PCC Task Group on Linked Data Best Practices
Final Report

Submitted to PCC Policy Committee

September 12, 2019

The recommendations in this document are intended to provide guidance to catalogers, application developers, and vendors seeking to enhance their MARC bibliographic data with linked data URIs. The recommendations outline best practices in the following areas:

- Avoiding ambiguity when associating MARC data with URIs
- Parsing MARC data when supplying authority URIs
- Coding subfields $0, $1, and $4
- Providing work identifiers in the 758 field

Linked data is still an area of experimentation for the PCC, and the work of the Linked Data Best Practices Task Group is only one part of the PCC's larger exploration of its implications for cataloging practice. For this reason these recommendations do not address *all* of the issues involved in providing linked data URIs in MARC. Among the issues still under investigation by the PCC are questions concerning choice among different vocabularies and how to record associations among vocabularies.

The recommendations offered here will provide a basis for practical experimentation by the PCC URIs in MARC pilot due to start in late 2019. The pilot is expected to provide feedback that may result in adjustments to these recommendations.

The PCC Policy Committee welcomes comments and inquiries concerning this document. Please address any correspondence to coop@loc.gov.

Introduction

The PCC Task Group on Linked Data Best Practices was formed by the PCC Policy Committee (PoCo) in June 2018. The group was charged to consider questions of best practice arising from two previous PCC efforts: MARC cataloging issues that had been identified by the PCC Task Group on URIs in MARC, and BIBFRAME implementation issues that had been identified by the PCC BIBFRAME Task Group. The BIBFRAME issues were subsequently removed from the charge when PoCo decided that the PCC LD4P2 cohort would be a better avenue to explore them. What was left was the not insignificant task of carrying forward the URI Task Group's pursuit of what Richard Wallis has called "linky MARC" as a transition strategy to linked data. The original charge for the Task Group is available [here](#).

MARC is a legacy format that has developed by accretion over a long period. It should not come as a surprise that it has inconsistencies or that is not uniformly well suited to serve as a carrier for linked data. It should also be stressed that good practices for providing URIs in MARC data cannot be arrived at simply by reading the MARC definitions. Indeed, in some cases the Task Group

recommends against using URIs even where subfields $0 and $1 have been defined. It is hoped that these recommendations will help practitioners to maximize the very real gains from using URIs in MARC data and to avoid some of the pitfalls.

The Task Group undertook its work in what is clearly still an evolving environment. The recommendations offered here should be regarded as provisional and will need to be reviewed periodically in the light of continuing developments. In the course of its discussions the Task Group came to realize that it would have to accommodate the practical realities of available editing tools as well as the cataloging profession's relative inexperience in working with data sources produced by outside communities. Other drivers will also influence library practice in this area. For example, it seems likely that discussions concerning a target specification for the Library of Congress's planned BIBFRAME to MARC converter will lead to some reevaluation of MARC practice even for native MARC cataloging.

For the same reasons it has not always been possible in these recommendations to give outright instructions that reflect a settled consensus of practice. In many cases what the recommendations seek to do instead is to provide guidance on making sound choices from the available options. The Task Group hopes that these recommendations will help create a space where meaningful experimentation can begin.

Use cases and audience

A question the Task Group frequently asked itself was which use cases the data was expected to support. Production use cases included enrichment of legacy data by machine applications, and creation of new records by cataloging practitioners. Use cases for consumption of MARC data with URIs included conversion of MARC to BIBFRAME, conversion of MARC to other linked data models, or use of MARC data by external stakeholders such as Google. The requirements suited to each use case can vary, and this is reflected in the recommendations, which are sometimes qualified according to the purpose that is envisaged for the data. For example, if MARC records are produced for conversion to BIBFRAME, that will influence the choice of predicates to embed in the MARC data. The Task Group's recommendations for $4 take this consideration into account. There are also practical ramifications in the short term. If catalogers know that their MARC data will be passed through LC's BIBFRAME conversion program, they can safely assume that a MARC geographic, language, or relator code that they enter manually will be correctly transformed into its corresponding URI upon conversion. This makes it unnecessary in this particular scenario to go to the added effort to look up a URI.

The utility of identifiers for label maintenance came in for much debate within the Task Group. It is a significant motivation for libraries to adopt the use of identifiers in the short to medium term, although it can be expected to become less important as adoption of linked data technologies progresses. But, to support headings maintenance, the MARC record needs to be structured with repeating fields so that a label can be associated unambiguously with its corresponding identifier. This imposes a requirement that, while easily met in some scenarios, may present an unnecessary obstacle in others. Again the recommendations acknowledge the demands of differing use cases by allowing some flexibility in acceptable practice. This may be seen particularly in the recommendations given in the MARC object table (see below).

The intended audience for these recommendations, then, includes not only cataloging practitioners but also application developers and policymakers. This document is also intended as a resource for libraries evaluating requirements for tool development or specifications for enhancement of legacy data.

Recommendations

The Task Group's recommendations fall into three parts: a set of general guidelines that applies to coding in all MARC fields, an analysis identifying the object in each MARC field, and a set of recommendations for best practice in specific fields and subfields. In addition, the Task Group makes a recommendation to reconsider a number of existing PCC policies.

I. General guidelines on using URI subfields

1. The draft recommendations presented here are for the use of URI subfields in MARC records in a shared environment, such as OCLC, SkyRiver, etc. The Task Group recognizes that individual institutions may have needs that require them to diverge from these recommendations in their local environment, or in workflows involving specific partners (such as Google Books).

2. The identifier given in $1 (Real World Object URI) must always be a URI. In $0 (Authority record control number or standard number) prefer the URI over string identifiers where the editing environment makes it practicable to provide it without significant risk of error. Prefer the string value if the editing environment does not make it practicable to enter the URI, if legacy applications necessitate use of the string, or if no corresponding URI is available.

3. Be careful to observe the distinction between $0 and $1. $0 gives identifiers associated with preferred labels from authority files, while $1 refers to the entity described, and is usually agnostic with regard to a preferred label. This distinction is further explained in the PCC URI group's FAQ. Catalogers seeking to determine the appropriate coding for a specific source should consult the PCC Formulating URIs document.

4. Refer to an authoritative source, such as the PCC Formulating URIs document, or the maintenance agency's own documentation, to determine the canonical URI to give in $0, $1, or $4. Do not copy the URI from the browser address bar. Where feasible, it is recommended that practitioners acquire an RDF URI from a SPARQL lookup, or validate the URI using a validator such as http://linkeddata.uriburner.com:8000/vapour or Vafu http://vafu.redlink.io/.

5. For sources (e.g., Wikidata) that do not specify a single preferred label, generally use a label available from that source that suits the purposes of the cataloging agency, e.g., a label in the agency's preferred language.

II. Identifying the MARC object

The following observations should be read in conjunction with this Task Group document:
[MARC object table: a field-by-field analysis of the bibliographic format](#)

MARC makes provision for URIs in a large number of places, but in most cases it is not stated in the definitions how URI subfields relate to other data in the same fields. In a few cases the subfield corresponding to $0 is named: in 033 Date/Time and Place of an Event, for example, $0 is explicitly related to $p for Place. But more often the relevant subfields are not specified. Without this information it is difficult to set down best practices for populating URI subfields or to develop tools to reconcile MARC data with available URIs.

MARC also presents a number of structural difficulties for identifying the object. These vary according to the field and generally fall into one of the following categories:

Repeated object subfields. In the example below, the label cannot be associated with its corresponding URI, and as a consequence, the URI cannot be used to facilitate headings maintenance. However, provided that the labels are disregarded, these fields still support unambiguous conversion to RDF triples. The lack of a means to associate an identifier in $0 with a label in $a would also become immaterial in a workflow where URIs were used as the sole means of generating and updating labels.

> 380 ## $a Novels $a Thrillers (Fiction) $2 lcgft
> $0 http://id.loc.gov/authorities/genreForms/gf2015026020
> $0 http://id.loc.gov/authorities/genreForms/gf2014026571

> *Example showing form of work given as both novel and thriller, with associated URIs.*

Multiple subfields implying different predicates. In these cases the URIs in the MARC field will not support conversion to RDF, because each predicate cannot be associated with its proper object.

> 370 ## $c United States $g Paris (France) $2 naf
> $0 http://id.loc.gov/authorities/names/n78095330
> $0 http://id.loc.gov/authorities/names/n79058874

> *Example showing a work associated with the United States, created in Paris, France. The URIs in this example are problematical for conversion and should not be included.*

Subfields that are meaningful only when given in combination. Often in these cases the field expresses a complex statement that cannot be represented by a single URI even if $0 or $1 are defined for that field. Sometimes order of subfields is also significant. An example is the 382 Medium of Performance field.

> 382 01 $a flute $n 1 $d piccolo $n 1 $d alto flute $n 1 $d bass flute $n 1 $s 1
> $2 lcmpt

This example represents a work for solo flute, doubling piccolo, doubling alto flute, and doubling bass flute. No URI should be given.

The nonspecificity of MARC field definitions about the implied object and ambiguities such as those introduced by the possibility of multiple objects are critical issues for implementation of URIs in MARC. In order to address them, the Task Group created the MARC object table cited above. It enumerates the subfields that denote the object represented by individual fields in the MARC bibliographic format and recommends best practices for coding each field.

MARC object principles

In order to do this analysis it was necessary to set down some working assumptions or principles about how URIs relate to MARC data. Below is an outline of the principles that the Task Group adopted in undertaking this task. Additional semantics not explicitly stated in MARC or in PCC documentation may be inferred in the context of specific implementations, but should not be assumed to be generalizable.

1. Each MARC field for an access point refers to one object.

    100 1# $a Dicks, Terrance. $0 http://id.loc.gov/authorities/names/n78057783

2. URIs are not given for portions of access points. Subfields implying an object other than that specified by the field as a whole (e.g., a name within a name-title access point) should not be given a URI within the same field. While systems may be able to parse out elements of a string internally and associate them with distinct URIs, the MARC format cannot itself convey discrete associations within the same access point. Only URIs corresponding to the full access point should be communicated when MARC data is exported.

    In the example below, $0 must refer to the work, not to the composer.

    700 1# $a Beethoven, Ludwig van, $d 1770-1827. $t Veränderungen über einen Walzer $0 http://id.loc.gov/authorities/names/n81127885

    In the example below, $0 must refer to the entire subject heading string, not to one or more partial components.

    650 #0 $a Gardening $x Equipment and supplies $x Marketing $0 http://id.loc.gov/authorities/subjects/sh85053091

3. When $0 and $1 are given together, the authority given in $0 has the Real World Object (RWO) given in $1 as its focus.

    100 1# $a Obama, Michelle, $d 1964- $e author.
    $4 http://id.loc.gov/vocabulary/relators/aut
    $0 http://id.loc.gov/authorities/names/n2008054754 $1 http://viaf.org/viaf/81404344

4. URIs in MARC fields must permit predicates to be associated unambiguously with their objects.

    a. For fields with multiple objects, URIs are given only if the predicate is constant.

URI(s) can be given:

380 ## $a Novels $a Thrillers (Fiction) $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2015026020
$0 http://id.loc.gov/authorities/genreForms/gf2014026571

URI cannot be given:

370 ## $c United States $g Paris (France) $2 naf

However, the following is permissible; for more detail, see "Best practices for mitigation" *(below):*

370 ## $c United States $2 naf $0
http://id.loc.gov/authorities/names/n78095330
370 ## $g Paris (France) $2 naf $0
http://id.loc.gov/authorities/names/n79058874

    b. For fields with multiple predicates, there must be only one object URI.

700 1# $a Lloyd-Jones, Hugh, $e editor, $e translator
$4 http://rdaregistry.info/Elements/e/P20338
$4 http://rdaregistry.info/Elements/e/P20346
$0 http://id.loc.gov/authorities/names/n50050624

5. The order of subfields may be prescribed by the relevant content standard or community best practices but unless otherwise stated in the MARC definitions the position of $0, $1, or $4 relative to other subfields is not significant.

6. The values in the subfields giving the RDF object for each field give the access point corresponding to the $0 for that field. The relationship of these subfields to a RWO identified in $1 is looser, since the linked data vocabularies cited in $1 typically do not provide a single preferred label. For this reason, the MARC object table mentioned above specifically addresses only $0.

Best practices for mitigation

Ambiguous cases involving multiple subfields each implying a distinct object can often be addressed by the expedient of repeating the field. In the 370 Associated Place field, for example, the content-bearing subfields in the bibliographic format are $c (associated country),

$f (other associated place), and $g (place of origin of work or expression). Giving each element in a separate occurrence of the field makes it possible to associate a URI unambiguously with each element.

370 ## $c Great Britain $f Reading (England) $2 naf

can also be expressed as:

370 ## $c Great Britain $2 naf $0 http://id.loc.gov/authorities/names/n79072763
370 ## $f Reading (England) $2 naf $0 http://id.loc.gov/authorities/names/n79023147

This method can also be applied in cases where object subfields are repeated. This allows URIs to be associated with their corresponding labels for label maintenance purposes:

380 ## $a Novels $a Thrillers (Fiction) $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2015026020
$0 http://id.loc.gov/authorities/genreForms/gf2014026571

can also be expressed as:

380 ## $a Novels $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2015026020
380 ## $a Thrillers (Fiction) $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2014026571

It may be noted that it is already accepted practice to repeat fields where each subfield draws from a different source vocabulary coded in $2. In these cases, it is unproblematical also to add a URI subfield:

344 ## $a digital $2 rdatr $0 http://rdaregistry.info/termList/typeRec/1002
344 ## $b optical $2 rdarm $0 http://rdaregistry.info/termList/recMedium/1003
344 ## $g surround $2 rdacpc $0 http://rdaregistry.info/termList/configPlayback/1004

However, this method is not available in all cases. In cases where legacy MARC data is enriched by a machine reconciliation process, the application may not support breaking down the field into separate occurrences or it may not be possible to do so reliably. In these cases there may still be value in adding URIs to support subsequent conversion to RDF.

Some fields, like the 382 Medium of Performance field mentioned above, resist breakdown into separate occurrences each with a single object and cannot be given a URI.

MARC codes and URIs

In some places, MARC defines coded values for use with specific fields. Examples include the geographical area code in 043 $a, and the MARC language codes in 041 $a and 377 $a. These codes are internal to the MARC specification and have a

corresponding URI in id.loc gov. The same is true of the MARC relator codes in 1XX/7XX $4. The LC MARC to BIBFRAME converter can be expected to recognize these codes and derive the correct RDF statements from them.

For use cases where a MARC record is created for the purposes of eventual transformation into BIBFRAME, it is sufficient to provide the MARC code without also giving the corresponding id.loc.gov URI. In these cases it is also generally safe to continue to give multiple codes within the same field where the MARC definition permits the subfield to be repeated. See the MARC object table for recommendations concerning specific fields.

Examples:

043 ## $a f-sa--- $a f-tz--- $a f-za---

041 1# $a eng $k ger $h swe

A note on precoordinated subject strings

Precoordinated subject strings such as those in LCSH present special difficulties in linked data. The subject heading in its entirety is intended to represent a single unit of meaning that is not reducible to its parts. To say that the subject of a book is "Law $x Study and teaching $z Indonesia" is to say more than it is about law, study and teaching, and Indonesia: it's about the study and teaching *of* law *in* Indonesia.

But while the individual concepts are enumerated in the vocabulary, the combination (unless explicitly established) is not. There is no URI that can be given in a 6XX $0 to represent the subject as a whole. Even if MARC did have a mechanism for associating each of the component parts with its own URI, it would still not be able to represent the relation among the terms in linked data terms. That is not an incidental omission in MARC, but a reflection of the fact that linked data representations have a richness that MARC structures can only very partially capture.

The PCC Task Group on URIs in MARC considered this issue at some length and arrived at the following conclusion:

> *After careful consideration, the Task Group recommends against providing URIs that represent only partial entities of a MARC field. Faceted vocabularies provide an alternative means to represent such concepts by post-coordination. If the entire concept is to be represented as a single semantic unit within the LCSH vocabulary, in our view that becomes an issue for the maintenance agency rather than for implementers.*[1]

---

[1] PCC Task Group on URIs in MARC Year 2 Report to PoCo (October 2017). https://www.loc.gov/aba/pcc/bibframe/TaskGroups/PCC_URI_TG_20171015_Report.pdf

III. Guidelines for specific fields and subfields

a. Order of URI and non-URI subfields

Note: one of the operating principles adopted by the PCC URIs in MARC Task Group is that the order of subfields carries no meaning. The guidelines set down here are intended solely to promote legibility of MARC fields by human operators.

Give $2 and the URI subfields after the textual ones in the following preferred order: $2, $4, $0, $1.

100 1# $a Obama, Michelle, $d 1964- $e author.
$4 http://rdaregistry.info/Elements/w/object/P10061
$0 http://id.loc.gov/authorities/names/n2008054754 $1 http://viaf.org/viaf/81404344

370 ## $i Setting: $f Wyoming $2 naf $4 http://id.loc.gov/vocabulary/relators/stg
$0 http://id.loc.gov/authorities/names/n79022108

700 1# $i Motion picture adaptation of (work): $a Austen, Jane, $d 1775-1817. $t Sense and sensibility. $4 http://rdaregistry.info/Elements/u/P60227
$0 http://id.loc.gov/authorities/names/no2009127349
$1 http://www.wikidata.org/entity/Q274744

Where different occurrences of $4 contain both MARC relator codes and URIs, give occurrences with MARC relator codes before the occurrences with URIs.

100 1# $a Chee, Alexander. $4 aut $4 http://rdaregistry.info/Elements/a/object/P50195
$0 http://id.loc.gov/authorities/names/n2001037216

386 ## $i Composer: $a Germans $2 lcdgt $4 cmp
$4 http://id.loc.gov/vocabulary/relators/cmp
$0 http://id.loc.gov/authorities/demographicTerms/dg2015060340

Observe existing conventions for subfield order in non-URI subfields whether or not URI subfields are also given. Names follow the order for authorized access points. Resource-to-resource relationships are given in a $i preceding the authorized access point for the resource, while agent-to-resource relationships are given in $e (or X11 $j) following the authorized access point for the agent.

$3 and $5 (not shown in examples here) retain their current positions respectively at the beginning and end of fields.

Examples:

710 2# $a University of Texas at Austin. $b Department of Anthropology, $e sponsor.
$4 http://id.loc.gov/vocabulary/relators/spn
$0 http://id.loc.gov/authorities/names/n2005077004

111 2# $a International Conference on Life and Its Origin $n (2nd : $d 2004 : $c Rome,
Italy), $j author. $4 http://id.loc.gov/vocabulary/relators/aut
$0 http://id.loc.gov/authorities/names/n2016246370
$1 http://viaf.org/viaf/183147118127126340376

700 1# $i Paraphrase of (work): $a Tippett, Michael, $d 1905-1998.$t Mask of time.
$4 http://rdaregistry.info/Elements/u/P60296
$0 http://id.loc.gov/authorities/names/no97078111

Note that the URI in $0 of the third example refers to the name/title authority, not to the name alone.

Catalogers should not feel compelled to reorder existing subfields.

b. MARC subfields $0 and $1

MARC proposal 2017-08 introduced a distinction between traditional identifiers and RWO URIs. A URI given in $0 usually refers to an authority file or controlled vocabulary, which typically serves the primary function of providing a preferred label for an entity. Pairing the authority URI with the label helps the identifier to perform its primary function of supporting the label. By contrast, $1 URIs refer directly to entities and are drawn from linked data sources that typically do not support a single preferred label.

Although $0 and $1 are defined in parallel throughout most of the MARC format, this difference in purpose has implications for the application of those subfields in library cataloging. These guidelines address the differing treatment needed for $0 and $1 in cataloging practice.

The guidelines reflect the realities of the current cataloging ecosystem. One of the main issues the group needed to consider was what to do in cases where several identifiers were available for the same entity. The prevailing sentiment in the task group was that mappings made among URIs - for example, between a URI from an authority file and a RWO URI from a non-library source - should be done at the vocabulary level rather than in the context of specific attributions made in a bibliographic record. This approach would also remove the need to assert equivalency among concepts in different vocabularies on a field-by-field basis. In reality, however, workflows for accomplishing this have yet to be widely incorporated in library cataloging practice. Some libraries have begun populating 024 in the authority format as a means of recording these equivalencies, while others see promise in external platforms, notably Wikidata, to serve as identifier hubs. (See the Identifiers section of the Wikidata entry for Ai Weiwei for an example.) This will be an important area for libraries to explore but best practice recommendations for this type of work are outside the scope of this report.

Users should consult the PCC Formulating URIs and PCC URI FAQ documents for guidance on choice of subfield ($0 or $1) and URI syntax for sources in general use by the library community. Users should be careful to use the canonical URI that designates the entity or authority, not simply the address for a web page.

For application of the following guidelines to specific fields, consult the MARC object table.

Subfield $0

1. The $0 should refer to the entire preferred label used in an access point.

2. Provide only one $0 containing a URI for each MARC object. This should be the URI for the authority providing the preferred label.

3. A $0 containing a URI may be given in addition to, or instead of, a $0 containing a non-URI identifier. If both a URI and non-URI identifier are given, they should identify the same authority and be from the same source.

4. For MARC fields taking multiple objects:
   a. It is *encouraged* to repeat the field for each $0 as long as context is not lost by doing so;
   b. It is *permissible* to give multiple $0 if the predicate is constant;
   c. It is *not permissible* to give multiple $0 if the field contains both multiple predicates and multiple objects.

Subfield $1

1. $1 containing a RWO URI may be given instead of, or in addition to, a $0 associated with the preferred label used in that field.

2. Since $1 is not generally associated with a preferred label, $1 may be repeated within the same occurrence of a MARC access point field, provided that the URIs given refer to the same RWO.

3. Generally, if a $1 contains a URI known to be mapped in a widely used external service (e.g. VIAF, Wikidata), additional URI(s) that are also found in those services need not be added.

   Examples

   100 1# $a Mitchell, Joni, $e composer.
   $0 http://id.loc.gov/authorities/names/n82108794

   100 1# $a Mitchell, Joni, $e composer. $1 http://id.loc.gov/rwo/agents/n82108794

100 1# $a Mitchell, Joni, $e composer. $1 http://isni.org/isni/0000000114765598

100 1# $a Mitchell, Joni, $e composer.
$0 http://id.loc.gov/authorities/names/n82108794
$1 http://id.loc.gov/rwo/agents/n82108794
$1 http://www.wikidata.org/entity/Q205721

100 1# $a Lin, Maya, $e architect. $2 ulan $0 (gettyulan)500001306
$0 http://vocab.getty.edu/ulan/500001306
$1 http://vocab.getty.edu/ulan/500001306-agent

111 2# $a International Computer Music Conference, $j author.
$0 http://id.loc.gov/authorities/names/n82224320
$1 http://www.wikidata.org/entity/Q4288208

257 ## $a France $2 naf $1 http://vocab.getty.edu/tgn/1000070-place
$1 http://www.bbc.co.uk/things/61ef4416-de68-49ff-9c97-e0779dafd9d2#id
257 ## $a Italy $2 naf $1 http://vocab.getty.edu/tgn/1000080-place
$1 http://www.bbc.co.uk/things/0021de37-b64a-46ac-a4bb-5bdbdf0908ec#id

340 ## $n large print $2 rdafs $0 http://rdaregistry.info/termList/fontSize/1002

344 ## $a digital $2 rdatr $0 http://id.loc.gov/vocabulary/mrectype/digital
344 ## $b optical $2 rdarm $0 http://id.loc.gov/vocabulary/mrecmedium/opt
344 ## $g surround $2 rdacpc $0 http://id.loc.gov/vocabulary/mplayback/mul

347 ## $a video file $2 rdaft $0 http://id.loc.gov/vocabulary/mfiletype/video

348 ## $a score $2 rdafnm $0 http://rdaregistry.info/termList/formatNoteMus/1007
348 ## $a part $2 rdafnm $0 http://rdaregistry.info/termList/formatNoteMus/1004

370 ## $i Setting: $f Grand Canyon (Ariz.) $2 lcsh $0
http://id.loc.gov/authorities/subjects/sh85056381
$1 http://www.wikidata.org/entity/Q118841

380 ## $a Popular music $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2014027009
$1 http://dbpedia.org/resource/Pop_music

385 ## $a Teenagers $2 lcdgt
$0 http://id.loc.gov/authorities/demographicTerms/dg2015060011

386 ## $a African Americans $2 lcdgt
$0 http://id.loc.gov/authorities/demographicTerms/dg2015060362
386 ## $a Women $2 lcdgt
$0 http://id.loc.gov/authorities/demographicTerms/dg2015060358

386 ## $a Sociologists $2 lcdgt
$0 http://id.loc.gov/authorities/demographicTerms/dg2016060033

600 00 $a Demeter $c (Greek deity) $x Cult.
$0 http://id.loc.gov/authorities/subjects/sh92002765

611 27 $a Olympic Games. $2 fast $0 (OCoLC)fst01408249
$0 http://id.worldcat.org/fast/1408249

630 07 $a Lord of the rings (Tolkien, J. R. R.) $2 fast
$0 http://id.worldcat.org/fast/1356106

650 #0 $a Animal welfare $x Religious aspects $x Buddhism.
$0 http://id.loc.gov/authorities/subjects/sh2019001155

650 #2 $a Measles $x epidemiology. $0 (DNLM)D008457Q000453
$0 http://id.nlm.nih.gov/mesh/D008457Q000453

650 #7 $a Dollhouses $x Collectors and collecting. $2 fast
$0 (OCoLC)fst00896450 $0 http://id.worldcat.org/fast/896450

651 #0 $a London (England) $x Social life and customs $y 19th century.
$0 http://id.loc.gov/authorities/subjects/sh85078225

655 #7 $a Zombie fiction. $2 lcgft
$0 http://id.loc.gov/authorities/genreForms/gf2018026019

655 #7 $a Life simulation video games. $2 olacvggt
$0 http://metadataregistry.org/uri/olac/1041

655 #7 $a Transcripts. $2 aat $0 http://vocab.getty.edu/aat/300027388

700 1# $i Television adaptation of (work): $a Maupin, Armistead. $t Tales of the city
$0 http://id.loc.gov/authorities/names/no2014054037
$1 http://viaf.org/viaf/308242532 $1 http://www.wikidata.org/entity/Q7679391
$1 http://worldcat.org/entity/work/id/1162587

710 2# $a I.M. Pei & Partners. $0 http://id.loc.gov/authorities/names/n79036327
$1 http://vocab.getty.edu/ulan/500033103-agent

710 2# $a Berliner Philharmoniker. $4 prf
$0 http://id.loc.gov/authorities/names/n81018318
$1 http://dbpedia.org/resource/Berlin_Philharmonic

730 0# $i Remake of (work): $a Lady for a day (Motion picture : 1933)
$0 http://id.loc.gov/authorities/names/no2016087749

$1 http://www.wikidata.org/entity/Q301649

$1 http://dbpedia.org/resource/Lady_for_a_Day

830 #0 $a Occasional papers of the California Academy of Sciences ; $v no. 161.
$0 http://id.loc.gov/authorities/names/n42034729 $1 http://viaf.org/viaf/183872531

Note that $0 and $1 URIs in 8XX series access points apply only to the title of the series, not also to the numbering.

c. MARC subfield $4

The MARC Advisory Committee (MAC) approved a proposal from the British Library in 2017 to redefine $4 in various MARC fields to accommodate not only the MARC relator codes that had previously been given in this subfield, but also URIs designating relationships. Typically these URIs would correspond to relationships expressed textually in $e or $i (or, in 111/711, $j) or as codes in $4. The Task Group on Linked Data Best Practices (LDBP) was charged with examining this and other issues outlined by the PCC Task Group on URIs in MARC in order to recommend best practices. The expanded definition of $4 raises a number of issues for best practice.

Please note that the recommendations below do not take into account the changes being made to RDA under the 3R project.

If approved, the following recommendations will entail several changes to existing PCC policy. These changes are noted at the end of the document.

1. The RDA relationship designator/element and the MARC relator vocabularies are both acceptable for use in PCC records.

   Reasoning: The PCC has not yet developed criteria to guide the selection of controlled vocabularies, or adopted an overall position on whether to prefer the RDA vocabularies or their MARC counterparts. The Linked Data Advisory Committee has initiated a discussion of the former issue, while the PCC BIBFRAME groups have made recommendations bearing on the latter. LDBP feels it is premature to make an unequivocal recommendation until those discussions are further advanced.

   Note that MARC currently defines codes only for agent-to-resource and not for resource-to-resource relationships.

   Example:

   245 14 The Iliad and the Odyssey of Homer / $c translated by Richmond Lattimore.

   Options:

   700 1# $a Lattimore, Richmond. $4 trl

700 1# $a Lattimore, Richmond. $4 http://rdaregistry.info/Elements/u/P60385
700 1# $a Lattimore, Richmond. $4 http://id.loc.gov/vocabulary/relators/trl
700 1# $a Lattimore, Richmond. $4 trl $4 http://rdaregistry.info/Elements/u/P60385
700 1# $a Lattimore, Richmond. $4 trl $4 http://id.loc.gov/vocabulary/relators/trl
700 1# $a Lattimore, Richmond. $4 http://rdaregistry.info/Elements/u/P60385
$4 http://id.loc.gov/vocabulary/relators/trl
700 1# $a Lattimore, Richmond. $4 trl $4 http://rdaregistry.info/Elements/u/P60385
$4 http://id.loc.gov/vocabulary/relators/trl

2. For use cases where the MARC relator rather than RDA relationship designators are desired, and the MARC record is produced for purposes of eventual conversion to BIBFRAME, use the MARC relator code in preference to the corresponding URI.

   Reasoning: These codes are internal to the MARC specification and have a corresponding URI in id.loc gov. The LC MARC to BIBFRAME convertor can be expected to recognize these codes and derive the correct RDF statements from them. Other conversion programs may also be able to take these mappings into account.

   Example:

   245 14 The Iliad and the Odyssey of Homer / $c translated by Richmond Lattimore.

   Option:

   700 1# $a Lattimore, Richmond. $4 trl

3. If a relationship URI or code is given, it is encouraged but not required to also provide the corresponding label ($e, $i, or $j depending on the field).

   Examples:

   700 1# $a Lloyd-Jones, Hugh, $e translator. $4 trl
   $4 http://rdaregistry.info/Elements/u/P60385

   700 1# $a Lloyd-Jones, Hugh, $e editor. $4 edt

   700 1# $i Translation of: $a Galilei, Galileo, $d 1564-1642. $t Dialogo dei massimi sistemi. $4 http://rdaregistry.info/Elements/u/P60244

4. The RDA relationships exist in constrained and unconstrained versions. If RDA relationships are used, PCC recommends using the constrained versions only in cases where it can readily be determined that the entities are consistent with constrained RDA domain and range definitions. In cases where compliance with RDA constraints is problematic, cannot be evaluated, or is otherwise in doubt, the unconstrained versions

should be used. In general, MARC data created with the objective of eventual conversion to BIBFRAME should use the unconstrained properties.

5. Multiple relationship designators (whether expressed as text, codes, or URIs) may be given in the same occurrence of a field provided that they apply to the same RDF subject and object. In cases of multiple objects, repeat the field. If textual relationship designators and their equivalent URIs are both given, it is not necessary to relate them through order of subfields. In addition, there is no expectation that any occurrence of $4 will have a corresponding value in $e or $j, or vice-versa.

Reasoning: MARC does not provide a way to relate an occurrence of $4 to a given occurrence of $e/$j, $i, or another $4 within the same field. However, as long as the field designates the same object, this is immaterial for conversion purposes.

Example:

245 10 Sophocles : $b Antigone, The women of Trachis, Philoctetes, Oedipus at Colonus / $c edited and translated by Hugh Lloyd-Jones.

Some options (all equally valid):

700 1# $a Lloyd-Jones, Hugh. $4 edt $4 trl
$4 http://rdaregistry.info/Elements/u/P60185
$4 http://rdaregistry.info/Elements/u/P60385

700 1# $a Lloyd-Jones, Hugh. $4 edt $4 trl
$4 http://id.loc.gov/vocabulary/relators/edt $4 http://id.loc.gov/vocabulary/relators/trl

700 1# $a Lloyd-Jones, Hugh, $e editor, $e translator.
$4 http://rdaregistry.info/Elements/u/P60185
$4 http://rdaregistry.info/Elements/u/P60385

700 1# $a Lloyd-Jones, Hugh, $e editor, $e translator. $4 edt $4 trl
$4 http://id.loc.gov/vocabulary/relators/edt $4 http://id.loc.gov/vocabulary/relators/trl

Examples:

245 10 $a Galileo on the world systems : $b a new abridged translation and guide / $c Maurice A. Finocchiaro.
700 1# $i Translation of: $a Galilei, Galileo, $d 1564-1642. $t Dialogo dei massimi sistemi. $4 http://rdaregistry.info/Elements/u/P60244
700 1# $i Abridgement of (work): $a Galilei, Galileo, $d 1564-1642. $t Dialogo dei massimi sistemi. $4 http://rdaregistry.info/Elements/u/P60223
700 1# $i Guide to (work): $a Galilei, Galileo, $d 1564-1642. $t Dialogo dei massimi sistemi. $4 http://rdaregistry.info/Elements/u/P60252

Repeated field with single $i and $4 in each; relationship designators and constrained URIs from RDA Registry

245 10 $a Galileo on the world systems : $b a new abridged translation and guide / $c Maurice A. Finocchiaro.
700 1# $i Translation of, $i Abridgement of (work), $i Guide to (work): $a Galilei, Galileo, $d 1564-1642. $t Dialogo dei massimi sistemi.
$4 http://rdaregistry.info/Elements/u/P60244
$4 http://rdaregistry.info/Elements/u/P60223
$4 http://rdaregistry.info/Elements/u/P60252

Single field with repeated $i and $4; relationship designators and constrained URIs from RDA Registry

386 ## $i Composer: $a French $2 lcdgt $4 cmp
$4 http://id.loc.gov/vocabulary/relators/cmp
386 ## $i Arranger: $a French $2 lcdgt $4 arr
$4 http://id.loc.gov/vocabulary/relators/arr

Repeated field with single $i and multiple $4; relator terms, codes, and URIs from MARC

386 ## $i Composer, $i Arranger: $a French $2 lcdgt $4 cmp $4 arr
$4 http://id.loc.gov/vocabulary/relators/cmp
$4 http://id.loc.gov/vocabulary/relators/arr

Single field with repeated $i and multiple $4; relator terms, codes, and URIs from MARC

386 ## $i Composer, $i Arranger: $a French $2 lcdgt
$4 http://id.loc.gov/vocabulary/relators/cmp
$4 http://id.loc.gov/vocabulary/relators/arr

Single field with repeated $i and multiple $4; MARC relator URIs and labels given

d. MARC field 758: work or resource identifier

MAC approved a proposal from the PCC URI Task Group in 2017 for a new field in the bibliographic format to accommodate resource identifiers. The MAC review noted that "the application of field 758 would need to be developed as a matter of best practice by the community". PoCo assigned the development of these best practices to the Linked Data Best Practices Task Group as one of the main items on its charge.

It is important to note that the field is agnostic with regard to the data model of the entity that is referenced. In particular, the URI given in 758 need not be for a FRBR or LRM work.

FRBR expression entities may be given in this field as well as FRBR work entities; but so may resource entities that do not conform to the FRBR model, such as BIBFRAME works.

The field will often be populated with identifiers generated by a machine process, but it can also accommodate identifiers associated with traditional access points. The recommendations below address both possibilities.

Two main kinds of use cases were presented in the MAC discussion paper and subsequent proposal. The first was instance-to-resource relationships, sometimes referred to (especially in an RDA context) as "primary" relationships. This was the use case that initially motivated the proposal. The second was resource-to-resource relationships, i.e. relationships of the kind given in Appendix J of the pre-3R RDA Toolkit. The Task Group recommends that implementation of 758 for resource-to-resource relationships be deferred until the PCC community gains experience with its use for instance-to-resource relationships. The Task Group's recommendations therefore address only instance-to-resource relationships.

1. Always give a value in $4 specifying the predicate to be used.

    For records intended for conversion to BIBFRAME, the bf:instanceOf property (http://id.loc.gov/ontologies/bibframe/instanceOf) may be used. In RDA the "has expression manifested" (http://rdaregistry.info/Elements/m/P30139) and "has work manifested" (http://rdaregistry.info/Elements/m/P30135) properties and their corresponding object and datatype subproperties serve an analogous purpose. However, the Task Group recommends using RDA relationships only in cases where it can readily be determined that the resource entity meets the applicable RDA range definition.

    Current LC-PCC policy (LC-PCC PS 17.0) is not to record primary relationships. The Task Group recommends changing this policy.

    Examples:

    100 1# $a Wasserstein, Wendy, $e author. $0 http://id.loc.gov/authorities/names/n79023585
    245 14 $a The Heidi chronicles : $b a play / $c by Wendy Wasserstein.
    758 ## $a The Heidi Chronicles $4 http://id.loc.gov/ontologies/bibframe/instanceOf $1 http://www.wikidata.org/entity/Q7739301

    Example above shows bf:instanceOf predicate. Note that the label in 758 $a is taken from Wikidata.

    100 1# $a Rooney, Sally, $e author. $0 http://id.loc.gov/authorities/names/no2017108521
    240 10 $a Conversations with friends. $l Spanish
    245 10 $a Conversaciones entre amigos / $c Sally Rooney.

758 ## $i Has expression manifested: $a Rooney, Sally. Conversations with friends. Spanish $4 http://rdaregistry.info/Elements/m/P30139 $0 http://id.loc.gov/authorities/names/no2018130067

Example above shows RDA "has expression manifested" predicate.

2. Generally do not give a 758 field without including either a $0 or a $1.

3. It is permitted, but not required, to give a label.

    Example:

    100 1# $a Lee, Harper, $e author. $4 http://id.loc.gov/vocabulary/relators/aut
    $1 http://www.wikidata.org/entity/Q182658
    245 10 $a Go set a watchman / $c Harper Lee.
    758 ## $4 http://id.loc.gov/ontologies/bibframe/instanceOf
    $1 http://worldcat.org/entity/work/id/2283978583

4. When identifying resources that stand in a primary relationship with the instance being described, give all the relevant identifiers in 758 regardless of whether they also correspond to an access point also present in the record. Optionally, give the identifier in $0 or $1 of the field containing the access point as well (e.g., to support authority control use cases).

    Examples:

    130 0# $a Beowulf. $l English $s (Osborn)
    $0 http://id.loc.gov/authorities/names/no2014100183
    245 10 $a Beowulf : $b a verse translation with treasures of the ancient North / $c by Marijane Osborn ; with an introduction by Fred C. Robinson.
    758 ## $a Beowulf. English (Osborn) $4 http://id.loc.gov/ontologies/bibframe/instanceOf
    $0 http://id.loc.gov/authorities/names/no2014100183 $1 http://viaf.org/viaf/309855255

    The 130 $0 given in the example above is optional.

    110 1# $a Brazil, $e enacting jurisdiction. $4 http://id.loc.gov/vocabulary/relators/enj
    240 10 $a Constituição (1988) $0 http://id.loc.gov/authorities/names/n88279508
    245 10 $a Constituição da República Federativa do Brasil : $b atualizada até a Emenda Constitucional no 95, de 15/12/2016 : novo regime fiscal / $c organização, José Luiz Tuffani de Carvalho.
    758 ## $4 http://id.loc.gov/ontologies/bibframe/instanceOf
    $0 http://id.loc.gov/authorities/names/n88279508

    The 240 $0 given in the example above is optional.

5. If asserting relationships to more than one resource entity, repeat the 758 field.

Examples:

100 1# $a Irving, Washington, $d 1783-1859, $e author.
$1 http://id.loc.gov/rwo/agents/n79005645
245 10 $a Tales of the Alhambra = $b Cuentos de la Alhambra : versión bilingüe inglés español / $c Washington Irving ; translated by Ana Merino.
700 12 $i Container of (work): $a Irving, Washington, $d 1783-1859. $t Alhambra.
$0 http://id.loc.gov/authorities/names/no2019118835
700 12 $i Container of (expression): $a Irving, Washington, $d 1783-1859. $t Alhambra.
$l Spanish $s (Merino) $0 http://id.loc.gov/authorities/names/no2018159769
758 ## $i Has work manifested: $a Irving, Washington, 1783-1859. Alhambra
$4 http://rdaregistry.info/Elements/m/P30135
$0 http://id.loc.gov/authorities/names/no2019118835
758 ## $i Has expression manifested: $a Irving, Washington, 1783-1859. Alhambra.
Spanish (Merino) $4 http://rdaregistry.info/Elements/m/P30139
$0 http://id.loc.gov/authorities/names/no2018159769

100 1# $a Williams, Tennessee, $d 1911-1983, $e author.
$4 http://rdaregistry.info/Elements/w/P10436 $1 http://id.loc.gov/rwo/agents/n79090096
240 10 $a Plays. $k Selections (2012)
245 10 $a Orpheus descending ; $b and, Suddenly last summer / $c Tennessee Williams ; introductions by Martin Sherman.
700 12 $i Container of (work): $a Williams, Tennessee, $d 1911-1983. $t Orpheus descending. $0 http://id.loc.gov/authorities/names/no2019123405
700 12 $i Container of (work): $a Williams, Tennessee, $d 1911-1983. $t Suddenly last summer. $0 http://id.loc.gov/authorities/names/no2019123048
758 ## $4 http://rdaregistry.info/Elements/m/P30135
$0 http://id.loc.gov/authorities/names/no2014024667
758 ## $4 http://rdaregistry.info/Elements/m/P30135
$0 http://id.loc.gov/authorities/names/no2019123405
758 ## $4 http://rdaregistry.info/Elements/m/P30135
$0 http://id.loc.gov/authorities/names/no2019123048

The first 758 is for the compilation. The second is for Orpheus descending*, and the third is for* Suddenly last summer*.*

100 1# $a Mussorgsky, Modest Petrovich, $d 1839-1881, $e composer. $4 cmp $4 http://id.loc.gov/vocabulary/relators/cmp
$0 http://id.loc.gov/authorities/names/n79127149
$1 http://id.loc.gov/rwo/agents/n79127149
240 10 $a Kartinki s vystavki; $o arranged $s (Ravel)
$0 http://id.loc.gov/authorities/names/no2019124048
245 10 $a Pictures at an exhibition = $b Tableaux d'une exposition = Bilder einer Ausstellung / $c Moussorgsky ; [orchestrated by Maurice] Ravel.

758 ## $i Has work expressed: $a Mussorgsky, Modest Petrovich, 1839-1881. Kartinki s vystavki $4 http://rdaregistry.info/Elements/e/P20231 $0 http://id.loc.gov/authorities/names/n80149485

758 ## $i Has expression manifested: $a Mussorgsky, Modest Petrovich, 1839-1881. Kartinki s vystavki; arranged (Ravel) $4 http://rdaregistry.info/Elements/m/P30139 $0 http://id.loc.gov/authorities/names/no2019124048

IV. Recommendations for review of existing PCC policies

The Linked Data Best Practices group recommends the following changes to PCC policy in bibliographic records:

Relationship designators

- Multiple occurrences of $i for resource-to-resource relationships in access point fields should be permitted. Current PCC policy is to repeat the field rather than the subfield. By contrast, PCC guidelines do permit repetition of $e to accommodate multiple agent-to-resource relationships. The Task Group does not see a clear justification for this divergence. This change of policy may necessitate changes to existing punctuation guidelines; these are not addressed here.

- There should be no prescribed order for $e/$j relationships in name fields. Current policy is to give them in WEMI order. However, the Task Group feels that this prescription places an unnecessary burden on catalogers, and the order will in any case be immaterial when MARC data is converted to RDF.

- Current policy is to indicate relationships using text in $i or $e/$j rather than through codes or URIs in $4. We propose that codes, URIs, or text all be permitted, and in any combination. In an ideal systems environment the metadata editor would enable URIs to be added easily and the discovery system would be able to obtain the corresponding labels for indexing and display. However, at present system capabilities vary greatly, and the Task Group considers that while use of URIs should be encouraged, cataloging practices also have to acknowledge the limitations of existing systems.

Primary relationships

- Current LC-PCC policy (LC-PCC PS 17.0) is not to record primary relationships. The Task Group strongly recommends changing this policy so that instance-to-resource relationships can be stated explicitly, as is necessary in the 758 field.

Issues not addressed

The Task Group's recommendations do not address issues in the following areas:

- *Use of URIs in authority records.* The use of URIs was not endorsed for NACO at the time the group was formed. Current LC guidelines say to "consult LC's Cooperative Programs Section before using subfield $0 or $1". The Task Group did not feel that it had the necessary communication channels in place to resolve this important issue and recommends that PoCo charge a group to pursue it.

- *The RDF subject represented by a MARC bibliographic record.* Although the implied subject of a field in a MARC bibliographic record is typically an instance or manifestation, this is not invariably the case. The MARC format does not currently provide the means to specify the subject of a statement explicitly. This limitation is probably intractable without introducing very radical changes of practice. For this reason the Task Group did not attempt to address it.

- *024 in authority records.* PoCo referred this issue, including best practices for use of the new $0 and $1 subfields, to a group convened by Paul Frank with representation from the Linked Data Advisory Committee, the Identity Management in NACO Task Group, and the Standing Committee on Standards. At the time of writing this group is planning a pilot project that will include examination of this issue.

- *$2 in name fields.* This MARC proposal had not been approved at the time the group was charged, but was approved by MAC in January 2019 and by the MARC Steering Group in March 2019. Implementation raises a number of issues that would require further examination, such as review of the MARC source code lists. This additional work was not considered practicable for this group within the available time frame.

- *Further MARC issues.* Apart from a successful fast track proposal to align the definition of $0 in 516 with the related one in 033, the Task Group considered MARC proposals to be out of scope for its work. However, one issue that arose in the Task Group''s discussions was the possibility of expanding the definition of $2 to accommodate URIs as well as MARC codes. The Task Group noted that the MARC source codes defined for $2 now have URI counterparts in id.loc.gov. The Task Group recommends that PCC task a suitable group to investigate such a proposal.

- *Endorsement and choice of vocabularies for PCC use.* The Linked Data Advisory Committee is leading exploration of this important topic.

Task Group Members

Chew Chiat Naun (Harvard University) (Co-chair)
Jennifer Baxmeyer (Princeton University) (Co-chair)
Bryan Baldus (OCLC)
Christine DeZelar-Tiedman (University of Minnesota)
Erin Dobias (Google)
Nancy Fallgren (National Library of Medicine (NLM))
MJ Han (University of Illinois at Urbana-Champaign)
Stephen Hearn (University of Minnesota)

Nancy Lorimer (Stanford University)
Honor Moody (Harvard University)
Terry Reese (Ohio State University)
Adam Schiff (University of Washington)
Jackie Shieh (Smithsonian Libraries)
Jodi Williamschen (Library of Congress (LC))

Acknowledgements