

The Library of Congress is one of the most important libraries in the nation, and in Chapter 5 of the new book, *Public Knowledge: Access and Benefits*, edited by the late Miriam A. Drake along with Donald T. Hawkins, Drake briefly recounts its early history, describes its mission to support Congress and “further the progress of knowledge and creativity for the benefit of the American people,” and reviews five of its major collections, including the Chronicling America newspaper project and American Memory.

The Library of Congress

Miriam A. Drake

The Library of Congress (LC), one of the nation’s great treasures, houses the world’s largest collections of books, manuscripts, newspapers, films and videos, maps, sheet music, and other materials. The legislation signed by President John Adams in 1800 transferring the seat of the federal government to Washington, D.C. from Philadelphia also provided for the establishment of a library for the exclusive use of Congress. The first library was burned in 1814 during the invasion of Washington by British troops, and shortly after the fire, President Thomas Jefferson donated his personal library to replace what had been lost. Since then, the library has continued to grow significantly, especially because of the requirement that a copy of every work registered for copyright must be donated to it.

The mission of the LC is to “support the Congress in fulfilling its constitutional duties and to further the progress of knowledge and creativity for the benefit of the American people.” It is an agency of the legislative branch of the U.S. government and provides its services through the following three major divisions:

- The **Congressional Research Service** (CRS; loc.gov/crsinfo/about) exclusively serves Congress and assists throughout the legislative process with services provided from the following research divisions:

2 Public Knowledge

- Domestic Social Policy
 - American Law
 - Foreign Affairs, Defense and Trade
 - Government and Finance
 - Resources, Science and Industry
- The **U.S. Copyright Office** (copyright.gov) carries out the functions mandated by Article I, Section 8 of the U.S. Constitution: “To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries. . . .” It is headed by the register of copyrights and was established by Congress in 1897 as a department of the LC. In FY2011, the Copyright Office processed more than 700,000 registration claims; its registration system is the world’s largest database of copyright ownership information, and with more than 45 million cards housed in 25,000-plus drawers, its card catalog is the world’s largest. (See copyright.gov/history/copy_lore.html for more fascinating information on copyrights.)
 - The **Law Library of Congress** (loc.gov/law/about) was established in 1832 in recognition of Congress’ need for legal materials. Today it has grown to more than 2.65 million volumes and is the world’s largest law library. Its collection covers nearly every jurisdiction in the world, so it is well-qualified to provide international research services through its foreign law specialists. It also has its own digital preparation and preservation facility.

Collections

Today the LC houses a large variety of materials and provides online access to a rich collection of resources, which has 158 million items—including more than 36 million books, more than 69 million manuscripts, 6 million

recordings and videos, 5.5 million maps, 7.1 million pieces of sheet music, the world's biggest newspaper collection, the largest rare book collection in North America, and more than 13.7 million prints and photographs that not only document and preserve the history and culture of the nation but also provide extraordinary resources for research and study. The manuscript collection contains exceptional and diverse resources for scholars such as an outstanding African-American collection and the papers of Clara Barton, Alexander Graham Bell, James Madison, Thomas Jefferson, Samuel F.B. Morse, and many others who contributed to the culture and history of the nation. The performing arts collection, another outstanding resource, includes coverage of music, dance, theater, recorded sound, and film. The LC's *Performing Arts Encyclopedia* (lcweb2.loc.gov/diglib/ihas/html/guide/guide-home.html) contains a guide to these important materials.

This chapter will describe five collections: electronic resources, the Chronicling America newspaper project, prints and photographs, the American Archive of Public Broadcasting, and American Memory. It is based on interviews, the collections' websites, and material provided by interviewees.

Electronic Resources

All of the LC's electronic resources are available to visitors, but due to licensing terms, some of the collection is not available online to users outside the library. Donna Scanlon, the LC's electronic resources coordinator, noted, "We have quite a variety of resources and cover all disciplines." Acquisition of electronic resources follows the library's collection policy guidelines: They supplement or duplicate print resources and fill in gaps. Scanlon added, "We run the gamut from U.S. publishers to international works in support of our international studies group. We have a lot of Russian content as well as content from China, Japan, South Korea, and Pakistan." She mentioned the breadth and depth of the collections and some of the treasures contained in them (the catalog is available at eresources.loc.gov) and said, "We want to encourage people to come in and use our resources. The coverage in terms of subjects and

4 Public Knowledge

time is awesome. For the most part feedback from users is very positive. Occasionally someone will have issues with one of the providers where digitized content is not as clear as it might be.” For example, one user commented on ShipIndex.org: “The paid version here is much more useful than the free web version and is really good for distinguishing between ships of the same name but of a different era or type of construction.” The user added, “Still working in the clipper era, approximately 1840–1900, with heavy emphasis on the 1850–1860; I can see that if I branch out into earlier periods, and start getting into the ships that don’t appear in *Lloyd’s [Register of Ships]*, the ship index will be even more useful.” Another user commenting on the 18th-century Parliamentary papers said that they contained “rich content for researching American history for the colonial period, the American Revolution, and the early republic. American history is inextricably linked to British history for this time period. [This collection is a] valuable resource for questions related to the American Revolution, such as the British Parliament’s reactions to the Declaration of Independence.”

Chronicling America and the National Digital Newspaper Program

One of the most interesting and important electronic collections at the LC is the Chronicling America project (chroniclingamerica.loc.gov), produced as part of the National Digital Newspaper Program (NDNP; loc.gov/ndnp). The program is dedicated to digitizing newspapers to provide free and open access to local and national news from the past and is a joint project between the National Endowment for the Humanities (NEH) and the LC. NEH funds awards to representatives in each state to digitize historic newspapers and send the resulting files to the library for inclusion in the collection. So far, the Chronicling America website contains about 9.6 million pages. Deborah Thomas, the NDNP’s digital projects coordinator, said, “Our goal is to reach all 50 states and the territories and to have them contribute content from their collections published between 1836 and 1922 and aggregate this content into one large searchable collection.” Thomas explained that each grantee provides

100,000 pages during the 2-year term of their grant (with some receiving additional follow-on grants to provide more content) and establishes an advisory council of journalism professors, educators, historians, and librarians to select what should be included. Digitizing is done locally to technical specifications provided by the LC. Content is then delivered to the LC for aggregating into the *Chronicling America* website. Grant recipients need to have experience with digital projects—as Thomas explained, “It is not a project for beginners. The technical specifications are fairly exacting in order to create a consistent, sustainable digital collection. Current participants—state universities, historical societies, or state libraries and archives—have strong newspaper collections in their states and a historic role in newspaper library management.”

The NDNP builds on a previous NEH-LC project, the U.S. Newspaper Program (USNP), which was carried out between 1982 and 2011. Thomas said, “That program was the first time there was a concerted effort to identify newspapers across the country, get them described and made available to libraries around the world, and selectively preserve them through microfilming.” One institution in each state headed the effort to identify all the newspapers in the state, both historical and present. The newspapers were cataloged and added to the Newspaper Union List, managed by OCLC, and are now part of the WorldCat database (oclc.org/worldcat.en.html). “The Digital Newspaper Program began in 2005,” Thomas noted. “We wanted to take advantage of all the work that had gone into the previous program and utilize microfilm for digitizing, include the bibliographic records from WorldCat and the holdings records for libraries, and incorporate them into the *Chronicling America* website.”

The next step will be to continue to expand coverage by geography and language, reaching all states and territories and capturing ethnic news in American-published papers. More than 200,000 pages from about 1,650 titles are available already in French, German, Italian, and Spanish from 37 states and territories. Thomas said, “We see continued growth moving to cover the entire nation and further addition of ethnic titles in a broad range of languages. We want to build a rich, in-depth, and representative collection of historical newspapers within our selected time period. There

6 Public Knowledge

are some technical reasons why we chose to limit newspapers to 1836–1922. For example, prior to the 1830s, newspapers were less about community affairs than political issues, influencing the structure of newspapers, the content, and the size of the type, and newspapers published after 1922 are currently often under copyright.” She commented on the role of newspapers and their value:

The most important aspect of the program is aggregation and what the aggregation of news sources can tell us about history. The stories you think are covered only in your town may be covered in many different locations which have different perspectives and different voices. The value of newspapers is that they represent the voice of their community. Having all these materials available in one place with such a broad range of coverage, deep coverage of one event can lead to significant developments in research and discoveries: what people are doing and thinking at the time of a given event. An excellent example of that variety is the death of Abraham Lincoln and the comparison between how people in the South reacted to it versus people in the North. Some of the reactions may be surprising.

The number of people using the Chronicling America site continues to grow. In 2013, the average monthly number of visits was approximately 350,000, and the average number of webpages viewed was approximately 2.9 million. But the numbers tell only part of the story: Users derive great value and information from the site. For example, in November 2013 a devastating typhoon in the Philippines destroyed the town of Tacloban. Thomas noted, “A few days after the typhoon, we had six times the normal usage—a dramatic increase. We eventually tracked the searches back to someone discovering a 1912 typhoon that also decimated Tacloban. About the same number of people were killed in each storm. The story was noticed on Facebook and went viral in the Philippines over the internet and in Philippine news sources for about a week.”

The Chronicling America site also supports the digital humanities with a master set of the content available for various kinds of uses. Thomas described one researcher at Northeastern University “who is tracking different stories that appear in different newspapers over time, sometimes years apart. The stories are essentially filler material distributed by the original publisher; other newspapers insert them as fun vignettes. Sometimes someone may cut out a part they do not like or change the wording. This research is best done with a large amount of data and computational analysis.”

Another researcher studied the effect of the press on public opinion and the spread of influenza in 1918. Thomas said, “He and partners at the National Library of Medicine were analyzing the data together. The end goal was to inform current epidemiologists how to utilize the press to find things that make a difference in an epidemic situation. Harvesting all the data in that six-month period allowed the researchers to look at what happened in various places and how the newspapers treated the events. The key aspect to this kind of research is making data open and available through standard protocols. One of the things researchers crave is reliable and authoritative datasets.”

Thomas also provided another type of research example: “Barry Popik, a contributor to *The Oxford English Dictionary* and a blogger, frequently refers to newspapers found in Chronicling America in his research into the origin of ‘Americanisms’ and integrates newspaper references with primary sources. His blog, The Big Apple (barrypopik.com), includes more than 200 entries citing Chronicling America source material. Examining the newspapers available through Chronicling America encourages a wide range of uses—there’s something for everyone in newspapers.”

Prints and Photographs

Newspaper journalism goes beyond words—it also encompasses photographs that provide visual images of people and events. Many images originally appearing in newspapers can be viewed using the LC’s Prints & Photographs Online Catalog (loc.gov/pictures), which contains about 14

8 Public Knowledge

million photographs, posters, cartoons, original drawings, and historical prints, as well as ephemera such as postcards and bumper stickers. Helena Zinkham, chief of the LC's prints and photographs division, said, "The oldest photograph is an 1839 self-portrait made by Robert Cornelius. It is the oldest selfie" (see Figure 5.1).

Zinkham pointed out that although the library collects items from around the world, the photographic collections are primarily American, with international coverage for areas where the U.S. is involved or where a major event has impacted the U.S. She added that the most frequently consulted and most popular collections are close to 200,000 photographs from the Great Depression, World War II, the American Civil War,



Figure 5.1 1839 self-portrait by Robert Cornelius, from the LC's Prints & Photographs Online Catalog

American architecture and landscape, and a group of about 2,000 photographs from the 20th century documenting the Russian Empire. The most famous photograph is Dorothea Lange's 1936 "Migrant Mother," of Florence Thompson and three of her seven children (see Figure 5.2).

Zinkham discussed photographs that showed views of the U.S.: "With 200,000 photos covering the entire United States, you can see the effects

of logging, what it is like to live in a city, what it is like to be poor, what it is like to work in a munitions factory coming out of the Depression, and the home front in World War II. There are very compelling faces of men, women, and children, as well as the physical landscape,” which includes depictions of where pipelines are laid, farmlands, and terrain showing what erosion looks like.

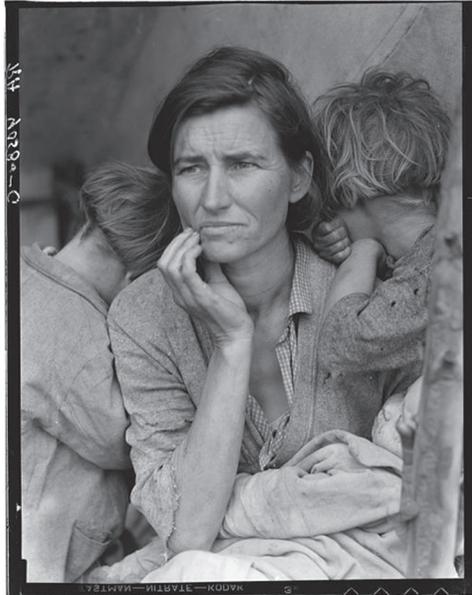


Figure 5.2 “Migrant Mother” photograph from the LC’s Prints & Photographs Online Catalog

The poster collection is a rich history of politics and commerce covering areas such as the performing arts, including the circus, theater, vaudeville, dance, and moving images. Political posters are another strength that includes propaganda as well as presidential campaigns and political activism throughout the world. Other posters show travel locations and advertisements for household products, health campaigns, and so on. Zinkham said that a tremendous baseball card collection is also available, and most of the cards are old. She noted, “The first card everyone asks about is Babe Ruth,” but the collection did not include a Babe Ruth card until 2014 (loc.gov/pictures/item/2014646962).

10 Public Knowledge

Digitizing the collection is prioritized to the most heavily used materials. Zinkham indicated that there is provision for prints and photographs to be scanned on demand. Some of the new collections are heavily rights-restricted and can only be used on-site. The library recently received “a wonderful gift of posters from Amos Kennedy, a well-known letterpress poster printer. They are copyrighted and we will not digitize them,” said Zinkham.

The LC is a founding contributor to the Flickr Commons (flickr.com/commons), to which more than 100 libraries around the world have contributed rights-free photographs. (The British Library recently contributed 1 million of them.) The goal is to share hidden treasures from the world’s public photography archives. Zinkham said, “It has been wonderful to have people from all over the world in our Flickr community weigh in and help tell us more about what is happening in a picture or the names of the people involved.”

Users of prints and photographs include the public, historians, filmmakers, teachers, and specialists in particular subjects. One large group of users is Civil War enthusiasts who may be reenactors, people who live in the area of a Civil War battle, or family historians seeking a relative. Artists, photographers, and printmakers study older works as sources of inspiration for the creation of new works.

The American Archive of Public Broadcasting

The creation of new works in entertainment and education goes beyond print and photographs and today relies increasingly on television, film, video, and moving images. Public television began as an educational medium in the 1950s. Now, the LC and station WGBH in Boston are collaborating on a project to digitize National Educational Television (NET) broadcasts, the predecessor of PBS (Public Broadcasting Service; it launched in 1970). The goal of the project is to create a permanent and centralized archive of U.S. radio and television public broadcasting, the American Archive of Public Broadcasting (americanarchive.org). The project was anticipated by the Public Broadcasting Act of 1967, which called for an archive. WGBH began to archive its material in 1979.

In 2007, when stations were changing over to digital mode, Congress allocated funds through CPB (Corporation for Public Broadcasting) for the digital transition. Alan Gevinson, special assistant to the chief of the National Audio-Visual Conservation Center (a facility for preserving audiovisual material) at the LC, said, “A lobbying group for public broadcasting stations thought that this would be a good time to digitize the analog material sitting on their shelves so that they could use it in the future.” Congress therefore allowed some of the digital conversion funds to be used for the digitization of older material. Shortly thereafter, CPB introduced the American Archive via a pilot program with Oregon Public Broadcasting that featured civil rights-related and veterans’ stories from World War II. It involved 24 stations and about 2,500 hours of material. Ken Burns, a director and producer of documentary films, was doing his series on World War II around that time, and the stations had a significant amount of material related to his project that was used in CPB’s archive.

The LC and WGBH have received a grant from CPB to further develop PBCore, the public broadcasting metadata dictionary that sets metadata standards for audiovisual material. The American Archive project funded about 120 stations to inventory their shelves and created about 2.5 million records. After that, CPB funded another project to digitize 40,000 hours of material and gave a grant to the LC to permanently preserve it. Much of the material that the stations selected was an excellent source of their local events and local history. In order for this project to grow in the future and for the material to be accessible to library and WGBH visitors, as well as on the web, it will be necessary to obtain rights to the material.

Gevinson discussed the technical problems associated with the analog-to-digital conversion: Some of the older material is in better condition than the videotaped material from 1971 and later, which is at risk because it degrades. Although the library has a state-of-the-art storage facility, much of the material is housed on shelves at local stations all over the country and is not properly stored. Once the material has been digitized, it needs to be periodically migrated and digitized again for archival and preservation purposes.

12 Public Knowledge

The library also has a large collection of 8,000 to 10,000 items covering NET and dating from around 1952 to 1972 that are waiting to be digitized—the material is old and most at risk of degrading. Gevinson observed, “By about 1964, commercial networks were putting less and less effort into producing documentaries, and NET viewed that as an opportunity for them.” The library is working on raising money for the project.

Public television and NPR (National Public Radio) have been responsible for some of the most important programs that informed, educated, and entertained the country. Many PBS programs were rejected by commercial advertisers, but they document our history and culture. However, programs produced for local viewing are not accessible to everyone. Audiovisual documentations of our history provide extraordinary resources for students, scholars, authors, and researchers. We also know that today’s students learn in different ways. Earlier generations learned primarily by reading, but some students today learn more effectively using audio, visual, and tactile models. The American Archive of Public Broadcasting will provide a rich treasure that will contribute to learning and resources for all.

American Memory

American Memory (memory.loc.gov/ammem/index.html) is a multimedia online resource on the history and culture of the U.S. It began with a pilot digitization program in 1990 that made selected collections available to 44 schools and libraries on CD-ROM. In 1994, the pilot concluded, but recipients of the CD-ROMs were highly enthusiastic about them and wanted more, similar resources. The National Digital Library Program was formed—aided by private-sector donations—to create an archive of historical materials from the LC’s rich collections. Congressional support and private-sector sponsorships ensured the continuance of the project. It reached its goal of making 5 million items available by 2000.

As a result of a \$2 million gift from Ameritech Corp., the LC sponsored the National Digital Library Competition between 1996 and 1999 (lcweb2.loc.gov/ammem/award/index.html) to enable libraries

and archives to create digital collections of their resources. Twenty-three winners had their collections incorporated into the American Memory project. Then librarian of Congress James H. Billington said, “Integrating these collections into our National Digital Library Program will not only enhance the depth and breadth of available materials, but also make it truly national in scope. . . .” It is particularly interesting to note the lessons that the 23 winners learned in preparing their digital entries (which are available on the competition’s website); they fell into the broad areas of reproduction formats, workflow and project management, staffing and skill requirements, intellectual access, and publicity and other outcomes.

American Memory now has more than 9 million items organized into 100-plus collections on subjects as diverse as African-American history, the Chesapeake Bay, Abraham Lincoln, Coca-Cola advertisements, and the Texas border. Materials run the gamut of formats, including documents, maps, photographs, and recordings. For example, there are about 65,000 papers from George Washington, which constitute the largest collection of his original documents in the world (memory.loc.gov/ammem/gwhtml/gwhome.html).

A special feature of American Memory is a Today in History section on its homepage that provides interesting tidbits and trivia for every day of the calendar. Another section of its homepage explains that many of the collections developed for American Memory have been migrated to new presentations and integrated into the library’s general collections.

Conclusion

The LC is not only a national treasure but a worldwide one. Although its overall mission, as reflected in its name, is to be the library for the U.S. Congress, it has a far broader appeal. A description of all of the LC’s resources is well outside the scope of this book; this chapter has focused on electronic resources and has highlighted some of the most useful ones for researchers, scholars, students, and anyone else interested in American history and culture.

About the Editors

Miriam A. (“Mimi”) Drake (1936–2014) was dean of libraries at Georgia Institute of Technology. She planned and implemented the Georgia Tech Electronic Library and specialized information services for faculty and students. She was the author or editor of more than 100 articles and four books, received several industry awards and honorary doctorates from Indiana University–Purdue University–Indianapolis (IUPUI) and Simmons College, and was president of the Special Libraries Association (SLA) and the Association of Information and Dissemination Centers (ASIDIC).

Donald T. Hawkins is an award-winning conference blogger and information industry freelance writer. After receiving his Ph.D. degree from the University of California–Berkeley in 1970, he worked in the AT&T Bell Laboratories Library Network before joining ITI as editor-in-chief of *Information Science & Technology Abstracts*. He was also secretariat for the Association for Information and Dissemination Centers (ASIDIC) and editor of the ASIDIC Newsletter and has published more than 350 articles. He is the editor of *Personal Archiving: Preserving Our Digital Heritage* (Information Today, Inc., 2013).

This chapter will appear in Public Knowledge: Access and Benefits, edited by Miriam A. Drake and Donald T. Hawkins. For more information visit <http://books.infotoday.com>.