Title
The Social Aspects of Identity Management Work

Permalink
https://escholarship.org/uc/item/2074b7cx

Journal
Technicalities, 40(3)

Author
Riemer, John J

Publication Date
2020-05-01

The Social Aspects of Identity Management Work

Technicalities Column

March 9, 2020

by John J. Riemer

Head, UCLA Library Resource Acquisitions & Metadata Services

Introduction

One book had sat in my subconscious for over a year, *Ethical Questions in Name Authority Control*, edited by Jane Sandberg.[1] Being philosophically inclined, I was waiting for a relaxing bit of quiet time to take in and weigh the various viewpoints. This column is not a book review, but an effort to engage with the some of the stimulating ideas found in the book.

The Power of Naming in Authority Work

The book's editor states that the performance of name authority work "gives catalogers a very specific type of power over the people they describe."[2] For very many years, the only visible product of name authority work was a name *heading*. That heading very often deviated from the form the person would record on a sheet of paper with a signature, particularly because of inversion of the name's elements. The rules called for selecting the best entry element, to facilitate the type of browsing seen in past printed telephone directories. Additional elements were added to the headings, since that text string was the sole means of differentiating similarly-named people.

The overall effort in authority work was never motivated by desire to issue pronouncements about individuals. It was aimed far more at connecting users to the resources they might be seeking. Normalizing the varying forms of names for the same person across various publications was an attempt to save the time of the reader.

I still remember the admonitions to catalogers to be mindful that an authority record was not meant to be a reference work. The product of the effort was more like a decision record, carrying just enough information to document the basis of the decision. Bibliographic records and classification got public visibility when a cataloger's work was done, but the authority data stayed in the back room, indirectly guaranteeing consistency in future cataloging data for additional works by repeat authors. If there was any power inherent in selecting among variant forms of name used by an author, it was a very modest amount.

With the emergence of "identity management," the perceived primary product of authority work is the identifier, not the text string of the heading.[3] A byproduct of this change is that it actually has become more possible for catalogers to honor requests to change the representation of a person's name in library metadata; we will increasingly be able to let a person curate a preferred name form in metadata displays. Under identity management, the identifier is the chief means for connecting users to information connected to a particular person, while the form of name provides confirmation of which person the identifier represents.

Authority Work as Storytelling

Is the addition of entity attributes to authority data, which commenced with the arrival of RDA,[4] tantamount to telling a story about a person? The new MARC data elements[5] cover such aspects as the place associated with person's birth, death, or residence; the area of expertise and the occupation in which the person has worked; the person's institutional affiliation; gender; the language used in the person's creative activity; and the demographic groups to which a person belongs.

For much of my career, catalogers have looked longingly at the prospect of seeing just a little bit more of the contents of an authority record getting included in an online catalog. Would it be too much to ask to have the variant forms of name included as "see references" in the online catalogs we offer to our users?

With the new entity attributes just starting to show up in authority data, there has been no systematic effort to retrospectively cover the past four decades of authority records that have accumulated in the LC Name Authority File. The different data elements are snippets of a person's life, but even that much data seems far from appearing in today's online catalogs for any searching or display. Those catalogers taking the time to include the data are doing so on the promise of its coming into useful service within a linked data environment.   The standalone Library of Congress Authorities database is one place the entity attributes can be viewed currently.[6]

A long-standing practice in the description of manuscript collections has been the inclusion of a "biography/history note" in the bibliographic record corresponding to a finding aid. The MARC 545 fields can contain a rather full-fledged life story,

composed by an archivist or cataloger.[7] These go well beyond the recording of bits
of factual information in an authority record!

With the new entity attributes in mind, Sandberg asks "Which stories should a
cataloger share with other library workers and patrons? How can catalogers know
that a story is not theirs to share?"[8] If libraries, seeking to develop a more diverse
collection, acquire material from outside the canon, why would they *not* want to tell
the story of an author, or at least identify the person?

Those data elements in separate fields, accurately recorded, are mere snippets of a
person's life and they have enormous potential to serve new discovery roles. When
I first got a chance to contribute authority data to the NACO program in the late
1980s, I noticed the unsystematic presence of what were called "subject-to-name
see also references," and some sad guidance to delete such data.[9]   As a serials
cataloger at the time working with corporate names, I thought it was useful to be
able to identify those that represented hospitals in Minnesota or student societies in
Austria. With linked data discovery needs ascendant, it was heartening to see
authority data like that make a stunning comeback.

Being able to answer a query about who are (some of) the 20th century women
composers seemed like such a compelling use case for those entity attributes. I
wrote in this space in 2012 about the Text Encoding Initiative work of humanities
scholars including not only mark up of proper names but also strongly desiring to
embed demographic data about those names. Saving the time of researchers, the
data could instead be referenced directly from our authority data.[10] If all of the

researchers at one's home institution were represented in the authority data, complete with the applicable entity attributes, think of how that might be used to harness and track the extent of local scholarly output—by department, discipline, or the entire campus.

Privacy Concerns

"Our desires to facilitate retrieval most effectively, to give intellectual credit appropriately, and our inclination to engage in increasingly invasive detective work to saturate the values and connections in an authority record, all generally come from an urge to serve users and maintain a rigorous and accurate cultural record."[11] It is odd to see the implication that something in which catalogers have long striven to gain excellence and which has gained libraries significant credibility is seen as a compulsion in need of being curbed. Its contrast can be seen in a tagline OCLC currently uses to describe its purpose, "Because what is known must be shared."[12]

Kazmer specifically faults the cataloging practice of tracking the relationship between a person's real name and the various pseudonyms they use. "The fundamental issue of violating authors' autonomy by revealing, connecting, or keeping persistent connections between multiple identities and names has the potential to affect everyone."[13] "In order to accurately reflect people's identities, NACO contributors should err on the side of contacting creators whenever there is a concern about privacy issues."[14] Collecting information about the context surrounding the creation of a work is quite intensive, only to have it sometimes lead to suppression of information in the end.   An effective way to create obscurity for a work would be to include it as part of a collection-level record for other material in

the same genre. If there is doubt about the appropriateness of handling the work of

a living person, at the same time there are productivity demands, it would always

be possible to leave it in a backlog until such time as the discovery context is more

certain.

Personal Details

Sandberg touts the need to ask "questions about how necessary is it to record

gender at all in various contexts."[15] "The literature critical of RDA's gender attribute

and its use in NARs raises legitimate concerns regarding personal privacy and

safety."[16]

The percentage of biographies in Wikipedia covering women subjects was an

incredibly low 17.67% as of 2018.[17] The Women Red Project, a volunteer effort

organized to address the imbalance took its name from the "redlinks" pointing to

non-existing articles in Wikipedia. The percentage of Wikipedia editors who are

women is even worse. First systematically surveyed by the Wikimedia Foundation

in 2011, the figure was only 9%[18] and strenuous efforts to increase the proportion

have not resulted in noticeable change. If the gender of the authors and subjects is

not even being recorded, how would we ever be able to measure the degree of

success of any efforts made to remedy the situation?

Martin writes, "If we wish to act ethically, we must acknowledge that the current

use of dates of birth in name authority records violates privacy ethics by publicly

exposing the personal information of living subjects of authority records without

those subjects' consent."[19]   Until recently, inclusion of a date was one of only two

available standard means of differentiating one person from another, the other being the addition/expansion of an initial. With the change to identity management, it becomes less necessary to blatantly feature a date of birth in the heading seen in catalog browse displays and access points. Differentiating data like this can reside in separate data elements in a record, displayable (or not) as warranted.

The various quotations I have included in this column all suggest that we would be able to achieve a consensus and define a best practice we could all adhere to in the library community, if we set our minds to it.   However, we are moving away from an environment where metadata creators work in a separate silo and can set expectations among our colleagues.   We no longer work in isolation.   There are many other parties we are expecting to collaborate with, in the form of building on the portions of metadata we can obtain from their efforts.

Conduct a search of Whitepages.com and you see the address and telephone number historically offered by phone directories; quite often you receive and added bonus of the person's approximate age, to assist with differentiation.   Ancestry.com also contains directories that often also provide birthdays along with birth years. The ISNI database[20] contains rights management data, where real names and pseudonyms are routinely connected, enabling royalty payments to reach the appropriate individual. The Internet Movie Database (IMDb) is a tool offering not only retrieval of titles of films and television programs, but also (authority record-like) data about actors, actresses, and directors. Wikidata will very likely continue to contain any entity attributes that NACO contributors might be convinced to stop recording.

Conclusion

We no longer work in isolation. We are not the only game in town. We cannot

realistically expect to set the all rules for the metadata we use. Walking away from

providing the metadata users expect to find amounts to unilateral disarmament.

People will just go elsewhere for what they need.

We are becoming part of an increasingly interrelated metadata ecosphere. In the

past year, I had the opportunity to contribute this segment of the PCC Position

Statement on RDA:

Part of the benefit of changing over to creating linked data natively is that we would be able to collaborate with others working in the same or a similar schema, such that we could frequently build on the metadata provided by others, as opposed to create it all ourselves in the library community according to standards that are relatively unique to our community. Strategically, we will come out much better if we go the route that enables lots of collaborating with others and frequent capitalizing on the efforts of others. We will end up with fuller metadata on a much greater array of resources. An implication of this practice is that we may well end up with metadata whose degree of RDA influence is significantly less than what we would have if we created the metadata largely ourselves or if we dictated that any collaboration with us had to involve RDA adherence.[21]

[1] Sacramento, California: Library Juice Press, 2019. ISBN 9781634000543
https://worldcat.org/oclc/1048278063

[2] Introduction by Jane Sandberg. *Ethical Questions in Name Authority Control,* page 2

[3] John J. Riemer and Violeta Ilik. "'Authority Control' see (also) 'Identity Management'." Part of the American Library Association LITA-ALCTS/CaMMS Authority Control Interest Group program, ALA Midwinter Meeting, Atlanta, GA, January 22, 2017.   https://connect.ala.org/viewdocument/lita-alctscamms-authority-control Slide set: https://connect.ala.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=eb4d9365-3554-4261-b6aa-6b35e1560654&forceDialog=0   (see definitions on slides 3-4)

[4] *Resource Description & Access: RDA*, developed in a collaborative process led by the Joint Steering Committee for Development of RDA (JSC). Chicago: American Library Association, 2010-

[5] Library of Congress. Network Development and MARC Standards Office. MARC 21 Format for Authority Data, 1999 edition. Headings – General Information http://www.loc.gov/marc/authority/ad1xx3xx.html (fields 368-378, 386)

[6] Library of Congress. Library of Congress Authorities. https://authorities.loc.gov/ For an example of a public display of the entity attributes for an individual, Ron Brookmeyer, see https://lccn.loc.gov/n87807315

[7] Compare the note in the MARC record for the Maxine Glorsky papers relating to Martha Graham, 1940-2019 (bulk 1944-1981) https://lccn.loc.gov/2017563970 (422 words) with the Biographical Note in the finding aid itself https://findingaids.loc.gov/exist_collections/ead3pdf/music/2020/mu020001.pdf (738 words).

[8] *Ethical Questions in Name Authority Control,* page 2.

[9] Library of Congress. Descriptive Cataloging Manual Z1, Update 2-3 (August 2009), page 5XX, 1: "Routinely delete all subject-to-name references (5XX …) when changing a record for another reason. (These references are no longer made; since existing references have not been maintained, they may not be in accord with current subject cataloging policy.)" https://www.loc.gov/aba/publications/FreeDCM/DCM_2009-0203.pdf

[10] John J. Riemer. "New Prospects for Library Authority Data." *Technicalities*, v. 32 no. 4 (July/August 2012) page 7. Online version viewed on Ebscohost Library Literature & Information Science Full Text.

[11] Michelle M. Kazmer. "Identity Theft: How Authority Control Undermines Women's Agency." In Jane Sandberg, editor. *Ethical Questions in Name Authority Control,* 2019, page 40.

[12] George M. Eberhart. "OCLC: Purpose, Place, and People: Company Holds IFLA Industry Symposium." American Libraries, August 15, 2016. https://americanlibrariesmagazine.org/blogs/the-scoop/oclc-purpose-place-and-people/

[13] Michelle M. Kazmer, page 40.

[14] Violet B. Fox and Kelly Swickard. "'My Zine Life Is My Private Life': Reframing Authority Control from Detective Work to an Ethics of Care." In Jane Sandberg, editor. *Ethical Questions in Name Authority Control,* 2019, page 17.

[15] Introduction by Jane Sandberg. *Ethical Questions in Name Authority Control,* page 3.

[16] Thomas A. Whittaker. "Demographic Characteristics in Personal Name Authority Records and the Ethics of a Person-Centered Approach to Name Authority Control." In Jane Sandberg, editor.

Ethical Questions in Name Authority Control, 2019, page 60.

[17] Melody Kramer. "Women in Red Is Changing Wikipedia's Coverage of Women, One Article at a Time." Wikimedia Foundation news release, 18 October 2018. https://wikimediafoundation.org/news/2018/10/18/women-in-red-wikiproject/

[18] "Gender Bias in Wikipedia," https://w.wiki/347

[19] Jennifer M. Martin. "When Public Identity Meets Personal Privacy: Ethical Considerations for the Use of Dates of Birth in Name Authority Records for Living Persons." In Jane Sandberg, editor. *Ethical Questions in Name Authority Control,* 2019, page 54.

[20] International Standard Name Identifier (ISNI). http://www.isni.org/

[21] Program for Cooperative Cataloging Policy Committee. "PCC's Position Statement on RDA." August 2, 2019 http://www.loc.gov/aba/pcc/rda/PCC%20RDA%20guidelines/PCC-Position-Statement-on-RDA.docx