

Getting More out of MARC with Primo: Strategies for Display, Search and Faceting

ELUNA
May 4, 2018

Kelley McGrath
University of Oregon

Lesley Lowery
Orbis Cascade Alliance

New LC Fields/Subfields: In recent years, many new fields and subfields have been introduced into MARC21 that can prove useful when leveraged in a faceted discovery interface. In this presentation, we will look at several of these, describing how we incorporated them into Primo and discussing the challenges we encountered.

LC Genre-Form Terms

French **poetry** \$y 20th century
English **drama** \$x Women authors
Children's **films**
American **literature** -- New York (State)
Songs (High voice) with piano

Page 2

New fields for non-genre/form info: In 2007 the Library of Congress (LC) began a project to create a new genre/form vocabulary with the intention of moving information describing what a resource is out of topical subject fields and into a separate genre/form field. Although, the terms in red were incorporated into this new vocabulary, something had to be done with the related information in black that was also part of the existing Library of Congress subject headings.

Creator Demographic Group

Library of Congress Demographic Group Terms (LCDGT)

English drama \$x **Women authors**

Refined by: Genre: Drama x Original Language: English x
Creator Demographic Group: Women x Creator Demographic Group: English x

Language: English
Subjects: English drama -- 18th century; English drama -- Women authors
Genre/Form: Drama; Drama; Acting editions.
Work Information: Time Period of Creation: Eighteenth century; Original language: English;
Creator Demographic Group: Women; English; Britons

Creator Demographic Group	
Men	(547)
Americans	(500)
Women	(226)
African Americans	(115)
Britons	(115)
English	(105)
Authors	(94)
Statesmen	(93)
Inventors	(86)
Scientists	(84)
Printers	(83)
Slaves	(74)
Politicians	(47)
Activists	(45)
Jews	(229)
Germans	(217)
University and college faculty memb	
Australians	(96)
Estonians	(74)
Journalists	(35)
Mexicans	(73)
Immigrants	(33)
Poets	(29)
Argentines	(63)
Jews	(28)
Brazilians	(62)
Clergy	(27)
Norwegians	(59)
New Zealanders	(41)
Greeks	(34)
Turks	(23)
Portuguese	(14)
Gogo (African people)	(2)


Page 3

To account for this information, LC created some additional new vocabularies, including LCDGT for demographic group terms. There is a new MARC field (386) to record creator demographic group terms, which we display and facet in Primo as shown here. This facet is straightforward to create, but there are many questions about best practices for choosing what data to put in this field.

Audience

DISPLAY: Combination of 3 sources

- 008/22 Target Audience (fixed field)
- 385 Audience Characteristics (controlled vocabulary)
- 521 Target Audience Note



Page 4


Demographic group terms for audience may be recorded in the new 385 field, which we also display and facet. In the audience PNX display field, we combine information from three different sources: the audience code in 008, the recently-created 385 field and the audience note field.

Audience

Forest essays, level 7-12.
Oregon Forest Resources Institute, issuing body.

Audience: Junior high school students; High school students

DISPLAY: 385 Audience Characteristics



Source: MARC, Field: 385, Ind1: [], Ind2: [], Subfield: Include, a

Conditions: No conditions Specified

Transformations: Define subfield delimiter, Behavior: MERGE, First delimiter: ;, First delimiter Space: After, Repeat Number: [], Remaining delimiter: []

1 The dictionary of the book : a glossary for book collectors, booksellers, librarians, and others
Sidney E. Berger, author.; Lanham, Maryland : Rowman & Littlefield, 2016

Check holdings Audience: Book collectors; Booksellers; Librarians

Print Book

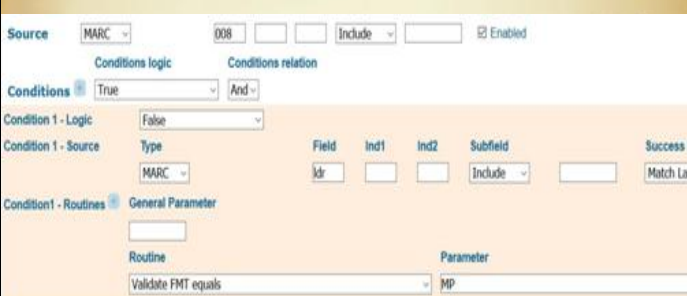
Page 5

Display of the 385 field is straightforward and the rule follows the pattern used in Primo for displaying subject fields.

Audience

DISPLAY: 008/22 Target Audience (FF)

*not used in all MARC record types



Source: MARC, Field: 008, Ind1: [], Ind2: [], Subfield: Include, Enabled

Conditions logic: True, Conditions relation: And

Condition 1 - Logic: False

Condition 1 - Source: Type: MARC, Field: 008, Ind1: 22, Ind2: [], Subfield: Include, Success If: Match Last

Condition 1 - Routines: General Parameter: Validate FMT equals, Parameter: MP

Conditions 2-3 (not shown) exclude mixed materials and serials

Page 6

Positions in MARC 008 can have different meanings depending on record format. The audience fixed field is not used in all record formats. In the map format, this byte has a completely different meaning. If maps are not excluded from this rule, you will get odd results, such as all of your quadrangle maps having an audience of preschool children. We therefore add conditions to exclude the record formats, such as maps, that do not include audience.

Audience

DISPLAY: 008/22 Target Audience (FF)

*don't display fixed field value if 385 exists; prevents likely duplicate terms

Condition 4 - Logic: False

Condition 4 - Source: Type MARC, Field 385, Ind1, Ind2, Subfield Include, Success If a

Condition 4 - Routines: General Parameter, Routine Input exists, Parameter

Page 7

We also added a condition to ignore the fixed field if the 385 field also exists. This is done to reduce redundancy in display and would be unnecessary if Primo de-duped PNX display fields.

Audience

DISPLAY: 008/22 Target Audience (FF)

*use mapping table to display human-friendly terms

Transformation: Take substring, Use mapping table, Replace string by string

Parameter: 22@@1, 008_22_audience, @,^,^,^,^,^

Action: MERGE, ;, After, Repeat Number, Remainder

Page 8

We then isolate the value of the audience fixed field and send it to a mapping table to turn it into a human-friendly value for display.

Audience

DISPLAY: 008/22 Target Audience (FF)

*use mapping table to display human-friendly terms

sourceCode1*	targetCode
a	Preschool children,Children
b	School children,Children
c	Preteens,Children
d	Teenagers
e	Adults
f	Specialized audience
g	General audience
j	Children

22 - Target audience (006/05)

- Unknown or not specified

a - Preschool	e - Adult
b - Primary	f - Specialized
c - Pre-adolescent	g - General
d - Adolescent	j - Juvenile
	- No attempt to code

Label	Vocabulary
Preteens	LC Demographic Group Terms

Pre-teens ; Preadolescents ; Preteenagers ; Tween


Page 9

This is our mapping table, which we use for display, search and faceting. Because we also use the textual terms from the LCDGT vocabulary for faceting, we chose terms that matched that vocabulary to improve collocation. The label for the value "c" in the MARC format is "pre-adolescent," but that is a cross-reference for preteens in LCDGT so we used "Preteens" as our label.

Audience

DISPLAY: 521 Target Audience Note

*Use indicators to preface note

1  **A ball for Daisy**
 Christopher Raschka, illustrator;
 Available at John E. Jaqua Law
 RAS BAL) and other locations

Audience: Preschool children; Children
 Preschool.
 Interest Age Level: Ages 3-7.

Source: MARC 521 1 Include a,b,3

Conditions: No conditions Specified

Transformations

Transformation	Parameter
Add to beginning of string	Interest Age Level: ^

Behavior: ADD

Page 10

Finally, we display the audience note. This field includes indicator values that mark narrower meanings. We add prefixes to the note display based on these indicator values where applicable, as shown here with 1st indicator 1 for interest age level.

Audience

SEARCH: mapped 008 display terms

FACET: lopsided results

Audience	
General audience	(13,989)
Adults	(9,546)
Children	(8,609)
Specialized audience	(687)
Teenagers	(660)
School children	(488)
Preteens	(453)
Preschool children	(198)
Music teachers	(5)

Page 11

We made sure to map the coded values from the fixed field to textual terms in the search index. Because of the way Primo weights facet values, the facet results are dominated by the fixed field terms and it is hard to find examples with LCDGT. It's also not clear how useful the general, adult and specialized audience values are since they are much less consistently used.

LCMPT

Library of Congress Medium of Performance Thesaurus for Music

- stand-alone vocabulary
- terminology to describe the instruments, voices, etc., used in the performance of musical works
- based chiefly on existing LC subject headings

alto saxophone	single reed instrument
baritone saxophone	reed instrument
saxophone	woodwind instrument
saxophone ensemble	wind instrument
	...
	performer

Page 12

Medium of performance is another type of information that is commonly combined with genre in LCSH. LC and the Music Library Association have developed a new vocabulary called Library of Congress Medium of Performance Thesaurus for Music or LCMPT. LCMPT was created as a new home for this information and the 382 field was added to MARC to support it.

MARC 382: Medium of Performance

382 01 \$a saxophone \$n 1 \$a piano \$n 1 \$s 2 \$2lcmt

\$a - Medium of performance (R)
\$b - Soloist (R)
\$d - Doubling instrument (R)
\$e - Number of ensembles of the same type (R)
\$n - Number of performers of the same medium (R)
\$p - Alternative medium of performance (R)
\$r - Total number of individuals performing alongside ensembles (NR)
\$s - Total number of performers (NR)
\$t - Total number of ensembles (NR)
\$v - Note (R)
\$3 - Materials specified (NR)

Page 13

As you can see by the number of subfields, 382 is designed to support more complex details about medium of performance than LCSH. It was also intended to be useful in a faceted environment.

Looking at the raw 382 field at the top of this slide, it is clear that some intervention is needed to make this new data usefully displayable and searchable in Primo.

Medium of Performance: Display

382 0 1 viola \$n 1 \$p clarinet \$n 1 \$a piano \$n 1 \$s 2 \$2 lcmt

\$a viola \$n 1 \$p clarinet \$n 1 \$a piano \$n 1 \$s 2

Option 1: Write separate rules for each subfield and then merge them

- Can treat subfields differently
- Cannot retain order of subfields; each type of subfield is processed sequentially

Option 2: Write one rule that takes multiple subfields as input

- Must treat all subfields the same way
- Retains order of subfields

Music Info: Duet: viola (1); piano (1) (alternate instrumentation: clarinet)

Page 14

Although the Primo norm rules provide powerful tools for manipulating MARC data, there are some limitations on what is possible, especially for complex fields with many subfields. There are two approaches that can be used to deal with fields that include more than one type of subfield.

The first approach is to process each type of subfield separately and then combine them. This allows you to treat specific subfields differently, but it doesn't preserve the order of the subfields. That is, you could get all the subfield a's followed by all the \$n's.

The second approach is to process the different types of subfields with a single rule. In this case, you preserve the order of the data, but you can't do something different to subfield a from what you're doing to subfield n.

We used a combination of both approaches to turn 382 into something eye-readable. I will now walk through how we got from the raw MARC at the top of this slide into the human-readable display at the bottom.

Medium of Performance: Display

382 0 1 viola #n 1 #p clarinet #n 1 #a piano #n 1 #s 2 #2 lcmpt

2
Duet

Source	Type	Field	Ind1	Ind2	Subfield
MARC		382			Include s

Transformation	Parameter
Use mapping table	no_musical_parts
Behavior	
First delimiter	
First delimiter Space	
Repeat Number	
Remainder	
Action	MERGE

Music Info: Duet: viola (1); piano (1) (alternate instrumentation: clarinet)

Page 15

The first rule that applies takes subfield s, a non-repeatable subfield, as input. Subfield s should contain a number representing the total number of performers. We use a mapping table, which I'll show later, to change numbers into words for display. Here the rule takes the input of 2 and maps it to the word Duet.

Medium of Performance: Display

382 0 1 viola #n 1 #p clarinet #n 1 #a piano #n 1 #s 2 #2 lcmpt

viola 1 1 piano 1
 viola; 1; 1; piano; 1
 ;_1 → _(1) viola (1) (1); piano (1)

Source	Type	Field	Ind1	Ind2	Subfield
MARC		382			Include a,b,e,n

Transformation	Parameter
Define subfield delimiter	;^
Replace string by string	;1@^(1)
Replace string by string	;2@^(2)

Music Info: Duet: viola (1); piano (1) (alternate instrumentation: clarinet)

Page 16

The second rule takes its input from multiple subfields. The starting input for this rule is the top red line that says "viola 1 1 piano 1" Notice that the order is preserved, but there's no way to identify which data came from what subfield.

We start with the transformation to define a subfield delimiter. We use a semicolon and space, which results in the second line where the data from each subfield is separated by a semicolon.

We then take advantage of the fact that we can identify numbers to remove the semicolon preceding any number and instead put the number in parentheses. This is a situation that is crying out for a regular expression. There actually is a transformation in the documentation that looks like it would be helpful here, but unfortunately, that function is broken. So we explicitly wrote a replace transformation for the numbers 1-9. This will fail if there are 10+ performers coded as playing the same instrument, but that is uncommon.

Medium of Performance: Display

382 0 1 viola #n 1 #p clarinet #n 1 #a piano #n 1 #s 2 #2 lcmpt

viola (1) (1); piano (1)
 viola (1); piano (1)
 : viola (1); piano (1)

Replace string by string ; 9@^(9)
 Replace string by string) (1)@
 Replace string by string) (1)@
 Replace string by string) (1)@
 Replace string by string) (1)@
 Replace string by string) (1)@
 Replace string by string) (1)@
 Replace string by string) (2)@

Behavior First delimiter First delimiter Space Repeat Number
 Action MERGE : After

Music Info: Duet: viola (1); piano (1) alternate instrumentation: clarinet

Page 17

We have another problem, though, which is that subfield n is used both for the number of performers in the main medium of performance statement and also after doubling and alternate instrumentation as shown here with the clarinet part. We have no way to distinguish these different meanings of subfield n, which is causing the multiple appearance of the number 1 after viola.

We undertake another series of find and replace actions to remove extra numbers. Because we cannot do this conditionally, we tried to include enough transformations to account for the vast majority of situations. However, it is still possible to find medium of performance statements in our catalog with extra numbers if you know where to look.

Finally, the merge action for this rule pre-pends a colon and space to the instrumentation if there is preceding text, such as duet.

Medium of Performance: Display

382 0 1 viola #n 1 #p clarinet #n 1 #a piano #n 1 #s 2 #2 lcmpt

clarinet
 (alternate instrumentation: clarinet)

Source	Type	Field	Ind1	Ind2	Subfield
MARC		382			p

Transformation	Parameter
Define subfield delimiter	;^
Add to beginning of string	(alternate instrumentation: ^
Add to end of string)

Behavior First delimiter First delimiter Space Repeat Number Remaining delin
 Action MERGE : Before

Music Info: Duet: viola (1); piano (1) (alternate instrumentation: clarinet)

Page 18

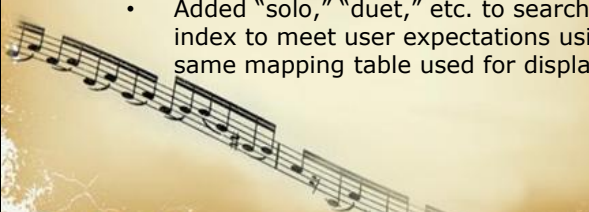
The last thing to deal with in this example is the alternate instrumentation in subfield p. Given our choice between preserving the order of the data and thus associating clarinet with the instrument that it is an alternative for and being able to process the subfield separately, we have chosen the latter. This enables us to separately label the alternate instrumentation.

Medium of Performance: Search

382 0 1 viola #n 1 #p clarinet #n 1 #a piano #n 1 #s 2 #2 lcmpt

Music Info: Duet: viola (1); piano (1) (alternate instrumentation: clarinet)

- "Duet" displays
- "Duet" does not exist in the MARC data
 - Added "solo," "duet," etc. to search index to meet user expectations using same mapping table used for display



Page 19

For the most part, we have just added the 382 subfields to the general search index without any intervention. However, we have included rules to index the words, such as duet and alternate instrumentation, that we generated for display. It turns out that at least some users do want to search by words like "solo."

Medium of Performance: Facet

382 \$a [medium]

382 \$b [soloist]


382 \$b + « (solo) »

Music: Medium of Performance

- mixed chorus (10)
- alto saxophone (9)
- piano (8)
- soprano saxophone (7)
- baritone saxophone (6)
- More options ▾

Refine Search

Include	Exclude	Music: Medium of Performance
<input type="checkbox"/>	<input type="checkbox"/>	mixed chorus (10)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (9)
<input type="checkbox"/>	<input type="checkbox"/>	piano (8)
<input type="checkbox"/>	<input type="checkbox"/>	soprano saxophone (7)
<input type="checkbox"/>	<input type="checkbox"/>	baritone saxophone (6)
<input type="checkbox"/>	<input type="checkbox"/>	tenor saxophone (6)
<input type="checkbox"/>	<input type="checkbox"/>	soprano voice (5)
<input type="checkbox"/>	<input type="checkbox"/>	soprano voice (solo) (5)
<input type="checkbox"/>	<input type="checkbox"/>	trumpet (5)
<input type="checkbox"/>	<input type="checkbox"/>	percussion (5)
<input type="checkbox"/>	<input type="checkbox"/>	bass voice (5)
<input type="checkbox"/>	<input type="checkbox"/>	double bass (5)
<input type="checkbox"/>	<input type="checkbox"/>	viola (4)
<input type="checkbox"/>	<input type="checkbox"/>	cello (4)
<input type="checkbox"/>	<input type="checkbox"/>	bass voice (solo) (4)
<input type="checkbox"/>	<input type="checkbox"/>	trombone (3)
<input type="checkbox"/>	<input type="checkbox"/>	tenor voice (3)
<input type="checkbox"/>	<input type="checkbox"/>	tenor voice (solo) (3)
<input type="checkbox"/>	<input type="checkbox"/>	orchestra (3)
<input type="checkbox"/>	<input type="checkbox"/>	instrumental ensemble (3)



Page 20

Currently, we have implemented three facets based on medium of performance in the 382 field.

The first one is a facet for instrument, voice or ensemble type. This includes everything from \$a medium of performance and \$b soloist. The terms from \$b are double-posted with the qualifier solo to help users find pieces that have solo parts for their instrument.

This seems pretty straightforward, but there are a couple ways this facet behaves that likely clash with user expectations or preferences.

Medium of Performance: Facet

Music: Medium of Performance

- piano (2,440)
- violin (2,269)
- cello (1,206)
- viola (652)
- clarinet (466)
- More options ▾

Results 1 - 20 of **2,145** for **nrrwg music** Sorted by: R

AND

Refined by: Music: Medium of Performance: piano x Music: Medium of Performance: violin x

1 **Complete music for solo piano & violin and piano**
Luigi Dallapiccola, 1904-1975, composer, arranger of music.; Maria Clementi, (Pianis Fanfoni, 1964- instrumentalist.; Leeuwarden : Brilliant Classics, 2015

● **Online access**

Access Online Details

Refine Search

Include	Exclude	Music: Medium of Performance
<input checked="" type="checkbox"/>	<input type="checkbox"/>	piano (2,440)
<input checked="" type="checkbox"/>	<input type="checkbox"/>	violin (2,269)
<input type="checkbox"/>	<input type="checkbox"/>	cello (1,206)
<input type="checkbox"/>	<input type="checkbox"/>	viola (652)
<input type="checkbox"/>	<input type="checkbox"/>	clarinet (466)

Results 1 - 20 of **2,564** for **nrrwg music**

OR

Refined by: Music: Medium of Performance: violin x piano x

1 **Complete music for solo piano & violin and piano**
Luigi Dallapiccola, 1904-1975, composer, arranger of music.; Maria Clementi, (Pianis Fanfoni, 1964- instrumentalist.; Leeuwarden : Brilliant Classics, 2015

● **Online access**

Access Online Details

Page 21

First of all, users probably want to find pieces that include all the instruments or voices that they select. That is, they want to do a Boolean AND search.

If they search by selecting terms consecutively, which is the only way to do it with the short list shown at the top of the slide, they will get an AND search.

However, if they select multiple terms at once from the longer “refine my results” list, Primo will perform an OR search. This makes sense from the standpoint that for most long lists, an OR search would be preferred, and it’s desirable for all the “refine my results” lists to exhibit the same behavior. But it probably isn't what users want from this particular facet most of the time.

Medium of Performance: Facet

Music: Medium of Performance

- piano (2,440)
- violin (2,269)
- cello (1,206)
- viola (652)
- clarinet (466)
- More options ▾

Results 1 - 20 of **2,145** for **nrrwg music** Sorted by: R

AND

Refined by: Music: Medium of Performance: piano x Music: Medium of Performance: violin x

1 **Complete music for solo piano & violin and piano**
Luigi Dallapiccola, 1904-1975, composer, arranger of music.; Maria Clementi, (Pianis Fanfoni, 1964- instrumentalist.; Leeuwarden : Brilliant Classics, 2015

● **Online access**

Access Online Details

3 **Piano concerto : Violin concerto**
Kurt Atterberg, 1887-1974, composer.; Dan Franklin Smith, instrumentalist.; Christlar instrumentalist.; B. Tommy Andersson, conductor.; Gävle symfoniorkester, instrumen p1999

● **Available at Knight Library Audio & Video Room (COMPACT DISC CX5963)**

Find in a library Details

orchestra, op. 7, E minor.
Format: 1 audio disc (72 min.) : digital, stereo ; 4 3/4 in.
Genre/Form: Concertos; Concertos; Compact discs
Additional subjects: Concertos (Piano); Concertos (Violin)
OCLC Number: 53306686
Alma IZ record number: 9986955601852
Alma NZ record number: 99172176930001451
Music info: *piano (1); orchestra (solos for: piano)
*violin (1); orchestra (solos for: violin)

Page 22

Another important thing to notice is that the facets are working against bib records for the whole resource being described so violin and piano just have to be somewhere in the record.

They don't have to be referring to the same piece as in this example which includes a solo piano concerto and a solo violin concerto.

Total Number of Performers

Refine Search		Music: Number of Performers
Include	Exclude	
<input type="checkbox"/>	<input type="checkbox"/>	2 performers (107)
<input type="checkbox"/>	<input type="checkbox"/>	4 performers (84)
<input type="checkbox"/>	<input type="checkbox"/>	1 performer (44)
<input type="checkbox"/>	<input type="checkbox"/>	ensemble (42)
<input type="checkbox"/>	<input type="checkbox"/>	3 performers (14)
<input type="checkbox"/>	<input type="checkbox"/>	5 performers (12)
<input type="checkbox"/>	<input type="checkbox"/>	14 performers (10)
<input type="checkbox"/>	<input type="checkbox"/>	7 performers (10)
<input type="checkbox"/>	<input type="checkbox"/>	11 performers (9)
<input type="checkbox"/>	<input type="checkbox"/>	10 performers (7)
<input type="checkbox"/>	<input type="checkbox"/>	6 performers (5)
<input type="checkbox"/>	<input type="checkbox"/>	20 performers (4)

Page 23

In addition to medium of performance, we also implemented a facet for total number of performers if that is given.

It doesn't attempt to count ensembles, but just lists the presence of an ensemble. Some of the larger numbers of individual performers should possibly be double-posted under ensemble, but we're not doing that right now.

Again, the facets are working against the bib so if you pick 1 performer and saxophone, there's no guarantee that all the results will include a piece for solo sax. However, the results are not as misleading as they might be thanks to relevancy ranking and the fact that you can only ever see the top twenty values in a given facet.

Number of Performers with Medium

Refine Search		Music: Medium of Performance Statement
Include	Exclude	
<input type="checkbox"/>	<input type="checkbox"/>	soprano saxophone (1); alto saxophone (1); tenor saxophone (1); baritone saxophone (1) (58)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (1); piano (1) (39)
<input type="checkbox"/>	<input type="checkbox"/>	saxophone (1) (26)
<input type="checkbox"/>	<input type="checkbox"/>	saxophone (1); piano (1) (18)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (1) (10)
<input type="checkbox"/>	<input type="checkbox"/>	soprano saxophone (1); piano (1) (7)
<input type="checkbox"/>	<input type="checkbox"/>	baritone saxophone (1); piano (1) (6)
<input type="checkbox"/>	<input type="checkbox"/>	saxophone (1); piano (1); bass (1); drums (1) (5)
<input type="checkbox"/>	<input type="checkbox"/>	saxophone (2) (5)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (1); orchestra (1) (5)
<input type="checkbox"/>	<input type="checkbox"/>	tenor saxophone (1); piano (1) (5)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (2) (4)
<input type="checkbox"/>	<input type="checkbox"/>	alto saxophone (2); tenor saxophone (1); baritone saxophone (1) (4)
<input type="checkbox"/>	<input type="checkbox"/>	saxophone (4) (4)
<input type="checkbox"/>	<input type="checkbox"/>	horn (1); piano (1) (4)

Page 24

In order to compensate somewhat for that weakness, we have also created a medium of performance statement facet, which allows users to specify exactly what combination of instruments or voices they are looking for.

Challenges for Collocation

Music: Medium of Performance Statement

piano (1) (3,748) ★
violin (1); piano (1) (801)
piano (2) (413)
piano (1); orchestra (1) (349)
clarinet (1); piano (1) (207)
piano (1); violin (1); cello (1) (163)
violin (1); cello (1); piano (1) (161)
piano (1); orchestra (128)
piano (100) ★
oboe (1); piano (1) (90)
piano (1); violin (2); viola (1); cello (1) (68)
piano (1); violin (1); viola (1); cello (1) (50)
violin (2); viola (1); cello (1); piano (1) (43)
violin (1); viola (1); cello (1); piano (1) (36)

Order matters:
piano (1); violin (2); viola (1); cello (1)
violin (2); viola (1); cello (1); piano (1)

Page 25

There are a couple places where this facet doesn't provide the collocation that users will expect. One is due to the fact that the MARC format allows you to omit the 1 for any instrument or voice that has only a single performer. Since the facets are looking at literal text strings, they see piano and piano followed by 1 as two different things. This is an old screenshot and we have since been able to resolve this issue by supplying a 1 following any medium of performance that doesn't have a number of performers specified.

The thornier issue is that the instruments or voices are not listed in a standard order, which again results in different strings that don't collocate. The two lines marked with arrows both contain the same instrumentation, just in a different order. This is unlike LCSH headings for medium of performance, where there are rules for ordering the instruments so that correctly-constructed headings do collocate. As far as I can tell, reordering the instruments is beyond the capability of Primo's data manipulation tools so the only solution would be to do some massaging of the data before sending it to Primo.

Relationship Designators

Personal main entry:
 100 1_ \$a Poulin, Richard, \$e author.

Corporate added entry:
 710 1_ \$a United States. \$b Department of State, \$e issuing body.

Personal added entry:
 700 1_ \$a Soares, Paul \$c (Writer on computer games), \$e author, \$e illustrator.

Page 26

Implementation of the RDA cataloging rules also led to new fields in MARC and increased usage of some existing fields and subfields. We've tried to find a way to put these to good use.

For example, we've received requests to make more use of relationship designators for searching and faceting in Primo.

Composers and Performers

Top results for operas

Composer

- Verdi, Giuseppe, 1813-1901 (2,913)
- Mozart, Wolfgang Amadeus, 1756-1791 (2,403)
- Wagner, Richard, 1813-1883 (2,283)
- Puccini, Giacomo, 1858-1924 (1,770)
- Rossini, Gioacchino, 1792-1868 (1,398)
- More options ▾

Performer

- Wiener Philharmoniker (546)
- Philharmonia Orchestra (London, England) (542)
- Domingo, Plácido, 1941-(493)
- London Symphony Orchestra (466)
- Callas, Maria, 1923-1977 (464)
- More options ▾

Page 27

We've created three new facets – one each for Composers, Directors, and Performers. Here are the top results from a search for operas.

Composers and Performers

Any record	
100/700	\$e = composer
100/700	\$4 = cmp

Record format = score (ldr/06 = c or ldr/06 = d)	
100	any
700 2 nd indicator 2	\$t (title) exists

Record format = musical sound recording (ldr/06 = j)	
100	240 (uniform title) exists
700 2 nd indicator 2	\$t (title) exists

Page 28

We rely on a combination of approaches to generate the composer facet. We use relator terms or codes where they exist, but many records lack these. For records on a score format, we include all 100 fields and names from analytic name-title access points under the assumption that these are all composers. For records coded as musical sound recordings, we only infer that the 100 represents the composer if a uniform title is also present in 240, but we again include all the names from name-title entries in 700 with 2nd indicator 2.

New Fields & Facets in Library Catalogs

Chicken or the egg?

- Incomplete recall and user expectations
- Retrospective conversion

"I got tired of cataloging for my grandchildren."

Page 29

Chicken-and-egg question or What comes first, field use or discovery tools?

New fields that aren't widely used result in incomplete recall of results when used as faceting values. This causes users to see them as inaccurate, and degrades trust. This situation makes institutions hesitant to develop facets for new fields - which makes their cataloging shops hesitant to spend time encoding them. This results in the new fields not being widely used. Rinse, repeat.

Retrospective conversion projects can increase the field use, but they sometimes use incomplete or inaccurate algorithms to batch process records, and can result in degradation of trust in the converted data.

Humpty Dumpty Problem

Refined by: Music: Medium of Performance: violin x Music: Medium of Performance: harpsichord x

A miscellaneous collection of instrumental music for the organ, orchestra, chamber, and harpsichord
George Frideric Handel, 1685-1759, composer; Friedrich Chrysander, 1826-1901, editor; William Babel, approximately 1690-1723, arranger of music; Deutsche Händel-Gesellschaft, binding body; Ridgewood, New Jersey, U.S.A.; Gregg Press Incorporated, 1965

Check holdings

Find in a library Details

Music info: organ (1); orchestra (1) (solos for: organ)
organ (2); orchestra (1) (solos for: organ)
organ (1); string orchestra (1) (solos for: organ)
Duet: viola da gamba (1); harpsichord (1)
violin (2); continuo
flute (1); continuo
instrument (1); continuo
orchestra (1)
string orchestra (1)
Solo: harpsichord (1)
Duet: harpsichord (2)
singer (1); instrumental ensemble (1)

Aggregates can be misleading:
Limiting results to medium of performance of harpsichord and violin will retrieve resources that include both those instruments, **even if they are not from the same piece**.
The medium of performance statement facet tries to compensate for this specific situation, but can't connect medium of performance with other facets, such as composer or musical key

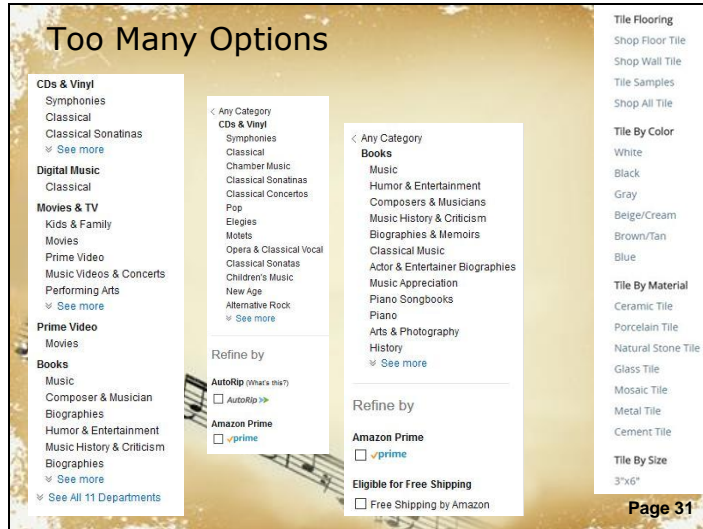
Page 30

Humpty Dumpty - aggregates of multiple works

Users limiting by multiple instruments will get results with those instruments present in the record - but that doesn't mean the instruments all appear in the same piece.

We've tried to improve this situation by providing the Medium of Performance Statement facet, but if the user is looking for a specific medium of performance statement in a piece by a specific composer or in a specific key, they're back to the aggregate problem.

This is a general problem for resources that include multiple discrete works, such as music compilations on CD and literary anthologies.



Using facets in a material-specific interface or Primo view

One reason that many libraries in our consortium haven't implemented these facets is that there are usability challenges with a long list of facets. Many of these fall "below the fold," making them easy for users to miss. Amazon and other retail sites often compensate for this problem by showing a shorter list of facets that are contextually linked to the user's search. This isn't possible in Primo, but one way to work around this limitation is to create a separate, material-specific interface.

There are some caveats to this approach:

You have to market that interface heavily to make sure users are aware of it.

It leads to using multiple interfaces instead of one.

How easy is it to switch between the general search interface and the material-specific one?"?

Will users get confused about which interface they're in?

Looking at the possibility of creating a format-specific interface segues us into...

Music-Specific Scope in Primo

Resource Type

eAudio & eVideo	(4,847)
Scores	(4,386)
Audio CDs	(3,976)

Primo Central Index (PCI) integration problem

5 **Symphony no. 9**
 Ludwig van Beethoven, 1770-1827, composer.; Friedrich Schiller, 1759-1805.; Soile Isokoski; Rosemarie Lang; Robert Gambill; René. Pape; Daniel Barenboim, 1942-; Deutsche Staatsoper Berlin. Chor.; Staatskapelle Berlin.; Hamburg, Germany : Teldec, 2000

Online access

[Access Online](#) [Details](#)

6 **The Dutch National Ballet**
 Hans van Manen, 1932-; Toer van Schaijk; Arthaus Musik (Firm); Nationale Ballet (Netherlands); Reiner Moritz Associates.; Germany? : Arthaus Musik; 200?

Online access

[Access Online](#) [Details](#)

Page 32

PCI integration with customized facets is not possible – this makes it difficult to use new facets to create format-specific scopes (and therefore views) for music or videos in Primo.

The lack of PCI integration creates some facet limitations for us, too. OTB facets always have to be present to pull in PCI resources, and they pull in things that we don't want to see (ex. Videos pulled into a music-specific view due to "Audiovisual" resource type used for PCI.)

Computational Constraints

- No true exploratory search or browse in Primo
- Incomplete recall in most cases

Page 33

There is no true exploratory search due to Primo [facet](#) limitations. These limitations lead to incomplete result sets in most cases. For example, contrast the behavior for:

Computational Constraints

Static facets

- Enumerated list (ca. 500 terms)
- Counts and displays all results

Resource Type	Original Language
Scores (102,585)	English (1,861)
Audio CDs (74,357)	German (1,774)
eAudio & eVideo (63,392)	French (928)
LPs (42,647)	Italian (741)
Audiocassettes (2,285)	Original Language
	Papuan (Other) (1)
	Songhai (1)
	Athapascan (Other) (1)
	Wolof (1)

https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Back_Office_Guide/100Facets Page 34

static facet = all results; relatively short list of enumerated values; marc language code list (close to 500 values)

https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Back_Office_Guide/100Facets

Computational Constraints

Refine Search

Include	Exclude	Genre
<input type="checkbox"/>	<input type="checkbox"/>	Scores (31,733)
<input type="checkbox"/>	<input type="checkbox"/>	Scores and parts (18,058)
<input type="checkbox"/>	<input type="checkbox"/>	Music (12,929)
<input type="checkbox"/>	<input type="checkbox"/>	Excerpts (12,912)
<input type="checkbox"/>	<input type="checkbox"/>	Songs and music (9,299)
<input type="checkbox"/>	<input type="checkbox"/>	Musical settings (7,954)
<input type="checkbox"/>	<input type="checkbox"/>	Chamber music (7,633)
<input type="checkbox"/>	<input type="checkbox"/>	Sound recordings (7,400)
<input type="checkbox"/>	<input type="checkbox"/>	Art music (6,112)
<input type="checkbox"/>	<input type="checkbox"/>	Parts (Music) (4,404)
<input type="checkbox"/>	<input type="checkbox"/>	Excerpts, Arranged (3,893)
<input type="checkbox"/>	<input type="checkbox"/>	Songs (3,454)
<input type="checkbox"/>	<input type="checkbox"/>	Piano music (3,043)
<input type="checkbox"/>	<input type="checkbox"/>	Arrangements (Music) (3,014)
<input type="checkbox"/>	<input type="checkbox"/>	Instrumental settings (2,452)
<input type="checkbox"/>	<input type="checkbox"/>	Sonatas (2,252)
<input type="checkbox"/>	<input type="checkbox"/>	Suites (1,991)
<input type="checkbox"/>	<input type="checkbox"/>	Orchestral music (707)
<input type="checkbox"/>	<input type="checkbox"/>	Motion picture music (517)
<input type="checkbox"/>	<input type="checkbox"/>	Instrumental music (292)

Dynamic facets

- Open-ended list
- Displays top 20* terms from records ranked in top 200* results
- Displays count (and retrieves) matching records from top 50,000 results

*Default values; can be increased somewhat, but trade-off on performance

Page 35

dynamic facet = displays only top 20 terms from top 200 results; hit count is from top 50,000 records only.

https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Back_Office_Guide/100Facets

Constraints on Manipulation

Norm Rules in Primo Back Office

- Browser-based GUI interface with hard-coded conditions and transformations
- There's no way to move just one rule of a group to a group in another rule set.
- OTB conditions/transformations can't always target really specific data strings.
- Some transformations are broken, which requires creative multi-step workarounds.

Page 36

Constraints on manipulating rules due to browser-based GUI with hard-coded conditions and transformations. We're *mostly* able to achieve our goals with the interface, but there are some problems that are real bugbears here.

Long page load times

No way to move just one rule of a rule group to the same group in a different set - you have to replace the entire group.

Conditions and transformations are hard-coded, and they are not always flexible enough to make really targeted changes and data manipulations.

Broken transformation – Replace Last RegExp By String

With this transformation, it would be possible to easily handle some of the end-of-string punctuation (particularly useful with name normalization) issues that we've encountered.

Without it, we're forced to use clunky multi-step workarounds, or just live with "lesser evils" in data display.

SF#00177731 (WWU) in Development status since 12/21/2015.

Video at
<http://www.screencast.com/t/8ohPUNMkNs2>. Feel free to reference this case number if you open your own! :-D

Primo Indexing Limitations

MARC field-specific search and browse

- These features have garnered some interest with our members
- We're limited to 50 local search indexes in Primo, and none will pull in PCI resources. OTB indexes have fields grouped together.
- Browse indexes are significantly more limited in number and data source.

Page 37

Primo indexing limitations

There is interest in our consortium in MARC field-specific search and browse indexes. These are not currently supported by Primo due to limitation on number of locally-configurable search fields and lack of PCI integration, which we've mentioned earlier. Using OTB indexes to get the PCI results incorporated means we can't have a tag-specific index. Browse indexes are very limited in number and function.

We could use lateral-linking fields to deliver partial functionality in this regard, but there are some critical...

Lateral Linking Limitations

Using lateral links could provide an alternative to facets, but...

- Limited to 9 locally-defined fields.
- Functionality isn't **quite** the same as faceting (executes new search rather than filtering current result set)

Subjects: Trio sonatas (Flutes (2), continuo) -- Scores and parts
Genre/Form: Sonatas; Scores; Parts (Music); Parts (Music); Scores; Sonatas
Music Info: flute (2); continuo
Numeric designation: FWV N:e2 (Küntzel)
Key: E minor
OCLC Number: 913199264

Page 38

Hyperlinking (Lateral linking) field limitations

Current limit in Primo is 9 laterally-linked fields. With the expansion of display fields to 250 in number, we'd love to have additional fields set aside for lateral linking. (Maybe the 13X and 23X fields? Or could Primo support a lateral-linking "toggle" for every local field? Then institutions could decide field-by-field what creates a link and what doesn't.)

Lateral links provide functionality that's similar to faceting, but not equivalent (Executing new keyword/exact search rather than limiting current results). But institutions may prefer this to creating

facets for certain metadata fields.
(Depends on local user preference and institutional needs).

https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Back_Office_Guide/100Facets



Moving (and Looking) Forward

We've been very successful in using new metadata fields to set up facets for Alliance member institutions' use.

As time goes on and these newer fields are more widely used, the facets will prove more accurate and more helpful.

- Stats on field use are available at <http://experimental.worldcat.org/marcusage/>

We can review facet use with Primo Analytics.

Page 39

In summary, we've been able to achieve a LOT with Primo's norm rules to expose new, more granular metadata to users for faceting. This will help both nascent and advanced researchers locate resources that can advance their scholarly work.

The increasing use of the new MARC fields makes this kind of flexible innovation more important as time goes on – and more useful. OCLC has a longitudinal study of field use going on at <http://experimental.worldcat.org/marcusage>. Check it out!

Ultimately, we can use either Primo or Google Analytics to assess the use of “clickable” elements like facets and hyperlinked display fields in Primo.

Moving (and Looking) Forward

We're investigating additional facets on request.

- Specific relationship designators
- Would a general "relationship" facet be useful?

We have lots of questions about Primo VE!

- Will Primo normalization rules still exist?
 - Will Primo VE have the flexibility of Primo for translating metadata for faceting?
 - If not, we'll need to advocate for this.

Page 40

Next up for us:

Receiving/reviewing requests for new facets, display fields, and search indexes. (e.g. using specific designators, a general relationship facet)

Will Primo VE allow institutions to take advantage of new metadata through a customization structure that's as flexible as NRs? Early indications are that it won't – but VE is very much a work in progress. Stay tuned, and be ready to advocate for this if it's important to you!

Further Reading

ALA Subject Access Committee's A Brave New (Faceted) World: Towards Full Implementation of Library of Congress Faceted Vocabularies (2017)
<https://tinyurl.com/ybu2vwdx>

Library of Congress information page for LCDGT
<http://www.loc.gov/aba/publications/FreeLCDGT/freelcdgt.html>

McGrath, Kelley. Getting More out of MARC for Music and Movies with Primo: Strategies for Display, Search and Faceting (OLAC 2017 poster)
<https://tinyurl.com/ya9a2eoz>

McGrath, Kelley. Using MARC Facets for Music with Primo: Strategies and Challenges (ALA Midwinter 2018 presentation)
<http://connect.ala.org/node/273985>

Music Library Association Music Discovery Requirements (2017)
http://www.musiclibraryassoc.org/mpage/mdr_IA

Music Library Association LCMPT Best Practices <https://tinyurl.com/yaip7sux>

Music Library Association LCGFT for Music Best Practices
<https://tinyurl.com/ycxfce7r>

Page 41

Primo Sandbox with Facets

http://alliance-primo-sb.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vid=UO_NRWG

Page 42

- ALA Subject Access Committee's A Brave New (Faceted) World: Towards Full Implementation of Library of Congress Faceted Vocabularies (2017) <https://tinyurl.com/ybu2vwdx>
- Library of Congress information page for LCDGT
- <http://www.loc.gov/aba/publications/FreeLCDGT/freelcdgt.html>
- McGrath, Kelley. Getting More out of MARC for Music and Movies with Primo: Strategies for Display, Search and Faceting (OLAC 2017 poster) <https://tinyurl.com/ya9a2eoz>
- McGrath, Kelley. Using MARC Facets for Music with Primo: Strategies and Challenges (ALA Midwinter 2018 presentation)
- <http://connect.ala.org/node/273985>
- Music Library Association Music Discovery Requirements (2017) http://www.musiclibraryassoc.org/mpage/mdr_IA
- Music Library Association LCMPT Best Practices <https://tinyurl.com/yaip7sux>
- Music Library Association LCGFT for Music Best Practices <https://tinyurl.com/ycxfce7r>
- http://alliance-primo-sb.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vid=UO_NRWG