Investigating the Identifier Ecosystem of the University of Colorado Boulder Faculty

Chris Long, University of Colorado Boulder
Presented at the ALCTS CaMMS Heads of Cataloging Departments Interest Group Meeting
January 27, 2020

CU Boulder's Experience with Identifiers

- Long-time NACO participants

- PCC ISNI Pilot
    - College of Engineering & Applied Science faculty & academic units
        - (ca. 250 faculty, 25 units)
    - School of Education faculty (ca. 50)
    - Added ORCID IDs to ISNI

CU Boulder's Experience with Identifiers

- Local Wikidata project
    - Women Poets of the Romantic Period (CU Boulder digital collection, ca. 250 poets)
    - Creating new Wikidata entities
    - Creating new NACO and ISNI records as needed
    - Adding existing NACO, VIAF and ISNI records to Wikidata

PCC URIs in MARC Pilot

- Pilot purpose: to further the PCC's linked data transition by "road testing" the conceptual work of PCC task groups

- Pilot activities will chiefly involve adding identifiers to bibliographic records and/or to NACO authority records.

- Pool limited to current CU Boulder faculty with existing NACO records that contain an affiliation with CU Boulder (179 persons)

- Enriching NACO records of current CU Boulder faculty with identifiers in MARC 024 $0 and/or $1 (VIAF, ISNI, ORCID, Wikidata)

Richness of Identifier Ecosystem for CUB Faculty

- VIAF – 100%

- ISNI – 84%
  - Some due to our previous ISNI pilot

- ORCID – 86%
  - Faculty Affairs promotion

- Wikidata – 29%

- 27% have an ISNI, ORCID, and Wikidata identifier

- 6% have no ISNI, ORCID, or Wikidata identifier

- But ... how *interconnected* is the identifier ecosystem for CU Boulder faculty?

Embedded Identifiers

Wikidata Entities

- LC NAF – 61%

- VIAF – 68%

- ISNI – 62%

- ORCID – 39%

- The most interconnected identifier system, but …

- the lowest incidence rate in this pool of faculty

Embedded Identifiers

ISNI Records

- LC NAF – 80%

- VIAF – 81%

- ORCID – 9%

- Wikidata – 35%

- Strong connections with LC NAF and VIAF, weaker connection with Wikidata, few ORCIDs

- Will ORCID rate improve with recent ability to link them to ISNIs?

Embedded Identifiers

VIAF Records

- LC NAF – 100%

- ORCID – 1%

- ISNI – 78%

- Wikidata – 64%

- Robust connections with ISNI and Wikidata, very low incidence of ORCIDs

Embedded Identifiers

NACO Records

- VIAF – 2%

- ORCID – 3%

- ISNI – 1%

- Wikidata – 0%

- Significantly less interconnected ...

  - Due to prior practice?
  - 024 moratorium?

Embedded Identifiers

ORCID Records

- ISNI – 2%

- Heavily used by many campuses to harvest research productivity, but the least interconnected and interconnectable.

- How much can this improve since ORCIDs are self-claimed?
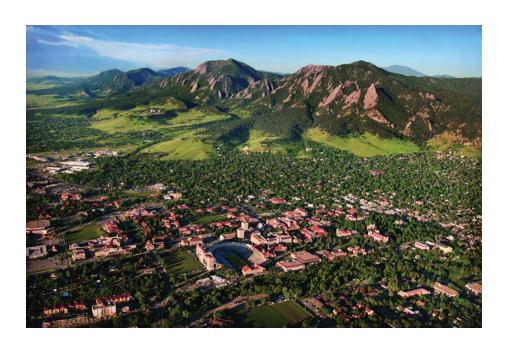
Where is Our Energy Best Directed?

Actively create and connect identifiers in multiple venues?

- If not us, then whom?
- We have the capability, but do we have the capacity?
  - CU Boulder as about 19,000 faculty

Focus energy on specific identifiers?

- Tend to one garden and let others tend to theirs?
- Can we rely on others to create the identifiers we want to interconnect?
- NACO and ISNI both have high barriers to participate and are labor intensive
- Wikidata has very low barrier to participate but best practices are in nascent stage

Thank you!

Chris Long, Associate Professor
Head of Monographic and Special Materials Cataloging
University of Colorado Boulder
chris.long@colorado.edu
ORCID: 0000-0003-3953-7940
ISNI: 0000 0001 3817 1225