

When a Consortium isn't enough, NEMLA, Fall 2018, Montreal, by Andrea Cawelti

Title slide

Greetings, fellow sheet music enthusiasts! My apologies for reading this presentation, but if I don't read, I wander hopelessly. Trust me, you're better off, and I'll try not to be too dull about it. We are living in a great age for sheet music enthusiasm. It seems that after hitting the snooze alarm for the last few decades, the world has awoken with a bang! to the promise of historical sheet music; from its topical texts, to its extraordinary illustrations. Today, I'm going to talk with you about some of the repercussions of this awakening, mostly centered around how we access sheet music, how music libraries and collectors make this possible, and most of all, about the Sheet Music Consortium, where it has been, and where it is going.

LC Vocal music for concert performance

First off, researchers today can discover online sheet music-based exhibits on a huge variety of subjects: ... from the blockbuster *Music for the Nation* exhibits and digital collections on the Library of Congress site, which embed detailed articles and essays into curated collections of sheet music resources ...

Music in the Civil War

... to exhibits created by specialized institutions like the National Museum of Civil War Medicine, which bring together songs from the North and South, concerning enslaved persons, pacifists, Carpetbaggers, you name it, while providing some historical context and analysis ...

University of Alabama moon blog

...to fascinating, smaller blogs tying sheet music into current events ...

Women, work and song, in 19th cen. France

...and right on to McGill's own Music Library exhibit on *Women, work, and song, in 19th cen. France*, with impressive historical context and essays ...

Women, work and song Contents

in English ... and French. Accessed through the lens of sheet music, this exhibit demonstrates precisely the power of the Wayback Machine that sheet music can provide us. (~~For those of you who are unfamiliar with the true power of the Wayback Machine, please watch more episodes of Mr. Peabody's Improbably History, in the cartoon Adventures of Rocky & Bulwinkle.~~) All things cultural: popular, gender, consumer, Diva, sociological, LGBTQ, economic, religious, material, ... I can go on, I have more! Sheet music records these cultures as they happen in a specific time and place, which is invaluable for research of all kinds. But I don't need to sell this concept to you, because you are Music Librarians, and you get this! McGill's curators have harnessed just this kind of documentary evidence, to build this excellent exhibit. But how did they find the music?

McGill Music Library sheet music coll.

McGill is lucky enough to have its own collection from which to create such informative exhibits, and in fact, most online material highlights scores from a home institution. But what if you don't have any interesting sheet music available to you? What if you, in Tinyville USA or Canada, want to research and/or create an exhibit, or a digital humanities project, but your local library doesn't have any sheet music? What is the road from there to McGill? Attempting to google "sheet music" will bring the unsophisticated user a massive result:

Google sheet music for sale

I got over a billion, in a confusing jumble of sheet music for sale, references to sheet music, unrelated sites, and eventually, some institutional sheet music collections and digital access sites. Trying WorldCat is even worse! Again, many researchers come to grief here, simply because of the way we catalog popular sheet music: often those words don't even appear on a sheet music record!

WorldCat results for "sheet music"

If users persist, and studies have indicated that some don't (in fact I remember watching a video a few years ago, about an early user research trial at Harvard, where a student confidently went right to WorldCat when asked to find something in (I *think* it was in our Emily Dickinson Collection but don't quote me on that). She froze, petrified, when the initial haul from her search was immense, and the results didn't appear on the first few pages) ... and check it out, I love the Ellora's Cave result at the bottom there: this is an early eBook erotic romance and fantasy publisher! Whaaaaat?

Google collections

Some find their way to the Sheet Music Consortium (cue major chord sound effect here) through one path or another. Some search engines regularly privilege the site, many librarians have added the Consortium to high-profile libguides, recommended it to researchers, and generally encouraged users to treat the Consortium as their first stop for all sheet music needs.

SMC contents

And for good reason. As you can see, the Sheet Music Consortium has a sizeable aggregated content of digitized scores. The Consortium also includes metadata records without scans, but their main goal is to build an open collection of digitized sheet music using the Open Archives Initiative Protocol for Metadata Harvesting. Now, I'm just taking this directly from the Consortium's own website: Harvested metadata about sheet music in participating collections is hosted by UCLA's Digital Library Program, which provides an access service via this metadata to sheet music records at the host libraries. Data providers have chosen to catalog their sheet music in different ways, but a large proportion of the scores in participating collections have been digitized, I'd guess about 75% based on non-scientific browsing, allowing users direct access to the music itself. Because there is no expectation of data conformity, names, titles, and subjects can appear in a wide variety of forms. While this allows for participation by collections using different cataloging guidelines, just as it should, we'll see later how this can impact the user experience. Laurie J. Sampsel wrote an excellent overview of the growth of the Consortium in the March 2007 issue of *Notes*, in case anyone wants more details.

Sheet music links spreadsheet

Please allow me a brief detour here, as a grim reminder of reality: without the Consortium or something like it, sheet music researchers would have a rough time of it. In preparation for this presentation, I began assembling sheet music resources across the internet, including collection material (both analog and digital), helpful links, and interpretive material. I was helped at once immensely, and frustrated and confused, by a list of links shared with me by Don Krummel, legendary music librarian, antiquarian music specialist, retired professor at the University of Illinois, and of the "American music printing and publishing" course at Rare Book School. He had updated the RBS course links he shared in 2006, and in the last 12 years, EVERY ONE of the 200+ links he shared, but two. TWO. had changed. Just keep this in the back of your mind as I progress here. And I had a list to begin with! I don't want to imagine beginning from scratch ...

For those of you are not familiar with the Consortium's interface, let's have a look:

SMC home page

Now, let's be clear from the start: this is an aged interface. We all get that. My own hair has been white for a long time, and I'm absolutely not one to toss out the old for the new, except perhaps on some particularly cold, clammy mornings. Work began on the Consortium in 2001, and the website went live in 2003. It was a marvel of its kind at that time, but that was a long time ago in tech years, which I believe at this point are even shorter than dog years. In 2009, the Institute of Museum and Library Services awarded a National Leadership Grant to the Consortium, which funded several updates to their backstage works, as well as to the user interface, but as you can see right there on the home page, news has been scarce for the last five years. The keyword search is not Boolean, but provides what we would understand to be an "OR" search, so if you try to search anything other than a title word, or a single composer name, you retrieve exponentially extraneous results.

SMC name browse

You do have the option to perform several species of browse searches, including title, subject, name, and date range, and in some cases these searches can be quite precise. One-shot wonders, for instance, might show up right where you want to find them. If all of the music you seek has been processed by a university library, or the Library of Congress, for instance, you are likely to find browse links right where you expect to see them. Once you have a composer or title you want to explore, you would simply click on the link to bring you to a list of associated records. But remember that lack of data conformity I mentioned earlier? This is where things can get complicated from a retrieval standpoint.

SMC browse problems

What if the institution holding the sheet music you want to see, lists their names in direct order? Or they include initial articles with their title metadata? Or they simply don't follow library conventions of standardizing names? On their own, none of these concepts are wrong: many collections and/or collectors follow their own guidelines; to exclude content from those which follow different conventions would be ridiculous, and shut out many unique and important resources. But these browse examples show how this simple decision has had complex repercussions as the Consortium has grown. Note particularly the Strauss situation. The obvious answer is to expand the ability to utilize Boolean keyword searching. The current system allows for an advanced keyword search, which overcomes some of the problems with the "OR" search, but can still prove to be highly idiosyncratic. Why do I care so much about this, apart from a thankless impulse to be a good citizen?

Pop music MARIA

I'm no popular American music expert, in fact, my specialization is in 18-19th century opera and vocal music. I suspect that most of you here have forgotten more about this subject than I know, but after writing and overseeing a grant to increase access to Houghton's "hidden" music collections, I had the opportunity to look through some of the huge holdings of sheet music we have at Harvard; those of you in NEMLA may remember Dana Gee's presentation on this project a few years ago. I fell in love with the humor,

WWI angel

The heart-breaking sentiment, I tell you, the rags on that soldier's feet made me tear up the first time I saw this one,

Too much Ginger

The brazen sexual innuendo,

Agawam

The local history,

Brandard

Performance documentation, like this cover, whose artist used to visit opera and ballet performances in order to sketch exact details to bring authenticity to his work,

Mavis

And sheer, ravishing, beauty represented by the covers. Shortly afterwards, I also fell in love with that Wayback quality, and ability to document all things cultural that I mentioned earlier. The sheet music collections backlog at Houghton Library number almost three hundred thousand scores. While our six-month grant allowed us to locate and accession all of our sheet music collections, which I'm embarrassed to report were until then, essentially living in an Indiana Jones-style backlog of unlabeled boxes ... we were only able to catalog about 300 scores individually during that time. Our current method of cataloging special collections materials is simply unsustainable when quantities are this large.

Mavis Hollis record

I'm fiercely proud of our detailed special collections-style cataloging, and the access it provides for our users. But every field takes time to generate, and those circled fields take extra time to research, and to create authority records where necessary. And I see no reference to "art deco" style here—yikes, someone fell down on the job. But when facing 300,000 scores, it would take an army of catalogers to create detailed records of this kind in our lifetime. And more problematic, is interesting library leadership in supporting such an initiative: I had enough trouble selling the idea of writing a grant just to accession the sheet music backlog! I'm sure that many of you are in the same boat, and please forgive me, as I'm well aware of folks' attitudes about Harvard whining about money and staff, but trust me, that largesse you all imagine doesn't make it to our shop. We are testing out a finding aid approach, as that seems to be the only way to go forward with such a small staff, and student help. I hate to do this with printed material like sheet music, but what is the alternative?

Santo Domingo excel

Here is an example of a spreadsheet for a collection in-process: the Ludlow-Santo Domingo Library Collection focusing on altered states. Its sub-series of fascinating and rare sheet music illustrates, both visually and textually, themes of drugs, dance, and sex. With some supervision and an iterative cataloging process, students will be trained to scan, then enter data in a spreadsheet, using controlled vocabulary which can then be converted into a finding aid, which will then be uploaded to ArchivesSpace. But while the material will be searchable in our own library database, it will be completely invisible outside of Harvard, with the exception of the collection record. What if there were a place which could ingest our specific sheet music metadata, and make it searchable along with other sheet music collections large and small? See where I'm going with this?

SMC data

Well, duh! If only we were able to work through at least some of our 200+ collections of sheet music, and then just let our spreadsheets do double duty by sending a copy off to form the basis of our institutional finding aid, and another copy off to the Sheet Music Consortium, bingo, our metadata would be visible to the world, and more importantly, available where researchers actually look, with the other data in the Consortium. But that doesn't seem to be the way metadata harvesting works in the Open Archives Initiative Protocol. Was I interested in putting in the time to figure all of this out, given all of the afore-mentioned problems with the Consortium's interface?

Bound for glory MLA

Then came the Portland meeting of the Music Library Association last year, where I was very much in a sheet music frame of mind. When Darwin Scott dropped into the Sheet Music Interest Group, as Princeton is

working on a project to scan some of their sheet music holdings, and asked if anyone knew what was going on with the Sheet Music Consortium, I jumped in and meddled unconscionably. With Darwin, Bob Kosovsky at the NYPL Music Division, Mary Kay Duggan, the queen of California sheet music, Jennifer Wochner, chair of our sheet music interest group, and some other interested parties, we contacted the UCLA Digital Library Program, and asked about their current situation regarding the Consortium. Were they interested, and/or in a position to upgrade the system, and perhaps solve a few of the above-mentioned searching issues?

UCLA LibGuide

I'm delighted to report that the librarians at UCLA responded immediately: they were indeed dedicated to the Consortium, and interested in making changes to the interface. Their first step was to produce detailed instructions which were aimed at users rather than at potential participants, as a stopgap between now and when they might realistically attend to additional upgrades. Now, some of these plans will require megabucks to implement, and a lot of staff time. But I'm happy to report that the UCLA Digital Library is discussing the following possibilities, and already working on some of the lowest-hanging fruit:

Additional planned upgrades:

1. Website Aesthetics
 - For instance, remove "Activity Stream" (already in process) and "News" feeds on home page
2. Outreach for collections not yet included
 - Some of you might want to consider participation going forward, if you have significant holdings
3. Update search
4. Update site architecture
5. Utilize a knowledgebase/discovery layer structure to incorporate it into other institutions easily
 - This has the potential to be grant supported
6. Incorporate music specific metadata standards, MEI and Music XML?

SMC/OCLC

This will all take time, but UCLA is well aware of how important the Consortium is in the larger picture. Like these poor ragged performers at the King's Theatre in 1791, who didn't have a license to present Italian opera or ballet, the Consortium may have been languishing. But with an upgraded interface, they will rise like a phoenix, and become the first stop for sheet music researchers everywhere. There is a lot of excellent sheet music metadata in Worldcat, but like Harvard, so many institutions don't catalog their collections individually. But what if processing sheet music were easier? What if we could even use a dash of Artificial Intelligence to help us with transcription, and searching authorized headings? Just imagine the power of the Wayback abilities, if we could get everyone's records together in the Consortium?

Kijas Carreno

Because if we are living in great age for sheet music, this is nothing to the current world of Digital Humanities. Harken back to all those cultures I listed at the beginning of this presentation, and ponder for a moment: if we were able to assemble enough metadata, like for instance, subject, place of publication, and date, what would a bit of computational analysis teach us? I have been floored by some of the DH projects I've seen in the last few years, like this visualization by our own Anna Kijas which she shared at MLA in Orlando, highlighting her work on Teresa Carreno. If catalogers begin adding subject heading besides music formats, as many in the Consortium already have, we could really do something!

BC visualizations

We haven't even cracked the surface of what is possible from a subject analysis standpoint. If we have the data available, there are a lot of brilliant folks who are poised to draw fascinating conclusions from it. The popular music alone would highlight trends that would keep pop culture gurus cogitating for years. Imagine analyzing American subject trends, rooted in place of publication, from 1818 to 2018? The mind simply boggles! This Geographic Distribution of Undergraduate Students resource created by Boston College, uses Tableau software to depict enrollment over time per state. Just visualize harnessing something similar to track Courtship themes, year by year, city by city? Discover which trends reliably foretell other trends?

Word cloud

Even a simple wordle, created in Voyant in less than 5 minutes from the top 50 subject and genre headings in our historical sheet music collection guides, shows us something. Melancholy, loneliness, and racism may be up there, but Love songs still top the mix by a large percentage. What would we learn from a larger, less "Harvardian" sample? Would Love songs lose their ascendancy from place to place, or year to year? How do world events affect Love songs? If we can harness the power of our sheet music Wayback Machines, we will be able to go forth and discover! So I encourage you all to ponder, because where there are music librarians, anything is possible. If you have a large collection of sheet music in your backlog, languishing without staff support, call me, drinks are on me; if you are a brilliant techie, and you have an idea about how AI might work to do some of the transcription so that catalogers can save their time for subject and authority analysis, get it out there. And most of all, if you use the SMC regularly, and have ideas which might help, or something to share which they could take to their \$\$ people and harness for fund or institutional support, tell them. If you are already in the Consortium and experiencing frustration and blood pressure issues, have patience, and hang in there. And if you are cataloging a collection right now, consider including subject headings other than music format. Every little bit helps.

Thank you!

I won't be able to answer everything, but can always get back to you later. Any questions?