Researcher Identifiers—What’s in a Name (or URI)?

Karen Smith-Yoshimura
Program Officer
## Authorship Trends, Issues, & Questions

<table>
<thead>
<tr>
<th>Trend</th>
<th>Potential Authorship Issues</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in number of coauthors</td>
<td>- ‘honorary’ authorship</td>
<td>- How to disambiguate author names?</td>
</tr>
<tr>
<td></td>
<td>- ‘ghost’ authorship</td>
<td>- How to communicate attribution in citation?</td>
</tr>
<tr>
<td></td>
<td>- disputes</td>
<td>- How to describe contributions to work?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- How to evaluate and predict impact?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Who is responsible?</td>
</tr>
<tr>
<td>Shift from academic publishing in books to journals</td>
<td>- loss of sole-author-book as a evaluation measure</td>
<td>- How to integrate name authority and researcher identifier systems?</td>
</tr>
<tr>
<td>Decreasing granularity of publications</td>
<td>- persistence of “nano” publication vs. authorship</td>
<td>- How to document authorship over substructure of work?</td>
</tr>
<tr>
<td>Dynamic documents</td>
<td>- version misattribution</td>
<td>- How to document authorship over time?</td>
</tr>
<tr>
<td>Increasing diversity in citable scholarly outputs</td>
<td>- citation cannibalization, overcounting</td>
<td>- How to cite data, software, presentations(?), blogs (?), tweets (?)</td>
</tr>
</tbody>
</table>
Scholarly output impacts the reputation and ranking of the institution

We initially use *bibliometric analysis* to look at the top institutions, by publications and citation count for the past ten years...

Universities are ranked by several indicators of academic or research performance, including... highly cited researchers...

Citations... are the best understood and most widely accepted measure of research strength.
A scholar may be published under many forms of names

Also published as:
Avram Noam Chomsky
N. Chomsky

Works translated into 50 languages (WorldCat)

Noam Chomsky
Linguist

Journal articles

Also published as:
Avram Noam Chomsky
N. Chomsky

Works translated into 50 languages (WorldCat)
One researcher may have many profiles or identifiers...

(from an email signature block)

Profiles:  [Academia](http://www.academia.edu) / [Google Scholar](https://scholar.google.com) / [ISNI](http://isni.org) / [Mendeley](http://www.mendeley.com) / [Microsoft Academic](http://academic.microsoft.com) / [ORCID](http://orcid.org) / [ResearcherID](http://www.researcherid.com) / [ResearchGate](http://www.researchgate.net) / [Scopus](https://www.scopus.com) / [Slideshare](http://www.slideshare.net) / [VIAF](http://viaf.org) / [Worldcat](http://www.worldcat.org)
Registering Researchers in Authority Files Task Group Members

• Micah Altman, MIT - ORCID Board member
• Michael Conlon, U. Florida – PI for VIVO
• Ana Lupe Cristan, Library of Congress – LC/NACO trainer
• Laura Dawson, Bowker – ISNI Board member
• Joanne Dunham, U. Leicester
• Amanda Hill, U. Manchester – UK Names Project
• Daniel Hook, Symplectic Limited
• Wolfram Horstmann, U. Oxford
• Andrew MacEwan, British Library – ISNI Board member
• Philip Schreur, Stanford – Program for Cooperative Cataloging
• Laura Smart, Caltech – LC/NACO contributor
• Melanie Wacker, Columbia – LC/NACO contributor
• Saskia Woutersen, U. Amsterdam

• Thom Hickey, OCLC Research – VIAF Council, ORCID Board
## Stakeholders & needs

<table>
<thead>
<tr>
<th>Role</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Researcher</strong></td>
<td>Disseminate research</td>
</tr>
<tr>
<td></td>
<td>Compile all output</td>
</tr>
<tr>
<td></td>
<td>Find collaborators</td>
</tr>
<tr>
<td></td>
<td>Ensure network presence correct</td>
</tr>
<tr>
<td></td>
<td>Retrieve other’s scholarly output to track a given discipline</td>
</tr>
<tr>
<td><strong>Funder</strong></td>
<td>Track funded research outputs</td>
</tr>
<tr>
<td><strong>University administrator</strong></td>
<td>Collate intellectual output of their researchers to fulfill funder or national mandates, internal reporting</td>
</tr>
<tr>
<td><strong>Librarian</strong></td>
<td>Disambiguate names</td>
</tr>
<tr>
<td><strong>Identity management system</strong></td>
<td>Associate metadata, output to researcher</td>
</tr>
<tr>
<td></td>
<td><strong>Disambiguate names</strong></td>
</tr>
<tr>
<td></td>
<td>Link researcher's multiple identifiers</td>
</tr>
<tr>
<td></td>
<td>Disseminate identifiers</td>
</tr>
<tr>
<td><strong>Aggregator (includes publishers)</strong></td>
<td>Associate metadata, output to researcher</td>
</tr>
<tr>
<td></td>
<td>Collate intellectual output of each researcher</td>
</tr>
<tr>
<td></td>
<td><strong>Disambiguate names</strong></td>
</tr>
<tr>
<td></td>
<td>Link researcher's multiple identifiers</td>
</tr>
<tr>
<td></td>
<td>Track history of researcher's affiliations</td>
</tr>
<tr>
<td></td>
<td>Track &amp; communicate updates</td>
</tr>
</tbody>
</table>
Systems profiled (20)
Where are researchers?

![Researchers graph](image)

Wild Guesses

Researchers
## Researcher Identifier ≠ Name Authorities

<table>
<thead>
<tr>
<th></th>
<th>Traditional Name Authorities</th>
<th>Researcher Identifier Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Stakeholders</td>
<td>Libraries</td>
<td>Publishers, Researchers, Funders, Libraries</td>
</tr>
<tr>
<td>Internal standardization/integration</td>
<td>Standardized and well integrated within libraries but new models are emerging</td>
<td>Fragmented. Some well-integrated communities of practice.</td>
</tr>
<tr>
<td>Organization</td>
<td>Primarily top-down, careful controlled entry from participating organizations</td>
<td>Varies: top down, bottom-up, middle out; often individual contributors</td>
</tr>
<tr>
<td>External integration</td>
<td>Very limited: High barriers to entry, few simple API’s</td>
<td>Varies, but more open. Some services offer simple open API’s; integration with web 2.0 protocols (e.g. OpenId)</td>
</tr>
<tr>
<td>Works Covered</td>
<td>Primarily books &amp; other works traditionally catalogued by libraries</td>
<td>Journal articles; Grants; Datasets</td>
</tr>
<tr>
<td>People covered</td>
<td>Authors and people written about represented in the library catalogs</td>
<td>Authors of research articles, fundees, members of research institutions – international</td>
</tr>
<tr>
<td>Key record criterion</td>
<td>Persistent and unambiguous identifier with a preferred label for the community served</td>
<td>Persistent and unambiguous identifier for an individual contributor</td>
</tr>
</tbody>
</table>
Some overlaps
Researcher Identifier Information Flow
Some emerging trends:

• Widespread recognition that persistent identifiers for researchers are needed
• Registration services rather than authority files as a solution for researcher identification
• Interoperability between systems is increasing:
  o ISNI & VIAF interoperability
  o ORCID and ISNI coordination
  o Research information system integration with ORCID, ISNI, VIVO
Early adopters

“More than a third of contributors in Books In Print have an ISNI” (ProQuest press release, 4 May 2014)
Adoption trends: Funders
Adoption trends: Universities

Assigning ORCID to authors when submitting electronic dissertations in institutional repositories

Pilot to automatically generate preliminary authority records from publisher files

Assigning ISNI identifiers to their researchers.

Assigning local identifiers to researchers

Enabling ORCIDS to be linked to university personnel profiles

Integrating ORCID into VIVO open source research profiling system, used by over 100 institutions
Key recommendations

**Researcher:** Get persistent identifier (earlier in career the better) and use it on all external communications.

**Librarian/University Administrator:** Assign persistent identifiers to authors if they don’t already have them.
- Retain traditional identifiers (e.g., VIAF IDs)
- Ensure ISNI or other ID for organization is accurate
- Advocate benefits and reasons for using and disseminating identifiers
Manage risks

• Environment is evolving
  o Funder mandates and policies are incomplete
  o No dominant business model
  o Incomplete adoption, no single comprehensive data source
  o Integration between classic and new name authority is lacking

• Researchers ...
  o will not drive change alone.
  o are sensitive to who controls their profile, and how information can be “corrected”.

• Incentive mechanisms, well‐timed nudges, setting norms with junior scholars, and establishing information feedback loops are critical.
Choosing identifiers

• Broad Researcher Identifiers: ORCID & ISNI
  o National mandates
  o Capabilities
  o Usage patterns

• Retain traditional identifiers: VIAF, NACO
  o Well supported in library systems
  o Primarily describe authors of books and similar works

• Be aware of community identifiers for local integration (e.g. ArXiV)
ISNI & ORCID

Complementary systems with two different approaches

**ISNI:** Consolidate data from multiple databases

**ORCID:** Researchers self-register

Share two goals:
1. Assign and share identifiers so both databases have only one identifier for a specific person.
2. Share publicly available metadata.

Coordination:
- ISNI allocated range of identifiers for ORCID’s exclusive use
- ORCID using ISNIs for organizations
- Developing interoperation: consult ISNI database during ORCID registration

From: ISNIs for researchers 2013-09
http://www.isni.org/filedepot_download/126/345
Report just published!

Plus supplementary datasets:

- Use case scenarios
- Functional requirements
- Links to 100 researcher networking and identifier systems
- Characteristics profiles
- Mapping of profiles to functional requirements
- Researcher identifier information flow diagram

Questions? Your plans?

http://oclc.org/research.html
Karen Smith-Yoshimura
Program Officer

smithyok@oclc.org
@KarenS_Y