

Disambiguating the Departed

Using the Genealogist's Tools to Uniquely Identify the Long Dead and Little Known

Chris Evin Long

The need to uniquely identify people with the same name will be as important in a linked data environment as it is in traditional library catalogs. Although older cataloging rules allowed multiple identities to share space in an authority record, the current rules are more stringent, requiring that all authorized access points for people to be disambiguated. While this task has been made much easier in recent years because of the amount of biographical material on the web, deceased and obscure people can pose a complex challenge. This is especially true for special collections materials, which are being given greater attention but which often deal with people who are long dead and little known. This paper builds on previous research in the use of online resources to perform authority work by providing an in-depth analysis of the genealogist's toolkit and examining how freely available online genealogical sources can be used to find the types of distinctive information needed to create unique access points for people.

The author suspects that many catalogers harbor a dirty little secret—authority work is their favorite part of cataloging! Their reluctance to admit it is understandable. Although its value to bibliographic control has been recognized for more than a century, authority work is seldom seen as the most attractive part of the metadata creation realm. It can be a time-consuming task and its benefit-cost ratio has sometimes been called into question. But catalogers are puzzle solvers at their core, and it is by creating authority records that their latent desire to play the detective is most fully satisfied.

As cataloging moves toward a linked data environment, authority work promises to take on renewed importance. Semantic ambiguity poses a challenge for computers, so the necessity of uniquely identifying people with the same name, recording the various forms of name that a person has used, connecting different bibliographic identities that a person has assumed, and collocating all of the resources related to that person is just as important in the Semantic Web as it is in traditional library catalogs. Personal name authority files like the Library of Congress (LC) Authority File fulfill this need and therefore will be key factors in linked data applications.

The research necessary to create authority records for people has become much easier over the last two decades because of the wealth of biographical material available on the web. This is particularly true for living people, for whom

Chris Evin Long (chris.long@colorado.edu) is Head of Monographic and Special Materials Cataloging in the University Libraries, University of Colorado Boulder.

Manuscript submitted February 12, 2016; returned to author April 13, 2016 to revise; revised manuscript submitted April 29, 2016; manuscript accepted for publication June 14, 2016.

publishers' data, email, personal and organizational websites, and social media provide many avenues for discovering the information needed to create an authorized access point that distinctively identifies the individual. Deceased people, in contrast, often pose a more difficult challenge. Although information about nationally prominent people from bygone days can be readily found in a host of sources, the cataloger must frequently rely on other strategies to uncover the evidence needed to uniquely identify the long dead and little known.

A convergence of circumstances in recent years has made finding facts to uniquely identify deceased people a more crucial task for catalogers. Libraries are striving to make their hidden collections more discoverable. Catalogers of all stripes, not just rare book specialists, are being tasked with making these resources accessible.¹ Many of the items in these hidden collections are older and local, and the people associated with these items are often obscure. In some instances they share a name with other people in the authority file. Under former cataloging guidelines, this did not pose an insurmountable problem. Undifferentiated personal name records were constructed to represent different people with the same name; multiple people could be described on the same authority record. However, in November 2013, LC and the Program for Cooperative Cataloging (PCC) agreed to no longer allow the creation of new undifferentiated personal name authority records or to add identities to existing ones. This change was precipitated by the recognition that the intermingling of identities on a single record precludes the population of new MARC fields that contain information specific to individual people, such as birth and death dates, gender, associated places and organizations, and occupations, and interferes with potential linked data uses of the authority records, such as their inclusion in the Virtual International Authority File (VIAF) and other Semantic web applications.² Now, catalogers who participate in the PCC Name Authority Cooperative Program (NACO) must create unique authorized access points for each person being established. So where can catalogers turn to find distinctive information about hard-to-find individuals, such as their birth or death dates, fuller forms of their name, and their occupations or professions?³

Genealogists are constantly vexed with the same basic problem, and the tools on which they rely can also assist the cataloger. This paper will examine the types of information available in online genealogical sources that can aid in solving the name ambiguity problem for individuals who are not included in standard reference sources. The focus is restricted to personal names, the geographical scope is limited to the United States, and because catalogers increasingly work in an environment of diminishing resources, online resources that are freely available are emphasized.

Literature Review

Numerous studies have documented the role that authority control has played in the bibliographic world over many decades. Although Charles Ammi Cutter did not use the term in his *Rules for a Printed Dictionary Catalog* in 1876, the concept of authority control is clearly there in his objects of enabling users to find a book by a known author and showing what the library has by a given author using the means of author-entry and attendant references.³ In the late nineteenth century, LC began creating authority cards that contained many of the primary elements used today, including the preferred form of name and variant usages of it. While authority control was not explicitly discussed in the 1961 International Conference on Cataloguing Principles, various working groups examined ways to effectively collocate the works of an author by choosing a uniform heading that tied together all the variant forms of names.⁴ Interest in the topic continued into the 1970s, during which many international authority control initiatives were pursued as national libraries and international organizations sought ways to reduce the work's expense by sharing the load.⁵

As card catalogs began to give way to their online counterparts in the 1980s and 1990s, the enhanced retrieval capabilities of the new catalogs exposed gaps in authority control as differences in headings became more noticeable to users.⁶ As most librarians saw a continuing need for authority control in an automated environment, some called it into question. Kilgour opined that future catalog design would obviate the need for authority control.⁷ Koel and Taylor questioned whether the expense of implementing certain aspects of authority control outweighed the benefits.⁸ Even in the more evolved online catalogs that are currently available, Ayres claims that authority control, while theoretically still valuable, does not work effectively because the reference structure in authority records is not supported or is underutilized.⁹ As catalogers anticipate how our work will change in the imminent linked data environment, other experts see a continuing need for the basic principle of collocation that authority control provides, even if the mechanisms we use to produce it change drastically.¹⁰

The emerging importance of special collections to the library's research mission and reputation is well chronicled. At a time when the relevancy of libraries is being questioned, Dooley and Luce note that "special collections and archives are increasingly seen as elements of distinction that serve to differentiate an academic or research library from its peers."¹¹ Yet access to these materials has been a long-standing concern. In its 1998 survey of special collections, the Association of Research Libraries (ARL) identified the backlog of uncataloged and unprocessed materials in special collections as one of the most serious challenges facing the profession. Libraries reported that significant portions

of their special collections had not been processed, even though staffing levels for special collections had increased or at least remained stable for the previous ten years.¹² Jones and Panitch estimated in 2004 that at the present rate it would take hundreds of years to make these hidden collections accessible.¹³

Hubbard and Myers' 2010 survey of rare books catalogers found that 97.8 percent of the respondents still reported a backlog at their institutions, and although Dooley and Luce's report of that same year found backlogs for some types of materials had decreased at more than half of the institutions surveyed, the size of the collections continues to grow.¹⁴ While this situation presents challenges for library staff, the real harm is to users. Under-cataloged collections result in inadequate intellectual access to those resources, particularly hindering the research efforts of some of libraries' most vulnerable users who lack the financial means to travel to other institutions, such as undergraduates, graduate students, and junior faculty.¹⁵ To tackle this problem, Mandel emphasized the importance of not only integrating technical services with special collections, but also having technical services assume the primary responsibility for cataloging special collections items, especially now that technical services staff are typically handling fewer standard materials.¹⁶ Research by Russell and Lundy showed that this has been happening, as more staff responsible for cataloging special collections material report to a cataloging department than to other units within the library.¹⁷

Dealing with special collections materials, though, is not for the faint-hearted, regardless of the cataloger's background. Even agreeing on a definition of what the term "special collections" means has proved challenging. Although a strict definition limits its usage to rare books, more often the phrase encompasses manuscripts, art, photographs, cartographic resources, and even microforms and audiovisual materials. Catalogers therefore must be prepared to deal with a diverse array of formats. Furthermore, these collections are often replete with unique, local materials that are unpublished or are ephemeral, and within which reside a host of personal names—authors of local histories, diaries, letters, manuscripts, etc., whose obscurity renders them no less valuable to researchers.¹⁸

Indeed, studies have shown that names of people and organizations found in archival and special collections materials are of great interest to scholars, particularly those in the humanities. Duff and Johnson found that historians collect names to aid their navigation of archival collections, relying heavily on them to identify relevant collections and to locate pertinent information within those collections. They highlight the importance of names to researchers by stating, "Collecting names may be a fundamental practice in historical research since the past is often interpreted through the activities of individuals or organizations."¹⁹ Wiberley further

found that the uniqueness of a name is of prime importance to scholars. His study of subject access in the humanities showed that singular proper terms (i.e., the names of unique entities such as people) are the most precise terms used by humanists and are therefore central to their work, emphasizing the value of distinguishing two or more people who share the same words in their names by their dates and places of birth and death.²⁰ Given the humanists' reliance on unique names, Wiberley concluded that nonexistent or inadequate authority control is a great disservice to them and "will impede humanists from access to the full range of sources relevant to their research."²¹

Providing this level of authority control for less well-known people as may be found in local special collections presents its own set of challenges. Catalogers have come to rely on the collective efforts of other libraries to assist them in uniquely identifying individuals, but Mandel points out that the uniqueness of materials in special collections, and by inference the people associated with them, makes it less likely that the cataloger can benefit from the cooperative efforts of others.²² This opinion is reiterated by Bradshaw and Wagner, who emphasized that the subject and name access delivered by large-scale cataloging cooperation may be inadequate for local needs such as these.²³ Catalogers creating authority records for the lesser-known, then, must rely on other means.

Fortunately, the Internet offers an abundance of sources of biographical information that can be used to create unique access points for people. Catalogers early on recognized the value of online resources to their work, particularly in the area of authority control. A brief survey conducted by Long in 1997, at the advent of the Internet's incorporation into the work of librarians, found that catalogers were using resources such as online phone books, email directories, and other libraries' online catalogs to resolve name conflicts and clarify ambiguous headings.²⁴ The Internet, though, did not prove to be a problem-free panacea for catalogers, as the impermanence of the web and the suspect nature of some of the information found there became increasingly apparent. In 2001, Russell and Spillane's examination of how catalogers were using online resources for name authority work showed that little had changed since Long's survey. The Internet was being used essentially the same way as before and although it had the potential to make authority work more productive and efficient, catalogers were often frustrated by the dubious reliability of certain websites and the limited amount of the information available to them (often just contact information), especially for older material.²⁵

The next year, though, Ellero demonstrated a way that web resources could be used in special collections authority work. Relying mainly on United States government websites, a team in the Claude Moore Health Science Library created a controlled list of 1,692 unique name entries for people

associated with their Philip S. Hench Walter Reed Yellow Fever Collection, an archive of books, articles, correspondence, photographs, and artifacts from the Yellow Fever Commission of 1900. While the names of well-known individuals were often easily found and standardized, the project showed that the web's reach could also extend to lesser-known people. The list remained a local one, as the team did not feel it fell within the scope of LC's Name Authority File's mission.²⁶ However, Ellero predicted, "As more and more institutions (i.e., libraries, archives, and museums) in the United States and around the world process special collections of unpublished materials on an analytic level and make these resources available on the Web, an enhanced and global system for authority records will become essential." She was also prescient in her observation that professions should be used as qualifiers to better identify individuals.²⁷

As the Internet matured, so did catalogers' employment of it. Marshall's 2012 article offered more sophisticated strategies for using the Internet to discover birth and death dates of lesser-known people, focusing on the utilization of online genealogical tools to accomplish this task. She examined many general genealogical websites and also delved into more specialized sources such as death indexes, tombstone inscription sites, family trees, obituaries, and local histories.²⁸ Marshall concentrated on finding birth and death dates with the goal of making them more useful to library users, but genealogical resources can provide a wider variety of information about individuals as well, information that can be invaluable to catalogers working under the current imperative to not create undifferentiated access points.²⁹ This paper covers some of the same ground as Marshall and others, and builds on their work by providing a more in-depth analysis of the genealogist's toolbox and examines how free online genealogical resources can be used to find not only birth and death dates, but also other prescribed types of distinctive information.

Resource Description and Access (RDA) 9.19.1.3–9.19.1.8 instructs catalogers about the attributes used to distinguish one authorized access point from another. They are

- date of birth and/or death (9.19.1.3);
- fuller form of name (9.19.1.4);
- period of activity of the person (9.19.1.5);
- profession or occupation (9.19.1.6);
- title of the person, including terms of rank, honor, or office (9.19.1.7); and
- other designation associated with the person (9.19.1.8).

Birth and death dates are given preeminence because the other attributes are to be used only if these dates are unavailable. While the meaning of categories such as fuller form of name and profession or occupation is self-evident,

others may need further explanation. "Period of activity" is a date or range of dates indicative of the period in which a person was active in his or her field of endeavor. "Other term of rank, honour, or office" can include terms associated with people of religious vocation, military ranks, or academic degrees. "Other designation" is a catchall category for attributes not covered by the other options and can include associations with other people, corporate bodies, and places. As will be shown, although genealogical tools are obvious choices for ascertaining a person's vital information such as birth and death dates, full name, and occupation, they also provide abundant material for discovering these other attributes as well when the vital information proves elusive.

Genealogical Sources Useful to Catalogers

Genealogical researchers employ a vast array of tools to search for facts about people's lives. *The Source: A Guidebook to American Genealogy* provides a comprehensive overview of the types of records genealogists use in their research, and this section summarizes each category of records listed in the book.³⁰ Not every resource, though, is a practical candidate for a time-strapped cataloger's attention. Lack of widespread availability, inconsistent data entry, and a paucity of the kind of information needed to disambiguate people disqualify some types of sources from consideration because the time spent searching outweighs the potential benefit. The focus, therefore, will be on the types of records that are most likely to aid catalogers in their quest for the information needed to create unique access points.

Vital records that document key life events such as birth, death, and marriage are likely the first place that catalogers will want to start searching. Birth and death records are obviously useful because they contain birth and death dates, the paramount attributes used to distinguish people, and they are often also sources of an individual's full name. Records in online death indexes such as the Social Security Death Index (SSDI) and the Online Searchable Death Indexes and Records website (www.deathindexes.com) also contain birth and death dates, but may have limitations in availability and time span coverage. Less obvious sources such as cemetery records, tombstone transcriptions, funeral home records, and church records also merit attention. These sources may contain not only birth, death, and name data, but sometimes occupational information as well. Marriage records in general are unreliable options because of lack of availability, occasionally falsified information, and clerical errors, with the possible exception of marriage licenses, which often contain useful personal information such as full name, birth date, and occupation. Court records, while often providing fascinating insights into a person's life, are largely also not of much assistance to the cataloger, with

the singular exception of probate records, which are likely to include a person's death date.

Censuses are among the most frequently used records by genealogists because of the importance of the information they contain, and are also likely candidates for the sleuthing cataloger. Early US censuses included relatively few details about individuals, but in 1850, census takers started gathering more information about age, sex, color, occupation, and birthplace. Although the questions asked varied from census to census, names, ages, and occupations were consistently collected, only the 1900 census asked for the month and year of birth. Consequently, census information can often provide fuller forms of names, occupations, and at least an approximate year of birth. Since US censuses are not released to the general public until seventy-two years after the census was originally taken, the latest one available is the 1940 census. The mortality schedules created using data collected in the 1850–85 censuses can also be used to determine month and year of death and occupation.

Newspapers are another abundant source of biographical information for the cataloger. Obituaries are the obvious initial step in newspaper research because they are often the only biography written about most people. Furthermore, for those born before the early twentieth century, obituaries fill in the gaps left by the spotty availability of official US government vital records. While birth and death dates are their primary attraction, obituaries often include full names and occupations. Birth, death, and funeral announcements do not afford the same breadth of evidence as obituaries, yet still typically include the birth and death dates that catalogers seek. It was also not uncommon for small town newspapers in earlier times to publish brief biographical sketches of prominent or newly arrived citizens.

Military records can be rich sources of information for the cataloger, especially for birth and death dates. Military pension records for the Revolutionary War and later wars usually provide this information, as do burial registers and lists at national cemeteries and military post cemeteries. World War I and World War II draft registration cards offer the added benefit of full names and occupations, although privacy laws restrict access to most WWII cards.

Immigration and naturalization records provide fascinating information for genealogists, but offer only a modicum of help to the cataloger. Still, they are worth searching if other approaches fail. Official US government passenger-arrival lists of immigrants are available from 1820 through the 1950s for most of the ports of the United States with customs houses. The amount of information collected on these lists varies greatly and they are relatively unreliable for names and ages, but the lists often include the person's occupation. Naturalization records also vary in the amount of personal data recorded, and often provide the added benefit of a date of birth.

Local and family histories and family trees can also be ripe with documentation for catalogers. In the nineteenth century, many communities published local histories and biographical sketches of area residents. In cases where official vital records of this period may not be extant, these local publications may be the only source of information of birth, death, and occupational information to help catalogers create distinctive access points. Family histories have long been a staple of genealogical research and are rich sources of biographical material, as are the ever-growing number of online family trees.

Business and organization records can be another source of useful information. In this category, occupational records are particularly noteworthy. Many occupational groups, especially clergy, legal professionals, physicians, and trade associations, created directories and biographical sketch books of their members that, in addition to indicating their profession, frequently included birth and death dates and full names. Occupational registries were often compiled by cities interested in a particular vocation and include similar information. Because they may indicate a term of rank or office (such as Reverend or PhD), or include associations with corporate bodies, occupational records can also offer suggestions for the final two categories of RDA attributes, "title of the person" and "other designation associated with the person." Similarly, city directories may list occupations and, beginning in the twentieth century, might also list a date of death for an individual who had died since the last directory had been compiled.

Finally, records for cultural groups should be considered when pertinent. Whereas information about individuals belonging to specific cultural groups can be found in vital records, censuses, military records, and other sources previously discussed, there are some group-specific resources worth noting. For African Americans, records of the Freedman's Savings and Trust Company (a.k.a. the Freedman's Bank) can deliver valuable evidence. The Freedman's Bank was created as a way for soldiers and ex-slaves to invest their money, and eventually included numerous branches in various parts of the country. While lacking important information like birth dates, the signature registers that were required to open an account typically included the person's occupation and place of residence. Native American ancestry research can be difficult, but there are resources that can help. From 1885 to 1940, agents or superintendents of Native American reservations were required to submit annual census rolls, although there is not a yearly census roll for every tribe. The rolls typically included the age or date of birth, and later rolls additionally recorded place of residence. Individuals who wanted to be classified as official members of a tribe had to complete an enrollment process, and these tribal enrollment records often contain fuller forms of names and death dates when applicable. The Jewish

community is very active in genealogy, and the JewishGen website (www.jewishgen.org) has databases of family trees, burial registers, and Holocaust victims and survivors.

Strategic Use of Genealogical Tools

Armed with knowledge of the types of records likely to be useful in creating unique access points for people, catalogers can then consider where to find them. Fortunately, such information is readily at hand. *GenealogyInTime Magazine* compiles an annual list of the top one hundred most popular genealogical websites, and the list shows the wide variety of online resources available to genealogists to assist them in their research. This list can also serve as a useful foundation for catalogers' name authority research.³¹ Broadly speaking, the websites fall into two different categories—free versus pay sites, and those offering a comprehensive array of records versus those concentrating in a specialized area. As stated in the introduction, this paper's focus is on freely available resources, but even after removing the for-pay options from the list, a multitude of both comprehensive and specialized sites remain. With so many options, catalogers may have difficulty determining which sources to try first. The remainder of the paper, therefore, will discuss practical strategies that time-starved catalogers can use to glean the most useful information in the least amount of time.

Tombstone Inscription Sites

Since birth and death dates are obvious ways to uniquely identify people and are given favored status in RDA, tools providing that information are logical places to start. Certain types of sources, though, can offer more expedient solutions than others. For individuals who died in 1961 or later, the SSDI (discussed below) is a good starting point, but for people living in an earlier era, tombstone transcription sites are often the best place for catalogers to begin their research. They usually include photographs of gravesites or transcriptions of gravestones from which birth and death dates can be harvested, and some sites also include obituaries. The ease of searching and the span of eras covered are major advantages of these sites, but their reliance on volunteers to contribute information limits their comprehensiveness and sometimes yields illegible photographs of markers. There is undoubtedly much overlap between the sites, and the cataloger may have to search multiple locations to find data on the individual in questions. Furthermore, databases of this size naturally pose some searching challenges. Individuals with rather unique names like the author's father Halleck Long do not require much effort to find, while those with more common names like his grandfather Samuel Long can prove more difficult. Furthermore, the information found may not be enough for the cataloger to adequately confirm

that it relates to the person being researched. Despite these obstacles, tombstone transcription sites are often a quick and easy way to locate birth and death dates. A review of the most useful ones follows.

Findagrave (www.findagrave.com)

Volunteer contributors have added 138 million grave records to this site. Gravestone photographs and transcriptions are the most common items found here, but obituaries and links to other family members are also sometimes added. The basic search interface is easy to use, covers locations both within and outside of the United States, and the option to limit by the state where the cemetery is located ameliorates the problem of searching for people with common names. Although Findagrave information is also part of the more comprehensive FamilySearch website discussed later, researchers must often navigate a dizzying array of search results in Family Search to find the subject. Findagrave's narrow focus, the size of its database, and the inclusion of obituaries combine to make this the cataloger's best initial option for simple birth and death information. Findagrave is also integrated into some of the more comprehensive websites described below, but those sites often house billions of records, so coming here first can avert the prospect of wading through a mass of irrelevant material.

Billion Graves (<http://billiongraves.com>)

Although this site does not in fact contain a billion records, it does check in at about 15 million. Free registration is required, but records for some individuals are only available to paid subscribers. The information on the site is mainly limited to tombstone photographs and transcriptions, and the number of available records makes this a good second option.

Interment (<http://interment.net>)

Whereas it does not contain as many records as Findagrave and Billion Graves, this site is a viable alternative if searches in the other two fail. The search interface is simple and covers the entire United States, and searchers have the option of browsing by state. Although links to state vital records and obituaries are available, they ultimately lead to subscription-based sites.

USGenWeb Tombstone Transcription Project
(<http://usgw-tombstones.org>)

This site's arrangement and limited search capability hinders its effectiveness for catalogers. It is arranged by state and county, and because there is no way to search across state

projects, the searcher must know where the individual was buried. The number of available records is also much smaller than other sites, likely making this the cataloger's last option in this category.

Death Indexes

If an individual cannot be found using the tombstone transcription sites, online death indexes are another option for quickly finding birth and death dates. The two major resources in this category are the SSDI and the Online Searchable Death Indexes and Records website. The SSDI offers the ability to do a nationwide search for basic birth and death information, but only for a very specific period. The Online Searchable Death Indexes and Records site contains a wider variety of resources and encompasses a much broader time span, but its arrangement of resources by state makes searching more challenging if the cataloger does not know the location of the person's death.

SSDI

The SSDI currently contains information about more than 94 million people who lived in the United States and had a Social Security number. Although the database officially goes back to 1934, virtually all of the people in it died after 1961, rendering it useful only for researching individuals who were alive in the mid-twentieth century and later. The information of primary interest to catalogers contained therein is names, complete birth dates, and the month and year of death. The SSDI is freely available from two main sources, FamilySearch (<https://familysearch.org/search/collection/1202535>) and GenealogyBank (www.genealogybank.com/gbnk/ssdi?kbid=9064&m=9), although the GenealogyBank site requires free registration. Both have user-friendly interfaces and allow searching and filtering by first and last name, approximate birth and death year, and geographic location. GenealogyBank offers more initial search options, and is only current through March 2011 as of the date of this writing. The FamilySearch site, in comparison, is current through February 28, 2014, and is therefore the better choice if the individual being researched is likely to have died recently. While the SSDI sometimes yields very quick results, searching for individuals with common names can be arduous if the cataloger does not know the state, county, or city of last residence.

Online Searchable Death Indexes and Records (www.deathindexes.com)

This site is a collection of links to websites containing death-related information such as death records, death certificate indexes, death notices and registers, obituaries, wills and

probate records, and cemetery burials. It is arranged by state and county, making it imperative that the cataloger know something about the subject's residence. Its primary appeal is that it provides one-stop shopping for a diverse set of resources that cover a broad swath of time, sometimes dating back as far as the early nineteenth century. The site includes links to numerous locally compiled obituary indexes that are not available in the more comprehensive genealogical sites. Even though coverage is hit-or-miss, this benefit is not to be overlooked given the paucity of freely available online obituaries. Since it is a collection of links instead of a database, its major drawbacks are the lack of inclusive searching capability and the intermingling of free and fee-based sources. Nevertheless, while not likely to be one of the quicker birth or death date options available, searching the Online Searchable Death Indexes and Records website is a worthwhile venture before pursuing other possibilities.

Obituaries and Newspapers

There are numerous sites where online newspapers and obituaries can be searched for free, but many of these charge fees to retrieve the article or obituary. There are also hundreds of state and county online newspaper collections, often covering brief ranges of times. Furthermore, whereas many sites provide access to recent obituaries, there are fewer in which the cataloger can find historical ones. Since the researcher could spend many hours locating and scouring collections separately, this section concentrates on sites that provide compiled national lists of free online newspapers and emphasizes those where obituary information can be actually retrieved, not simply searched.

Chronicling America (<http://chroniclingamerica.loc.gov>)

Almost seven million digitized historic American newspapers can be researched on this site. The collection includes papers from 1836 to 1922; those published after December 23, 1922, are not available due to copyright restrictions. The site offers a simple search box, and the advanced option that allows users to limit by state, newspaper, and date range and perform Boolean-like word searches is the better choice. The number of accessible sources, search options, and time span covered make this an ideal first choice for the cataloger searching for people living in the mid-nineteenth and early twentieth centuries.

Google News Archive
(www.news.google.com/newspapers)

Once the premier online newspaper site, Google News Archive was shut down for many years because of complaints and threats from newspaper publishers. This recently

resurrected site contains about 2,000 scanned newspapers and while not as extensive as *Chronicling America's* inventory, it includes some newspapers dating back to the 1700s and many small town newspapers. Newspapers are listed alphabetically but cannot be searched individually; the site provides only the option to browse the entire collection. It offers no advanced search features, and users cannot restrict by location or date. The archive search tends to retrieve older articles more reliably than the more general web search option. Faulty optical character recognition and poor scan quality further hinder the researcher, making names and events sometimes unfindable and the newspaper articles themselves difficult to read. Still, Google News Archive covers a broader time span for some papers than *Chronicling America* and can be a good second choice when the latter yields no results.

Ancestor Hunt Newspapers! Page
(www.theancestorhunt.com/newspapers.html)

This site is a gathering of links to collections containing over 12,000 historical US papers arranged by state, then by city or county. Since the site is a list of collections rather than a collection per se, there is no way to comprehensively search all of the newspapers represented within it, nor even to search all those within a state, although some counties have compiled obituary indexes for all the papers within their area. Links to other projects like *Chronicling America* and *SmallTownPapers* are often available. Although useful as a means to discover the availability of online newspapers in a given state, the variegated nature of the listed sites provides little uniformity in either search capabilities or coverage. Another major drawback for the cataloger is the frequent inclusion of sites that allow free searching but which demand payment to retrieve the obituary.

New York Times Archives (<http://query.nytimes.com/search/query?srchst=p>)

Although the *New York Times* is a local paper, its scope is also national, including its obituaries. The archives from 1851 to 1922 and 1987 to present are free, but the intervening years are not. If the person being researched is famous enough to have an obituary published in the *New York Times*, it is likely that the cataloger can readily find information in other places. Nevertheless, it could prove to be a useful site.

SmallTownPapers (www.smalltownpapers.com)

This site contains more than 250 small town newspapers, some dating back to the mid-1800s. Users can browse by title or by state and can search within individual newspapers.

Chronological coverage varies greatly, but *SmallTownPapers* does contain newspapers for locales not included in the larger sites.

Obituary Central (www.obitcentral.com)

Although the obituary listings in this site are arranged by state and county, a statewide obituary index search is available for all states. Most obituaries are from the late 1990s until present, though, so *Obituary Central* may only be useful for recently deceased individuals.

Obit Finder (www.legacy.com/Obituaries.asp?Page=ObitFinder&CoBrand=Legacy)

This site contains obituaries for more than 1,000 US and Canadian newspapers, but since its coverage extends back only to the early 2000s, it has little utility when searching for people living in earlier times.

Local Histories

Local histories became popular in the later nineteenth century, in part spurred on by the 1876 centennial. In addition to historical information about the county, city, or town, they usually contained a biographical section that profiled area residents and frequently included genealogical information such as birth and death dates, place of birth, and occupations. Despite claiming to contain an egalitarian mix of the local citizenry, inclusion in the historical account was often dependent on the willingness to pay a subscription fee, and the genealogical information gathered from subscribers was seldom verified. Consequently, the accuracy of the facts found in these local histories can be called into question. Furthermore, it is necessary for the researcher to know the subject's place of residence to effectively locate a relevant history. Nevertheless, these sources can capture information about people who for whatever reason cannot be found in the other types of records already examined.

There has been a tremendous push in recent years to digitize local histories. These are scattered across the web, but some of the best places to do a more concentrated search for them are the Internet Archive (<https://archive.org/index.php>), Google Books (<https://books.google.com>), and Online County Histories, Biographies and Indexes (www.genealogybranches.com/countyhistories.html). The Internet Archive has the advantage of containing only freely available resources, but since it includes a wide variety of media formats, searches should be limited to text to effectively navigate the site for local histories. Once a book is found, there are numerous versions that can be full-text searched to find the person in question. Searches in Google Books, on the other hand, must be limited to free Google e-books to

efficiently wade through the morass of unavailable content. Online county histories, biographies, and indexes is a state-by-state guide to local histories and biographical indexes, some of which are online, but its reach does not extend nearly as far as the other two sites.

Family Tree Websites

Websites that host online family trees can be another overlooked source of information for the cataloger. Contributors submit the results of their genealogical research to these sites, and a successful search can yield at minimum birth or death dates. Three of the top free sites are WikiTree (www.wikitree.com), FamilyTreeNow (www.familytreenow.com), and Crestleaf (<http://crestleaf.com>). Of these, however, only WikiTree is a viable tool for catalogers. It contains information on more than 11 million people. Name searches are the only option and although the user is allowed to match by date, there is no way to limit by geographic area, which is a major drawback for a database of this size. Furthermore, information for some individuals is marked as private and is not viewable. Family tree information in FamilyTreeNow and Crestleaf, on the other hand, is often buried amid other types of records (census, birth, death, marriage, divorce, etc.). These additional sources are no doubt helpful to the budding genealogist, and are more easily searched in other applications and obscure any unique family tree content that might exist, making these sites bad bets for the cataloger.

Genealogical Warehouses

The increased interest in genealogical research by professionals and amateurs alike has driven the creation of many websites that house enormous quantities of genealogy information. The dizzying array of records stored in these warehouses, sometimes numbering in the billions, provides a one-stop shopping experience accompanied by easy-to-use search interfaces. Unfortunately, most of these sites are hidden behind paywalls. The free options discussed below, however, are all highly ranked sites and should be part of the sleuthing cataloger's arsenal.

FamilySearch (<https://familysearch.org>)

FamilySearch is sponsored by the Church of Jesus Christ of Latter-day Saints. The site requires free registration and contains more than 4 billion names from all over the world. Researchers can search records, genealogies, and family history books. In the records section, name queries can be connected with geographic areas and life events such as birth, marriage, residence, and death. Searches can also be restricted by record type, including birth, marriage, death, census and residence, immigration and naturalization,

military, probate, and others. Findagrave data are also incorporated into the search results. The historical record collections included here extend back to pre-1700, although the vast majority covers the year 1800 onwards. The section on genealogies has listings for hundreds of millions of people containing information drawn from user-submitted genealogies and the church's International Genealogical Index. The Family History Books section contains more than 200,000 publications that are contributed by several partner institutions; these resources must be searched separately by institution. Because of the comprehensive search capability across many types of records combined with the ability to limit searches by numerous facets, FamilySearch could easily serve as the cataloger's first research option.

USGenWeb Archives (<http://usgwarchives.net/search/searcharchives.html>)

This online archive is part of the larger USGenWeb Project (<http://usgenweb.org>), the work of volunteers striving to provide free genealogical research sources for every state and county in the US. Each state has a separate page and is organized by county. As might be expected of a volunteer-based work, the type, quality, and quantity of information available in the project varies widely from state to state. The project's major drawback is the lack of a national search option; queries must be conducted at the county level, requiring knowledge of the subject's residence that the cataloger may lack. Enter the Archives, created in recognition of the fact that much genealogy data cannot be limited to a single county or state. The Archives' primary advantage is its national search engine, which can be limited by state and document type to more narrowly focus the search. Again, the types of available records differ depending on the state, but the researcher might expect to find data from a host of sources including vital records, biographies, family histories and Bibles, obituaries, tombstone inscriptions, and census, church, court, immigration, land, military and occupational records. Although FamilySearch is a superior option, the presence of unique resources like family histories and Bibles in USGenWeb Archives makes it a viable alternative should the former prove unhelpful.

ArchiveGrid (<https://beta.worldcat.org/archivegrid>)

ArchiveGrid is a free beta site developed by OCLC that includes more than 4 million records describing archival materials gathered from more than 1,000 cultural heritage institutions. Most ArchiveGrid records are culled from MARC records in the OCLC database, although some are drawn from finding aids contributed by participating agencies. Many of the records contain biographical information, not only about the primary subject of the collection but also

of people associated with him or her. While the amount and kind of information varies greatly, birth and death dates, fuller forms of names, occupations, and places of residence are common. ArchiveGrid continues to grow, and because many archival and special collections deal with people who are only locally known, it has the potential to become a great discovery tool for catalogers.

Ancestry.com (www.ancestrylibrary.com)

Despite the earlier declaration that only freely available resources would be examined, this fee-based site is worthy of discussion because it is often included in libraries' ProQuest subscriptions, sometimes unknowingly even to the librarians themselves. Ancestry.com consistently ranks as the top genealogical website, and with good reason. Its scope is enormous, containing 15 billion resources from almost 10,000 record collections that span the globe and extend coverage back to the 1600s, making it by far the largest genealogical resource. There are more than 3,000 US collections and a complete listing of all available collections is provided. The offerings run the usual gamut of

- census and voter lists;
- birth, marriage, and death records;
- military records;
- immigration and travel records;
- newspapers and publications;
- family histories;
- court records; and
- city directories, organizational directories, and church histories.

Even with its billions of records, though, Ancestry.com cannot claim to deliver universal coverage of genealogical data. For example, information about the author's father Halleck Long can be found in FamilySearch's 1940 census and the GenealogyBank Obituaries that is not retrieved in Ancestry.com, even though the two databases contain essentially the same records. This discrepancy might be the result of differing algorithms.³² Nevertheless, the unparalleled robustness of Ancestry.com's sources should prompt wise cataloger-detectives to scour their ProQuest offerings to see if they have access and consider it as a first option in their quest.

Conclusion

Catalogers are on the precipice of, as Schreuer calls it, a "transformative revolution" in the way we describe resources.³³ Whether we are teetering or standing firm, it is hard to say. The work of catalogers in a linked data environment will evolve in ways not clearly perceived at the moment, but

it will undoubtedly involve a continuing need to uniquely identify people, whether they are the creators, contributors, or subjects of the works associated with them. Recent research has shown how the web has made the cataloger's task of discovering biographical information much easier for living or well-known people. Under the current cataloging rules, though, all authorized access points for people must be disambiguated, whether the individuals are living or dead, famous or little known. It can be a particularly difficult chore in cataloging special collections materials, which is often the realm of the obscure. This paper has shown how genealogists have paved the way to success for the cataloger-detective through a variety of freely available online research tools. Although personal name authorized access points dominate authority files, people are not the only agents that need to be uniquely identified. Further investigation therefore is needed to explore ways in which distinctive information about organizations, families, meetings, and jurisdictions can be uncovered.

References

1. Beth M. Russell, "Special Collections Cataloging at a Crossroads: A Survey of ARL Libraries," *Journal of Academic Librarianship* 30, no. 4 (2004): 295.
2. John J. Riemer and E. Schreuer, *The Future of Undifferentiated Personal Name Authority Records and Other Implications for PCC Authority Work*, accessed April 15, 2016, www.loc.gov/aba/pcc/Undiff%20Personal%20NARs%20Discussion%20Paper%20March%202012.doc.
3. Charles A. Cutter, *Rules for a Printed Dictionary Catalogue* (Washington, DC: Government Printing Office, 1876), 10.
4. Pino Buizza, "Bibliographic Control and Authority Control from Paris Principles to the Present," *Cataloging & Classification Quarterly* 38, no. 34 (2004): 117-33.
5. Barbara B. Tillett, "Authority Control: State of the Art and New Perspectives," *Cataloging & Classification Quarterly* 38, no. 3-4 (2004): 23-41; Marie-France Plassard, "IFLA and Authority Control," *Cataloging & Classification Quarterly* 38, no. 3-4 (2004): 83-89.
6. Buizza, "Bibliographic Control and Authority," 124.
7. Mitch Friedman, "A Conversation with Frederick C. Kilgour," *Technicalities* 1 (1981): 5.
8. Ake I. Koel, "Bibliographic Control at the Crossroads: Do We Get Our Money's Worth?," *Journal of Academic Librarianship* 7, no. 4 (1981): 220-22; Arlene G. Taylor, "Authority Files in Online Catalogs: An Investigation of Their Value," *Cataloging & Classification Quarterly* 4, no. 3 (1984): 1-17.
9. F. H. Ayres, "Authority Control Simply Does Not Work," *Cataloging & Classification Quarterly* 32, no. 2 (2001): 49-59.

10. Philip Evan Schreur, "The Academy Unbound: Linked Data as Revolution," *Library Resources & Technical Services* 56, no. 4 (2012): 234; Jinfang Niu, "Evolving Landscape in Name Authority Control," *Cataloging & Classification Quarterly* 51, no. 4 (2013): 418.
11. Jackie M. Dooley and Katherine Luce, *Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives* (Dublin, OH: OCLC Research, 2010), 9.
12. Judith M. Panitch, *Special Collections in ARL Libraries: Results of the 1998 Survey Sponsored by the ARL Research Collections Committee* (Washington, DC: Association of Research Libraries, 2001), 8, 51.
13. Barbara M. Jones and Judith M. Panitch, "Exposing Hidden Collections," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 5, no. 2 (2004): 86.
14. Melissa A. Hubbard and Ann K. D. Myers, "Bringing Rare Books to Light: The State of the Profession," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 11, no. 2 (2010): 139; Dooley and Luce, *Taking Our Pulse*, 9.
15. Barbara M. Jones, "Hidden Collections, Scholarly Barriers: Creating Access to Unprocessed Special Collections Materials in America's Research Libraries," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 5, no. 2 (2004): 89.
16. Carole Mandel, "Hidden Collections: The Elephant in the Closet," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 5 (2004): 110.
17. Russell, "Special Collections Cataloging," 295; M. Winslow Lundy, "Use and Perception of the DCRB Core Standard," *Library Resources & Technical Services* 47, no. 1 (2003): 16–27.
18. ARL Working Group on Special Collections, *Special Collections in ARL Libraries: A Discussion Report from the ARL Working Group on Special Collections* (Washington, DC: Association of Research Libraries), 19; Linnea Marshall, "Using Internet Resources to Research Dates of Birth and Death of Relatively Obscure Individuals for Inclusion in Name Authority Records," *Cataloging & Classification Quarterly* 50, no. 1 (2012): 18.
19. Wendy M. Duff and Catherine A. Johnson, "Accidentally Found on Purpose: Information-Seeking Behavior of Historians In Archives," *Library Quarterly* 72, no. 4 (2002): 493.
20. Stephen E. Wiberley Jr., "Subject Access in the Humanities and the Precision of the Humanist's Vocabulary," *Library Quarterly* 53, no. 4 (1983): 423.
21. *Ibid.*, 432.
22. Mandel, "Hidden Collections," 109.
23. Elaine Beckley Bradshaw and Stephen C. Wagner, "A Common Ground: Communication and Alliance between Cataloguer and Curator for Improved Access to Rare Books and Special Collections," *College & Research Libraries* 61, no. 6 (2000): 530.
24. Chris Evin Long, "The Internet's Value to Catalogers: Results of a Survey," *Cataloging & Classification Quarterly* 23, no. 3–4 (1997): 65–74.
25. Beth M. Russell and Jodi Lynn Spillane, "Using the Web for Name Authority Work," *Library Resources & Technical Services* 45, no. 2 (2001): 73–79.
26. Nadine P. Ellero, "Panning for Gold: Utility of the World Wide Web for Metadata and Authority Control in Special Collections," *Library Resources & Technical Services* 46, no. 3 (2001): 80.
27. *Ibid.*
28. Marshall, "Using Internet Resources to Research Dates of Birth and Death," 17–32.
29. *Ibid.*, 30.
30. Loretto Dennis Szucs and Sandra Hargreaves Luebking, ed., *The Source: A Guidebook to American Genealogy*, 3rd. ed. (Provo, UT: Ancestry, 2006).
31. "Top 100 Genealogy Websites of 2015," *GenealogyIn-Time Magazine*, accessed February 2, 2016, www.genealogyintime.com/articles/top-100-genealogy-websites-of-2015-page01.html.
32. "Clash of the Titans: Ancestry.com vs FamilySearch.org. And the winner is . . . ?," Brian Sheffey, accessed February 2, 2016, <https://genealogyadventures.wordpress.com/2014/01/07/clash-of-the-titans-ancestry-com-vs-family-search-org-and-the-winner-is>.
33. Schreur, "The Academy Unbound," 227.

Appendix. List of Genealogical Websites Reviewed

Comprehensive Genealogical Websites

FamilySearch (<https://familysearch.org>)
 USGenWeb Archives (<http://usgwarchives.net/search/searcharchives.html>)
 ArchiveGrid (<https://beta.worldcat.org/archivegrid>)
 Ancestry.com (www.ancestrylibrary.com): may be included in a ProQuest subscription

Tombstone Inscription Sites

Findagrave (www.findagrave.com)
 Billion Graves (<http://billiongraves.com>)
 Interment (<http://interment.net>)
 USGenWeb Tombstone Transcription Project (<http://usgw-tombstones.org>)

Death Indexes

SSDI: available through FamilySearch (<https://familysearch.org/search/collection/1202535>) or GenealogyBank (www.genealogybank.com/gbnk/ssdi?kbid=9064&m=9)
Online Searchable Death Indexes and Records (www.deathindexes.com)

Obituaries and Newspapers

Chronicling America (<http://chroniclingamerica.loc.gov>)
Google News Archive (www.news.google.com/newspapers)
Ancestor Hunt Newspapers! Page (www.theancestorhunt.com/newspapers.html)
New York Times Archives (<http://query.nytimes.com/search/query?srchst=p>)
SmallTownPapers (www.smalltownpapers.com)

Obituary Central (www.obitcentral.com)
Obit Finder (www.legacy.com/Obituaries.asp?Page=ObitFinder&CoBrand=Legacy)

Local Histories

Internet Archive (<https://archive.org/index.php>)
Google Books (<https://books.google.com>)
Online County Histories, Biographies and Indexes (www.genealogybranches.com/countyhistories.html)

Family Tree Websites

WikiTree (www.wikitree.com)
FamilyTreeNow (www.familytreenow.com)
Crestleaf (<http://crestleaf.com>)