PCC Strategic Themes Discussion Summary
November 2014 (rev. December 2014)

Background

In preparation for the face-to-face PCC strategic planning session in November 2014, the PCC Policy Committee (PoCo) commissioned three white papers:

- *Strategic Directions for the Program for Cooperative Cataloging* by Philip E. Schreur

- *The PCC and Continuing Resources* by Regina Romano Reynolds

- *Name Authorities in Transition: Implications for the PCC* by the PCC Advisory Committee on Initiatives (ACI)

From July to October, PoCo held themed-based discussions on its email list, focusing on the ideas from the papers. Rather than discussing each paper in its entirety, topical email threads were used to focus the discussion. This document is a summary of those discussions. Please note that statements in this summary reflect personal, institutional, and evolving opinions of representatives to PoCo, and are not necessarily PCC-endorsed statements.

1. PCC & ROI

- **Core questions:** When the PCC was formed, the ROI was clear--we all saved by contributing to the pool of quality cataloging available to all. As we move to a future PCC, the work changes and so should our expression of the value of membership. How do we find meaningful measures of value? What can we measure? Is what we currently count still of value? How should it change? How should the PCC best present its ROI in the 21st century? What metrics should be used to justify program participation for both members and administrators?

- Should we consider looking at the concept of Value on Investment (VOI)? Perhaps we need to focus on the VOI of PCC membership, instead of ROI. Here's an article out of the US Chamber of Commerce about VOI - http://institute.uschamber.com/move-from-roi-to-voi/.

- What is the return on the investment for PCC of having vendors rely on the CONSER database to manage their resources, and market their products?

2. PCC & RDA

- **Core questions** (from Schreur white paper): "Given all this, what is the relationship between the PCC and RDA? Can PCC make a clear statement on the place of RDA within its programs or should the association with RDA as a core part of our branding be tempered?"

- If there is an inherent conflict between RDA's emphasis on transcription, and the linked data world's insistence on identifiers, will PCC need to propose RDA changes? When it

comes down to conflicts between RDA and linked data, how will PCC make preferences, choices, decisions?

- o As we approach a linked data environment, and as we work with more digital resources, we need to focus more on identification and relationships as opposed to description and certainly not transcription. What the cataloger transcribes from an online resource today may be entirely different tomorrow. The main role of the serials cataloger is to record the essence and constants of the publication, not the accidents and variables. This role, true even in the print world, is all the more needed, I believe, in the highly volatile world of digital resources. RDA has gone in exactly the opposite direction.

- o As we move from records to an open graph of data, what does RDA mean in that context? As statements describing a certain resource get added or deleted over the years, what guarantee do we have that any of them would still conform to RDA? Is it possible to define a certain collection of statements as being RDA conformant and protect them over the years? If it were possible to mark RDA conformance at the statement level, what would it mean if 33% of the statements had the provenance of "RDA"? Does that make the collective RDA? Do I need 40%? 50? 60%?

- Continued PCC insistence on RDA exclusively could be a barrier to collaboration across other organizations, publishers, identity/researcher profile management agencies, rights management organizations, or parts of the world that won't adopt RDA.

- The creation of data according to RDA is very expensive. Many, once traditional cataloging departments, are responsible for the creation of metadata for an exponentially expanding set of resources. The percentage of data created according to RDA is becoming smaller and smaller but still absorbs a high proportion of the budget.

- Several people commented (particularly in the thread about continuing resources) that RDA has resulted in catalogers adding more data to records, not less. Is this really a step backwards, or a step in the direction of cataloger judgment? The BSR and CSR have always been advertised as floors, not ceilings, so how critical should we be that people are actually choosing to describe things more fully than the floor?

3. PCC & BIBFRAME

- What is the business case for BIBFRAME? Are there other options to explore? When we declare that PCC "supports" BIBFRAME, what exactly are we supporting?

- How will we define PCC's "collaboration" with BIBFRAME? What are the expectations of PCC's role? Should we go beyond expressing and committing to PCC support for testing and evaluation?

- PCC may need to consider a multi-pronged effort to: (1) analyze the BSR and CSR requirements in light of BIBFRAME (i.e., what would we mandate, recommend, etc.); (2) develop a PCC BIBFRAME profile for community testing; (3) develop a PCC BIBFRAME editor (LC's is too LC-centric).

- Facilitating learning, testing, and evaluation of BIBFRAME:

- o Provide a venue to help people get started;
  - o Perhaps use something like Google hangouts (or something else lightweight and informal) to having discussions, chat, ask questions. Wouldn't need to be restricted to PCC participants.
- NLM's BIBFRAME experimentation has revealed that there will be a need for libraries to transform local authority files into BIBFRAME. LC and Zepheira have made a start in mapping MARC bibliographic data to BIBFRAME, but there has been no similar work to map authority data to the BIBFRAME vocabulary. Is this a task that PCC could take on for the community and ease the burden of migration for individual libraries?
- Issues with BIBFRAME itself:
  - o The current vocabulary has focused too much on converting existing MARC records, rather than thinking about the actual basic needs for bib data. We should streamline the BIBFRAME vocabulary to the bare minimum needed among the shared communities of librarians, publishers, museums, archivists, etc. and use existing schemes like the RDA element set and PRESSoo to augment for the resources that need those elements to be properly described. There may be no need to re-create all the MARC and RDA elements in BIBFRAME, which is what some see as the current approach.

4. PCC & Linked Data

- RDA has a strong emphasis on transcription of data, which in a linked data environment is a hindrance – what does this mean for PCC standards that need to adhere to both RDA and BIBFRAME and other linked data applications/standards?
- Where can we rely on identifiers? What domains have not yet evolved to develop unique identifiers?
- Need to understand the ramifications of the fact that we don't create identifiers for persons, we create them for controlled strings.

5. Authorities beyond LCNAF

- **Core questions:** The ACI white paper includes recommendations to (1) develop guidelines for the use of VIAF vocabularies to authorize name entities, and (2) develop a process for evaluating, endorsing, and providing guidance for the use of name vocabularies beyond VIAF.
  - o Is the value in time-savings worth the trade-off? What would this change mean for the BIBCO and CONSER 'brands,' which are currently trusted to have NACO authorized headings? Would this move bring us additional international partners?
- Extend active use of VIAF in PCC records as a start. It is redundant to re-establish a name according to RDA when an authority record exists for the same name established by another authoritative agency (e.g. a national library), even if the other agency establishes the name differently from the way it would be done under RDA guidelines. Why can't we accept the work that the other agency has done, adding a variant access point for the RDA form of the access point to what the other agency has

done? The identity management tasks and collocation are the really valuable aspects of authority control; the text string of the authorized access point is less important.

- We are already doing this with subjects – FAST, or any other established thesaurus is permitted in PCC records, as long as it is identified.

- Our current authority records manage authorized text strings. In RDF, the focus is on representing the persons themselves and the issue of authorized text strings of various flavors is secondary. As we move towards identity management, we'll need to revisit the purpose and structure of our authority data to make sure that it can support this new need.

- Need to have authority vendors on any future PCC TF on non-LCNAF authorities.

6. Broader Participation in Authority Creation

- **Core questions:** Recommendation 3 from the ACI white paper: "Significantly expand the ranks of those who can create identifiers/contribute authority data".   They propose two ways of expanding: (1) expand NACO participation with a "NACO lite" level of participation, and (2) endorse the creation of a separate, parallel authority file to which non-NACO members could make contributions

    o What are the pros/cons of these options?   Should we do both, or neither?   Other thoughts on these recommendations?

- On the first recommendation, one might want members who could just contribute certain kinds of names (perhaps just personal names, for example), or "simple" names (although how to do define that I'm not sure). The second one is problematic. Making catalogers check two separate authority files is not very efficient, and is bound to create duplication of names. And how would those non-NACO members be trained to be sure they understand NACO guidelines and RDA?

- ISNI is already providing a way of expanding the ranks of those who do create authority data/ and the scope of what names are controlled. ISNI has workflows for regular updates to ISNI members. NACO will be part of that network by default once ISNIs have been loaded to the NACO file. In essence ISNI is an expanded NACO, both through VIAF and through data from other sources, such as Digital Author Identifier (DAI) network in the Netherlands, university repositories (e.g. La Trobe University, Australia) Rights Management Organisations and soon many more from a variety of content sectors requiring an ID for names in their metadata. In Europe there is increasing take up among national libraries joining or actively planning to join the ISNI network as a way of proactively working with a maintained linked authority file, using ISNI rather like NACO itself as a shared authority file. VIAF overlaps significantly and provides much of the core value in ISNI, but ISNI provides workflows, notifications, and maintenance of its clusters as authoritative IDs.

    o More active engagement by NACO members with ISNI would enable us to better address the inevitable policy issues that come up with trying to link authorities and define a common ID number. The basis for the BL's involvement in ISNI has always been to find a way of expanding the scope and reach of authority control through wider collaboration beyond the library sector. The key issue that ISNI is working on is how to make the system work like NACO or NACO re-imagined.

- o If PCC works in tandem with ISNI, we may not need to provide a NACO lite option; ISNI could be the de facto parallel authority file for non-NACO members.
- There already is a "NACO lite" in place; NACO has employed the method of allowing for the coding of 008/33 to "preliminary" and "provisional" to allow less than "full" NACO NARs into the NACO file. LC has employed these methods for machine-generated records for an OCLC-LC-Music (MDAR) project, for the addition of the Dance Heritage Project NARs, as well as for others. We could also accept other projects to do the same.
- OCLC and SkyRiver have local authority files to which non-NACO participants could contribute and these could then be brought into the LC/NACO Authority File by NACO members as needed. Should this be explored before we expend an enormous amount of time building a training program for something called NACO lite?
- OCLC response: OCLC does not have a local authority file to supplement the LC/NACO file. We've talked about moving in that direction several times over the years. Mostly recently it has resurfaced in the context of WorldShare Management Services and the use of WorldCat as the institution's catalog. That brings with it, even for smaller institutions, the need for some additional authority records. We have also talked a bit about trying to organize NACO funnel projects to cover this need but haven't been able to figure out how to staff such an effort.
  - o Any of these possibilities would require development, which would have to be prioritized. We'll be starting the process to plan for FY16 projects in mid-November.
  - o One additional comment from my personal perspective: in my experience, putting something in a separate files only works if one can clearly define the contents of the file in a way that the potential user of the file knows "when I should search here." I don't think describing an authority file as containing "other authority records that are not part of the NACO file for various reasons" would be that kind of clear definition.
- The cost of some level of training would be a factor and perhaps here's where the PCC could play a role. Training is the PCC's strong suite and while previously it's been tied to PCC membership, perhaps it could be broadened (since we're broadening things) to include the non-NACO community, especially since the resident expertise for creating and maintaining authorities does live within PCC.
- Are there opportunities to crowd source enhancement of the authority files, especially to communities of specialization?

7. PCC & Continuing Resources / Journal vs. Article-level Metadata

- **Core questions:** If the journal article is the "object of desire," do we still need detailed cataloging for journals at the title level, or would something less work for management of titles? How does the move to increased batch processing for cataloging change how we think about the need for CONSER records? Linked data has potential for new ways of managing title changes and other relationships between serial titles, but the

FRBR and BIBFRAME models don't address serial needs well. Do we need to consider PRESSoo (http://www.issn.org/the-centre-and-the-network/our-partners-and-projects/pressoo/) as an alternative for serials?   How would this interoperate with BIBFRAME?

- If we are to shift to incorporating things like metadata for articles, or at least enhancing them, what changes would it mean for us? Is it possible to create an extremely simple template that is still RDA compliant? Would we want to do that? How could we scope the work to make it even approachable? Take responsibility for the articles our own institution creates?

- While article metadata may take on increased importance, the management of titles may not be declining in importance. The publication as a whole continues to matter to others outside libraries: decisions on what gets bundled into aggregations; ratings like ISI Impact Factor; the grouping of peer reviewers onto editorial boards; the vision and approach of an editor; the stance on open access; the brand/prestige of where a faculty member manages to get published, etc.

- Quality of article-level metadata coming from other agencies varies widely. If catalogers were to somehow contribute to article-level metadata, could it be licensed to vendors, just as the CONSER database is already? Would cataloger contributions at the article level even be sustainable, given the volume of articles?

  - Also, the specificity of article topics means that subject analysis requires a higher level of expertise, especially in the sciences, than is required to do at the journal title level (a level of expertise in subject areas that might be difficult to realize, certainly it is already the case at LC).

  - NLM's experience as an indexing agency has shown that while catalogers may be capable of creating indexing metadata, the volume is large (NLM indexers do over 750,000 articles per year) and instituting authority control for personal and corporate names at this time would be very difficult. NLM did a study a few years ago, and only a very small percentage of article authors had entries in the NAF. Book authors and article authors (at least in medicine) have very little overlap. If ORCID ever actually takes off and authors can self-identify with a reliable ID, the situation will be different and article data might then fit well into the BIBFRAME model.

  - Within the ISNI system IDs have been assigned to a quarter of a million authors of articles in journals from the British Library's Electronic Table of Contents service. The ISNI Assignment Agency is working on ways of increasing levels of assignment to authors of articles across multiple data sources.

  - But, providing article level metadata does not have to mean human "cataloging" necessarily. HighWire Press does automate subject analysis of articles for customers, which they prefer to the human-assigned subject terms. If we registered all of our faculty with ISNI and used the identifier, did automated subject analysis, and some brief transcription would that be enough? We need to get beyond thinking things necessary to the university are out of scope and focus on how we might accomplish them with an ever growing tool set.

- Evolve PCC's relationship to ISSN:
  - Streamline journal description to be closer to what ISSN is doing.
  - One idea to explore is CONSER libraries helping to assign ISSN. The U.S. ISSN Center already adds ISSN data elements to CONSER records – it would require relatively little extra training (beyond a high level of serial expertise that CONSER catalogers already have) plus review for a period of time in order for an experienced CONSER cataloger to assign ISSN to U.S. materials. In anticipation of this possibility, Regina has requested that a new ISSN system in development at the Library of Congress be able provide the technical means for this kind of partnership to work.

- Do we need to differentiate between needs of journals/articles, and other CRs like newspapers, which may not have the same levels of access below main title?

8. Changing Cataloging Culture

- RDA's reliance on cataloger judgment has not been met with a corresponding change in the corporate culture of cataloging. The corporate culture of cataloging, for the most part, has been rule-bound, with examples and interpretations for every odd exception that ever occurred (or that someone thinks might possibly occur) and instructions in writing for the one and only "right" or "correct" way to do it. Ironically, I think we were moving toward exercising judgment with the "floor, not a ceiling" approach of the AACR2 CONSER Standard Record, but for some unexplained reason, we seem to have lost that momentum. RDA was also heralded as the code which supported user tasks, but a lot or emphasis seems to have shifted back to the "more is better" philosophy of catalogers when it comes to description.