Everything You Wanted to Know About "Linky MARC"*

*And Were Definitely *Not* Afraid to Ask, Being Librarians

Abby Dover
Linked Data Librarian
Northwestern University

MARC to Linked Data: More Possibilities - Pt. 2 of
MARC Formats Transition IG session, March 15, 2024

What is "linky MARC"?

- Linky MARC = MARC records that contain http URIs (Uniform Resource Identifiers) for entities
- Coined by Richard Wallis
- Uses the $0 and $1 subfields – not to be confused with URLs in $u

This is linky MARC:

380 ## $a Novels $2 lcgft $0 http://id.loc.gov/authorities/genreForms/gf2015026020 $1 http://www.wikidata.org/entity/Q8261

Not this:

856 42 $3 Essays from annual reports $u http://woodrow.mpls.frb.fed.us/pubs/ar/index.html

What is linky MARC for?

Linky MARC is *not* linked data. Rather, it's a (temporary) bridge – a hybrid form meant as an interim solution until we can fully transition to linked data-land.

Linky MARC can...

- Help with headings maintenance
- Prepare for conversion to RDF/BIBFRAME
- Bring external data into the discovery environment

What's the difference between $0 and $1?

$0 = Authority record control number or standard number

- Authorities are **documents describing a thing, not the thing itself**
    - Has a preferred label
    - Contains metadata about the heading, such as provenance, revision history, etc.
- Can be expressed either as a control number (n 80017868) or URI (http://id.loc.gov/authorities/names/n80017868)
- Sources tend to be published by or endorsed by library standards organizations
- Mix of legacy and linked data encodings

Examples of sources that use $0:

LCSH

LCGFT

LCDGT

LCNAF published as authorities (URIs styled like: http://id.loc.gov/authorities/names/n50024879)

AAT - Art & Architecture Thesaurus

RBMS Controlled Vocabularies

MeSH (Medical Subject Headings) RDF

What's the difference between $0 and $1?

$1 = Real World Object URI

- A Real World Object is **the thing that an authority describes**
- Are generally label-agnostic
- Specialized for linked data applications and formulated according to linked data conventions
  - globally unique
  - persistent
  - resolves to RDF-encoded description
- Source could be from the library community or third-party
- Approved for use after $0

Examples of sources that use $1:
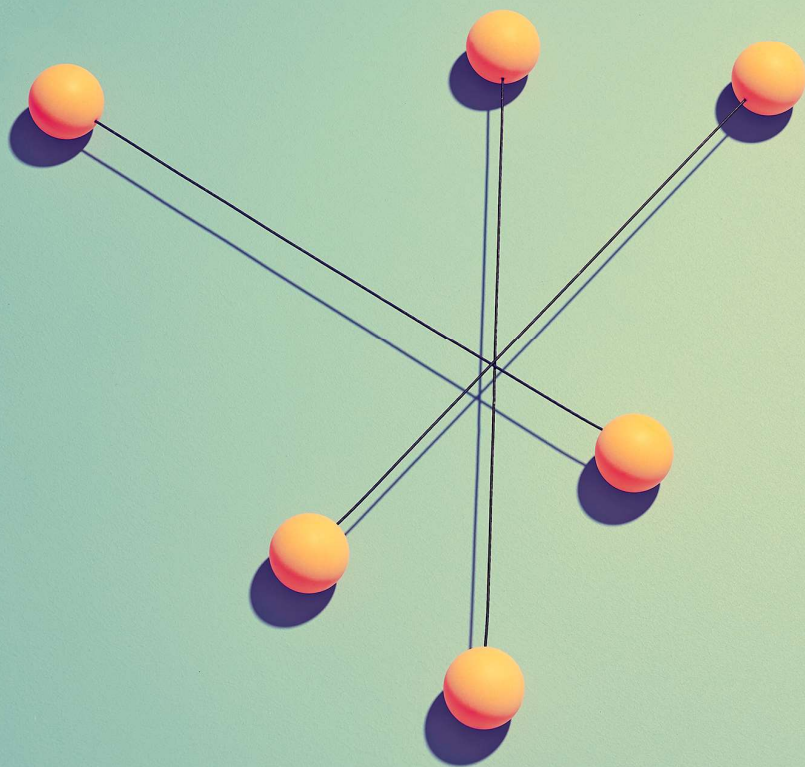
    Homosaurus
    VIAF
    Wikidata
    ORCID
    ISNI
    GeoNames
    WorldCat Entities
    LCNAF published as Real World Objects (URIs styled like:
    http://id.loc.gov/rwo/agents/n50024879)

What are some guidelines for using linky MARC?

Three key reference documents (even if you're not a PCC library):

- PCC Task Group on Linked Data Best Practices Final Report (2019)

- URI FAQ (2018, revised 2023)

- Formulating and Obtaining URIs Guide (2020)

What are some guidelines for using linky MARC?

Best practices:

- NACO authority records are not meant as hubs or URI reconciliation services, so it is not required to add URIs in the 024 field (But if you do, keep to 5 or less and only add if providing additional info)

- URIs should not be given for portions of access points – must refer to the *entire* string

- No limit to the number of URIs strictly speaking, but in practice, keep to one $0 and/or one $1

- Should limit the number of sources to 1 or 2 – mapping equivalencies across vocabularies should not be done at the bib record level. Using external "hubs" like Wikidata is an option.

- Formulating a URI isn't a matter of just copy-pasting the URL – consult a guide, don't try to wing it

Where can I get some extra help?

Tools

- Vapour (linked data validator) http://linkeddata.uriburner.com:8000/vapour

- Vafu http://vafu.redlink.io/

- MARCEdit MARCNext

- LOD/OpenRefine

- Custom SPARQL scripts

- Gary Strawn's Authority Toolkit: newer versions have URI functionality

Vendors

- MARCIVE, Backstage, and other library vendors have services that will add URIs to MARC records

Why have I been seeing so many more URIs in WorldCat lately?

- As of December 2023, OCLC started adding WorldCat Entities URIs for persons, places, and events to the 100, 600, 647, 651, 700 or 758* fields, and expanded to URIs for Works in late January 2024.

- There's currently an "Insert WorldCat Entity" feature in Record Manager; Connexion will receive this functionality later this year.

- Why WorldCat Entities and not a "hub" like Wikidata? – WorldCat Entities is associated with your libraries' actual holdings.

- Note that in fields that can be controlled by OCLC Connexion software (1XX, 6XX, 7XX, 8XX), $0s are removed when the field is controlled. However, $1 is retained. So – to control or not to control?

*the 758 (Resource Identifier) is a (relatively) new MARC field that records an identifier (typically a URI) for either the resource described in the bib record, or a resource to which it is related (with relationship info in the $i)

What should I consider before jumping in?

- To what extent does your ILS/LMS accept URIs in MARC records? Do they impact display or the indexing of headings?

- Since multiple URIs are not encouraged, what sources are best to use?

- How do you maintain quality control?

- What about pre-coordinated strings, such as LCSH?

- Do you need to be doing all this *right now*, or can you wait?

References consulted

Chew Chiat Naun (2020) Expanding the Use of Linked Data Value Vocabularies in PCC Cataloging, Cataloging & Classification Quarterly, 58:3-4, 449-457, DOI: 10.1080/01639374.2019.1705953

PCC URIs in MARC Pilot NACO subgroup (2020) 024 Best Practices Guidelines https://www.loc.gov/aba/pcc/naco/documents/NACO-024-Best-Practices.pdf

PCC Task Group on Linked Data (2019) Best Practices Final Report https://www.loc.gov/aba/pcc/taskgroup/linked-data-best-practices-final-report.pdf

PCC Task Group on URIs in MARC (2020) Formulating and Obtaining URIs: A Guide to Commonly Used Vocabularies and Reference Sources. https://www.loc.gov/aba/pcc/bibframe/TaskGroups/formulate_obtain_URI_guide.pdf

PCC Task Group on URIs in MARC (2018, revised 2023) URI FAQs https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI%20FAQs.pdf

PCC URIs in MARC Pilot (2021) Resource Documents & Tools https://wiki.lyrasis.org/pages/viewpage.action?pageId=178881300

Philip E. Schreur & Amy J. Carlson (2020) Bridging the Worlds of MARC and Linked Data: Transition, Transformation, Accountability, The Serials Librarian, 78:1-4, 48-56, DOI: 10.1080/0361526X.2020.1716584

Wallis, Richard (2018) MARC and Beyond: Our Three Linked Data Choices. Paper presented at: IFLA WLIC 2018 – Kuala Lumpur, Malaysia – Transform Libraries, Transform Societies in Session 113 - Information Technology. https://library.ifla.org/id/eprint/2124

Thank you!

QUESTIONS, THOUGHTS,
COMMISERATIONS?

CONTACT ME AT
ABBY.DOVER@NORTHWESTERN.EDU