

Researcher Identifiers, Metadata and Scholarly Communication in the Linked Data Environment:

The Implication for University of Tennessee, Knoxville.

Anchalee Panigabutra-Roberts

apanigab@utk.edu Twitter @apanig #OSC2018

Introduction

The move to linked data and the popularity of ORCID gained attention from the library community in the U.S. and worldwide. The shift from the national authority control to the global and web-based identity management with linked data capability has been explored (PCC Task Group on Identity Management in NACO, 2016).

Researchers continue to use different platforms to distribute their scholarly and creative works and to establish their researcher profiles (Tran & Lyon, 2017; Smith-Yoshimura, K. and others, 2014) with their identifiers in multiple systems. For an institution to collect researchers' information for the archive, the metadata for such collection needs to be robust and inclusive of the metadata elements from these diverse systems.

This study aims to evaluate different researcher identity management systems and their metadata in order to design the metadata model that can accommodate the data and metadata elements from these systems.

The Question

What should be the metadata model for researcher identity profiles in the linked data environment, if to be implemented at the University of Tennessee, Knoxville?

The Study's Population

Researcher identifiers systems used in the United States with the following selection criteria:

- 1) Academic and other researcher profiles
- 2) Multidisciplinary systems
- 3) Relevant to the United States, with international coverage in their scopes
- 4) Based on diverse types of library materials
- 5) Both professional and self-registered services (Panigabutra-Roberts, 2015),
- 6) Both closed and open data platforms
- 7) Excluding discipline-specific systems

Selected Researcher Identifier Systems:

LCNAF, VIAF, ISNI, ORCID, SCOPUS, ResearcherID, Google Scholar, Microsoft Academic, ResearchGate, Academia.edu, Mendeley, LinkedIn, Symplectic Elements and PIVOT.

Data Collection and Analyses

- I. In Excel spreadsheets, metadata elements in the selected researcher identifier systems were compiled and analyzed for the key metadata elements.
- II. Each system and its metadata elements were also evaluated based on the FAIR principles (Wilkinson, M. D. et al., 2016); from each system's purpose, its organization/corporate ownership and business model, its type of service, data access, data creation process, sources of data, metadata guidelines, APIs/interoperability, linked data capability, citation metrics and altmetrics, social media to other special features.

Figure 1.
Researcher ID Services by Type

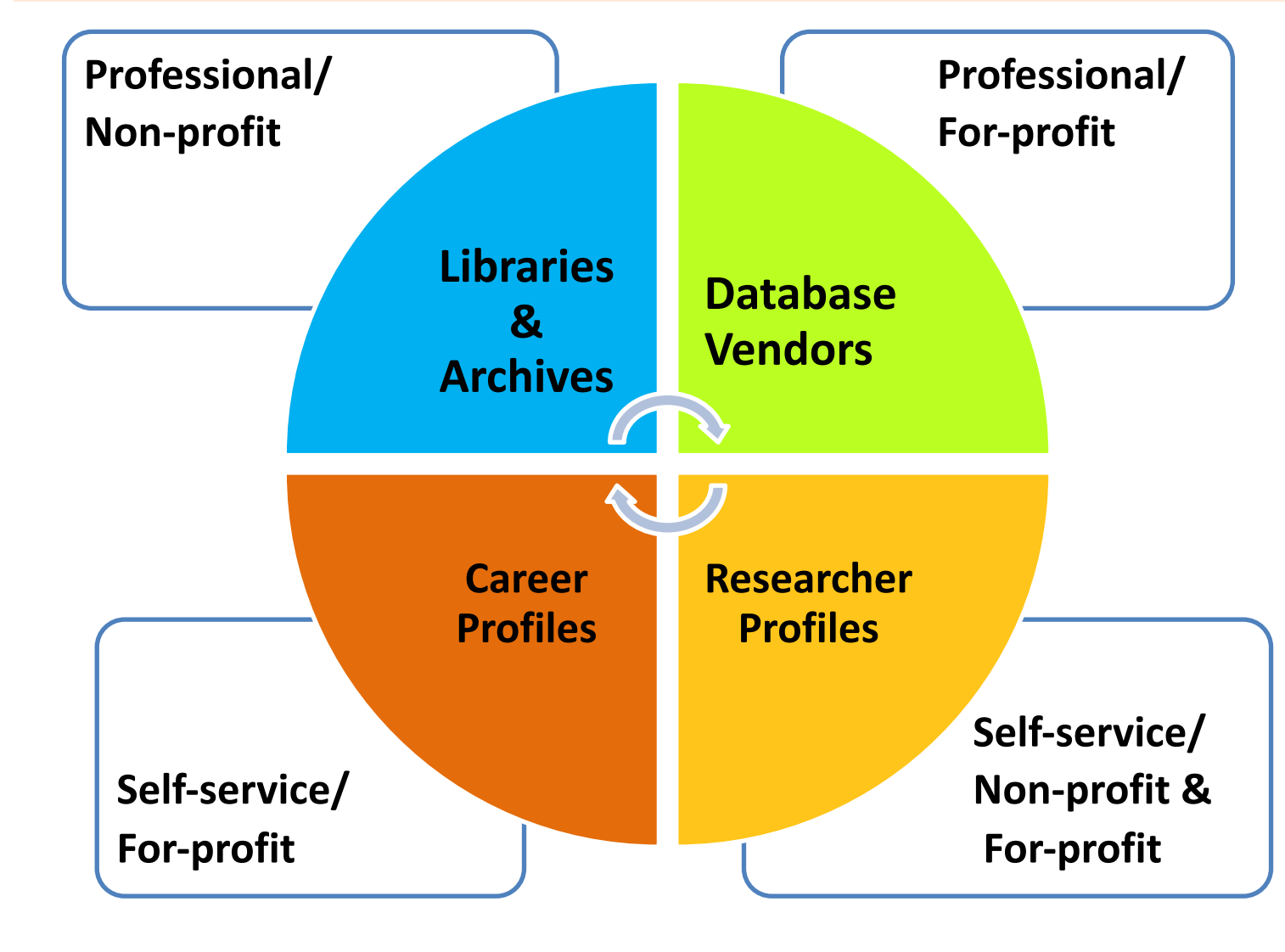
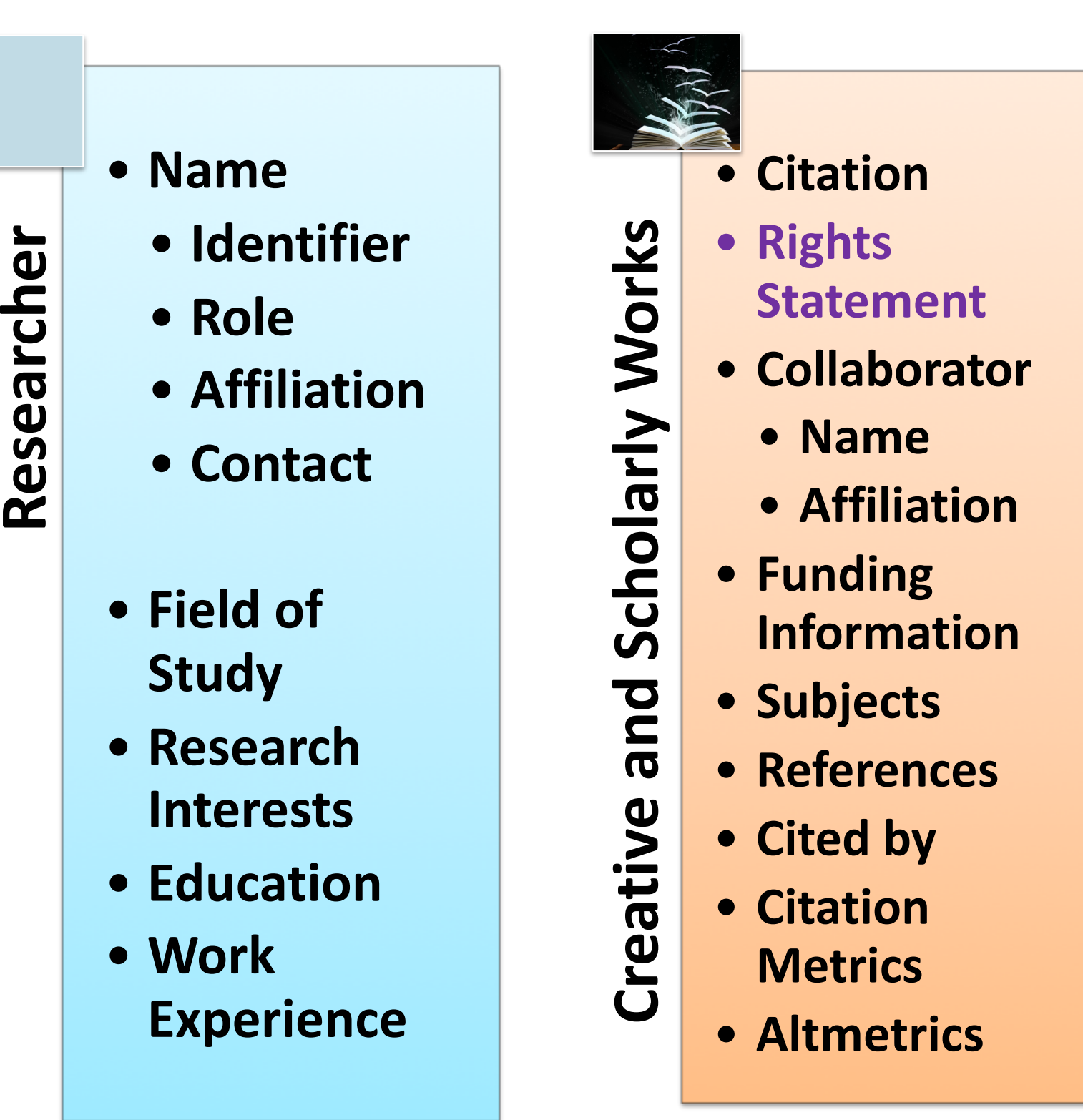


Figure 2.
Key Metadata Elements



Preliminary Findings

- I. The metadata elements in each system vary based on the type of service (professional or self service), the type of organization hosting the service and its business model (not-for-profit or for-profit model) (figure 1). These typologies are needed to consider:
 - a. Data access and rights (of the researchers for their data within these systems and of libraries and archives for archiving these data)
 - b. Openness of data based on the FAIR principles (Wilkinson, M. D. et al., 2016)
 - c. Interoperability among systems (their agreements on data sharing)
- II. Access to the services also varies:
 - 1) Open data (no log-in required)
 - 2) Semi-open data (public view on the web with login to access the full data)
 - 3) Completely closed to the service subscribers only.
- III. Linked open data services are only available in the systems hosted by OCLC (LCNAF, VIAF and ISNI).
- IV. Other systems are cross-linked via APIs (ORCID, SCOPUS, Mendeley, ResearcherID, Elements and Pivot).
- V. Citation counts, metrics and/or social media features are only available in some for-profit services (SCOPUS, ResearcherID, ResearchGate, Google Scholar, Microsoft Academic, Mendeley, Academia.edu and Elements).
- VI. Combined among the selected systems, the top-level metadata elements are in figure 2.

Questions to Explore Further:

1. How do the business models of researcher ID systems impact researchers' and their institutions' rights to access, use and preserve the profiles and associated data?
2. How can rights statements be applied to researcher profiles?

Copyrights

Template © Copyright Colin Purrington
Content © Copyright Anchalee Panigabutra-Roberts
2018

Literature cited

LC PCC. (2016). *Charge for PCC Task Group on Identity Management in NACO*. Washington, D.C. Retrieved from <https://goo.gl/7Jmwk6>
Panigabutra-Roberts, A. (2015). *Researcher Identity Management in the 21st Century Networked World: A Pilot Study of American University in Cairo Faculty*. Semantic Web in Libraries Conference. Hamburg, Germany.
Smith-Yoshimura, K., Altman, M. and others. (2014). *Registering Researchers in Authority Files*. Dublin, Ohio: OCLC

Research. Retrieved from <https://goo.gl/Avu23u>
Tran, C. Y. & Lyon, J. A. (2017). Faculty Use of Author Identifiers and Researcher Networking Tools. *College & Research Libraries*, 78(2): 171-182. <https://doi.org/10.5860/crl.78.2.171>
Wilkinson, M. D., Dumontier, M. and others. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Nature*, 3: Urn:issn:0028-0836.

Acknowledgments

Conference Travel Support From:

- University of Tennessee, Knoxville. Office of Research and Engagement. Scholarly Activity and Research Incentive Funds. Foreign Travel Fund.
- University of Tennessee Libraries.