

Western Washington University

From the Selected Works of Casey A. Mullin

February 2, 2018

Deriving Faceted Terms from Library of Congress Subject Headings for Music

Casey Mullin, *Western Washington University*
Gary Strawn



Available at: <https://works.bepress.com/casey-mullin/11/>


Deriving Faceted Terms from Library of Congress Subject Headings for Music

CHALLENGES AND POSSIBILITIES

Casey A. Mullin
Western Washington University
(Outgoing) Chair, MLA Vocabularies
Subcommittee

Gary Strawn
Northwestern University

MLA Annual Meeting, Portland, OR
Friday, February 2, 2018

A solid orange horizontal bar spanning the width of the slide at the bottom.

Overview

Why?

- Why is this necessary?
- Why now?
- Why is this so difficult?

How?

- Algorithm and associated OCLC toolkit

Who, when and what('s next)?

- Public “beta” release of program and documentation
- More testing
- Future of LCSH practice

Q & A

Prologue

- What won't be covered
 - History and development of faceted vocabularies
 - How to apply the vocabularies in current cataloging
- Parallel developments in the broader library metadata community
 - ALA/ALCTS Subcommittee on Genre/Form Implementation
 - (Reconstituted in 2017 as Subcommittee on Faceted Vocabularies)
 - White paper "A Brave New (Faceted) World" (July 2017)
 - <https://alair.ala.org/handle/11213/8146>
 - Recommendation no. 3: "**Retrospective implementation** of faceted vocabulary terms using algorithms developed, vetted, and tested by expert communities"
 - Recommendation no. 4: "Display and granular indexing of all faceted data, including (but not limited to) MARC bibliographic fields 046, 370, 382, 385, 386, 388 and 655 (or their equivalents in other encoding standards)"

Why is this necessary?



Our legacy: LCSH

What it does well

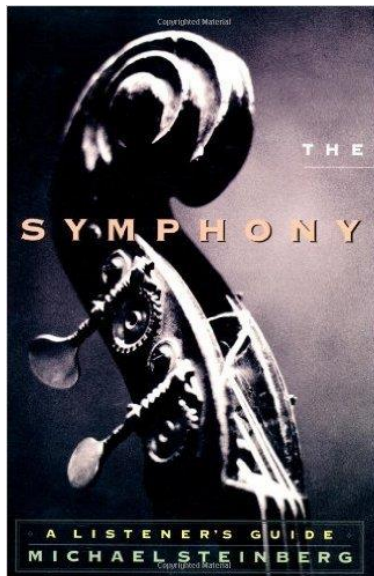
- Alphabetic arrangement
- Systematic breakdown of a topic through subdivisions
 - Example: [New York \(N.Y.\)—History—1951-](#)
- Terminology kept (mostly) up to date
- Widespread implementation (thus ready availability in cataloging copy)

Our legacy: LCSH

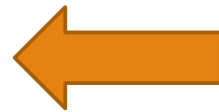
What it does *not* do well (for music)

Distinguishing between topical and
form/genre/medium uses

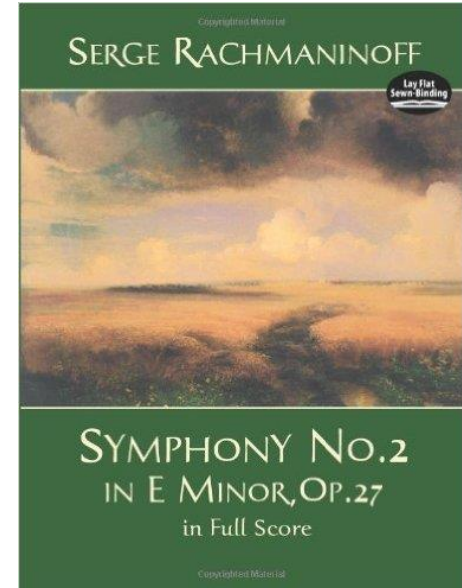
Aboutness vs. *Is-ness*



LCSH: “Symphony”



LCSH: “Symphonies”

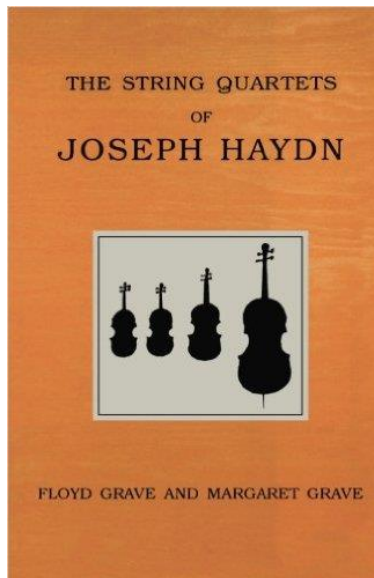


Our legacy: LCSH

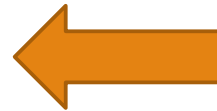
What it does *not* do well (for music)

Distinguishing between topical and
form/genre/medium uses

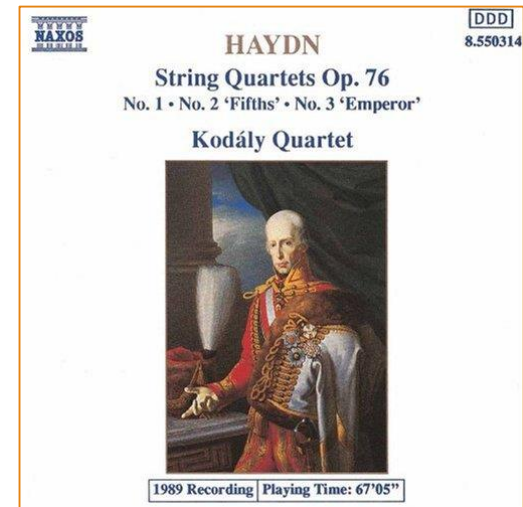
Aboutness vs. *Is-ness*



LCSH: “String quartets—
History and criticism”



LCSH: “String quartets”



Our legacy: LCSH

What it does *not* do well (for music)

Lack of granularity

1 player: Sonatas (Piano)

5 players: Suites (Bassoon, clarinet, flute, horn, oboe)

9 players: Variations (Piano, violins (3), violas (2), cellos (2), double bass)

10 players: Instrumental ensembles

Our legacy: LCSH

What it does *not* do well (for music)

Lack of granularity

Singer and guitar:

- Songs with guitar

Soprano or tenor and piano:

- Songs (High voice) with piano

Baritone and string quartet; Mezzo-soprano and two violas:

- Songs (Medium voice) with string ensemble

Alto and SATB chorus:

- Choruses (Mixed voices), Unaccompanied

Our legacy: LCSH

What it does *not* do well (for music)

Disparate attributes combined semi-opaquely in one string of text

- Overtures (Violin and piano), Arranged—Scores and parts.
- Songs, French.
- Rock music—Norway—2011-2020.
- Children's songs—Juvenile sound recordings.
- Jews—Music.

Our legacy: LCSH

What it does *not* do well (for music)

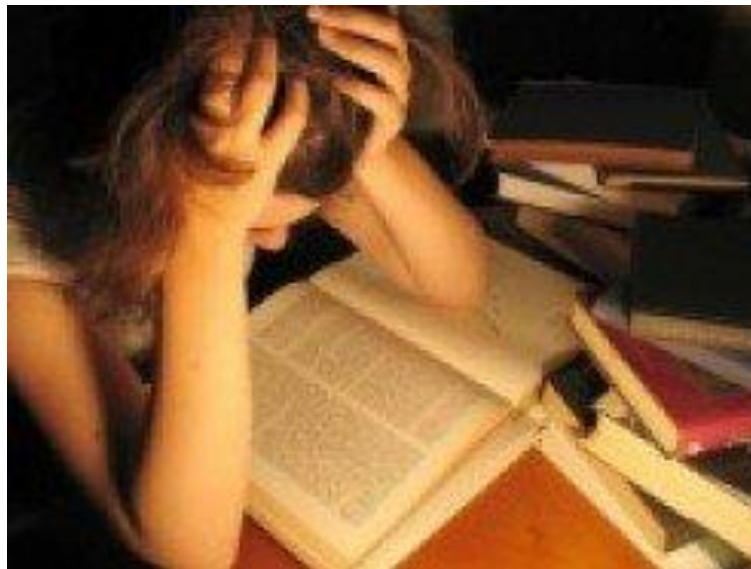
Difficult to learn

- Implied medium of performance
- When to include form subdivisions (e.g., “—Scores”)
- Order of instruments
 - “(1) keyboard instruments, (2) wind instruments, (3) plucked or hammered stringed instruments without a keyboard, (4) percussion, electronic, and other instruments, (5) bowed stringed instruments, (6) unspecified instruments, and (7) continuo. Within each numbered category, give the instruments in alphabetical order with the exception of bowed stringed instruments: give these in score order, high to low, basing the order on the range of each instrument as described in standard music reference sources. ” (SHM H 1917.5)
- And on and on...

Our legacy: LCSH

The \$64,000 question:

If LCSH for music is this difficult for *librarians* to use and comprehend, what hope do our end users have?



A new way forward

Faceted access with purpose-built vocabularies (and MARC fields)

Library of Congress Medium of Performance Thesaurus for Music ([LCMPT](#))

- MARC field 382

Library of Congress Genre/Form Terms for Library and Archival Materials ([LCGFT](#))

- MARC field 655

Library of Congress Demographic Group Terms ([LCDGT](#))

- MARC fields 385 and 386

Geographic area ([LCSH](#) and [LC/NACO Authority File](#))

- MARC field 370

Chronology

- MARC fields 046 and 388

Language

- MARC fields 008, 041, 546

A new way forward

Faceted access with purpose-built vocabularies

Benefits:

- True thesauri, built to NISO standards
- Post-coordinated (no subdivisions!) (compare FAST)
- Easy to navigate the hierarchy and select among terms
- Specific metadata carrier elements for each facet
- Significant reduction of Western art music bias
- Vocabularies freely available on the web
- Linked Data friendly!
- **Implementation well underway in current cataloging in the US**

Why retrospective implementation? (Why now?)

Catalogers are currently doing double work, and this is not sustainable.

Manual enhancement of legacy data does not scale.

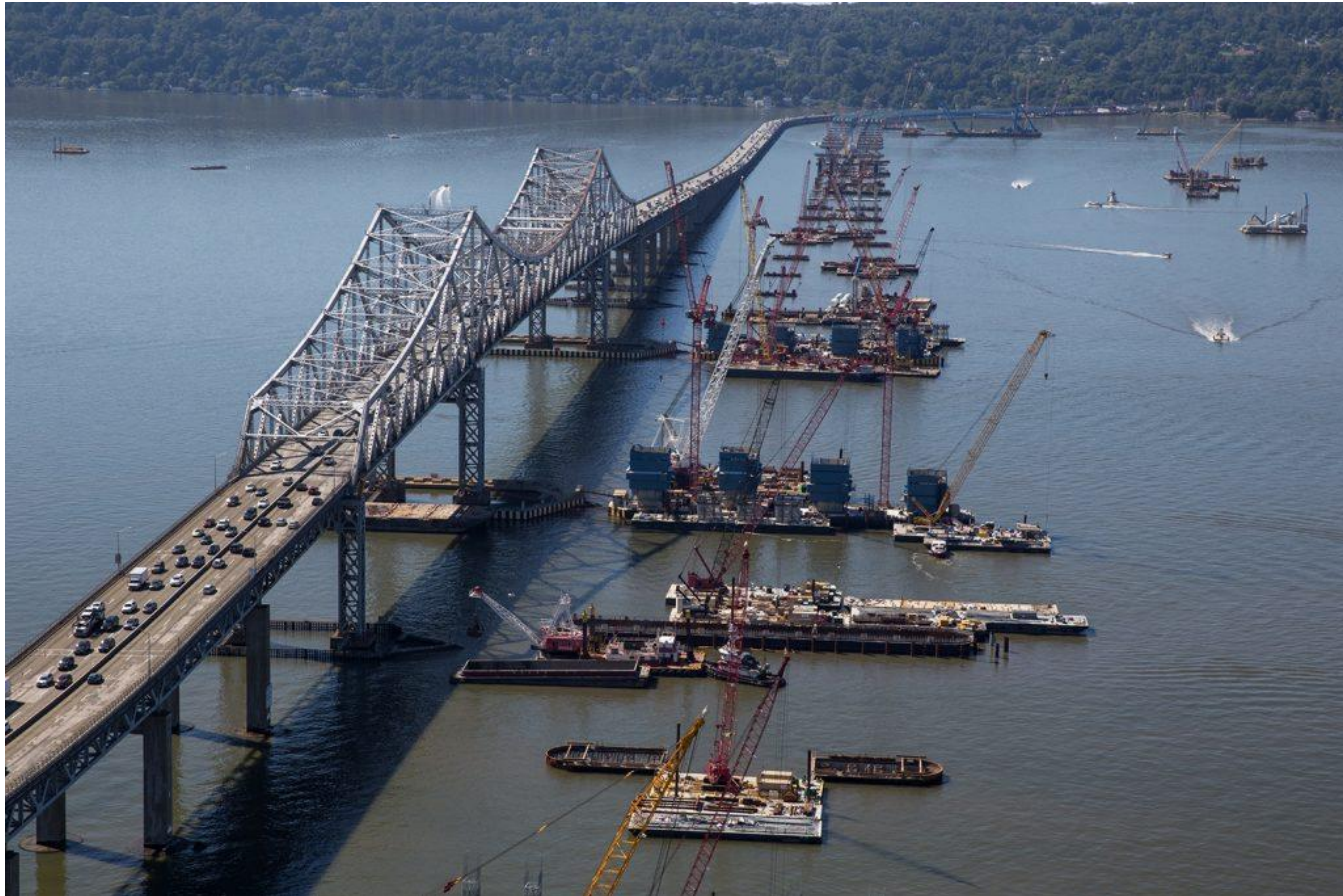
- As of January 2018, WorldCat has ~4.2 million bibliographic records for scores and ~6.6 million records for audio recordings
- As of January 1, 2017, 382 field used in 143,203 records in WorldCat (~8%)
- Local catalogs: even less? (YMMV)

The vision of enhanced discovery for our users will not be fully realized until a critical mass of our metadata carries faceted terms.

But does this mean we should delay current implementation until we achieve retrospective implementation? **No.**

- Adding faceted data in current cataloging provides the ideal laboratory for exploring the possibilities of what can be done retrospectively.
- The juxtaposition of new and old practices constitutes an evidence base that will inform decisions about the future of LCSH for music.

Why retrospective implementation? (Why now?)



MLA Vocabularies Subcommittee derivation project: Where we are now

Partnership between MLA VS (editorial leadership) and Gary Strawn (programming and documentation)

LCSH-to-LCMPT algorithm

- 2014-2015: developed, preliminary testing done

LCSH-to-LCGFT/LCDGT/etc. algorithm

- 2016-2017: developed, preliminary testing done

Combined algorithm and OCLC toolkit

- Early 2017: analysis of “non-productive” LCSH headings
 - Testing corpora: Northwestern University and Syracuse University’s music bib records
 - Main question: why is the algorithm **not** generating what we think it **should**
- Late 2017-present: testing of OCLC toolkit on individual bib records in real time
 - Main question: why is the algorithm generating what we think it **shouldn’t**

Retrospective implementation: Parameters

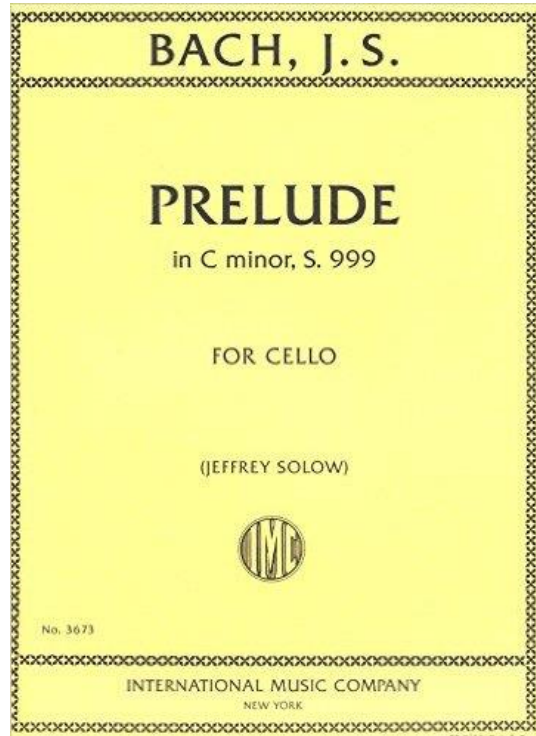
- Each LCSH genre/form/medium heading should beget, at least, one LCMPT and/or one LCGFT term
 - LCDGT or LCSH terms generated in some cases
- Selective MARC composition and format of music codes are handled, where they fill in gaps in LCSH terminology
- Algorithm should work equally well for score and sound recording records
- Terms in machine-generated fields must be conformant to LCGFT and LCMPT
- Machine-generated fields should ideally be conformant to MLA best practices
- Duplicate fields should be removed
- LCSH headings retained for the time being
 - Not “conversion”

Retrospective implementation: Why is this so hard?

Many LCSH headings are easily (and losslessly) repurposed into faceted data, but complexities and limitations abound.

- Implicit medium of performance in LCSH (e.g., “Symphonies”, “Monologues with music”)
- Differences in vocabulary
 - LCSH (“mixed voices”) vs. LCMPT (“mixed chorus”)
 - LCSH (“Canons, fugues, etc.”) vs. LCGFT (“Canons” or “Fugues”)
 - Completely new terms in LCGFT (“Art music”)
- Identifying solo performers
- True “topical” headings: how to distinguish?
 - By presence of certain modifiers (e.g., “Songs and music”, “History and criticism”)
- Videorecording records?
- Current approach generally treats a single 650 or fixed field code at a time.

Retrospective implementation: Example of an “easy win”*



008/18-19 pr
008/20 l (lower-case “el”)
650 _0 Cello music, Arranged.

begets

382 01 cello \$n 1 \$s 1 \$2 lcmpt
655 _7 Preludes (Music) \$2 lcgft
655 _7 Arrangements (Music) \$2 lcgft
655 _7 Scores. \$2 lcgft

*And there are lots of easy wins!

Retrospective implementation: More difficult situations

1) Lack of granularity in existing LCSH heading: an instrumental ensemble with more than 9 players

650 _0 Wind ensembles.

“Lossless” but generic LCMPT output:

382 01 \$a wind ensemble \$e 1 \$t 1 \$2 lcmpt

Ideal, detailed LCMPT output requiring human intervention:

382 01 \$a flute \$n 2 \$a oboe \$n 2 \$a clarinet \$n 2 \$a bassoon \$n 2 \$a
bass clarinet \$n 1 \$a trumpet \$n 2 \$a trombone \$n 1 \$a tuba \$n 1 \$s 13
\$2 lcmpt

Retrospective implementation: More difficult situations

2) Lack of granularity in existing LCSH heading: a song for soprano and piano

650 _0 Songs (High voice) with piano.

“Lossless” but generic LCMPT output:

382 01 \$a high voice \$n 1 \$a piano \$n 1 \$s 2 \$2 lcmpt

Ideal, specific LCMPT output requiring human intervention:

382 01 \$a **soprano** voice \$n 1 \$a piano \$n 1 \$s 2 \$2 lcmpt

Retrospective implementation: More difficult situations

3) Multiple LCSH headings required to bring out various aspects

1. 650 _0 Viola with instrumental ensemble.
2. 650 _0 Sextets (Piano, violins (2), violas (2), cello)

Redundant[?] LCMPT output:

1. 382 01 \$b viola \$n 1 \$a instrumental ensemble \$e 1 \$r 1 \$t 1 \$2 lcmpt

(brings out soloist aspect but obscures make-up of ensemble)

2. 382 01 \$a piano \$n 1 \$a violin \$n 2 \$a viola \$n 2 \$a cello \$n 1 \$s 6 \$2 lcmpt

(lists all instruments but soloist aspect is lost)

Ideal LCMPT output:

382 01 \$b **viola** \$n 1 a piano \$n 1 \$a violin \$n 2 \$a viola \$n 1 \$a cello \$n 1 \$s 6 \$2 lcmpt

(single 382 field accomplishing both objectives)

Retrospective implementation: More difficult situations

4) Limitations of scope in existing LCSH heading: a sonata for string quartet

650 _0 String quartets.

[not valid in LCSH: 650 _0 Sonatas (String quartet)]

“Lossless” but incomplete LCGFT output:

655 _7 Chamber music. \$2 lcgft

Ideal, complete LCGFT output requiring human intervention:

655 _7 **Sonatas**. \$2 lcgft

655 _7 Chamber music. \$2 lcgft

Retrospective implementation: More difficult situations

5) Ambiguous meaning of religion demographic group term

650 _0 Buddhist hymns.

Easy LCGFT output:

655 _7 Hymns. \$2 lcgft

But where does “Buddhist” go?

~~385 __ Buddhists \$2 lcdgt~~ (not necessarily for a Buddhist audience)

~~386 __ Buddhists \$2 lcdgt~~ (not necessarily created or performed by Buddhists)

~~650 _0 Buddhists.~~ (not really “about” Buddhists or Buddhism)

Is a new MARC field needed for “associated demographic group [that’s not necessarily a subject, intended audience, creator or contributor]”?

Retrospective implementation: More difficult situations

6) LCSH form subdivision that changes the meaning of the base heading

650 _0 Motion picture music \$v Excerpts.

LCGFT output:

655 _7 Motion picture music. \$2 lcgft

655 _7 Excerpts. \$2 lcgft

But...

650 _0 Operas \$v Excerpts.

LCGFT output:

~~655 _7 Operas. \$2 lcgft~~ (an album of opera arias does not constitute an opera)

655 _7 Excerpts. \$2 lcgft

Retrospective implementation: More difficult situations

7) The LCSH chameleon “Arranged”

650 _0 Viola music, Arranged. *(base heading describes the **arranged** medium)*

LCMPT output:

382 01 \$a viola \$n 1 \$s 1 \$2 lcmpt

But...

650 _0 Suites (Piano) \$v Excerpts, Arranged. *(base heading describes the **original** medium)*

LCGFT output:

655 _7 Excerpts. \$2 lcgft

~~655 _7 Suites. \$2 lcgft~~ *(a single extracted movement is not a suite)*

LCMPT output:

~~382 01 \$a viola \$n 1 \$s 1 \$2 lcmpt~~ *(not possible to infer the **arranged** medium from this heading)*

Retrospective implementation: More difficult situations

8) LCSH headings and MARC codes without a direct LCGFT/LCDGT equivalent

Solution: map to a broader LCGFT term

008/18-19 nc [Nocturne] → 655 _7 Art music. \$2 lcgft

008/18-19 ft [Fantasia] → 655 _7 Art music. \$2 lcgft

650 _0 Romances (Music) → 655 _7 Art music. \$2 lcgft

650 _0 Frottolas (Music) → 655 _7 Part songs. \$2 lcgft

650 _0 School music → 655 _7 ??? \$2 lcgft + 385 __ ??? \$2 lcdgt

- LCGFT “Music” might be the best option
- LCDGT “School children” too narrow (only elementary/primary level)

650 _0 Feminist music → 655 _7 ??? \$2 lcgft

Retrospective implementation: More difficult situations

9) LCGFT headings without a direct LCSH equivalent

Covered by MARC format of music and composition codes

Preludes (Music)

Fugues

Canons (Music)

Scores (*applies more universally than LCSH \$v Scores does*)

Not covered at all, e.g.

Opera adaptations

Impromptus (Music)

Live sound recordings

Concert etudes

Starting from what?

- Committee-written documents
- Spreadsheet
- Is this programmable?
- Rewritten and reformulated during programming

What's the outcome?

- Implementation of committee reports
- Tool for OCLC Connexion you can use
- Tool = macro that activates a DLL
- Assigned to button or keystroke
- Installation and operation are trivial
- No visible interface; no options

Today ...

- Describe how tool works
(Internal operation)
- Examples *supposed* to be non-controversial
- I'm not a music LCSH expert
- Impossible to include all details/exceptions
- Documentation

Simple example

Before:

650		0	String trios (Baryton, viola, cello)
-----	--	---	--------------------------------------

After:

382	0	1	baryton #n 1 #a viola #n 1 #a cello #n 1 #s 3 #2 lcmt
650		0	String trios (Baryton, viola, cello)
655		7	Chamber music. #2 lcgt

What records are considered?

- OCLC MARC21 bibliographic records
- Leader/06=c,d,j (notated music; music recording)
- *Not* Leader/06=i (non-music recording)

What elements are considered?

- LCSH (6XX, indicator 2=0)
Punctuation sometimes important
Normalized text matching
- 008/18-19 and 047 (655 only)
Form of composition
- 008/20 and 008/21 (655 only)
Format of music; Music parts

Other information available to tool

- Spreadsheet, with series of lists
- LCMPT configuration file
- 370: LC/NACO AF, LCSH accessed directly as needed

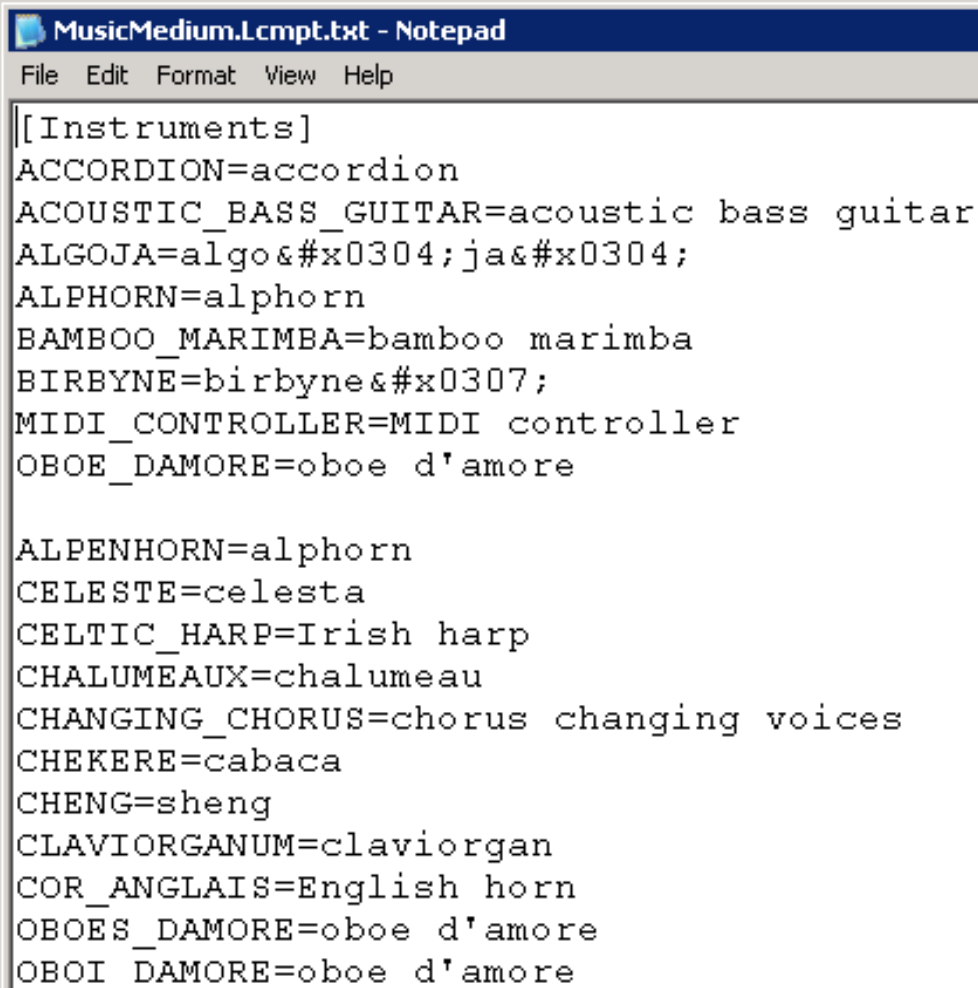
What gets produced?

- 046 EDTF; ISO 8061
- 370 LC/NACO; LCSH
- 382 LCMPT
- [384] LCDGT
- [385] LCDGT
- [388]
- 655 LCGFT

LCMPT configuration file

- List of authorized terms
- Variant terms
- Terms in the 'ensemble' hierarchy
- Static; updated periodically
- Incorporated in new installation package
- Implies: refresh your installation

LCMPT configuration file



The image shows a screenshot of a Notepad window titled "MusicMedium.Lcmpt.txt - Notepad". The window contains a list of instrument names and their corresponding sound bank names, separated by equals signs. The instruments are listed in two groups, one under the "[Instruments]" header and another group below it. The sound bank names are in lowercase and some include special characters like apostrophes and accents.

```
[Instruments]
ACCORDION=accordion
ACOUSTIC_BASS_GUITAR=acoustic bass guitar
ALGOJA=algo&#x0304;ja&#x0304;
ALPHORN=alphorn
BAMBOO_MARIMBA=bamboo marimba
BIRBYNE=birbyne&#x0307;
MIDI_CONTROLLER=MIDI controller
OBOE_DAMORE=oboe d'amore

ALPENHORN=alphorn
CELESTE=celesta
CELTIC_HARP=Irish harp
CHALUMEAUX=chalumeau
CHANGING_CHORUS=chorus changing voices
CHEKERE=cabaca
CHENG=sheng
CLAVIORGANUM=claviorgan
COR_ANGLAIS=English horn
OBOES_DAMORE=oboe d'amore
OBOI_DAMORE=oboe d'amore
```


382 versus 655 and all else

- Two separate paths
- 382 written several years ago
- 655, etc. added more recently
- Explained in two parts

382 generation

- Parse text of 650 subfield \$a (and \$v)
- Chiefly program code
- Spreadsheet tables
- Configuration file for LCMPT

382 techniques

- First word (plus more)
- Construction of \$a
- Subfield \$v (*Hymns*)
- Spreadsheet list 12, for entire \$a text

	A	B
1	650 \$a	382
80	salon orchestra music	\$a instrumental ensemble \$2 lcmpt
81	satirical songs	\$a singer \$2 lcmpt
82	sea songs	\$a singer \$2 lcmpt
83	sequences music	\$a vocal ensemble \$2 lcmpt
84	shepherds songs	\$a singer \$2 lcmpt

382: First word (plus more)

- Ballads[, ...]
- Intabulations (...)
- Monologues with music
- Octets [(...)]
- String <trios, etc.> [(...)]

382: Construction of \$a

- ... music [...]

382	0	1	baritone horn #n 1 #s 1 #2 lcmt
650		0	Baritone music.

382	0	1	violin #n 2 #s 2 #2 lcmt
650		0	Violin music (Violins (2))

- ... with ...

382	0	1	#b piano #n 1 #a harmonica ensemble #e 1 #r 1 #t 1 #2 lcmt
650		0	Piano with harmonica ensemble.

382: complicating factors

- \$a: Irrelevant qualifiers following comma
..., Juvenile
- \$v: Blocking texts
Instrumental settings; Excerpts; Analysis
- \$v override
2-piano scores; Vocal scores with ...

382: Incorrect and non-current

382	0	1	flute #n 17 #s 17 #2 lcmpt
660		0	Flute music (Flutes (17))

382	0	1	flute #n 2 #a piano #n 1 #s 3 #2 lcmpt
660		0	Flute music (Flutes (2)) with piano.

655, etc. generation

- Mostly based on spreadsheet
- Most tables used as given
- Some tables implemented in code

List 1: \$a construction

	A	B
1	650 \$a	655
2	[Instrument] and [instrument] music	\$a Chamber music \$2 lcgft
3	Ballads, [Language]	\$a Ballads \$2 lcgft
4	Brass [ensemble word implying number] (...)	\$a Chamber music \$2 lcgft
5	Cantatas, Sacred (...)	\$a Cantatas \$2 lcgft \$a Sacred music \$2 lcgft
6	Canons, fugues, etc. (...)	\$a Art music \$2 lcgft
7	Piano music (Pianos (x))	\$a Chamber music \$2 lcgft
8	Piano music (Pianos (x), x hands)	\$a Chamber music \$2 lcgft
9	Polkas (...)	\$a Polkas (Music) \$2 lcgft

List 4: \$v construction

	A	B
1	\$v	655
9	Methods ([Style])	\$a Methods (Music) \$2 lcgft \$a [Style]
10	Studies and exercises ([Style])	\$a Studies (Music) \$2 lcgft \$a [Style]
11	Textbooks for [Language] speakers	\$a Textbooks \$2 lcgft
12	Textbooks for foreign speakers \$v [Language]	\$a Textbooks \$2 lcgft
13	Video recordings for [Language] speakers	\$a Video recordings \$2 lcgft

Table of styles

	A	B
1	650 \$a parenthetical qualifier	655
2	alternative rock	\$a Alternative rock music \$2 lcgft
3	barrelhouse	\$a Blues (Music) \$2 lcgft
4	big band	\$a Popular music \$2 lcgft
5	bluegrass	\$a Bluegrass music \$2 lcgft
6	blues	\$a Blues (Music) \$2 lcgft
7	blues rock	\$a Blues-rock music \$2 lcgft
8	boogie woogie	\$a Boogie woogie (Music) \$2 lcgft
9	bop	\$a Bop (Music) \$2 lcgft
10	celtic	\$a Celtic music \$2 lcgft
11	country	\$a Country music \$2 lcgft

More lists for 655

- List 2: \$a plus \$v/\$x

	A	B	C
1	Tag	6XX \$a	655
2	650	Animated films \$x Sound effects	\$a Sound effects recordings \$2 lcgft
3	650	Apostles \$v Songs and music	\$a Sacred music \$2 lcgft

- List 3: entire \$a

	A	B
1	650 \$a	655
20	Alma maters (Songs)	\$a Alma mater songs \$2 lcgft
21	Alternate takes (Sound recordings)	\$a Alternate takes (Sound recordings) \$2 lcgft
22	Alternative country music	\$a Alternative country music \$2 lcgft

- List 6: Leader/06 plus entire \$v

	A	B	C
1	Leader/06	\$v	655
11	cdj	Anecdotes	\$a Anecdotes \$2 lcgft
12	cdj	Apologetic works	\$a Apologetic writings \$2 lcgft
13	cdj	Atlases	\$a Atlases \$2 lcgft

More lists for 655

- List 7: Leader/06 plus 008/20

	A	B	C
1	Leader	008	655
2	06=c,d	20=a	\$a Scores \$2 lcgft
3	06=c,d	20=b	\$a Scores \$2 lcgft
4	06=c,d	20=c	\$a Scores \$2 lcgft
5	06=c,d	20=d	\$a Vocal scores \$2 lcgft
6	06=c,d	20=e	\$a Scores \$2 lcgft
7	06=c,d	20=g	\$a Scores \$2 lcgft

- List 8: 008/18-19 or 047

	A	B
1	Code	655
2	cc	\$a Chants \$2 lcgft
3	cg	\$a Concerti grossi \$2 lcgft
4	cp	\$a Polyphonic chansons \$2 lcgft
5	cz	\$a Art music \$2 lcgft

Worked-out example: 382

Starting with:

650		0	Concertos (Baritone with brass band) #v Solo with piano.
-----	--	---	--

After considering \$a:

382	0	1	#b baritone horn #n 1 #a brass band #e 1 #r 1 #t 1 #2 lcmpt
-----	---	---	---

But after considering \$v:

382	0	1	baritone horn #n 1 #a piano #n 1 #s 2 #2 lcmpt
-----	---	---	--

Worked-out example: 655

Starting with:

650		0	Concertos (Baritone with brass band) #v Solo with piano.
-----	--	---	--

List 1:

	A	B
1	650 \$a	655
21	Concertos (...)	\$a Concertos \$2 lcgft
22	Duets (Unspecified instrument and [ins: \$a Chamber music \$2 lcgft	

List 6:

	A	B	C
1	Leader/06	\$v	655
235	cd	Solo with piano	\$a Scores \$2 lcaft \$a Arrangements (Music) \$2 lcaft
236	cd	Solo with pianos	\$a Scores \$2 lcaft \$a Arrangements (Music) \$2 lcaft

Worked-out example: the end

Before:

650		0	Concertos (Baritone with brass band) #v Solo with piano.
-----	--	---	--

After:

382	0	1	baritone horn #n 1 #a piano #n 1 #s 2 #2 lcmpt
650		0	Concertos (Baritone with brass band) #v Solo with piano.
655		7	Scores. #2 lcgft
655		7	Arrangements (Music) #2 lcgft
655		7	Concertos. #2 lcgft

370 field

370			#g Finland #2 naf
650		0	Chamber music #z Finland.
655		7	Chamber music. #2 lcgft

An interesting question

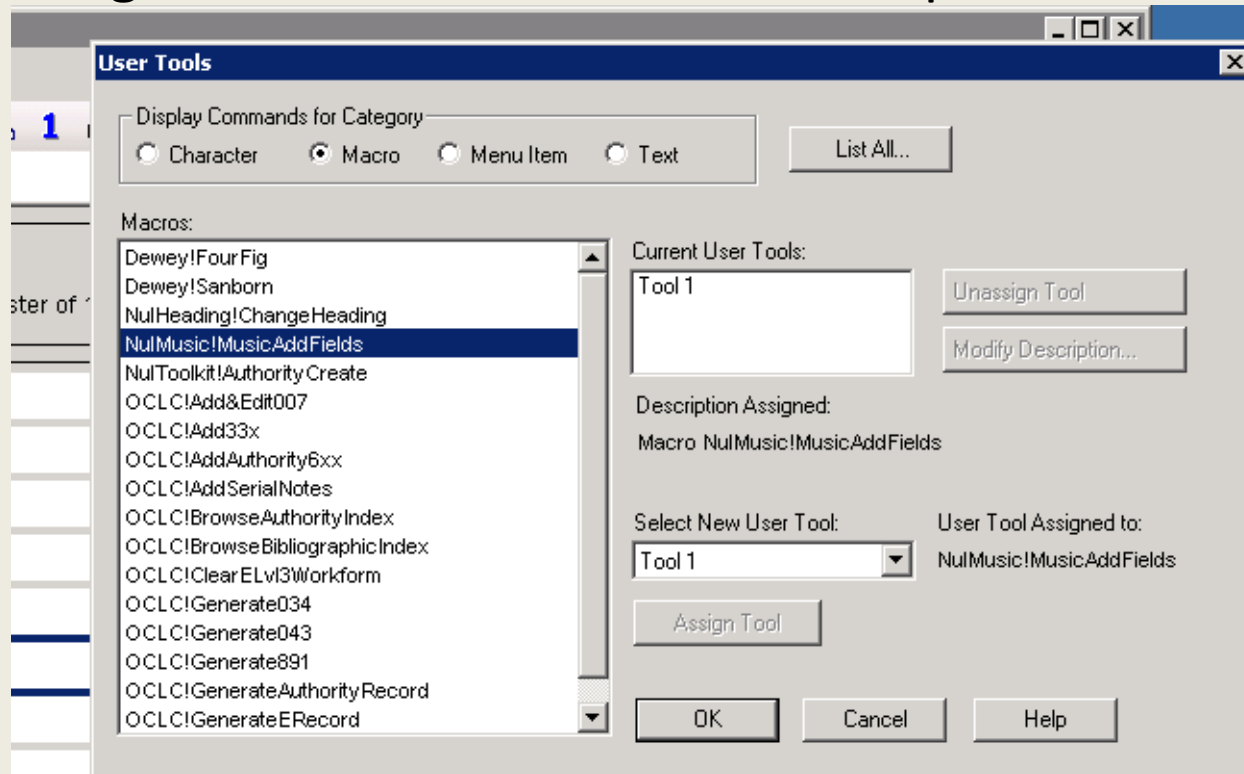
- Can a now-redundant LCSH be deleted?

How do I get this?

- Program documentation is at
<http://files.library.northwestern.edu/public/Music382/documentation/>
- E-mail to
mrsmith@northwestern.edu
casey@mullingroup.com

Installation

- Unzip file
- Run setup program
- Assign shortcut or button to macro part of tool



And?

- Tool ready for broader use
- Probably not blow up
- Not finished or perfect
- Suggestions/complaints/problems welcome
- Include OCLC #
- Try it!

What's next: Testing, testing and more testing

1) Read the algorithm and OCLC toolkit documentation

Who: **You!!**

Send feedback to MLA

Watch the CMC Blog for more information

2) Install and test the OCLC toolkit

Who: **You!!**

Install page: <http://files.library.northwestern.edu/public/Music382/>

How-to demonstration (on CMC YouTube channel):
https://www.youtube.com/watch?v=haO_8Bgj0hQ&t=1s

3) Test the algorithm on batches of bibliographic records

Who: individual libraries (including LC and **you!!**) and OCLC

4) Monitor community testing and feedback, revising algorithm regularly

Who: MLA Vocabularies Subcommittee

What's next: Improving the algorithm

Monitoring (and perhaps proposing) new terms in the LC vocabularies.

Combinatorial and “fuzzy” logic

- Mining text in 500 fields (example: “Sonata for 3 bassoons, 2 cellos and 8 double basses”)

Linking faceted data fields to other fields describing the same specific work(s) in a compilation

- Dare we discuss \$8?

Adding a marker to the MARC record indicating machine-generated data has been added.

- MARC field 883 is a possible method

Enhancing authority data

Applying algorithm to other encoding standards (MODS, BIBFRAME, etc.), including incorporating URIs

- See also PCC Task Group on URIs in MARC (<https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html>)

What's next: The future of LCSH

LCSH likely to remain in score and sound recording records for some time to come

- The algorithm may never be perfect
- Discovery systems need time to catch up

What about works about music?

- Not all LCSH music form/genre/medium headings are topic-friendly (though many are)
- LC will (we expect) eventually cancel headings deemed inappropriate or unnecessary for topical use
 - (Hint: Excellent future MLA-CMC project!!)

What's next: The future of LCSH

Types of topical music headings

- Topical only, per scope note and SHM
Opera (vs. form heading *Operas*)
- Usable as topical or form
Symphonic poems (presence of \$x History and criticism signals topical use)
- Usable as topical or medium
Violin music

How granular should topical music headings be?

- Headings qualified by language
Songs, French
- Headings that combine form/genre and medium
Concertos (Piano) (might be worth keeping)
Suites (Bassoon, clarinet, trombones (2), double bass, percussion) (maybe not)
- Headings qualified by Sacred/Secular
Cantatas, Sacred (or postcoordinate *Cantatas with Sacred music?*)

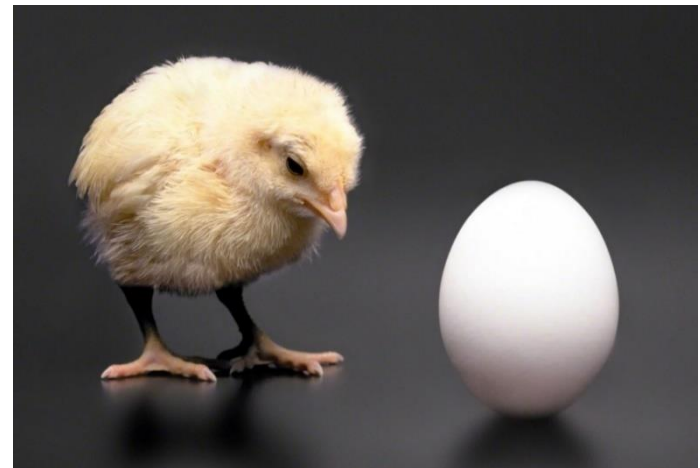
What's next: About those discovery systems

“Chicken-and-egg problem”

1. ILS managers reluctant to “switch on” faceted data functionality without sufficient instance data
2. (Some) agencies reluctant to implement vocabularies in current cataloging without assurances they will be displayed and indexed

Possible solutions

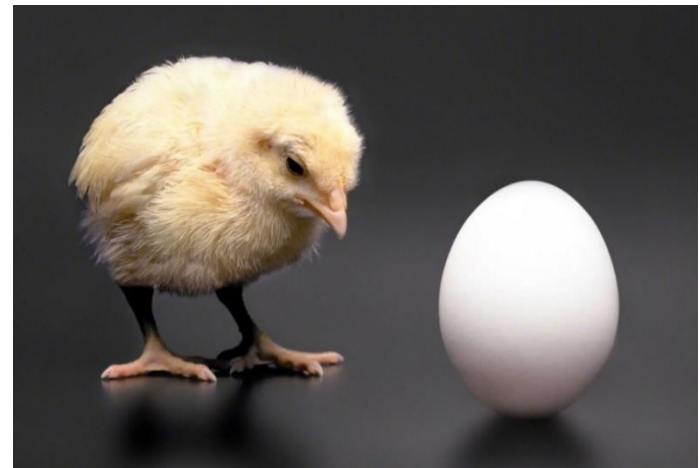
1. Retrospective implementation, best practices for indexing and display and user testing
2. Tools for streamlining vocabulary implementation (e.g., OCLC toolkit)



What's next: About those discovery systems

Resources for faceted vocabulary champions

- Music OCLC Users Group Reference, Discovery and Collection Committee (<http://musicoclcusers.org/resources/discovery-reference-collections/>)
- MLA/MOUG Joint OCLC WMS System Interest Group
- Other MLA System Interest Groups
- MLA Music Discovery Requirements (http://www.musiclibraryassoc.org/page/mdr_es)



Acknowledgments

MLA VS project group (2014-2018)

Casey Mullin, Western Washington University (Chair, MLA VS)

Jennifer Vaughn, Syracuse University

Nancy Lorimer, Stanford University (Chair, MLA Genre/Form Task Force (concluded 2016))

Matt Ertz, University of Louisville

Kirk-Evan Billet, Peabody Institute

Morris Levy, Northwestern University

Nurhak Tuncer, City Colleges of Chicago

Jeff Lyon, Brigham Young University

Ralph Hartsock, University of North Texas

Kyle Shockey, Library of Congress

Marty Jenkins, Wright State University

Special thanks to Gary Strawn!

Questions??



CASEY@MULLINGROUP.COM

MRSMITH@NORTHWESTERN.EDU

Thank you!!!