Stringing Us Along

From Traditional Authorities to Identity Management in Library Technical Services

Amber Billey
Systems & Metadata Librarian
abilley@bard.edu
@justbilley

NYTSL Fall 2017
Stephen A. Schwarzman Building
New York Public Library
Raise your hand...

How many of you perform authority work?

How many of you participate in NACO and contribute to the LC NAF?

How many of you create NARs or maintain lists of names outside of the LC NAF?

And...so then what are some of the reasons you needed to create a list of names outside of the LC NAF?
My authorities story.
Palisot de Beauvois
Keys are key.
and then…
Academic Libraries
Bibliographic Data

Authority Data
*SOME* GLOBAL HEADINGS CHANGE
Issues with the current system (LCNAF + NACO)

- NARs are complex
- It requires extensive training to make them
- The research for a single NAR can be time consuming
Issues with the current system (LCNAF + NACO)

- Creating NARs for the LCNAF requires proprietary and expensive software
  - Limits participation
  - Reduces the number of names
  - Inhibits access to collections and therefore further hides already hidden collections
Issues with the current system (LCNAF + NACO)

- Limited to only a single preferred label for an identity
- Cannot discretely declare different languages for labels
- Based on redundant workflows and catalogers’ (biased) judgement
  - Why not leverage existing data?
  - Why not use publisher’s author questionnaires? Let the individuals describe themselves?
  - Why recreate metadata that already exists in another datasource?
Same silo, different farm.
PCC Task Group on Identity Management in NACO
The PCC Vision, Mission and Strategic Directions document, 2015-2017, Strategic Direction 3 is: “Provide leadership for the shift in authority control from an approach primarily based on creating text strings to one focused on managing identities and entities.”

The PCC is well positioned to be a leader in this effort to provide a framework for coordination, consultation, and education in the realm of identity management.
Authority control manages access to entities by authorizing a specific form of name or other term for access use.

Definitions

**Authority control** manages access to entities by authorizing a specific form of name or other term for access use.
Definitions

Identity management operates by associating a registered identifier with characterizing data which specify a single identity or identified entity. It is distinguished by its focus on differentiating entities through the use of identifiers. Differentiation of names is of secondary importance and may be accomplished by assigning numbering to or concatenating identifying information with instances of the same name in presentation contexts.
Use Case #1

Title (Goal): Direct Search for Information about a Name

Primary Stakeholders: Public, Internal Users

Scope: Discovery

Priority: High
Use Case #2

Title (Goal): **Indirect Search for a Name/Discover experts in specific fields**

Primary Stakeholders: Public, Researchers

Scope: Discovery, reconciliation against external data sources

Priority: Medium
Use Case #3

Title (Goal): **Reputation Management Support/Scholarly Citation Harvesting**

Primary Stakeholders: **Administrators, Researchers**

Scope: Data gathering, metrics

Priority: High
Use Case #4

Title (Goal): Genealogical research

Primary Stakeholders: Genealogists

Scope: Discovery

Priority: High for archives and public libraries; Medium (low?) for academic libraries
Use Case #5

Title (Goal): Connecting Experts

Primary Stakeholders: Public, academic, economic developers, business leaders, researchers, and administrators

Scope: Discovery

Priority: Medium?
Use Case #6

Title (Goal): Researcher profiles

Primary Stakeholders: Academics, marketing, student groups

Scope: Discovery

Priority: High
Use Case #7

Title (Goal): Digitization Crosswalks

Primary Stakeholders: Internal users

Scope: Authority, Content Standards

Priority: Medium
Use Case #8

Title (Goal): Remixing Data for Digital Objects

Primary Stakeholders: Internal users

Scope: Authority control, batch processing, ingesting data

Priority: Medium
Use Case #9

Title (Goal): Record and discover creative and intellectual relationships

Primary Stakeholders: Public, researchers

Scope: Authority control, reconciliation against external datasets

Priority: Medium
The PCC-ISNI Pilot Study (aka NACO “Lite”)
What is ISNI?

- International Standard Name Identifier
- ISO Standard (ISO 27729)
- Part of a family of international standard identifiers, e.g. DOI, ISAN, ISBN, ISRC, ISSN, ISTC, and ISWC.
- Acts as a “bridge” identifier
- Made up of 16 digits. One hundred thousand billion possible combinations.
  - [http://www.isni.org/0000000107791034](http://www.isni.org/0000000107791034)
ISNI’s mission

“The mission of the ISNI International Authority (ISNI-IA) is to assign to the public name(s) of a researcher, inventor, writer, artist, performer, publisher, etc. a persistent unique identifying number in order to resolve the problem of name ambiguity in search and discovery; and diffuse each assigned ISNI across all repertoires in the global supply chain so that every published work can be unambiguously attributed to its creator wherever that work is described.”

http://www.isni.org/
Let’s have a Party!

“A Public Identity is the name by which a Party is publicly known.”

- A Public Identity might have variant spellings, but all variants represent the single Public Identity.
- A Party might have several Public Identities. Public Identities relating to a single Party will be issued separate ISNIs.
- Links between multiple Public Identities shared by a Party can only be linked if authorized by the Party.
How does ISNI work?

- ISNI is administered by the ISNI International Agency which is governed by the ISNI board and administered by OCLC.
- ISNIs are issued through Registration Agencies (RAGs).
- RAGs pay a fee to access and edit the ISNI database. This fee covers ISNI’s operation costs.
- RAGs interface between users and the ISNI International Agency.
- Any business with a proven interest in the scope of ISNI can become a RAG.
- ISNI holds public records of over 8.3 million identities.
- The ISNI database is a cross-domain resource, contributed to by 29 institutions and databases, and 40 major national and research libraries.
The PCC is a RAG!

(Umbrella) Registration Agency for ISNI

https://wiki.duraspace.org/display/PCCISNI/PCC+ISNI+Pilot+Home
PCC-ISNI Pilot

Participating Institutions:

- Brown, University of Chicago, University of Colorado-Boulder, Columbia, Cornell, Harvard, Michigan State, Stanford, Stony Brook, Texas A&M, UC Davis, UCLA, and University of Wisconsin-Madison...and sort of now Bard!

Goals:

- To develop an understanding of ISNI tools and systems, to create PCC documentation and training for its members, and to put in place member supports for cost-effective scaling of broader PCC involvement in ISNI.
PCC ISNI Pilot

Joint Areas of Focus:

● API review/testing
● Batch processing template/workflow
● Documentation & training
● ISNI maintenance best practices
● Workflow & tools best practices
PCC ISNI Pilot

“The experience of having PCC metadata librarians working in ISNI would generate information for all parties to understand the characteristics (tools, services, workflows, training & documentation) necessary for a sustainable, affordable and effective long-range PCC membership agreement beyond the pilot.”

https://wiki.duraspace.org/display/PCCISNI/PCC+ISNI+Pilot+Home
Metadata Required for an ISNI of a Public Identity

- Name of the Public identity
- Date and place of birth and or death (or registration and dissolution for legal entities)
- Class and Roles as defined by the RA. Classes define the repertoire (such as Musical, Audio-Visual, Textual) and Roles can be Author, Performer, Publisher, Illustrator, etc.
- Title of reference to a creation
- A URI (or URL) providing a link to more detailed information about the Public Identity.
ISNI runs test and production databases
Opening “home” screen
You can add new identities for organizations or persons/fictional characters

<table>
<thead>
<tr>
<th>View of WebCat functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Here is the list with available WebCat forms. Please pick the appropriate form for the record you want to create or edit.</td>
</tr>
<tr>
<td>Add an identity: organisation</td>
</tr>
<tr>
<td>Add an identity: person or fictional character</td>
</tr>
<tr>
<td>Identifier</td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>Name</td>
</tr>
<tr>
<td>Dates</td>
</tr>
<tr>
<td>Creation class</td>
</tr>
<tr>
<td>Creation class</td>
</tr>
<tr>
<td>Creation role</td>
</tr>
<tr>
<td>Creation role</td>
</tr>
<tr>
<td>Work</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Instrument</td>
</tr>
<tr>
<td>Contributed to or performed</td>
</tr>
<tr>
<td>Name variant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Field</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Related person</td>
</tr>
<tr>
<td>Relation</td>
</tr>
<tr>
<td>Relationship type</td>
</tr>
<tr>
<td>Relationship dates</td>
</tr>
<tr>
<td>PPN</td>
</tr>
<tr>
<td>Related person</td>
</tr>
<tr>
<td>Relation</td>
</tr>
<tr>
<td>Relationship type</td>
</tr>
<tr>
<td>Relationship dates</td>
</tr>
<tr>
<td>PPN</td>
</tr>
<tr>
<td>Related organisation</td>
</tr>
<tr>
<td>Relation</td>
</tr>
<tr>
<td>Relationship dates</td>
</tr>
<tr>
<td>PPN</td>
</tr>
<tr>
<td>Language</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Nationality</td>
</tr>
<tr>
<td>External Information</td>
</tr>
<tr>
<td>General note</td>
</tr>
</tbody>
</table>

Sample new identity continued
<table>
<thead>
<tr>
<th>ISNI</th>
<th>0000 0004 6235 0348</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Billey, Amber Michelle</td>
</tr>
<tr>
<td>Dates</td>
<td>born 1982-11-29</td>
</tr>
<tr>
<td>Creation class</td>
<td>text</td>
</tr>
<tr>
<td>Creation role</td>
<td>author</td>
</tr>
<tr>
<td>Related names</td>
<td>Drabinski, Emily (co-author)</td>
</tr>
<tr>
<td></td>
<td>Roberto, K. R. (co-author)</td>
</tr>
<tr>
<td>Titles</td>
<td>What's Gender Got to Do With It? : A Critique of RDA Rule 9.7</td>
</tr>
<tr>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>Sources</td>
<td>PCC</td>
</tr>
</tbody>
</table>
Benefits of ISNI

- One key for them all
- ISO Standard
- Utilizing existing metadata sources and organizational structures
ISNI is a MASTER KEY
Possible drawbacks of ISNI?

- Could OCLC buy it?
- Still “Pay to Play” barrier to participation
- No open way to query the data
- ILS software will need to utilize identifiers instead of strings
  - URIs in MARC
New silo, different kind of farm?
“Wikidata is a free and open knowledge base that can be read and edited by both humans and machines.”
Acknowledgements

- PCC Identity Management in NACO Task Group Members
- NYTSL
- Colleagues at Columbia
- You! You’re the best!
Thanks!

Questions?

abilley@bard.edu