Charge for PCC Task Group on Identity Management in NACO
Last updated March 31, 2016

Background & Issues:

Expressed in the *PCC Vision, Mission and Strategic Directions Report, 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.

While traditional library authority control has not been supplanted in library discovery systems, the use cases for identity management activities as part of or apart from authority work have grown in recent years, as have the systems that enable it. These changes in our environment have resulted in institutions having to forge this new ground on their own, often leading to duplicated effort, lack of shared best practices, and therefore compromised interoperability. Although VIAF (Virtual International Authority File: viaf.org) and ISNI (International Standard Name Identifier: isni.org) have made great strides to integrate entity identifiers from various sources, there is still much work that can be done to further their objectives for productive identity management, making maximum use of expertise from within libraries and expert communities outside of libraries.

There are a number of identity management issues that would benefit from community exploration. Most importantly, there is the overarching issue of "strings" versus "identifiers." In RDA, there are a number of ways to create unique "strings," but the community has recognized the need to shift focus to identity management. While some have expressed this as a need for definition of a more inclusive "NACO Lite" approach, there are many use cases for unique identifiers where there is no interest in or need to establish authoritative name forms. The coming years will be a transitional period as systems begin using identifiers in their delivery of optimized search results for users and library staff. As strings lose their function as the primary bearers of "uniqueness" but continue to have value as human-readable information, what is necessary to facilitate separating display needs from back-end functionality?

There is also the issue of the multiplicity of identity communities, including cultural institutions, rights management organizations, academic institutions, and libraries. Different systems with various requirements have been developed to meet these needs, for example, ORCID, NACO, etc. Additionally, there are local authority files held and maintained by many institutions to serve their internal needs that are not being shared with the broader community. What infrastructure must be put in place by libraries so that we can work most effectively together: minting and sharing identifiers, linking local identifiers to globally established ones, and creating metadata enrichment lifecycles that enable broad sharing of identity management activity, whether part of authority work or not.

In addition to developing ways for humans to perform their work in a coordinated manner, it will be critical for identity management activities to be performed with machine matching in mind as systems need to share data through automated means. With natural limitations to automated matching, how and under what circumstances will unresolved, provisionally assigned, or incorrectly linked identities be addressed? What role does the PCC have in providing documentation and education on how various authority and identifier systems interoperate, and therefore what librarians can do to ensure the best automated outcomes?

Finally, directly involving the subjects (e.g., authors, researchers, publishers) of identity management in the process of stewarding their profiles in global identify management platforms is emerging as an issue facing libraries. Librarians are receiving comments from authors requesting information or corrections to their "identities" as established in OCLC WorldCat Identities, ISNI, etc.

Charge:

Reporting to the PCC Steering Committee, the Task Group on Identity Management in NACO is charged to:

1. Investigate and analyze the reasons for local authority creation within the library community.
2. Investigate whether establishing guidelines or best practices for NACO identity management is feasible and furthers the goals of the PCC.
3. Investigate and educate the community about how components of our current ecosystem function and interoperate.
4. Investigate and lead PCC discussion on issues regarding the differences between authority control and identity management.
5. Analyze the issues regarding getting direct input from entities whose identities are being managed.
6. Identify use cases where library authority data can be put to new non-library uses, which PCC activity could respond to.

Suggested work plan:

Each item in the charge has issues or approaches specific to it. In the following work plan, each item in the charge is expanded to include key questions, approaches, and actions to be taken.

1. Investigate and analyze the reasons for local authority creation within the library community. These issues may include:
    a. What are the barriers to creating authorities in NACO?
        i. What are ways to lower these barriers for broader participation in NACO? Is this desirable and should these records be distinctly coded?
        ii. Are there ways to ingest data via automated processes from "local" work streams? Is this desirable and should these records be distinctly coded?
    b. What communities outside of the library community are creating local authorities?
        i. What particular needs do they serve and how can we ingest or interoperate with their data?
    c. What lessons can be drawn from the Linked Data for Production (LD4P) grant activity PCC has partnered with and many of its members have participated in?
2. Identify areas where PCC policies and guidelines should more explicitly respect and encourage the use and reuse of data from expert communities and non-library partners. Investigate whether establishing guidelines or best practices for NACO identity management is feasible and furthers the goals of the PCC.
    a. In collaboration with the PCC Standing Committees, develop best practices for authority/identity record creation, amending or complementing NACO practices, keeping in mind interoperability, flexibility, automated- versus human-creation, and ease of use. (For example, current systems such as VIAF use keywords in the 670 (Source data found) to match incoming data. Therefore, the addition of at least one source work for each identity could be helpful for interoperability.)
    b. Guidelines should consider how to create a metadata lifecycle that allows institutions to contribute entity creation data at all states of completeness, so that the intellectual effort expended in baseline work is shared as a foundation on which other institutions can build when their local needs require fuller entries or when additional information becomes known about an entity. With a philosophy of "no duplicated effort", protocols should maximize sharing and allow metadata enrichment over time.
        i. Develop or modify existing data elements to provide clear status information and identify the level of confidence associated with data to allow building upon prior work with clarity.

    c. How can the output of identity management activities feed into NACO when authority work is needed, while also being shareable when establishing identities is sufficient for the needs at hand?

3. Investigate and educate the community about how components of our current ecosystem function and interoperate. These issues may include:
   a. Are there guidelines that the PCC/NACO participants should follow to ensure smooth interoperability with VIAF, ISNI, etc.?
   b. Consider how best or whether to involve the PCC in the clean-up work that is generated by automated matching processes that interoperate between identity management systems (e.g., the "Possible Matches" that are left unassigned as provisional ISNIs in the back-end database)

4. Investigate and lead PCC discussion on issues regarding the differences between authority control and identity management. This work should have an aim to educate both the community and partner organizations outside of the PCC/library community.

5. Analyze the issues regarding "user" input into identity management systems. These issues may include:
   a. Educating communities on how to research and supply data about their identities.[1]
   b. Assess how to incorporate user-contributed data.
   c. Are there models in current user interfaces in identity management systems such as OCLC WorldCat Identities, ISNI, etc.?

6. Identify use cases where library authority data can be put to new non-library uses, which the PCC activity could respond to.
   a. Survey the community and the literature for examples of new uses being made of library authority data.
   b. Collect the details of the use cases and share in clearinghouse fashion on PCC's website.
   c. Assess what changes in PCC practices would better facilitate those uses and enable additional uses.

Roster:

Member	Affiliation	Email
John Riemer, chair	UCLA	jriemer[at]library.ucla.edu
Amber Billey	Columbia	ab3167[at]columbia.edu
Michelle Durocher, PoCo representative	Harvard	durocher[at]fas.harvard.edu
Paul Frank, PCC NACO	Library of Congress	pfrank[at]loc.gov

---

[1] As background, see recommendations by Jisc CASRAI-UK Organisational Identifiers Working Group related to involving organizations in the establishment and maintenance of their OrgIDs.

Stephen Hearn	Minnesota	s-hear[at]umn.edu
Violeta Ilik	Northwestern	violeta.ilik[at]northwestern.edu
Jennifer Liss	Indiana	jaliss[at]indiana.edu
Andrew MacEwan	British Library	Andrew.MacEwan[at]bl.uk
Erin Stalberg	Mount Holyoke College	estalber[at]mtholyoke.edu

The Task Group will call on community experts as needed for advice or to serve on subgroups.