Chapter 16

Authority Work as Outreach

Tina Gross and Violet B. Fox

Introduction

Unlike any other resource, the National Authority File (NAF) is the result of decades of research conducted by librarians from the Library of Congress and those participating in the NACO (Name Authority Cooperative) program. The exacting standards in place for creating and editing name authority records have ensured that the quality of the data in the NAF is consistently high, especially when compared to the varying quality of bibliographic data in OCLC's WorldCat.

However, librarians have not necessarily been interested in collaborating with those outside the library world. In the process of creating personal name authority records, NACO participants usually do not attempt to contact people for whom authorized names are being established, even when the person is known to be alive and contact information is readily available. NACO training does not include much information about contacting creators; the training material available at https://www.loc.gov/catworkshop/courses/naco-RDA/index.html (last updated in May 2018) does not mention it. Emailing creators is framed as a last resort, to be done only when additional information (such as year of birth or middle initial) is needed to differentiate that person from others with the same name.

This chapter argues in favor of reframing authority work as a collaboration not just between librarians, but between librarians and the subjects of authority work. By reframing our work in this way, we shift our role from agents following opaque rules and gathering information on other people to partners mutually working towards accurate information.

Contacting Creators in the Later Years of AACR2

Before the adoption of RDA, when catalogers participating in NACO contacted a creator whose name they were establishing, it was usually because there was a conflict to break. If the creator's name was the same as that of one or more people already in the authority file, and no information to add to the heading was present in the resource being cataloged or readily available on the web, the cataloger would typically attempt to contact the creator to ask for such information.

Another common dilemma that would prompt a cataloger to contact a creator was when they could not determine whether the creator of the work in hand was the same or different from a previously established personal name in the authority file. As one of this chapter's authors once wrote to such a creator:

> In addition to "Cell tower deaths" and "An optimist in Haiti" (which are yours, yes?), the following works are attributed (I believe incorrectly) to the same Travis Fox: A 2007 golf instruction video titled "How to beat the bogey man," and a 2005 book on hypnosis titled "You want me to look where?" Those are not you, correct? They don't seem to fit in with your other work, but hey, you never know (in doing similar queries, I have sometimes been surprised to have the person confirm that they actually were indeed one and the same.)[1]

The need to contact creators occasionally to ask for conflict-breaking information was mentioned in NACO training, but how to go about doing it was not usually covered. Many catalogers felt the need for

1. Tina Gross, e-mail message to Travis Fox, July 26, 2012.

training and guidance, such as in Ohio's NACO funnel, where "catalogers mentioned the need for help with researching and contacting authors. This was eventually covered informally with their reviewer. Building that into the training would have saved time, and empowered them to take action quickly."[2]

Catalogers who contacted creators on a frequent basis often developed scripts or templates to use, which gave a basic explanation of library authority control practices as background to the request for information. These templates typically endeavored both to establish the legitimacy of the request (that it was a genuine request coming from a real library, not some kind of phishing attempt) and to introduce the creator to the basics of name disambiguation. They have been a topic on cataloging listservs on several occasions, with catalogers discussing ways of getting creators to respond and sharing their templates.[3] Examples are included in **Appendix I** at the end of this chapter.

In the authors' experience of contacting creators via email, the majority responded promptly with the requested information, often expressing gratitude to the library for adding their work to its collection or for making the effort to verify their name. Sometimes catalogers would receive an unusually enthusiastic or helpful response. For example, one of this chapter's authors emailed the producer of a film about Ghanaian dance to ask about contacting the featured dancer/choreographer to confirm the preferred form of his name. The producer responded by phone almost immediately to provide a short lesson on honorifics used in Ghana.

Occasionally, catalogers received responses that doubted the need for authority control with questions along the lines of "Wouldn't someone looking for me easily be able to distinguish me from the imposters?"

2. Melanie McGurr, Catherine Mason, and Michael Monaco, "Public and Academic Library Cataloging Collaboration in Ohio's NACO Funnel Project," *Cataloging & Classification Quarterly* 51, no. 1-3 (2013): 76.

3. Autocat list, "Name inquiries that work!" Listserv thread beginning February 27, 2008, https://listserv.syr.edu/scripts/wa.exe?A2=AUTOCAT;b5cdafae.0802D and PCCList, "Writing to authors to disambiguate their names," Listserv thread beginning June 11, 2013,

Unfortunately, errors that can result from the absence of systematic disambiguation are not difficult to demonstrate. The following email to a creator outlines one example of a tangle of conflicting metadata found on commonly visited reference sites:

> There's a lot more confusion in cases like this than you'd think. For example, Wikipedia has a page on American Primitive... You're listed as the producer, and your name is hyperlinked, but if someone clicks on your name, they're taken to the page for the Australian painter of the same name. In the IMDB, the record for you indicates that you were a writer for SNL… I couldn't be sure if that was true or not - it didn't really seem to fit with the rest of your body of work, and it seemed a possibility that you were being confused with the other James Egan who's an SNL writer (but he spells his name Eagan).[4]

It takes time and effort to identify disambiguation problems, but in most cases using such issues to make the case for authority control would not require much additional time, because the cataloger would have already encountered them when researching the name and attempting to complete the name authority record.

Under AACR2, it was possible to include a creator's contact information, work history, gender, birthday, or other personal information in a personal name authority record. However, it could only be recorded in free text form in a note (not encoded in dedicated fields), and was not a common practice.

The Move to RDA

Changes to personal name authority work that came with the transition from AACR2 to RDA can be put in two broad categories. First, RDA rules provide for many additional methods of disambiguating a personal name, making it much easier to break a conflict with information provided within a resource being cataloged. Second, personal name authority records in RDA contain significantly more biographical

4. Tina Gross, e-mail message to James Egan, March 12, 2008.

information about creators. Because of the expanded role for authority records put forward by IFLA with the adoption of the Functional Requirements for Authority Data (FRAD), "the original scope of a name authority record was broadened from simply that of a carrier of an authorized heading or access point to a description of an entity with the development of an expanded list of attributes that can be included in name authority records."[5]

Additional methods of disambiguation available in RDA have reduced the need to contact creators for information. AACR2 rule 22.19 and the associated Library of Congress Rule Interpretation did allow a personal name to be qualified with a "descriptive phrase" that could be an occupation/profession in order to break a conflict, but only as a last resort to avoid creating an undifferentiated record. In RDA, adding the occupation or period of activity to an authorized access point to differentiate otherwise identical names is a routine practice. It is almost always possible to identify a piece of information to use as a qualifier, so catalogers rarely need to contact creators to ask for information to disambiguate their names.

While the increased flexibility in disambiguation under RDA has meant that catalogers don't need to contact creators as often as before, the proliferation of attributes that RDA calls for recording in personal name authority records raises issues that should compel catalogers to contact creators more frequently. These attributes include associated place (place of birth, country associated with the person), address, field of activity, affiliation (workplaces and organizations), gender, and language.

The practice of attempting to determine and record a creator's gender has drawn a great deal of concern and criticism. Catalogers attempting to do so run the risk of misidentifying a person's gender based on false assumptions, or of outing a person as transgender without their knowledge or consent, thus actively causing harm.

5. Kelly J. Thompson, "More Than a Name: A Content Analysis of Name Authority Records for Authors Who Self-Identify as Trans," *Library Resources & Technical Services* 60, no. 3 (2016): 140, doi: 10.5860/lrts.60n3.

In 2014, Billey, Drabinski, and Roberto explained how the RDA rule on recording gender and its interpretation by LC codified a mistaken idea of gender as a fixed binary in which all people experience gender as straightforward and fixed (or changing only in linear fashion).[6] Interestingly, their argument includes an account of cataloging a work and creating a personal name authority record in the presence of its creator, and how that interaction impacted the record.

> But when we reached MARC tag 375, the field for the RDA element for gender, the author was confused about why that information mattered. This author did not feel comfortable disclosing and codifying gender in the authority record… Interestingly, this author was a cisgender woman with a name that our culture commonly reads as female; "cisgender" is a term used to refer to non-transgender people who identify with the gender that they were assigned at birth. Most catalogers would simply transcribe the gender as female from a glance at the title page, the author's biographical information, or physical appearance. But the author explained that gender was simply not an important aspect of the work in question, or the body of work the author intended to produce. Indeed, the author expressed hope that one day gender would no longer be a social marker. It made no sense that the Library of Congress (LC) would be interested in the author's gender; why would LC care about that! Once the cataloger explained that 375 was an optional field, we decided together to omit that information.[7]

This is a good illustration of how a cataloger cannot successfully anticipate or deduce a creator's views about what information about them should be shared. When consulted, a creator might express strong preferences that would otherwise remain unknown.

Gender stands out as the most-discussed potentially harmful attribute that can be included in RDA personal name authority records without the person's consent, but things such as name variations, date of birth, work affiliation, and address could be information that a person might not want to disclose or have permanently recorded as data that is publicly

6. Amber Billey, Emily Drabinski, and K. R. Roberto, "What's Gender Got to Do with It? A Critique of RDA Rule 9.7," *Cataloging & Classification Quarterly* 52, no. 4 (2014): 412-421, doi: 10.1080/01639374.2014.882465

7. Ibid., 413.

accessible online. The fact that personal name authority records can now take on the character of a miniature biography or CV is an indication that the person who is the subject of such a record should, whenever possible, be made aware of its contents and have input in what is included.

Authority Work as Outreach

Catalogers and metadata librarians are often "behind the scenes" workers; they may or may not have occasion to interact with library users, and users are often not aware of their efforts. Authority work presents a unique opportunity for catalogers to engage in user outreach and to demonstrate the value of the library's work in a tangible way. Just one email to a creator can establish that libraries see their work as important and relevant, and highlight the fact that librarians are working to represent information about them accurately and facilitate access to their work.

Although creators might not be familiar with terminology like authority control or disambiguation, they care about getting credit for the work that they've created. It is easy enough to explain the value of authority work using examples of bad metadata, such as that often found for less prominent authors on Amazon.com. Creators are interested in people being able to easily find their work, and the value that authority control provides can be viewed as just as worthwhile to creators as, for example, the promotion of popular books by public services librarians.

To assist librarians contributing to the NAF in reaching out to creators, the Program for Cooperative Cataloging (PCC) could create a landing page which provides information about the goals of authority work and examples of how the information collected by librarians would be used in disambiguating identities. The page could contain frequently asked questions about who uses the data collected and why certain data is more useful than others.

A landing page hosted at loc.gov would provide authority; after all, we often ask for personal information (such as birthdates and fuller forms of names). The fear of identity theft may be one reason when

creators do not respond to librarians' requests for information. Being able to provide a link to a website which explains why librarians are looking for specific information may help convince creators to share their information. Of course, this landing page should fully explain that the information given will be made public and used in tandem with other linked data databases. Only when equipped with this information will creators be able to give fully informed consent when providing their personal data. In addition, a landing page could provide more detail on the purpose of a library authority record than can be explained in a short message; that is, a NAR does not necessarily document everything someone has ever created, but shows the range in topics that a person has produced within, as well as collocating differing forms of names for the same person.

When an institution has a collection development focus on materials in particular foreign languages or from particular locations, it may or may not be able prioritize employing catalogers with corresponding advanced language skills. It is not uncommon for catalogers to possess "bibliographic knowledge" of several languages, which allows them to catalog resources in those languages but to not be sufficiently fluent to feel comfortable contacting creators about their name authority records. This situation can cause a discrepancy between intended institutional priorities and the reality of readily making materials available to users. While this is a larger issue than the creation of name authority records, having help in the form of a landing page explaining concepts in multiple languages would help catalogers at many levels of expertise.

Even more important than explaining why librarians are collecting personal information, it is crucial for LC and PCC to provide a way for people to correct the information found on an authority record. Currently, there is no easy way for creators to know who is collecting information about them or how they might go about removing or correcting information. Upon finding erroneous information about themselves in a library catalog, they may try to contact their local library, which may or may not have the staff resources to address their concerns or to truly understand where the information in their catalog is coming

from. An easily found form could help creators feel they have recourse when the information that libraries have about them is incorrect.

As a proliferation of data is linked among various sources, inaccurate data is nearly impossible to corral and correct once it has been distributed online or published. Contacting creators when initially creating an authority record is the most effective way to ensure that the information about them is accurate and consists of data which they feel is appropriate to share with the world. The purpose of this outreach work would not be simply to collect additional information to break conflicts but to ask the creator to confirm that the information they want to be present is available and that no information they want removed is kept. Information included or excluded according to creator's wishes should be noted in the NAR, and the provenance of data vetted by the creator themselves should be made explicit.

Catalogers or administrators may balk at the extra time and effort involved in emailing most or all the creators they establish authority records for. However, equipped with appropriate email templates and an explanatory landing page, there would be little need for back-and-forth emails. Catalogers could create spreadsheets where they keep track of who they have emailed and create records either when they hear back from a creator or after a certain amount of time has passed. There would be no need to keep a physical item from moving forward in processing during the wait; cataloging could be completed and the relevant information collected on the same form or as a pending draft in the bibliographic utility's save file. In the case that the workload of creating authority records is larger than a librarian would want to contact authors for, librarians can prioritize the work according to their institution's needs; for example, prioritizing contacting local creators or those creating work within the institution's collection specializations. Other priorities for creating authority records might include members of marginalized populations, such as people whose gender does not fit within a traditional gender binary, as the potential for harm in authority work may be higher than with cisgender people.

We suggest that contacting creators should be a recommended best practice, not necessarily a task required for every resource cataloged. Cataloging departments and librarians should develop policies based on their own resources and institutional priorities.

Related Efforts and Developments

In a recent survey, 66% of catalogers indicated that they use social media (such as LinkedIn, Facebook, Twitter or Instagram) to find information when creating authority records.[8] Viewing a creator's social media output is obviously not the same as contacting them directly, but the fact that catalogers are making it a regular practice to perform research for authority work on social media sites designed to enable users to connect with each other is significant. It suggests that the means to contact creators is frequently identified, even if it is not used to do so.

As discussed elsewhere in this volume, the Legislative Library of the Legislative Assembly of Nunavut and the National Library of Israel have developed approaches to authority work that involve contacting creators as a routine practice. The NLI provides forms for catalogers to send to creators to fill out, and then invites them to review the finished authority record in the OPAC or discovery layer. Creators who come into contact with the library for any reason are routinely asked "to review their authority record and contact us with corrections and more information."[9]

The OAQ (Online Author Questionnaire), developed by Harvard University Library, "is a library-hosted web application that automates publishers' processes for gathering author data and allows them to share a subset of that data with libraries to support the timely creation of

8. Ilda Cardenas, "Catalogers Cyberstalking Authors: Relationship Between Cataloging and Social Media," paper presented at ALCTS CaMMS Cataloging Norms Interest Group program at the American Library Association Midwinter meeting, Denver, CO, February 10, 2018.

9. Ahava Cohen, e-mail message to Violet Fox and Tina Gross, June 21, 2018.

discovery metadata."[10] Publishers receive support to use a customizable interface that facilitates asking authors to provide biographical and career information by filling out a web-based questionnaire. In turn, the library receives the data provided on the questionnaire as structured output to automatically create personal name authority records for those authors.

Efforts such as those put forward by the Nunavut Legislative Library, the NLI, and Harvard University Library are exemplary in making it simple for creators and publishers to contribute information which makes identification easier. Further study may be warranted to discover whether creators adequately understand how the data they provide is used by libraries.

Shifting Assumptions

When cataloging traditionally published books, librarians assume that the form of name on the title page is the author's preferred form. With non-book formats, decisions have been made about where to source the NACO-preferred form of name, but those decisions are somewhat arbitrary, and do not necessarily reflect an author's preference. As libraries move to collecting (and cataloging) less traditional materials, librarians often no longer have a title page or a clear preferred source of information about the work they are cataloging. When contact information is available, creators should be contacted to ensure their preferred form of name is established correctly.

While the move to linked data will place less emphasis on the label of the entity represented by an authority record, the issue of determining the human-readable form of name will not disappear. The shift will make name changes easier to deal with, as the URI will remain unchanging while the label can change as frequently as is required or requested. A desired future for authority control might look like an ORCID identity,

10. Online Author Questionnaire, accessed June 22, 2018, http://library.harvard.edu/oaq#.

where creators can claim their record and make edits to the publicly accessible label.

Regardless of how the technology involved handles the data of personal name records, with very few exceptions, the person who is being described should be viewed as the ultimate arbiter of any decisions to be made regarding the information collected about them. But in order to empower creators, we must first give them information about what data is collected about them. Contacting creators at the point of name authority record creation is the most effective way to ensure that authority work is truly a collaborative partnership.

Bibliography

Billey, Amber, Emily Drabinski, and K. R. Roberto. "What's Gender Got to Do with It? A Critique of RDA Rule 9.7." *Cataloging & Classification Quarterly* 52, no. 4 (2014): 412-421. doi: 10.1080/01639374.2014.882465.

Cardenas, Ilda. "Catalogers Cyberstalking Authors: Relationship Between Cataloging and Social Media." Paper presented at ALCTS CaMMS Cataloging Norms Interest Group program at the American Library Association Midwinter meeting, Denver, CO, February 10, 2018.

Harvard Library. "Online Author Questionnaire." http://library.harvard.edu/oaq#.

McGurr, Melanie, Catherine Mason, and Michael Monaco. "Public and Academic Library Cataloging Collaboration in Ohio's NACO Funnel Project." *Cataloging & Classification Quarterly* 51, no. 1-3 (2013): 72-81.

Thompson, Kelly J. "More Than a Name: A Content Analysis of Name Authority Records for Authors Who Self-Identify as Trans." *Library Resources & Technical Services* 60, no. 3 (2016): 140-155. doi: 10.5860/lrts.60n3.

Appendix I
Templates for Requesting Information from Creators

Template 1:

I am a cataloger at [Institution name], and I am cataloging [resource title]. Because there are multiple other people with the name [creator's name] in the Library of Congress name authority file, I need additional information to distinguish you. Will you let me know the year of your birth and/or your middle name?

This information will be used to contribute to the international authority record for your name and will help ensure that you get proper credit for the works you've contributed to.

Template 2:

Subject line: question about [title of work by creator]

My name is [XX], and I am the [position title] at [institution]. I am currently adding "[title]" to our library's collection and I'm writing to ask if it's possible to get a bit of information from you.

One of the things libraries do to make library catalogs more reliable and accurate is to create a unique "name heading" or access point for each author or creator, so that searchers can find all of their work without mistakenly retrieving things by other people with the same name. (The function is somewhat similar to "disambiguation" in Wikipedia, or the roman numerals used after names in the Internet Movie Database when there are multiple people with the same name, if you're familiar with either of those.)

I need additional information to establish such a unique heading for you. There are already several other people named [XX] in WorldCat, the global catalog of library collections where the record will be, and

so in order to distinguish you from all of the others, I need to add an additional piece of information to the heading.

The most common and simple way to do this is to add your year of birth, but a middle name or middle initial can also be used. If you would be willing to provide this information, it would be very helpful.