Repurposing RDA’s Descriptive Standards to Facilitate Humanities Research

Making a Case for Howard University’s “Portal to the Black Experience” and Similar Neotraditional Research Tools

Seth M. Kronemer and Andrew T. Sulavik

Seth M. Kronemer (smkronemer@law.howard.edu) is Archivist, Howard University School of Law and Associate Project Director, and Andrew T. Sulavik (andrew.sulavik@howard.edu) is Head, Metadata and Resource Description Services, Howard University Libraries and Principal Investigator and Project Director, the Portal to the Black Experience Project, Washington, DC.

historians need a great deal of imagination, invention, and time to devise research strategies, cull large repositories of digital and print resources, and locate the materials needed to substantiate their work. It is the job of librarians and archivists to demonstrate imagination and invention as they design, develop, and upgrade ancillary research tools that assist scholars; for humanities’ research methodologies are not static (as some may think) but dynamic, and in need of being ceaselessly adapted to serve changing needs.

Modern historians, and scholars from other humanities disciplines, conduct research, and then analyze, interpret, and describe aspects of past and present society. The most time-consuming but essential of these functions involves researching and analyzing data. Although the recent explosion of digital humanities resources has made some aspects of research easier, the resultant analysis and (at times) access have become unmanageable for scholars who now heavily rely on the growing number of databases and other digital materials.9

In 2005, Carole Palmer observed that humanities researchers view Internet and local digital collections “as one big digital blur of information, quite separate from personal or physical library collections.”3 The situation has only grown far blurrier for those humanities researchers who depend upon electronic resources to support their research.

CHAING AND THE GROUNDED THEORY APPROACH

Although different subjects within the humanities require different research methodologies, most humanities researchers (for the purposes of this article we are primarily discussing PhD-level students and faculty researchers) continue to follow a well-trod path of steps.4 In his grounded theory approach, David Ellis reduced this path to six tasks.5 In deriving this theory he surveyed “academic staff [researchers] from the departments of psychology, economics, economic and social history, geography, sociology, education and continuing education, and prehistory and archaeology” at the University of Sheffield.6 The tasks described in Ellis’s theory have not changed over time, although the approach of researchers to each task has shifted in light of the new digital research environment. The task of “chaining” merits our attention here.

Chaining, as defined by Ellis, involves “following chains of citations or other forms of referential connection between material[s].”7 Typically, researchers follow a paper trail to the leading authorities on the subject and their documented contributions. The next step is to review all the authorities that those scholars cited, read their works, and plod one’s way back until the chain of relevant sources is exhausted. This focus on the leading authorities on a subject decreases the time and effort a researcher must spend on preliminary research. The review of secondary materials also often rewards the researcher’s effort by revealing new primary materials, a growing amount of which now reside in digital repositories.

Researchers also participate in chaining by contacting authorities in their field through the social network of learned and professional societies. This can lead to additional contacts and referrals to other authorities in the field. For example, a researcher may be investigating the degree to which Samuel Johnson’s Dictionary was a creative work or a mere compilation of earlier works. For more to the point, the article provides the names of leading scholars who might provide more information.8 These names can also lead to multiple works of these authors.

Chaining relies primarily on the individuals who have written about a particular subject or witnessed historical events. Both are important to researchers, although the artifact—whether an article, book, or a primary document—is secondary to the process. The authority of the work derives from the individual who produced the work. It has never been a practice of the library profession to utilize this traditional approach of discovering the most authoritative sources when creating library catalogs, but instead it has built its descriptive metadata standards around three galvanized areas: subject, title, and author.

The Internet has become a transformative research tool that has changed the search for authorities and their works.
In the two decades since Ellis defined chaining that task has largely been replaced by netchaining. As termed by Suzana Sukovic, netchaining “combines aspects of networking, chaining, browsing, and Web surfing in a new pattern. [It] is about establishing and shaping online information chains that link sources and people.” Not only can netchaining lead to cited works with just a click of a mouse, but, according to Sukovic, “online chaining can widen to include communication with the author, whose contact details appear as part of the reference or the linked e-text.” As Sukovic concluded, “Netchaining is an important way of gathering information by following broad and unpredictable information paths.”

The Internet is a superior tool for chaining, one that is far better at the task than online library catalogs. To more fully understand netchaining as a research task, however, it is important to review how changes in the Internet are giving researchers ever greater access to information on persons through hyperlinks and other data related to those persons.

**THE SYMBIOSIS OF LIBRARIES, ARCHIVES, AND SOCIAL NETWORKS**

The cultural phenomenon of social networking via the Internet has altered expectations of users, including scholarly researchers. Online social networking allows the Internet user to find persons, and to identify their interests, institutional affiliations, circle of friends, and place of residence and employment. The impact of social networking and its emphasis on connecting persons and relators to those persons has not been lost on librarians, archivists, authors or their publishers. In libraries, this heightened awareness is evident in the adoption of a new descriptive metadata standard: Resource Description and Access (RDA). In archives, there is a corresponding metadata standard in the Encoded Archival Context—Corporate Bodies, Persons and Families (EAC-CPF). Only a few of these prototype research tools are utilizing these standards on the Internet. One research project, the Social Networks and Archival Context (SNAC) project, “has developed a social network web application that has helped the project earn the nickname ‘Facebook for the dead.’ The comparison with the ubiquitous social networking site is not without some utility. Social networking sites like Facebook enable users to articulate and make visible their social networks.”

Artists, authors, and their publishers have collaborated on creating descriptive metadata standards for persons. Organizations such as Open Researcher and Contributor Identification (ORCID) and the International Standard Name Identifier (ISNI) create disambiguated data about persons; in combination with unique, persistent identifiers, such data authoritatively relate the individuals to their works. Thus the cultural phenomenon of social networking, with its emphasis on persons and on relators to persons and their creations, has heightened user awareness and expectations of the benefits to be derived from person-centered searches.

While traditional humanities research was dependent on subject and bibliographic searches through the use of subject headings in catalogs and printed subject area indexes, contemporary researchers may now begin their research by identifying a key individual and then directly viewing works by or about him, related online archival finding aids, and relevant digital archival collections (i.e., netchaining). This process requires accurate, authoritative, and disambiguated data that allow researchers to effectively and rapidly identify the chain of authors/creators and their related creations of books, articles, and archival collections.

**THE IMPORTANCE OF PROSOPOGRAPHY AND LATERAL CHAINING**

We have thus far concentrated on research chains moving in primarily a “vertical” direction toward individuals with greater authority in a subject area (see figure 1). In emerging online research environments, the power to link related persons “laterally” is becoming more fully realized in web-based tools that are useful for many research methodologies, such as prosopography. Prosopography, the study of related groups of individuals, is “about what the analysis of the sum of data about many individuals can tell us about the different...
types of connections between them, and hence about how they operated within and upon the institutions . . . of their time.” As K. S. B. Keats-Rohan noted in 2000, while in “the 1970s and 1980s it was still possible to get a history degree in Great Britain without having heard of prosopography, nowadays, it would be would be unpardonable for a graduate student to be unaware of the work of prosopographers.”

Keats-Rohan suggested that the growing importance of prosopography may be due in part to new technologies such as “the relational database [which] is ideally suited to prosopographical research, and has provided an invaluable stimulus to such research.” The adoption and application of RDA and EAC-CPF standards have stimulated the development of web-based tools that promote and exploit prosopography. It is important to note that relational databases can facilitate lateral chaining through faceted searching. (See figure 2 for a visualization of the operation of lateral chaining in a web-based research tool.)

**RDA AND NAME AUTHORITY**

RDA emerged from an international conference in 1997 on the future of the Anglo-American Cataloging Rules (AACR2). Attributes and relationships in the Functional Requirements for Bibliographic Records (FRBR) model were the foundation for this metadata standard, which the Library of Congress adopted and implemented in March 2013. The
Repurposing RDA’s Descriptive Standards to Facilitate Humanities Research

standards provided instructions and guidelines for formulating and amending name authority records (NARs) within the MARC format. The descriptive authority data in these records differs from the preceding AACR2 standard through the addition of 3XX fields, identifying relationships or attributes about the authority (person). Examples include the following:

Associated Place (370)
Field of activity (372)
Associated Group (373)
Occupation (374)
Gender (375)
Creator/Contributor Characteristics (386)

The authority data in these fields is authoritative because all information emanates from published sources, and is cited in a specific field provided for that purpose (i.e., field 670, “Source Data Found”). In addition, each field may be populated with controlled terms, which provide additional value for collating faceted searches. These new fields provide the structured authority data that have potential applications for prosopographical research and the development of new relational database tools that can be created for that purpose.

WHAT ARE NARS?

There are many, varied name authority records. The focus here is the structure and description of personal names in a name authority record. A MARC 21 formatted name authority record contains descriptive metadata, governed by RDA standards that regulate their use. Librarians create and edit these records under the auspices of the Library of Congress’s Name Authority Cooperative Program (NACO). The Library of Congress staff and other members of NACO also review and curate the records. Under AACR2 rules, name authority records for personal names required the minimal amount of information necessary for disambiguation. However, RDA guidelines added additional descriptive fields that went well beyond what was truly required for mere disambiguation; these included relators that anticipate the growth of linked data and use of the records within web-based environments.

Here is an example of recently populated RDA personal name authority record in the MARC 21 format:

1001 Locke, Alain, d 1885–1954
370 Philadelphia (Pa.) b New York (N.Y.) e U.S.
373 Howard University a Harvard University a Hertford College
374 Educators a Philosophers a Authors 2 lcsh
375 male
377 eng
386 African Americans 2 lcldgt
4001 Locke, Alain LeRoy d 1885–1954
670 Britannica online, May 9, 2012 (Alain Locke; b. Sept. 13, 1886 Philadelphia; d. June 9, 1954, New York City; American educator, writer, and philosopher, best remembered as the leader and chief interpreter of the Harlem Renaissance. Graduated from Harvard University (1907), Locke was the first black Rhodes Scholar, studying at Oxford and the University of Berlin. He received his Ph.D. in philosophy from Harvard (1918). For almost 40 years, until retirement in 1953 as head of the department of philosophy, Locke taught at Howard University, Washington, D.C.)
670 Alain Locke, 2005: p. 11–12 (Alain LeRoy Locke was born on 13 September 1885 in Philadelphia, Pennsylvania, not in 1886, as commonly thought. For reasons that have eluded historians, Locke always represented his year of birth as 1886)
670 Alain L. Locke, 2008 p. 5 (Alain Leroy Locke was born in Philadelphia on September 3, 1885)
670 African American National Biography, accessed via The Oxford African American Studies Center online database, July 27, 2014: (Locke, Alain Leroy; literary
critic, philosopher; born 13 September 1885 in Philadelphia, Pennsylvania, United States; graduated from Harvard University (1907); attended lectures at Hertford College, Oxford and at the University of Berlin (1910–1911); joined the Howard University faculty in 1912; received his doctorate in Philosophy from Harvard in 1918; edited a special edition of the magazine Survey titled the Survey Graphic; died 09 June 1954 in New York, New York, United States)

Currently, NARs reside in siloed library catalogs where the new 3XX fields are typically not indexed nor searchable. It is not clear how the next generation of library systems will utilize these records. However, in the next section, we describe a prototype that can make these records fully usable and accessible to humanities researchers: Howard University’s Portal to the Black Experience.

PORTAL TO THE BLACK EXPERIENCE: BUILDING A RELATIONAL DATABASE WITH NARS

In December 2012, Howard University received a grant from The Institute of Museum and Library Services (IMLS) to build a proof of concept relational database that used name authority records with RDA descriptive metadata. The Howard University prototype is an open source, ancillary research tool that will contain more than two thousand individual names, structured biographical information about
Repurposing RDA’s Descriptive Standards to Facilitate Humanities Research

As shown in figure 8, Portal to the Black Experience received its inspiration from SNAC, which the designers integrated into the portal’s search results, and is part of its vertical chaining functionality. However, unlike Portal to the Black Experience, which derives its content from RDA name authority records, SNAC is informed by EAC-CPF records. The SNAC web prototype includes a much more sophisticated cyber-infrastructure that provides radial graphs (showing a visualization of the links between individuals included in SNAC), biographical data, associated resources, and associations with other individuals. The other notable project from the University of Virginia is the People of the Founding Era (PFE), an Internet ancillary research tool with extensive documentary and visual information on persons who lived at the time of the founding of the United States. The self-described goal of PFE was twofold: “one is biographical; the other is prosopographical. These important and complementary approaches allow the user to discover a complex and rich set of offerings.”

Perhaps the most recognizable prosopographical feature of this portal is that the database includes primarily men and women of African descent; the minority of persons not belonging to this category are found in the “Other” category. As the number of NARs in this prototype grows it has the potential to become an ever more valuable tool.

The prototype also facilitates vertical chaining of a particular individual and his works. After a researcher identifies a person of interest, the portal provides a landing page that contains links to that individual’s works located in libraries and archives. The link to the WorldCat Identities page informs the researchers about what library materials (books, articles, music, and movies) exist either by or about this person in nearby libraries. As the example in figure 5 illustrates, the WorldCat Identities page lists works in libraries by and about Alain Locke. The WorldCat Identities page also allows the user to locate the closest library that holds a particular item via geospatial technology. Thus, a search conducted in Washington, DC on February 8, 2016, for *The New Negro: An Interpretation*, resulted in the list shown in figure 6. For those who want to locate archival materials related to Alain Locke, the prototype offers two resources, either ArchiveGrid as seen in figure 7, or SNAC as seen in figure 8. It is important to note that the vertical and lateral functionalities of the portal were based on two research tools, both devised at the University of Virginia, that were the antecedents to this prototype.

**NEOTRADITIONAL HUMANITIES RESEARCH TOOLS**

Research universities are challenging academic librarians and archivists to develop tools and services that help them realize their mission. Academic librarians must develop innovative discovery platforms that improve searching and location, and also anticipate future research needs. In the field of humanities, prototypes are helping researchers work their way through essential, traditional methodological research tasks.

Developing such ancillary research tools is a daunting task, since it not only requires a knowledge of library
services in the digital era, but also requires competencies, or collaboration with those who have competencies, in website design, database development, and knowledge management. Moreover, until metadata librarians and archivists start regularly creating the more robust sort of records envisioned in the adaption of RDA and EAD-CPF, populating the record fields needed by these tools will take a large investment of time. However, as more open-source tools, such as the Portal to the Black Experience become available, and more robust records become the norm, the potential of repurposing the code of existing tools, and utilizing already robust records, should somewhat mitigate the amount of effort necessary to create this sort of neotraditional research tool.

It is our contention that the humanities research tools described in this article are both at once new and traditional; thus, they are “neotraditional.” What are new are the tools that contain controlled terms and sets of authoritative, structured biographical information. When combined with a modern cyber-infrastructure, these tools allow researchers, for the first time, to rapidly identify relationships between persons, their institutional affiliations, race, gender, and works. Innovative tools such as these have stimulated the adoption of and overall approach to prosopographical research. What are traditional are the essential tasks that humanities researchers must still perform; most notable is chaining. These new research tools incorporate chaining and netchaining in their functionality to a degree unimaginable five years ago. Further progress of the linked data movement on the Internet will have a profound effect on these tools and will extend them well beyond their self-contained current limits. Indeed, netchaining may likely morph into what might someday be called “linked-chaining.”

If we recall Dr. Johnson’s comment that the writer of history “has facts ready to his hand; so there is no exercise of invention . . . some penetration, accuracy, and coloring will fit a man for the task, if he can give the application which is necessary,” we suggest that historians and other humanities researchers must continue to give application (a narrative or interpretation) to the facts.29 However, the emerging research tools begin to make those facts, even in this age of abundant data, truly ready to hand in great quantity, and with penetration and accuracy.

References and Notes


1. See also Rambler no. 122, where he writes “that no writer has a more easy task than the historian.” Samuel Johnson, The Rambler, vol. 3 (London: J. Parsons, 1793), 83.

2. Analysis is complicated by the volume of information available in digital sources that must now be shifted and weighed, not only for relevance but also for authenticity. Access to many digital humanities materials is often only available to researchers who have use rights for balkanized and exorbitantly priced subscription databases.


4. Since the digital humanities research environment is broad and can accommodate many levels of users and a wide range of humanities disciplines, it useful to more narrowly define which categories of researchers we refer to herein. Andy Barrett’s study of humanities students at the PhD level is useful here. It shows that PhD-level researchers are largely already comfortable with the “detective-like approach to information seeking, involving browsing [and] citation chaining,” i.e., chaining (Andy Barrett, “The Information-Seeking Habits of Graduate Student Researchers in the Humanities,” Journal of Academic Librarianship 31 no. 4 (2005): 330). So for the purposes of this article, all references to “researchers” primarily refers to PhD-level students and academic faculty conducting research in the liberal arts and social sciences.


6. Ibid., 474

7. Ibid., 482.


9. For example, Demaria and Kolb followed standard scholarly practices and placed the deeper scholarship in annotated footnotes directing their readers to best accounts and sources, such as in footnote 28, where they state, “For the best accounts of the amanuenses see Allen Reddick, The Making of Johnson’s Dictionary, and Eugene J. Thomas, A Bibliographical and Critical Analysis of Johnson’s Dictionary with Special Reference to Twentieth-Century Scholarship” (unpublished doctoral thesis, University of Aberystwyth, 1974); we are indebted to both studies” (Johnson’s Dictionary,” 33) The reader now has two names, Allen Reddick and Eugene J. Thomas, from which to continue chaining. Indeed Allen Reddick, in addition to his book, has published many articles on Johnson’s Dictionary.


11. Ibid.

12. Ibid., 278.


15. Both ORCID and ISNI not only disambiguate names, but more importantly they provide valuable relators to the persons, such
as the titles of their works, their publishers, etc.

17. Ibid.
18. Ibid., 6.
20. The 386 field is the newest relator field in a NACO record, appearing in autumn 2015. “The Library of Congress is developing a new vocabulary entitled *Library of Congress Demographic Group Terms* (LCDGT). This vocabulary is to be used to describe the creators of, and contributors to, resources, and also the intended audiences of resources.” These terms are applicable for recording a person’s nationality or ethnic group. “Library of Congress Accepts Demographic Group Term Proposals, Publishes Draft Demographic Group Terms Manual,” Library of Congress, accessed February 1, 2016, www.loc.gov/catdir/cpso/lcdgt-acceptance-manual.html.
21. “RDA has many controlled vocabularies. Only a few of the vocabularies are closed (e.g., content type; media type; carrier type; mode of issuance). Most of the vocabularies are open, you can either supply your own term as needed, or suggest a term be added to the vocabulary (or do both).” Salman Hader, “RDA FAQ,” *RDA Blog*, section 7, accessed February 25, 2016, http://resourc edescriptionandaccess.blogspot.com/p/rda.html#Vo1LiIrKUK.
24. The authors of this article conceived of and directed the project that developed this prototype. The project was made possible through an Institute of Museum and Library Services National Leadership Grant for Libraries (grant number LG-07-13-0294-13). A developmental version of the portal is available at “Portal to the Black Experience,” Howard University, accessed May 9, 2016, hugo.wrlc.org.
25. The prototype was in part inspired by the “Black Authors Index,” a unique research tool created by Dorothy Porter, the first curator of Howard University’s Moorland-Spingarn Research Center. During her forty-three-year tenure (1928–71), Porter began an in-house practice that continued until the 1990s of recording racial, ethnic, and national information about black authors in the Research Center’s Card Catalog. As a result, Moorland-Spingarn’s catalog includes information on black authored or edited works printed on 14,122 catalog cards, listing 6,921 authors, editors, and composers, along with the titles to their works. The practice of creating black authors cards appears to have ended during the early 1990s. Many of the names contained in these cards may be found in the portal.
26. In this portal, the term “black” has been used to describe the peoples native to sub-Saharan Africa, and of the African diaspora including African Americans, Afro-Brazilians and other peoples formerly of sub-Saharan Africa. Those that are not a part of the aforementioned groups are categorized under the term “other.”
27. “ArchiveGrid includes over four million records describing archival materials, bringing together information about historical documents, personal papers, family histories, and more. With over 1,000 different archival institutions represented, ArchiveGrid helps researchers looking for primary source materials held in archives, libraries, museums and historical societies.” “ArchiveGrid,” WorldCat, accessed February 25, 2016, https://beta.worldcat.org/archivegrid/