From Strings to Things:
Transitioning from Authority Control to Identity Management at Stanford University Libraries

Joanna Dyla, Head, Metadata Development Unit, Stanford University
Hilary Thorsen, Metadata Librarian for Humanities, Stanford University

Abstract
The cataloging community has arrived at a pivotal time of transition from traditional authority control to identity management. Originally, authority control relied on text strings to differentiate names and direct users to their authorized forms. However, with the rise of Linked Data, URIs (Uniform Resource Identifiers) can fulfill the role of entity differentiation aiding in needed disambiguation. The transition comes at a time when the PCC (Program for Cooperative Cataloging) was already looking for ways to make NACO (Name Authority Cooperative Program) more inclusive to institutions who wanted to contribute, but did not have the resources to meet the program's rigorous requirements. To involve more organizations, the PCC and the ISNI International Agency decided to establish a PCC umbrella membership in ISNI to be available to PCC members as a part of its cooperative programs.

The PCC embarked on a yearlong pilot project with ISNI to experiment with ways to reduce the burden on member institutions as they include ISNI (International Standard Name Identifier) creation in their metadata workflows and investigate use cases. Professional catalogers and non-catalogers in the Metadata Department at Stanford University Libraries have been participating in the project by creating ISNIs for persons in three categories not under authority control in our discovery environments-- authors and advisors of electronic theses and creators of electronic games.
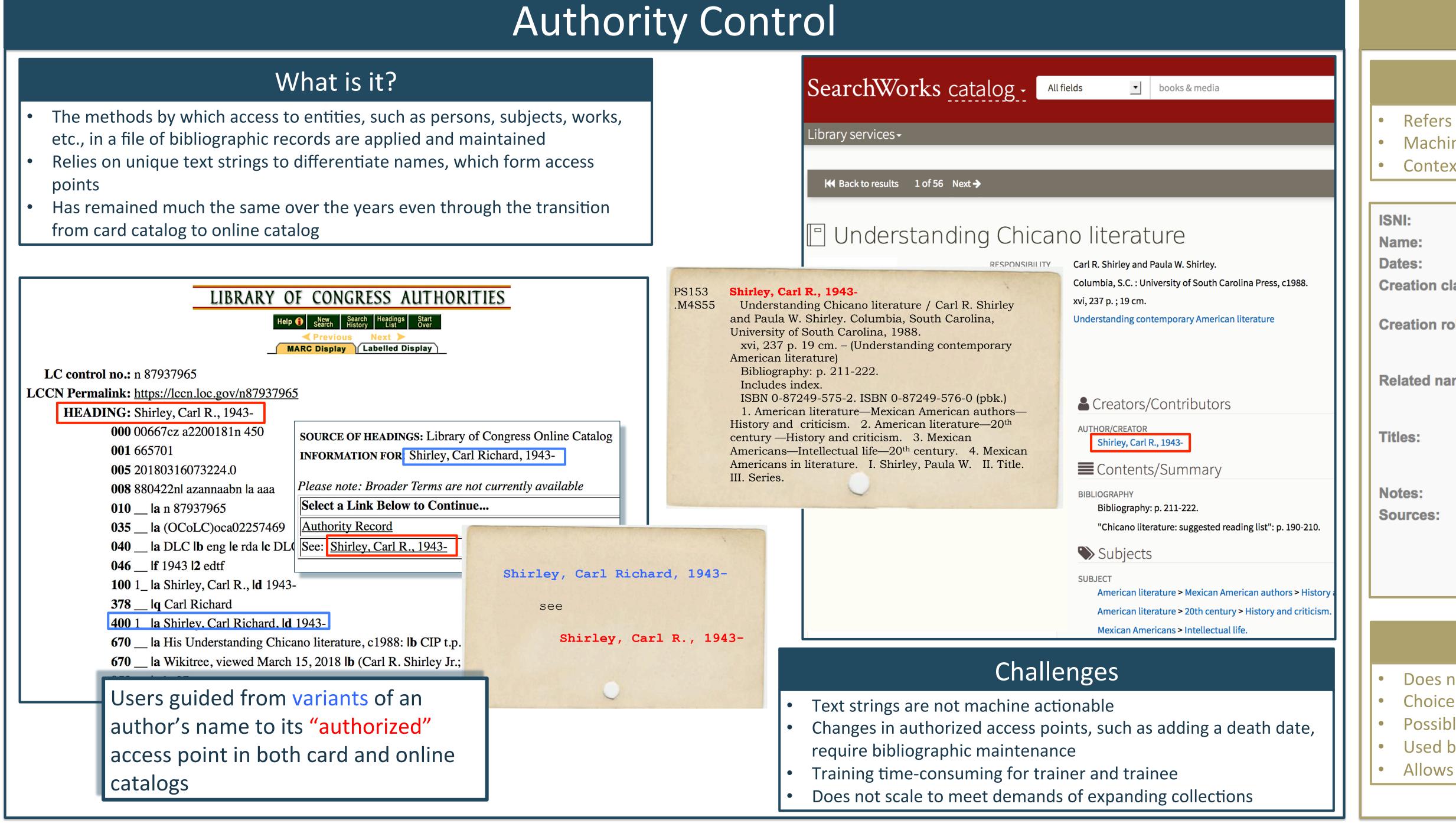
We surveyed participants shortly after the pilot began to gain a preliminary assessment of how effective the training materials were, how much time participants spent creating ISNIs, and their overall feedback about the process. Participants' responses indicated that the available training materials were adequate for self-instruction regardless of cataloging experience, but enhancements to the ISNI web interface, tools, and overall database maintenance would be needed to increase efficiency of ISNI creation in comparison to NACO. Participants envisioned creating ISNIs in the future for digital collections with large batches of names and for names that do not qualify for NACO.
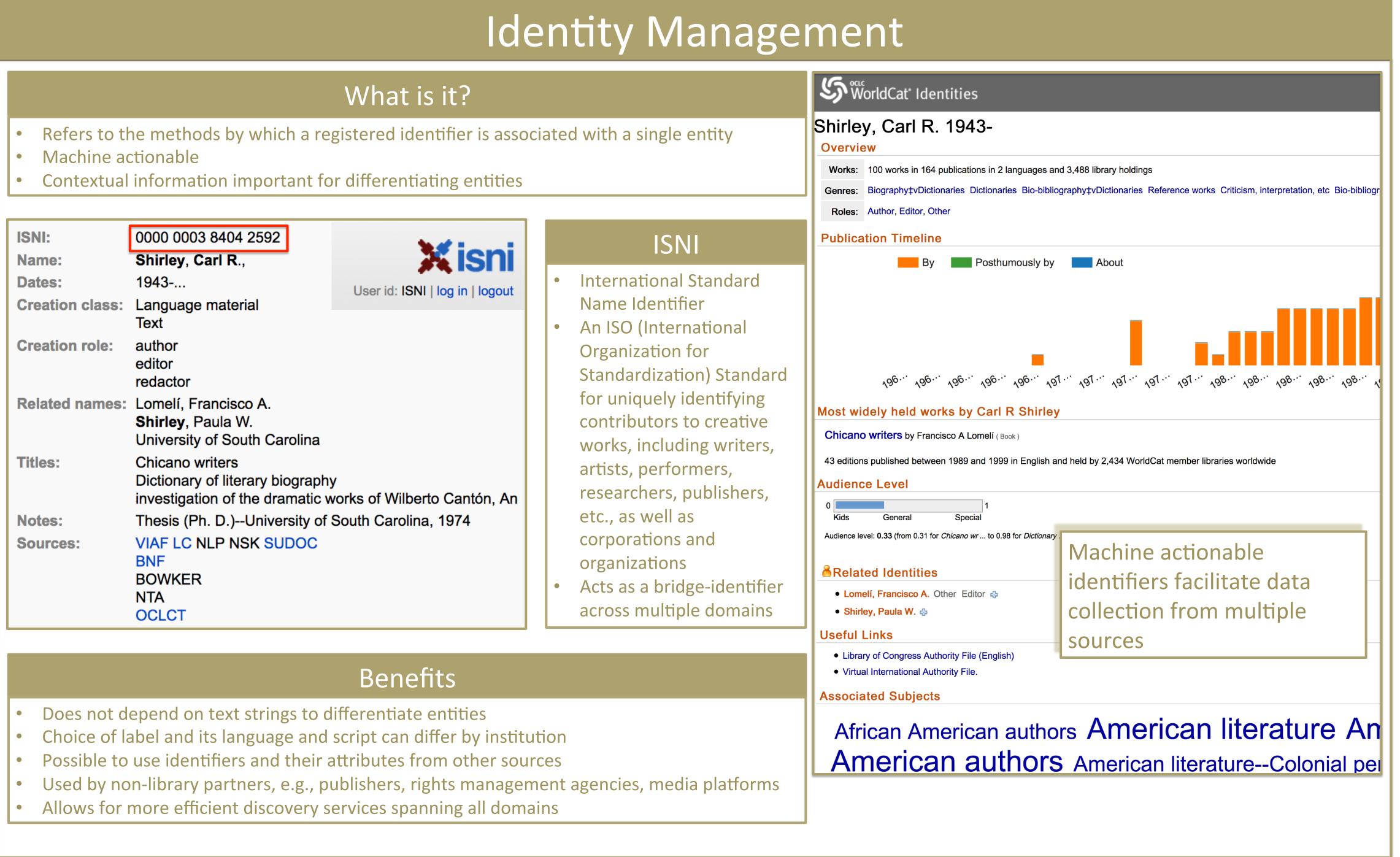
From Strings to Things:
Transitioning from Authority Control to Identity Management at Stanford University Libraries
Joanna Dyla & Hilary Thorsen

Since the inception of the modern library, authority control has been an integral part of library catalogs, but has remained largely unchanged. Recently, the cataloging community has begun transitioning from authority control to identity management to allow library users to take advantage of Semantic Web technologies. Stanford University Libraries has been participating in a pilot project through the Program for Cooperative Cataloging (PCC) to experiment with assigning ISNIs (International Standard Name Identifiers) to personal names as an alternative to creating NARs (Name Authority Records) through NACO (Name Authority Cooperative Program).

Authority Control

What is it?
- The methods by which access to entities, such as persons, subjects, works, etc., in a file of bibliographic records are applied and maintained
- Relies on unique text strings to differentiate names, which form access points
- Has remained much the same over the years even through the transition from card catalog to online catalog

Users guided from variants of an author's name to its "authorized" access point in both card and online catalogs

Challenges
- Text strings are not machine actionable
- Changes in authorized access points, such as adding a death date, require bibliographic maintenance
- Training time-consuming for trainer and trainee
- Does not scale to meet demands of expanding collections

Identity Management

What is it?
- Refers to the methods by which a registered identifier is associated with a single entity
- Machine actionable
- Contextual information important for differentiating entities

ISNI
- International Standard Name Identifier
- An ISO (International Organization for Standardization) Standard for uniquely identifying contributors to creative works, including writers, artists, performers, researchers, publishers, etc., as well as corporations and organizations
- Acts as a bridge-identifier across multiple domains

Machine actionable identifiers facilitate data collection from multiple sources

Benefits
- Does not depend on text strings to differentiate entities
- Choice of label and its language and script can differ by institution
- Possible to use identifiers and their attributes from other sources
- Used by non-library partners, e.g., publishers, rights management agencies, media platforms
- Allows for more efficient discovery services spanning all domains

PCC-ISNI Pilot at Stanford

Pilot Details
- Present day testing for future PCC involvement with ISNI
- Year-long (July 2017-July 2018)
- Public wiki: https://wiki.duraspace.org/x/m5s2BQ
- **Stanford participation in ISNI pilot**
- Early experimenters began in August 2017
- Full participation began in November using the ISNI web interface

PCC/Stanford Pilot Goals
- Expand PCC membership by removing some of the cost and training barriers that have prevented institutions from participating in the past
- Compare NACO catalogers' experiences with those without previous authority control experience
- Determine the extent of formal training needed vs. self-training
- Evaluate the effectiveness of PCC-created training materials to help in self-training, such as a series of short videos
- Establish the batch template and batch workflow processes that will be adopted as the common template within the PCC umbrella membership

Survey of Participants
- Survey sent out via email: February 7, 2018
- Final response received: March 12, 2018

Participants
- 8 catalogers (7 NACO catalogers, 1 non-NACO)
- 3 non-catalogers

Training
PCC training videos
- 10 watched; 1 did not
- 8 found them useful; 1 did not (1 N/A)

PCC-ISNI documentation
- 9 used; 2 did not
- 6 found it useful; 2 found somewhat useful; 1 did not (1 N/A)

Time

- Time per ISNI varies (from 1-60 minutes)
- Average: 12.76 minutes
- For some names, e.g., duplicates and common names, additional time required to enhance the initial data to change the ISNI status from provisional to assigned

Preliminary Conclusions
- Available training videos and documentation are adequate for self-instruction for catalogers (NACO and non-NACO) and non-catalogers, eliminating the need for more formal, time-consuming training
- Initially, no apparent time savings in comparison to NACO; could be improved with enhancements to tools and further hands-on experience
- Further time savings could be achieved by immediately assigning ISNIs to all PCC contributions
- NACO catalogers did not see ISNIs as a viable replacement for creating individual NARs at this time, mostly due to the lack of an effective interface, tools, and the overall maintenance of the database
- Participants see potential in creating ISNIs for large batches of names, especially for large digital collections
- ISNI makes it possible to provide URIs for names that do not qualify for NACO, e.g., authors of articles
- ISNI makes it possible to provide URIs for entities in finding aids not already under authority control

From Strings to Things: Transitioning from Authority Control to Identity Management at Stanford University Libraries

PDF of poster https://stanford.box.com/s/khvyjkt6xjj727v0ovjlc6mgnitrnhc6

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

ISNI website http://www.isni.org/

- ISNI search http://www.isni.org/search

Library of Congress

- Online catalog https://catalog.loc.gov/
- Authorities http://authorities.loc.gov/

WorldCat Identities search http://www.worldcat.org/identities/

Joanna Dyla: jdyla@stanford.edu

Hilary Thorsen: thorsenh@stanford.edu

Stanford | LIBRARIES