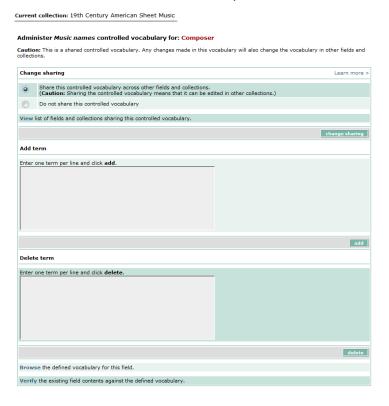


Figure 5. CONTENTdm Verify command

Ideally the Verify command delivers zero results:

Current collection: 19th Century American Sheet Music

Verify: Music names for: Composer

The field contents conform to the defined controlled vocabulary

Figure 6. Verify results

Results indicate a mismatch, telling us one of two things: 1) an error exists in the metadata, or 2) metadata exists that has not been added to the controlled vocabulary. When we achieved zero results from the Verify command, that field's clean-up was completed. To this day we run the Verify command on fields with controlled vocabularies on a monthly basis as a method of quality control.

Once all name fields were cleaned up, we combined the four separate controlled vocabularies into a single shared controlled vocabulary containing all the names used in the four fields. Using a shared controlled vocabulary has proven helpful in streamlining data entry for these fields, because some names appear in multiple name fields. For example, the composer of a work might also have written the lyrics to that work, and might have arranged other works.

As mentioned earlier, the Topic field was originally called Subject. We renamed the Subject field Topic, hoping to make it clearer that this field is used only for songs about a specific topic. We removed all non-topical terms from the controlled vocabulary, corrected terms in the controlled vocabulary as appropriate, then cleaned up metadata using the "Find & replace" function. It is possible that some records that would benefit from the Topic field are lacking it. In addition to making sure that correct topical terms were used, this metadata clean-up required removing numerous instrumentation and genre terms from the Topic field of records. Because of the magnitude of this problem, we were unable to move those terms into the Instrumentation field of their records, a task that would have required record-by-record editing. The information was simply deleted so that the Topic field would be clean across the database. This is an excellent example of a less than ideal decision made when the amount of inconsistent inherited metadata was simply too large to fix. What we did fix was the general integrity of the database.

In contrast to the problem described above with the Instrumentation field, we were able to transfer valuable metadata from the original Cover Design field to the Illustration field because it was of a manageable amount, although the transfer of metadata did have to be done one record at a time. Our Public Services Music Librarian reviewed the contents of the Cover Design field to determine what metadata should be retained. This metadata was moved to the Illustration field and the Cover Design field subsequently deleted.

Once metadata clean-up was completed, we knew that every detail wasn't as perfect as it might have been if we had designed things from scratch, but we also knew that the database was in stunningly better condition than when we had begun.

The Future

As pleased as we are with the current state of the 19th-Century American Sheet Music Collection, we have plans and hopes for future improvements. Most immediate is making the collection accessible via the Sheet Music Consortium, "a group of libraries working toward the goal of building an open collection of digitized sheet music using the Open Archives Initiative:Protocol for Metadata Harvesting (OAI:PMH)." [2] More labor intensive, but nevertheless attainable, goals are to bring names into compliance with LCNAF and topics with LCSH, which would bring the collection in line with our local digital collection practice. The importance of these goals lies in the fact that the UNC-Chapel Hill Library plans to make CONTENTdm collections available via Endeca, which offers relevance-ranked keyword search results as well as faceted navigation built on existing library metadata. It is highly undesirable to include collections lacking the authority control applied to names and subjects in our Millennium-based library catalog. Although our CONTENTdm collections will not actually reside in Millennium, the distinction between Millennium and Endeca is not apparent to our users, nor should it be. An Endeca search retrieves integrated results, providing access to a variety of the Library's collections, not just those cataloged in the MARC format. A current example of such resources available via Endeca is our EAD finding aids for special collections. A final hope for improving the sheet music collection is to populate the Instrumentation field in all records.

When the scanning and cataloging of all 131 volumes of sheet music is completed, we will turn our attention to improving an imperfect but highly usable and consistent collection. Inherited projects like the 19th-Century American Sheet Music Collection present many challenges, some that cannot be resolved in an ideal manner, but the collections are nonetheless valuable. As heir to such a project, we strove to improve the weakest parts of the project and to ensure quality from the present day forward.

End Notes

[1] Lois Schultz and Sarah Shaw's Cataloging Sheet Music (Scarecrow Press and Music Library Association, 2003) was helpful in forming some of our data dictionary's field definitions.

[2] Sheet Music Consortium: About the Project. [cited 2010 March 10]. Available from: http://digital.library.ucla.edu/sheetmusic/OAIProject.html

Acknowledgments

I would like to acknowledge several UNC-Chapel Hill colleagues for their help and advice with various aspects of this article:

- Lynn Holdzkom (Head of Special Collections Technical Services)
- Ericka Patillo (Ph.D. student, School of Information and Library Science)
- Jill Sexton (Information Infrastructure Architect)
- Tim Shearer (Head, Applications Development Team)
- Diane Steinhaus (Public Services Music Librarian)

About the Author

Renée McBride is Head of the Special Formats & Metadata Section of the Resource Description & Management Dept. at the University of North Carolina at Chapel Hill University Library. Her previous positions include Technical Services & Arts Liaison Librarian at Hollins University, and Humanities & Music Cataloger at UCLA.

Appendix A

GENERAL RULE: If data for a particular field does not exist, **do not enter** any information in the field. **Do not** enter information such as **unknown**, **n/a**, etc. The single exception is the Title field, which is mandatory. See Title field instructions below for guidance.

CONTROLLED VOCABULARY: When this is listed as the Vocabulary Source, it refers to CONTENTdm's Controlled Vocabulary for the field in which you are working.

TRANSCRIBE: When this is listed as the Vocabulary Source, it means transcribe exactly what you see in the work.

If name consists of only initials, a surname, or a

MULTIPLE ENTRIES IN A FIELD: When you have more than one entry in a field, connect the entries with ;space, e.g. piano; voice; violin (possible entry in Instrumentation field), English; French (Language field), Standard pieces for piano; Twelve standard pieces for piano (Other Titles field).

DIACRITICS DISPLAY: If your diacritics are not displaying correctly, do the following in Firefox: View/Character Encoding/Unicode (UTF-8) In Internet Explorer: View/Encoding/Unicode (UTF-8)

Field	Obligation	Comments	Content Examples	Vocabulary Source
Identifier	Mandatory			System supplied
Composer	Required if available	Personal & corporate names only; no functions, e.g. arranger,	Bellini, Vincenzo	Controlled vocabulary Title page or caption
Lyricist	Required if available	personal or corporate. Function is sometimes part of a corporate name, in which case it may be included. Do not use Anon. or Anonymous. Personal names Surname, Forename, Title (as applicable) Prepositions and articles follow forename. Titles/abbreviations to use Col., Dr., Hon., Maj., Miss, Mme., Mrs., Rev., Sir As others arise, ask before entering. Do not use Esq./Esquire. Use Mr. only when no forename or initial given. Spell out abbreviated forenames unless spelling is	Edw. L. Balch, Music Printer Endicott & Co. Browne, Miss Neukomm, Sigismund, The	Controlled vocabulary Title page or caption
Arranger	Required if available			Controlled vocabulary Title page or caption
Lithographer, Engraver or Artist	Required if available		Chevalier Pinna, Joseph de Suppe, Franz von Hache, Theodore von la Becket, Gilbert a, Mrs. Barton, J., Maj. Blackwood, Price, Mrs. Drake, Mr. Kontski, Ant. de (we aren't sure how to spell out Ant.) A Lady of New York Lefebure-Wely M.M.O.	Controlled vocabulary Usually title page or caption; engraver information may also be found at bottom of title page or on the last page of music. Consult the appropriate controlled vocabulary list for your person or corporate body. A correct entry from the list may not be an exact match to what is on your piece of music, explicitly differently in different publications, or your composer's name may have the forename initialized or
		Geo.	Abel, J.L.	abbreviated on your piece of musi

Koch, H.A.R.,

If you are uncertain whether your

		phrase, transcribe exactly, beginning phrase with articles if they exist. No spaces between adjacent initials. If more than one name in a field, connect with ;space Do not use any punctuation to indicate questionable information, i.e., do not use [brackets], (parentheses) or "quotation marks." Include diacritics if you always find the name with diacritics. If the name appears both with and without diacritics, do not include diacritics. Information about the entities in these name fields goes in the Additional Information field.	Dr. Caudill, William M.; Walker, William	person or corporate body is a match to a name in the controlled vocabulary, or as to how to enter the name, please consult with supervisor. Manually enter a person or corporate body only if your entry is not in the list. For these four names fields, there may be one and only one correct entry. This form of name will be in the controlled vocabulary. If you encounter a new name, you must add it to the controlled vocabulary.
Performer	Required if available	Phrase about a performer of the work	Sung by Miss Taylor As sung by Miss Matthews As performed by the Germania Musical Society	Transcribe from title page or caption (information located above first page of music)
Title	Mandatory	Title proper: subtitle Capitalize only first word of title, unless: 1) title contains proper names or words commonly capitalized, e.g. I, Christmas 2) title is in German, in which case nouns are capitalized Include definite and indefinite articles at beginning of title. Do not include non-title information, e.g. edition, composer, arranger, lyricist, publisher, dedication, etc. If a work does not have a title, supply one based on the information you have. The Title field is mandatory and must contain metadata. Do not enclose title in brackets. When you have to supply your own title, add in the Additional Information field the note: Title supplied by cataloger. If a file contains more than one work, enter multiple titles connected with ;space between titles. See next column for example.	12 standard pieces for piano : floating on the wind Alice : romance The devoted Land meiner seeligsten Gefiehlen Sechs Sonatinen in fortlaufender Schwierigkeit Come where the fountains play : La favorita; March masquerade	Transcribe from title page, or caption if no title page
Caption Title	Required if available	Title at top of first page of music Enter only if different from Title. If caption contains lengthy other title information, enter in this field only the first part of the title; enter any further title information in the Other Titles field.	Floating on the wind Alice, where art thou?	Transcribe from caption
Other Titles	Required if available	Title not covered by previous title fields, or portion of title that might reasonably be searched	Standard pieces for piano (a portion of the title one might search) Twelve standard pieces for piano (offers searching of numeral spelled out) Alice, ou donc estu? (title given in French after the primary English title on caption)	Anywhere in score
Edition	Required if available	Edition statement	10th ed. Fourth edition New editions carefully revised	Transcribe from title page, or first page of music if not given on title page

Publisher	Required if available	Primary publisher only First publisher listed. Enter publisher name exactly as it appears in publication. Be careful not to confuse publisher information with that about engravers/lithographers/printers. No spaces between adjacent initials. Publishers' names may appear in different forms and languages on different pieces of music. Use the form that appears on the work you are cataloging.	A.E. Blackmar & Bro. D.P. Faulds Lee & Walker B. Schott B. Schott & Sons B. Schott's Sons Le fils de B. Schott	Transcribe from title page, or first page of music if not given on title page. Consult the controlled vocabulary for your publisher. Manually enter your publisher only if it is not in the controlled vocabulary. If you encounter a new name, you must add it to the controlled vocabulary.
Publisher Location	Required if available	Primary publication location only First location listed. For U.S. locations, give city and state, separated by ,space, using two-letter US Postal Service abbreviations, available at: http://www.usps.com/ncsc/lookups/ usps_abbreviations.html For non-U.S. locations, use city and country in its English form, separated by ,space. For the name of the city, use the form found in the controlled vocabulary list, or if not in the list, the form in English, or as it appears on your piece of music if the English form cannot be determined. If you don't know the state or country, enter only the city name. If only the state is given, enter the state name spelled out. If no place of publication at all is given, do not enter any information in this field.	Baltimore, MD Philadelphia, PA Berlin, Germany Brussels, Belgium London, England Mainz, Germany Kentucky Brunswick (state unknown)	Use title page, or first page of music if not given on title page, for source of information. Consult the controlled vocabulary for your location. Manually enter your location only if it is not in the controlled vocabulary. If you encounter a new location, you must add it to the controlled vocabulary.
Publication Date	Required if available	Year of publication Give in Arabic numerals. Give year only; do not include text about the year of publication or copyright.	1870	Transcribe from title page, or first page of music if not given on title page Sometimes given as part of copyright statement, e.g. Entered according to Act of Congress in the year 1870.
Publisher's Number	Required if available	A numbering designation assigned to an item by a music publisher, appearing on the title page, first page of music, or at the bottom of various pages of music. It may include initials, abbreviations or words in addition to numerals. Do not include Pub. pl. in your statement. This often exists and is not a meaningful part of the publisher's number.	G.A. 153 953.9 4924 No. 10014	Transcribe from anywhere in the score
Pages	Mandatory	Number of pages in the score Give in Arabic numerals. Enter only a numeral. If the pages are numbered, give the printed number of pages. This number often includes the title page and title page verso as pages. If the pages are unnumbered, count the pages, not including the cover, title page, and title page verso in your numbering.	4 67	Score
Instrumentation	Mandatory	Instruments for which the work is composed This field will usually contain multiple entries, which are to be connected by ;space You may list instruments in any order you like. List every instrument represented in the score,	piano; violin; harp flute; piano guitar piano; organ; voice piano; chorus piano; chorus; high voice	Score, title You are free to use terms in any order, in accordance with the instructions given in the Comments column.

		including optional instruments, whether named in the score or not. Sometimes an instrument is not named in the score, e.g. in the case of solo piano. Use chorus for all choral works; do not specify voices. Use piano , <i>not</i> pianoforte or piano forte. If for a single instrument, use only the name of the instrument, <i>not</i> "solo [instrument]." E.g. use piano, flute, voice, violin, etc. Use Arabic numberals when numbers are needed, e.g. 4 voices. Do not use hyphens, e.g. 4 hand piano. If there is a phrase about instrumentation that you think is useful, transcribe it in the Additional Information field.	soprano; tenor; piano (voice parts are specified) 2 voices; piano (voice parts are not specified) 4 voices; piano 2 pianos 4 hand piano 8 hands, 2 pianos chorus; piano; middle voice 2 tenors; bass; guitar chorus; piano; soprano; alto high voice; low voice; piano 2 high voices; piano	
Торіс	Required if available	The topic that a vocal work is ABOUT Use this field only when you are dealing with a song that is clearly about a topic. Do not enter instrumentation or musical genre in this field. If there are multiple terms, connect with ;space and begin each new entry with a capital letter.	American Civil War Battle of Prague (1757) Berry picking Children Frederick II of Prussia Love Minstrel shows Old age Patriotism Sailing Sea coast; Seashells Sunrise Unrequited love	Title and/or text of song Determine term(s) to use in this order: 1) Controlled vocabulary: if a term exists that covers your topic, use it. 2) Wikipedia: a) if term exists in singular form and plural is clearer, use plural form (e.g. Children, Minstrel shows). b) if you are redirected from a term that more accurately describes your topic than the term to which you are redirected, use the more accurate term (e.g. Berry picking redirects you to Berry; use Berry picking). 3) Webster's Third New International Dictionary, Unabridged If unsure what term to use, please ask before making decision.
First Line of Text	Required if available	First textual phrase of vocal works Limit to the first grammatical phrase of the text, often made clear through punctuation such as a comma at the end of the phrase. Enter in the primary language of the song if there is more than one language.	Oh! dearly do I love to rove among the fields of barley (first line of song entitled Bonnie Charlie)	Transcribe
First Line of Chorus	Required if available	First textual phrase of the chorus, or refrain, of vocal works Often indicated by the term "Chorus" at the beginning of this section of the music.	Yes, my love, we're growing old (first line of chorus of work entitled You are always young to me)	Transcribe
Language	Required if available	Languages of text in vocal works that contain languages other than English Do not use this field for works with only an English text. You may list languages in any order you like. Multiple entries separated by ;space If English is one of the languages, do include it in the list of languages. If there is information about languages that you	English; French English; French; Italian	Score

		want to include that does not fit the limitations of this field, enter it in the Additional Information field.		
Illustration	Required if available	Describe illustrations such as scenes, portraits, photographs. Can be as succinct or lengthy as necessary to provide accurate description. Do not describe designs such as calligraphic swashes, geometric or abstract designs, decorative designs.	Venice at night. Several domed buildings in the background, across a broad canal with two gondolas floating. Several men and women stand on a landing near the water's edge. Steps lead up to a building, with another lady standing on a balcony. Color lithograph of 3 women by shore, rainbow on horizon. Trees and rocks surround them. Portrait of Jenny Lind.	Score
Inscription	Required if available	Handwritten information on or inside cover, on title page, or on caption page If illegible, do not transcribe. If partially legible, transcribe and explain illegible part in [brackets], e.g. [illegible], [cut off] This is usually information about who is being given the music, often as a gift. May be easily confused with Previous Owner, so be careful. If an inscription applies to the volume as a whole, enter that inscription only once, in the record for the first piece of music from that volume.	Presented to Miss Ada J. Coit by Reinhold F. Hunt Dear Little One, from [cut off] A Monsieur Sigismond Benedict	Transcribe from cover or inside of cover, title page, or caption page
Previous Owner	Required if available	Name of person who owned the print music Sometimes comes from print volume as a whole, or on or inside cover, on title page, or on caption page of individual pieces of music. If illegible, do not transcribe. If partially legible, transcribe and explain illegible part in [brackets], e.g. [illegible], [cut off] May be easily confused with Inscription, so be careful. If it is clear that the same Previous Owner information applies to multiple songs in a volume, include it in every applicable record even if the name is not written on each individual piece. E.g. the volume has pieces published together as a single collection, with the owner's name appearing only on the title page of the collection. He/she clearly owned each of the pieces within that collection.	Eliza C. Anderson Belle Hannah McGehee, Burleigh, near Milton N.C.	Transcribe, usually from cover or inside of cover, title page, or caption page; check entire score
Marginalia	Required if available	Handwritten information within score Be careful not to confuse with Inscription and Previous Owner. If illegible, do not enter information. If multiple entries, separate by ;space	Performer's annotations; On the recto to p. 1 in pencil: Gallopade Quadrille Ossia bars added to final measures X's or hash marks next to pieces in the advertisement	Score; free text description of marginalia; transcribe if appropriate
Dedication	Required if available	Printed dedication statement	Most cordially dedicated to Mrs. Archibald Robertson (of Philadelphia) To Mrs. Harriet E. Griffin (Albany N.Y.)	Transcribe from title page or caption

			Music composed & dedicated to Mrs. Gardiner of Rochester N.Y.	
Dealer's Stamp	Required if available	Dealer or distributor information Hand-stamped or on an affixed label. Usually on title page.	Geo. B. Mitchell successor to Zogbaum & Co Music Dealer Savannah Ga Gaines, Riches & Co., Petersburg	Transcribe
Price	Required if available	Printed price Usually on title page.	50 cents \$2.00 nett	Transcribe, omitting words like price.
		Omit words like price . Spell out the word cents for the cents symbol.	50 cts. Net 2 Fr. One dollar	
Defects	Required if available	Describe any substantial defects in the printed music	Second page torn and repaired with tape Pages missing	Score
Volume	Mandatory	Numbering of print volume in which work is found The numbering in the Volume field must match exactly the numbering in the Volume field's controlled vocabulary. For		Print volume and Controlled vocabulary
		volumes that have yet to be cataloged, only the base volume number is in the controlled vocabulary, e.g. Old Series 102. When you catalog a new volume, add the appropriate song number to the base volume number, e.g. Old Series 102,4. Enter this information in both the Volume field and the controlled vocabulary.	III,31 New Series XIV,9 Old Series 10,1 White I, 22 XXXV,8	
Additional Information	Optional	Important information that does not fit into preceding fields Any information you think will be helpful in describing and accessing the work may go into this field.	English lyrics printed beneath French and Italian. The popularity of this song has induced persons in Philadelphia, Baltimore, and New York, to publish music with the title of the 'Carrier dove.' The publisher of this song would respectfully remind purchasers: that the genuine copy has the Imprint Inside cover page provides an extract from the France Musical testifying to the power of Gottschalk's music. Flute part is in different key than piano and vocal music, unless intended for a transposing instrument. Title page caption: This lighthouse was entirely swept away on the night of April 16th 1851. The assistant keepers Joseph Wilson & Joseph Antone were drowned, being the only persons in it at the time of the accident.	Score; transcribe or free text as appropriate
Unknown	DO NOT USE	Because this field contains a lot of useful information in previously input records, we are retaining the field and making it searchable, but hidden from public view. We are NOT using this field for currently input records.		DO NOT USE
Туре	Mandatory	Constant data is: Sheet music Use for every digital object whether or not this	Sheet music	Use controlled vocabulary for this field.

		information appears elsewhere in the metadata.	This will always be the content of this field.	
Repository	Mandatory	Entity responsible for publishing the digital object on the web Use for every digital object whether or not this information appears elsewhere in the metadata.	University of North Carolina at Chapel Hill. Music Library This will always be the content of this field.	Use controlled vocabulary for this field.
Digital Collection	Mandatory	Name of the digital collection of which the digital object is a part Use for every digital object whether or not this information appears elsewhere in the metadata.	19th-Century American Sheet Music This will always be the content of this field.	Use controlled vocabulary for this field.
Host	Mandatory	Constant data is: University of North Carolina at Chapel Hill Use for every digital object whether or not this information appears elsewhere in the metadata.	University of North Carolina at Chapel Hill This will always be the content of this field.	Use controlled vocabulary for this field.

Subscribe to comments: For this article | For all articles

2 Responses to "Look What We Got! How Inherited Data Drives Decision-Making: UNC-Chapel Hill's 19th-Century American Sheet Music Collection"

Please leave a response below, or trackback from your own site.

1. CBS Bibliotek Blog – Innovation & Ny Viden » Blog Archive » Nyt nummer af The Code4Lib Journal, 2011-04-14

[...] Look What We Got! How Inherited Data Drives Decision-Making: UNC-Chapel Hill's 19th-Century Americ... [...]

2. Gabriel Farrell, 2011-04-20

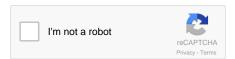
The use of Wikipedia for the topics was a great idea. I imagine they were easier to check and input for those doing the cataloging, and were also more useful for information seekers.

Leave a Reply

Name (required)

Mail (will not be published) (required)

Website



Log in

This work is licensed under a Creative Commons Attribution 3.0 United States License.