



LIBRARY OF CONGRESS
Office of the Inspector General



Library Services

The Library Collects Extensively but Faces Increasing Challenges in Processing, Controlling, Storing, and Making Accessible All It Collects

Report No. 2013-SP-102
September 2013

FOR PUBLIC RELEASE



UNITED STATES GOVERNMENT

LIBRARY OF CONGRESS

Memorandum

Office of the Inspector General

TO: James H. Billington
Librarian of Congress

September 30, 2013

FROM: Karl W. Schornagel
Inspector General

SUBJECT: *The Library Collects Extensively but Faces Increasing Challenges in Processing, Controlling, Storing, and Making Accessible All It Collects*
Report No. 2013-SP-102

This transmits our final report summarizing the results of the Office of the Inspector General's evaluation of the Library's framework of policies and procedures supporting the development of collections and the logistical impact of these policies and procedures. The Executive Summary begins on page *i* and complete findings appear on pages 20 to 46.

Although no recommendations were made, we provided a draft of this report to the Office of the Librarian and Library Services for their review and comment. The full text of their response is included as an appendix to the report. The report will be available to the public.

We appreciate the cooperation and courtesies extended by the Library during this evaluation.

cc: Deputy Librarian
Associate Librarian for Library Services

▶▶ TABLE OF CONTENTS

▶▶ Executive Summary	<i>i</i>
▶▶ Background	1
▶▶ Objectives, Scope, and Methodology.....	16
▶▶ Findings	20
I. The Library has been Meeting Its Top Priority of Continually Obtaining a Universal Collection of Both Analog and Digital Materials	20
II. The Library is Not Able to Timely Process the Inflow of Materials It Acquires–Current Unprocessed Arrearage is Approximately 28 Million and Has Grown Almost 50 Percent Since 2000	27
III. Growing Magnitude of the Collections Puts an Overwhelming Demand on Library Storage Space That Comes at a Cost.....	34
IV. Other National Libraries Face Similar Constraints but Have Addressed Them Differently	41
▶▶ Appendix A: Criteria for Selecting Viable Trusted Partners.....	48
▶▶ Appendix B: Benchmarking of National Libraries	53
▶▶ Appendix C: Digital Preservation.....	57
▶▶ Appendix D: Current Status <i>LC21: A Digital Strategy for the Library of Congress</i>	66
▶▶ Appendix E: Cataloging Priorities	69
▶▶ Appendix F: Management Response.....	71

▶▶ EXECUTIVE SUMMARY

The Library of Congress' (Library) mission 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. To accomplish this mission, the Library must acquire, preserve, store, and make accessible materials from around the world that it deems relevant to the American people, Congress, and the federal government. Founded in 1800, the Library of Congress is the nation's oldest federal cultural institution, holding more than 155 million items on some 838 miles of shelves. These items include books, manuscripts, maps, prints and photographs, printed music, sound recordings, moving images, and microforms. Each workday, the Library receives about 15,000 items and adds approximately 11,000 items to its collections. In addition to its three Capitol Hill buildings and its National Library Service for the Blind and Physically Handicapped facility in Washington, DC, the Library operates six overseas offices and stores collection material in purpose-built facilities in Maryland and at the National Audio Visual Conservation Center in Culpeper, VA (where more than 90 miles of shelving house its collections of nearly 5 million film, video, and sound recordings, together with extensive modern facilities for the acquisition, cataloging, and preservation of all audiovisual formats). The Library's digital collection also includes media such as archived Web sites, e-books, and tweets.

In the past decade, experts thought that with the internet/digital explosion, the material to collect would significantly shift from analog (tangible)¹ material to digital. However, this shift has not occurred. While the volume and types of material have significantly increased, cuts in staff and appropriated funds have hindered the Library's ability to accomplish other critical parts of its mission—that of processing, preserving, properly storing, and providing access to all its collections.

In response to a 2006 Office of the Inspector General (OIG) report that recommended the Library reduce the amount of analog material it acquires to alleviate issues surrounding a lack of space and damage, the Library stated that acquisitions are a top priority that is inescapable in the Library's pursuit of

¹ For the purposes of this report, analog is defined as non-digital material.

fulfilling its mission. The objective of this evaluation was to assess the framework of policies and procedures supporting the development of Library collections and assess the logistical impact of these policies and procedures.² We found that the Library is successful in its mission to obtain a universal collection of material for future generations. However, without relief or substantially adjusting its priorities, the Library is confronted with several critical challenges in its processing and storage of materials. These challenges become increasingly difficult to manage with the recent reductions in budget and staff. Further, these challenges make it difficult for the Library to meet one of its critical mission components: to make its collections accessible. To determine how other national libraries addressed such challenges, we benchmarked the Library with two other national libraries and found that these national libraries approached similar challenges differently. Our findings are discussed below.

The Library has been meeting its top priority of continually obtaining a universal collection of both analog and digital materials

Based on its “canons of selection” developed in the early 1940s, the Library’s analog and digital collections have grown almost 22 percent since 2003.³ In fiscal year (FY) 2012, the Library acquired more than three million items through purchase, gift, copyright deposit, and exchange and transfer from other government agencies. The U.S. Copyright Office alone forwarded more than 600,000 copies of works to the Library’s collections in FY 2012, most of which were in analog format.⁴ The Library adds reprints and new editions to its literary works, such as *Wuthering Heights* (with 105 copies in various editions) and *The Adventures of Tom Sawyer* (with 128 copies in various editions). The Library purchased approximately 737,000 U.S. and foreign items in 2012, of which more than 508,000 were foreign items, including some from

² This evaluation did not assess whether the items collected were the appropriate material to collect.

³ The collections have grown from 127 million to more than 155 million and at its current pace will grow to over 178 million by 2020.

⁴ Works that are published in the United States are subject to mandatory deposit with the Library. The owner of copyright or the owner of the exclusive right of publication in the work has a legal obligation to deposit two copies (or in the case of sound recordings, two phonorecords) in the Copyright Office for the use or disposition of the Library of Congress.

third-world countries. Further, the Library has the largest collection of Russian, Chinese, Japanese, and Korean materials outside of those countries. Half the Library's book collections are in languages other than English, representing 470 languages.

At its National Audio Visual Conservation Center, the Library acquires, catalogs, preserves, and provides access to a collection of audiovisual formats and has collected more than 1.3 million film, television, and video items and nearly 3 million sound recordings containing music, spoken word, and radio broadcasts. In 2012, to help guide it in its effort to collect and preserve relevant material, the Library issued its National Recording Preservation Plan as a result of the National Recording Preservation Act of 2000. The Act required the establishment of a Board that will assist the Library in the development of a coordinated national sound recordings and collections policy, among other duties.

Although the majority of the Library's collections are in an analog format, the Library is also increasing its digital collections. For example, as of September 2012 more than 37 million original source analog items have been digitized since 1990, and 233 billion tweets were collected through February 2013. The Library started its digital collections via the American Memory project, begun in 1990, which provides a digital record of American history and now features more than 100 thematic collections online. In 2000, the Congress appropriated monies to the Library for the creation of the National Digital Information Infrastructure and Preservation Program (NDIIPP), and as of 2012, the Library had 200 partners in 47 states and 39 nations working cooperatively to develop standards and preserve at-risk digital content.

Our audit work focused primarily on analog material. We intend to review digital issues in more depth in subsequent audit work.

The Library is not able to timely process the inflow of materials it obtains for its collections, hindering accountability, security, and accessibility of collections

As of September 30, 2012, there were more than 28 million analog and digital items unprocessed in the Library, an increase of almost 50 percent since FY 2000 (not including tweets). The unprocessed arrearage currently represents 18 percent of the Library's total collections. Collections of manuscripts makes up more than half the arrearage. Some items, received as far back as the 1980s, have not been processed, and other items are so unique—e.g., Balinese manuscripts written on palm leaves—that they require specialized knowledge to process. The Library states that it must acquire material as it becomes available because it may not be available again, or because its importance may only be known at a later date.

Most unprocessed material is not available to researchers, and the bulk of the material lacks adequate collection controls, including bibliographic, inventory, and security controls. Bibliographic and inventory controls allow the Library to account for and track materials as well as connect users with the Library's resources. Security controls, such as marks and labels, are applied to newly acquired materials to help protect the collections from theft. The lack of these controls increases the risk of material being stolen, lost, or forgotten. Due to the absence of these controls over unprocessed materials, we could not design tests to determine the extent to which the risks exist. However, we observed some instances where unprocessed material was not adequately secured, and other materials that we attempted to locate could not be found. With a significant portion of the Library's collections in arrearage, addressing the logistical challenges of accountability and security for these items is key to properly safeguarding and making them accessible to the public, an important component of the Library's core mission.

We were not able to determine a clear cause for the arrearage. One factor could be the reduction in staff while the material collected continues to increase. Other factors contributing to the arrearage could be the lack of: (a) a digital preference for copyright deposits; (b) archivists, preservationists, and catalogers—with knowledge in various languages and

experience working with rare, old, fragile, or unique material; (c) a system-wide strategy that marries its collections efforts to available staff and proper storage; or, (d) a coordinated national collections effort and policy, similar to that developed under the National Recording Preservation Plan.

The Library faces a challenge in its mission to properly store its collections

The Library's most prized storage facility is the historic Thomas Jefferson Building, opened in 1897, where many of America's Heritage assets are stored. However, the Jefferson Building currently houses upwards of a million overflow volumes, far exceeding its capacity. This excess is causing premature damage and deterioration to some of the items, and requiring staff to periodically shift materials to avoid structural damage to the building due to excessive weight. According to the Architect of the Capitol, the agency that manages the Library's buildings, the Library's three Capitol Hill buildings are collectively at 110 percent of their capacity.

Preservation of Library materials requires appropriate temperature and relative humidity; different formats have different ideal temperatures and humidity set points. Sophisticated, highly secure storage buildings (modules) were authorized by the Congress, with the first module opened in 2002. However, only four units have been constructed. Even if funding is received and the fifth module is built, the resulting space would not accommodate the overflow. The Library estimates that it needs a new module every two years for its ever-growing collections of books and bound periodicals. The Library leased a temporary facility in 1975, but the facility is not adequate for the Library's security and preservation needs. The Library paid \$51 million in rent over the lease term (which is the equivalent of approximately three new modules) and plans to rent for another 10 years, paying another \$36 million—the cost equivalent of two additional modules. The Library identified an additional temporary solution in a National Archives and Records Administration facility in Illinois (to store analog material that is digitized), but it only accommodates a small part of the overflow, and it

is not easily accessible.⁵ As the Library adds an average of more than 3 million items, including an average of 447,000 books and periodicals, to its collections annually, the shortage of space will continue to grow.

Other national libraries faced similar reductions in funding and space issues but have addressed them differently

The British Library and Library and Archives Canada (Canadian Library) realized that “staying the course” and “going it alone” in their collection strategy were not viable options because of the economic and budgetary environment and because it was no longer achievable or appropriate to universally collect due to the increasing magnitude of published material. Our benchmarking with these libraries indicated that both institutions:

- *revised their collection strategy.* The British Library will not expend significant funds toward material that is “adequately achieved” in the country of origin, will not acquire in areas well served by other libraries, and focused its approach on a smaller number of subject priorities. The Canadian Library recently adjusted its acquisition strategy to focus on Canadian heritage material and established a more collaborative approach with provincial and territorial libraries to build complimentary collections.
- *recently increased their storage capacity.* The British and Canadian Libraries received funding for storage facilities that enabled the discontinuation of a number of lease-held facilities and increased its storage capacity to accommodate future acquisitions.
- *are transitioning to digital deposit as a preferred method for copyright material.* Going forward, the British Library’s preference is digital rather than analog for its material that is similar to the Library’s copyright deposits. Also, the British Library has placed greater emphasis on attracting external

⁵ In its response, the Library stated that its plans have now been expanded to go beyond the original category of materials sent to the Illinois facility. In addition to the existing issues associated with this off-site storage, the storage of non-digitized material will have associated access costs that may affect an already constrained budget.

investments and fundraising to supplement funding for library initiatives (and has secured an agreement with Google to digitize 250,000 out-of-copyright books from the British Library collections at Google's expense). Digital is now the Canadian Library's default choice in the acquisition of material.

Options for the Congress and Library to Consider

The Library acquires materials at rates that exceed its ability to timely process, make accessible, and safely store them. Results of this imbalance are evident by the issues we discuss in this report. We are not making recommendations in this report because some solutions may require congressional action, but we are presenting short- and long-term considerations. According to the Library, "the acquisition policies of the Library are what have made it—and what will keep it—the greatest repository of knowledge in history and a source of national pride. Changing acquisition policies will fundamentally change the nature of the Library of Congress."

Congress may decide to maintain the current acquisition policies and collection rate because it is of vital importance to the Library, Congress, and the Nation. Nonetheless, there is an imperative for the short-term that Congress should consider:

1. The Library needs resources to process and obtain sufficient, adequate facilities to store the current arrearage and overflow. The issues that we highlight in our report are symptomatic and, if resources are not provided for processing and storage capacity, will increasingly impair the Library's ability to meet its congressional mission.

For the long-term, some options Congress may want to consider are to:

2. Require the Library to create a system-wide strategy that marries future collections efforts to its ability to timely process, properly store, and make accessible the material it acquires. This strategy may require Congress to establish changes in how the Library collects, such as only digitally for newly copyrighted

materials (similar to our benchmark libraries), and the amount it collects in any given year; it may require the Library to aggressively pursue additional private monies to provide for housing and processing its collections, rather than totally relying on congressional funding; and it may require other tactics such as expanding the Library's internship programs with colleges and universities in preservation and cataloging.

3. Create a board similar to that required in the National Recording Preservation Act of 2000 that can develop, among other things, a coordinated national acquisitions and collections policy and "trusted partnerships" with organizations in the public and private sectors, similar to arrangements that other national libraries have made.

Management Response and OIG Comments

In its formal response to our draft report (see Appendix F), Library management agreed with our comments while specifically citing its disagreement with certain comments and items presented.

The Library agreed with our finding that it has been meeting its top priority of continually obtaining a universal collection of materials. The Library stated that sustaining a collection that meets the needs of Congress and the American people is essential and is its pre-eminent consideration. The Library agreed that it was not able to timely process the current arrearage of 28 million items, a 50 percent increase of its unprocessed arrearage since 2000. Management attributed delays in processing acquisitions to "primarily a matter of resources." However, management disagreed with our narrative regarding the arrearage in manuscripts and feels that our report does not take into account that many of these collections will be processed at the macro level, and that processing times will often vary even with a consistent level of processing resources. We agree that processing times will vary, and in our report, we reflect examples of the difficulties in processing the materials. The arrearage numbers used in our report are taken directly from the Library's annual reports to Congress; the Library does not report data on processing

times of the various types of material it collects, and our evaluation did not attempt to perform such time and motion studies. However, we did analyze the trend of material in arrearage since 2000; regardless of manuscripts' proportion of the arrearage, the arrearage amount continues to increase while the resources to process it diminish. We acknowledge the Library's initiatives to manage the arrearage cited in its response, but there is no evidence to assure that any or all of them will materially affect the current arrearage trend, and the initiatives will not change the imperative for the short term that Congress may want to consider.

The Library agreed that the growing magnitude of the collections puts an overwhelming demand on the Library's storage space and that it comes at a cost. However, the Library further stated that a number of examples cited in the report, including photographs, mischaracterized the impacts of delayed processing and were not representative of overall conditions at the Library. There are many effects from the delay in processing and our report is not meant to be representative of conditions at each building and division. For example, at Ft. Meade we did not observe any adverse storage conditions; on the other hand, the Architect of the Capitol has stated that the three Capitol Hill buildings are collectively over capacity and requested that excess materials stored on booktrucks be moved around to prevent compromising the structural integrity of the bookstacks. Finally, in assessing OIG benchmarking with the British Library and Library and Archives Canada, the Library stated it is acting on most of the areas we identified, but management does not believe that changing collections development policies is a responsible option for the Library and that reductions in other Libraries' collections scopes impose a greater reliance on the Library of Congress to fill the gaps created by those reductions.

Library Services Oral Comments to Discussion Draft

Prior to issuing our draft report for the Library's comments, we conducted a meeting with Library Services' senior management to obtain their feedback. Overall, Library Services agreed with the general message of the report and felt that it positively characterized the issues addressed. Library Services stated that any change to the comprehensiveness of its collections and collection policies would be detrimental to

the Library's worldwide reputation and leadership; the imperative is to explore collaboration and storage alternatives. Library Services specifically disagreed with the two national libraries we selected for benchmarking—the British Library and Library and Archives Canada. Library Services stated that the two national libraries were facing fiscal difficulties and were, therefore, not appropriate to benchmark. Library Services recommended that we benchmark smaller libraries operating more effectively in specific areas, such as digitization. We believe these libraries were appropriate benchmarks because they are two large, national libraries that face the same fiscal challenges of the Library of Congress, with similar roles and responsibilities to their nation.

Library Services added that, in the following areas, further research and analysis would assist the Library in addressing the issues of storage, partnerships, digitization, and the allocation of diminishing resources.

- Research acquisition and digitization efforts made by other federal agencies and identify how the Library can leverage those efforts.
- Identify lessons learned on digital preservation and assess the cost effectiveness of digitization as compared to analog storage.
- Analyze the benefits of establishing relationships with state and local governments and identify partnerships the Library could establish across the nation to make its materials more accessible.
- Research the benefits of additional overseas offices, such as in China.

We acknowledge the importance of these areas and will consider and evaluate these items as we develop future audit plans.

▶▶ BACKGROUND

The Library of Congress' Mission

The Library of Congress' (Library) mission 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. In 1870, Congress centralized all U.S. copyright registration and copyright deposit activities in the Library. This congressional action placed upon the Library the full responsibility for collecting and maintaining the record of the nation's cultural and intellectual heritage and preserving the nation's creativity.

The Collections Policy Statements govern the Library's collections development and acquisition efforts. They provide the policy framework to support the Library's responsibilities to serve the Congress as well as the U.S. Government as a whole, the scholarly community, and the public. The policies provide a plan for developing the collections and maintaining their strengths. Additionally, they set forth the scope, level of collecting intensity, and goals sought by the Library to fulfill its service mission.

The Library is a broad collector, covering virtually every discipline and field of study, including the entire range of different forms of publication and media for recording and storing knowledge, with the exception of technical agriculture and clinical medicine (where it yields to the National Agricultural Library and the National Library of Medicine, respectively). The Library's goal is to formulate statements that are sufficiently inclusive to ensure this broad coverage, yet specific enough to serve the particular needs of the Library's stakeholders.

The Library has been developing its body of Collections Policy Statements since the mid-20th century and has based its formulation on three broad fundamental principles, or "canons of selection"⁶ which are:

- The Library should possess all books and other library materials necessary to the Congress and

⁶ The Canons of Selection were developed by Archibald MacLeish, the Librarian of Congress in the early 1940s.

- the various officers of the federal government to perform their duties;
- The Library should possess all books and other materials (whether in original form or copy) which record the life and achievement of the American people; and
 - The Library should possess in some useful form, the records of other societies, past and present, and should accumulate, in original or in copy, full and representative collections of the written records of those societies and peoples whose experience is of most immediate concern to the people of the United States.

Overview of Library Funding and Staffing Levels

The Library relies primarily on congressional appropriations to support its programs and operations. In addition to appropriations made directly to the Library, other government agencies (for example, the Architect of the Capitol and U.S. Capitol Police) use congressional appropriations and other financing sources to provide support to the Library. These appropriations go directly to those agencies and include funding for construction, structural care and maintenance of the Library's buildings and grounds, and collections security services. In addition, the Library receives donations from the public in the form of gifts and trusts; in 2012, the Library received \$35 million in donations.⁷

In FY 2012, the Library received appropriations totaling \$587.3 million, 6.57 percent lower than in FY 2011. Figure 1 shows the Library's total appropriated dollars for FY 2006 through 2012.

⁷ Of the \$35 million in donations, \$27 million were donations of property and services, and \$8 million were donations of cash or securities.

Fiscal Years 2006 - 2012 Appropriated Dollars							
Dollars in Thousands							
Appropriation	2006	2007	2008	2009	2010	2011	2012
National Library	\$385,509	\$381,247	\$388,461	\$412,680	\$439,801	\$431,785	\$413,743
Copyright Office	22,429	22,662	5,321	18,277	20,864	17,829	16,137
Congressional Research Services	99,907	100,786	102,344	107,323	112,490	111,017	106,790
Books for the Blind & Physically Handicapped	53,905	53,614	66,923	68,816	70,182	68,046	50,674
Total	\$561,750	\$558,309	\$563,049	\$607,096	\$643,337	\$628,677	\$587,344
Percentage of Change	-	-0.61%	0.85%	7.82%	5.97%	-2.28%	-6.57%

Figure 1: Library of Congress appropriations for FYs 2006-2012.
 Source: Library of Congress, Office of the Chief Financial Officer.

Over a six-year period, the number of full-time equivalent personnel (FTEs) decreased from 4,302 to 3,746—a decrease of 556 FTEs or 12.9 percent. See Figure 2.

Fiscal Years 2006 - 2012 Authorized FTEs							
Appropriation	2006	2007	2008	2009	2010	2011	2012
National Library	2,915	2,849	2,632	2,594	2,492	2,492	2,492
Copyright Office	530	523	475	475	475	475	475
Congressional Research Services	729	705	675	675	675	651	651
Books for the Blind & Physically Handicapped	128	128	128	128	128	128	128
Total FTEs	4,302	4,205	3,910	3,872	3,770	3,746	3,746
Percentage of Change	-	-2.25%	-7.02%	-0.97%	-2.63%	-0.64%	0%

Figure 2: Library of Congress authorized FTEs for FYs 2006-2012.
 Source: Library of Congress, Office of the Chief Financial Officer.

Despite the reductions in funding and FTEs, the number of materials acquired for the Library’s collections has increased. Over the same six-year period, the Library’s collections have increased by 15 percent from 134.8 million to 155.3 million items.

Acquisition and Collection Activities and Applicable Resources

Library Services, a unit within the Library, is responsible for collection activities. Its mission is to “acquire, organize, preserve, secure, interpret, make accessible, and sustain for the present and future use of the Congress and the Nation a comprehensive record of American history and creativity and a universal collection of human knowledge.” There are five directorates under Library Services: Collections and Services, Preservation, Partnerships and Outreach Programs, Technology Policy, and Acquisitions and Bibliographic Access (ABA). These directorates perform mission functions in support of the Library’s goals and priorities.

For this report, there are three directorates that play a critical role. The Collections and Services Directorate is responsible for developing the Library’s collections in all languages, subjects, and formats; organizing, managing, and storing the collections materials; serving requested items to patrons; providing on- and off-site reference/information services through specialist librarians; and supplying intellectual, curatorial, and expert knowledge for, among other things, acquisitions. The Preservation Directorate coordinates and oversees all activities throughout the Library relating to the preservation and physical protection of Library materials.

The ABA acquires materials for the Library in all formats—books, periodicals, maps, manuscripts, music, prints, photographs, recorded sound, videos—and in all subjects from all over the world. It is also responsible for the cataloging of physical and digital materials for the Library’s collections and making them accessible to the Congress, on-site researchers, and Library Web site users.⁸ ABA’s acquisitions and cataloging functions are performed under the following divisions:

⁸ There are exceptions to the materials the ABA divisions catalog. Special format materials, such as maps, pictures, three-dimensional objects, and manuscripts are cataloged by the respective collection division that made the recommendation to acquire the material.

- African, Latin American & Western European Division⁹
- Asian & Middle Eastern Division¹⁰
- U.S./Anglo Division¹¹
- Germanic & Slavic Division¹²
- U.S. & Publisher Liaison Division¹³
- U.S. General Division¹⁴

Generally, the Library's activities involved in collecting, processing, preserving, storing, and providing access to its collections are:

Collecting—As established in 2 U.S.C. § 131, the Library of Congress is composed of collections "united under authority of law" and added to "from time to time by purchase, exchange, donation, preservation from publications ordered by Congress, acquisition of material under the copyright law, and otherwise." This basic law, together with the mandatory deposit provisions of the Copyright Act, provide the foundation for the various methods the Library uses to build what the Library Strategic Plan describes as "a universal collection of knowledge and the record of America's creativity."

⁹ The African, Latin American, and Western European Division acquires (via purchase and non-purchase), processes, and catalogs materials from all sub-Saharan African, Iberian, Latin American, South American, and Caribbean countries, and from France, Belgium, Luxembourg, the Netherlands, and Italy and processes and/or catalogs some materials from the Rio de Janeiro and Nairobi Offices.

¹⁰ The Asian and Middle Eastern Division acquires (via purchase and non-purchase), processes, and catalogs materials from Asia and the Middle East and processes and/or catalogs materials from the Cairo, Jakarta, Islamabad, and New Delhi Overseas Offices.

¹¹ The U.S./Anglo Division acquires (via purchase and non-purchase), processes, and catalogs materials from the United States, Australia, Canada, Ireland, New Zealand, the United Kingdom, and Oceania; catalogs rare materials; accessions and disposes of multiple copies of government documents; and coordinates major gifts to the Library, as well as the acquisition of special collections.

¹² The Germanic and Slavic Division acquires (via purchase and non-purchase), processes, and catalogs materials from Scandinavian, Baltic, and Germanic countries; Russia; and East Central and Southeast Europe.

¹³ The U.S. and Publisher Liaison Division catalogs Copyright and Cataloging In Publication (CIP) materials and materials coming in through the International Standard Serial Number (ISSN) application process; provides CIP and ISSN infrastructure support; and catalogs law materials.

¹⁴ The U.S. General Division catalogs Copyright and CIP materials and provides Dewey Decimal Classification infrastructure support.

To assist in its collection efforts, Library Services reinstated the Collection Development Office (CDO) and officer position in 2012. This office will directly support the Library's goal to acquire and maintain a universal collection of knowledge and the record of America's creativity to meet the needs of Congress, researchers, and the American public.¹⁵ It will ensure that the Library's analog and digital collections reflect the breadth and depth of knowledge published in all media, languages, and regions of the world. CDO staff will work closely with those in acquisitions and recommending units to achieve the Library's collection-building mission.

Once the material to collect is identified, the Library collects the analog and digital material in several ways. It can purchase on the open market, including internationally via the Library's six overseas offices—Kenya, Egypt, India, Indonesia, Pakistan, and Brazil. It also receives material via its copyright responsibilities. Private donations are another source of collections for the Library, such as Bob Hope's joke collection or vintage films that are purchased, preserved, and stored through the generous donation of the Packard Humanities Institute. And in 2010, the Library entered into an agreement with Twitter to preserve tweets. In recent testimony to Congress the Librarian stated, "[the Library]...is trying to assess...the intrinsic value of...this material..." The scope of the usefulness of this data is not known at this time, but it does pinpoint one of the key collection/acquisition challenges—what to collect immediately in anticipation that it may be important years from now.

In 1998, the Librarian of Congress commissioned the Computer Science and Telecommunications Board of the National Academy of Sciences to conduct a study to provide strategic advice concerning the information technology path the Library should traverse over the coming decade. In July 2000, the committee issued the results of its study in a report titled, *LC21: A Digital Strategy for the Library of Congress* (see Appendix D).

¹⁵ The creation of the Collection Development Office had not been finalized as of the date of this report.

Overall, the LC21 study committee was emphatic in its belief that the Library continue to play a vital role in documenting and preserving the history of American creativity and in building a collection with worldwide scope. However, it opined that the Library could not proceed as before and that the Library was not planning far enough ahead to enable it to act strategically and coherently. The report concluded that “[based on the events] in the last half-century, forces such as the explosion of publishing, the rapid expansion of education and higher education, globalization, and ever-growing funding for many kinds of research...no single library—not even the Library of Congress—can today collect and deliver comprehensively, if ever it could, the world’s most important literature and information sources.” The study committee recommended that the Library put in place mechanisms to systematically address the infrastructure required for it to “collect” digital materials.

In 2000, Congress appropriated \$100 million¹⁶ for the establishment of the National Digital Information Infrastructure and Preservation Program (NDIIPP) to develop a national strategy and network of partners to collect, preserve, and make available significant digital content, especially information that is created only in digital form (or, born digital), for current and future generations to build a national repository of digital materials. The Library was directed to develop a strategic plan, in collaboration with other federal and non-federal entities, to identify a national network of libraries and other organizations with responsibilities for collecting digital materials that will provide access to and maintain those materials.

The 2003 approved NDIIPP plan established four goals:

- develop a national preservation network;
- develop a content collection plan that will grow a national collection and preserve important at-risk content;

¹⁶ Beginning in FY 2001, the Congress made \$100,000,000 of no-year money available for NDIIPP. However, \$54,078,000 of unobligated no-year money was subsequently rescinded, leaving NDIIPP with a net enacted budget of \$45,922,000. Additional annual funds in the amount of \$27,674,000 were appropriated for FYs 2008-2013. This provided a combined funding of \$73,596,008 for NDIIPP.

- build a shared technical platform for network preservation; and
- develop recommendations to address copyright issues and create a legal and regulatory environment that encourage incentives to preservation.

NDIIPP established and engaged hundreds of partnerships to focus on the collection and preservation of at-risk digital content. The National Digital Stewardship Alliance is an out growth of the program that will continue to advocate preservation standards and practices. Collaborative efforts include Viewshare.org, a platform that helps users create interactive interfaces to digital collections; BagIt, a file packaging format designed to support disk-based storage and network transfer of arbitrary digital content; Lots of Copies Keep Stuff Safe (LOCKSS), open source digital preservation tools for libraries and publishers to preserve and provide access to digital content; and the DuraCloud pilot, a partnership of major research institutions and libraries that explored cloud technologies and the use of large format files. NDIIPP has harvested more than 29 petabytes¹⁷ of Web site information (equivalent to 29,696 terabytes), and it is expected to grow exponentially.

In-processing—Each workday, the Library receives approximately 15,000 digital and analog items and adds about 11,000 items to its collections. The amount of time required to catalog the material depends on the material received. Cataloging is done in the order of the priority assigned to the material (see Appendix E). At each stage in the process, items of highest priority will be processed before any others. For example, Priority 1 - Titles requested by Members of Congress or their staffs, by agency heads or higher officers of the Executive Branch, by Supreme Court Justices, or by division chiefs or higher officers of the Library of Congress, and Pre- and Post-Publication Cataloging in Publication titles. Material that is not cataloged is “arrearage” and is stored waiting to be processed based on the assigned priority. Some items waiting to be cataloged, preserved, and properly stored date back to the 1980s.

¹⁷ A petabyte is a unit of information equal to one quadrillion bytes, or 1,000,000,000,000,000 bytes.

Once books and serials material are received on-site, reviewed, and selected for the Library's collections, the material is labeled,¹⁸ tagged,¹⁹ and stamped.²⁰ This process includes part of the Library's bibliographic and inventory controls over its material. The Library's special collections divisions are responsible for processing and cataloging non-print materials, such as maps, manuscripts, motion pictures, and prints and photographs.²¹ Generally, these are not labeled, tagged, and stamped, but each Division has its own procedure for processing.

The Library has two types of tracking systems—manual and automated—for its processed collections, depending on when the items were processed. A system is critical for controlling the acquisition, cataloging, circulation, and inventory of Library material. It provides access to researchers and the public to identify, search, and locate Library resources. In the absence of bibliographic controls the collections are at risk of theft, loss, or being forgotten. Materials processed before 1999 were recorded in manual and automated card catalog systems maintained in the various Library Services divisions. The manual card catalog system is still in place but is gradually being automated.

In 1999, the Library implemented the Integrated Library System (ILS) cataloging and circulation modules, the Online Public Access Catalog (OPAC), and the acquisitions and serials check-in modules. Through use of a shared bibliographic database, the ILS integrated all major Library Services functional areas, such as acquisitions, cataloging, serials management, circulation, inventory control, and reference.

¹⁸ Labels are placed on material to provide identifying information to help track items and pieces as they are processed from the point of entry to their final destination within the Library.

¹⁹ Tagging is the process of inserting a magnetic security strip into the material.

²⁰ A Library of Congress property stamp with permanent ink is applied to material.

²¹ If deemed necessary, materials requiring binding and shelf preparation are forwarded to the Preservation Directorate, Binding and Collection Care Division.

The ILS system allows the Library to create and maintain bibliographic records for all types of library materials held in its collections (books, serials, computer files, music (scores and musical sound recordings), non-music sound recordings, cartographic materials, manuscripts, and visual materials (still and moving image). It also allows the Library to know the items held in its collections, where each trackable item resides, and to provide responsive service when requests are received for items from the general collections.

Although there are various means by which the Library acquires material (i.e., purchase, gifts, and copyright), generally there are three main records attached to an item in ILS:

- A bibliographic record contains information describing a book, serial, etc. In many instances, it is established with the creation of a purchase order or through Cataloging in Publication (CIP).²² The intended custodial location is defined when a purchase order is established.
- A holdings record contains information for the bibliographic item (i.e., number of copies owned and where they are held). Holdings records are dependent on the bibliographic records and require the intended custodial location be identified. This record is attached to the purchase order. If multiple copies are held, a new holding record is created for each location (i.e., Ft. Meade or Landover).
- An item record which holds the barcode information is then attached to the holdings record.

Preserving—To maximize its investment in the materials it acquires, the Library has a preservation program that involves several steps over a number of years by many Library staff, including those with expertise in library science, computer science, and materials science. Library Service’s Preservation Directorate divisions—Binding and

²² CIP data is a bibliographic record prepared by the Library of Congress for a book that has not yet been published. The publisher includes the CIP data on the copyright page to facilitate book processing for libraries and book dealers.

Collections Care, Conservation, Preservation Reformatting, and Preservation Research and Testing—are responsible for ensuring long-term access to the intellectual content of the Library’s collections. The Preservation Directorate accomplishes this directly through conserving, binding and repairing, reformatting, testing, and educating staff and users. The current FY 2013 Preservation Directorate budget is \$18.4 million, a \$2.5 million decrease from FY 2012.

The preservation challenges and Library expertise required are extensive. Each format demands its own approach to preservation. For example, much of the mass-produced paper of the 19th century has a high-acid content that deteriorates the paper. By immersing the paper in a liquid suspension of magnesium oxide, the acids can be neutralized without damaging the integrity of the item. A treated or deacidified book (or loose sheet), if kept in cool storage, is predicted to last up to 1,500 years, and the Preservation Directorate plans to deacidify 250,000 books and 1,000,000 loose sheets annually as part of a 35-year, one-generation plan. On the other hand, magnetic tape will inevitably deteriorate in decades, not centuries. The solution is to digitize the sound on the tape at a high enough standard that the human ear cannot distinguish between the sound on the analog original and its digital copy. The digital file is then stored on a server and regularly checked for bit loss,²³ preserving the sound well beyond the shelf-life of the original tape.²⁴

The preservation of born-digital items, such as a Web site, presents an entirely different set of challenges from those involved in safeguarding physical materials. By their nature, Web sites are short-lived and evolving, often with multi-media contents. Since 2000, for example, the Web Archiving Team has been harvesting Web sites related to U.S. Federal elections. Practices the team employed in preserving these Web sites helped in preserving social networking sites such as Twitter, YouTube, and Facebook. Alternative methods, such as digitization, may be used for

²³ Bit loss is commonly defined as the corruption of the smallest possible amount of digital information in a file or data set. Bit loss may occur during transmission, copying, or during data storage.

²⁴ Reference: Library of Congress, *Library Services Journal*, Vol. 3, No. 1.

collection items that are considered more valuable and/or endangered.²⁵ Providing digital surrogates²⁶ for fragile materials greatly reduces handling that could put valuable collection items at risk. The Library's preservation strategy continues to evolve as preservation science progresses. Significant investments in materials, supplies, capital assets, and highly trained specialists comprise the resources used in the Library's preservation efforts. More information on digital preservation can be found in Appendix C.

Storing—The Library's collections were originally housed in the U.S. Capitol. The U.S. Congress later appropriated funds for construction of a new building, now the Thomas Jefferson Building, and relocated there in 1897. Since then, the Library has expanded into two additional buildings in the Capitol complex, the James Madison Memorial Building and the John Adams Building, as well as a number of government-owned and leased off-site facilities in Virginia and Maryland, and at a National Archives and Records Administration (NARA) facility in Illinois.

Ft. Meade Collections Storage Modules Serves As the Primary Climate Controlled Repository

As the collections continued to expand, the Library recognized that additional storage would be necessary. In 1995, Ft. Meade, Maryland was identified as a site on which the Library could develop a series of collections storage modules (modules), as needed, in succeeding years. The Architect of the Capitol's master plan for Library facilities at Ft. Meade provided for a total of thirteen storage modules and other facilities to be constructed over time. The first module was opened in 2002. The plan for the modules is at least eight years behind schedule.

The relocation of some collections from Capitol Hill to Ft. Meade has relieved pressure at the Library's main campus. To date, four modules have been built at Ft. Meade. Modules 1 and 2 are configured as repositories for

²⁵ Reference: Library of Congress, Library Services Journal, Vol. 3, No. 1.

²⁶ A digital surrogate is a digital copy of an item.

books while modules 3 and 4 are allocated to special format materials. Module 1 was filled in 2005, and module 2 is also currently full. Modules 3 and 4 were filled in 2012. All modules have similar environmental controls and security arrangements; however, the interiors have been fitted with different storage container systems to support the various formats. Module 5 was planned to open in 2005 and would have provided 17,000 square feet of storage and held 2.2 million items. Presently, Congress has not funded the construction of module 5. The cost of constructing module 5 was last estimated at about \$17 million. For the project to proceed, only funding for construction remains to be secured. According to the original Master Plan, module 8 would have been constructed in FY 2011.

The National Audio Visual Conservation Center Is the State of the Art Facility for Films, Television Broadcasts, and Sound Recordings

The Library's National Audio Visual Conservation Center (Packard Campus) in Culpeper, VA opened in 2007 and was designed for the acquisition, cataloging, storage, and preservation of the nation's collection of moving images and recorded sounds. The state-of-the-art facility houses the largest and most comprehensive collection of American and foreign-produced films, television broadcasts, and sound recordings. The Packard Campus offers extensive capabilities and capacities for the preservation and reformatting of audiovisual media and is a long-term digital storage archive. Different collection formats demand different preservation approaches. In addition to preserving the collections of the Library, the Packard Campus was designed to provide preservation services for other archives and libraries in both the public and private sector.

Providing Access—The Library's mission is not fulfilled unless its acquired collections are made accessible to the public. The Library's collections can be accessed:

- onsite—where a visitor can physically access material through the reading rooms and research centers, and

- online—to view a growing collection of digitized photographs, manuscripts, maps, sound recordings, motion pictures, and books, as well as "born digital" materials such as Web sites.

The Library is also the home of the National Library Service for the Blind and Physically Handicapped that provides access to braille and audio materials circulated to eligible borrowers through its partnerships with regional and sub-regional libraries.

Additionally, the Library makes use of a multitude of outreach programs, collaborations, and partnerships to bring its collections to the American people. A few examples of this are the Library's National Book Festival, an annual gathering of authors and activities that promote literacy, cultural preservation, and preserving digital culture; the "Gateway to Knowledge," a traveling exhibit designed to bring the riches of the Library of Congress to the heartland of America; and the Summer Teacher Institute workshop where educators work with Library specialists to learn to access and explore the Library's digital collections and resources in their classrooms.

Project One Initiative Was Created to Advance the Library's Web Presence

In 2010, the Library recognized that its Web presence²⁷ was essential to advancing the Library's mission and meeting all of its strategic goals. The Librarian appointed the Deputy Librarian to chair the Web Governance Board to create a Web strategy that identified three core areas of the Library's Web presence: the National Library, information by and for Congress, and Copyright. The goal of the Web strategy, known as Project One, is to "build and manage a world-class, user-centered Web presence that truly reflects the breadth, quality, expertise, and authority of the Library." The strategic goal of Project One is to create a cohesive Web presence that ensures that the Library is one

²⁷ The Library's "Web presence" includes the Library's Web sites, online services, and software applications (e.g., interactive kiosks, iPhone Apps), as well as third-party channels used for the delivery of Library content and services (e.g., Flickr, iTunes, Facebook, Twitter, YouTube).

unifying institution, with no difference in accessibility to its resources regardless of the manner in which the public gains access (in-person or Web).

The Web strategy aims to create a user-friendly environment for researchers to search and browse, find what they need, and understand what they have found. Currently, a search for a Civil War event may result in an incomplete representation of all of the items in the collections. The Library's new page design will offer links to connect to related areas of interest. For example, a map, a manuscript, and a sound recording associated with the Civil War will be included in a set and accessible for the user.

In 2012, the Library launched the re-designed and updated LOC.gov page to improve the user experience. The page also includes links to legislative and copyright information, the Library's social media sites, and services provided by the Library. Congressional information is available thorough Congress.gov (beta Web site) which will eventually replace THOMAS.gov, launched in 1995.

►► OBJECTIVES, SCOPE, AND METHODOLOGY

The objective of this evaluation was to assess the framework of policies and procedures supporting the development of Library collections and the logistical impact of those policies and procedures. The evaluation did not assess whether the items collected were the appropriate materials to collect.

At the onset of this evaluation, the Associate Librarian for Library Services requested the OIG's assistance in identifying (1) criteria for selecting viable trusted partner(s) and (2) best preservation standards for the Library's digital materials. To satisfy Library Services' first request we consulted the *Trustworthy Repositories Audit & Certification (TRAC): Criteria & Checklist*, which is used for assessing an organization's viability, financial sustainability, and procedural accountability (Appendix A). To address Library Services' second request, we reviewed and compiled best practices/standards used by NDIIPP, the National Archives and Records Administration (Appendix C), and the British and Canadian Libraries (Appendix B). We benchmarked the British and Canadian Libraries to determine how other national libraries accomplished their missions while managing reductions in staffing and funding (Appendix B).

To understand the Library's collection development policies and procedures we reviewed laws, regulations, directives, and policies that apply to the acquisition of materials for the Library's collections. We reviewed the LOC [Library of Congress] Facilities Plan Collections Storage 2010 and the Library Services Collection Management Plan, 2011-2016 to identify the Library's long-term and short-term collection storage plans. We also conducted interviews with Collections & Services officials to gather information on the collections storage problems.

To gather information on the General & Special Collections (G&S) Divisions' collection development activities, we conducted interviews with division chiefs and reference librarians, as well as a walk-through of the G&S divisions to

observe the management of the collections materials.²⁸ We did not assess the sufficiency of scope or clarity of the Library's canons of selection.

We conducted interviews with Acquisition and Bibliographic Access (ABA) division chiefs responsible for acquiring and cataloging materials for the collections to understand how materials awaiting processing are tracked. To identify the number of unprocessed print and pictorial materials, maps, and rare books throughout the Library, we reviewed the ABA and Collections & Services division's annual reports. We also conducted interviews and walkthroughs of the ABA and Collection & Services divisions to identify unprocessed materials.²⁹ In FY 2005, Library Services decided to omit unprocessed print material, maps, pictorial materials, and rare books from its arrearage figures. Those unprocessed materials remaining became labeled work on hand. Therefore, in 2005 the Library began reporting only the number of unprocessed manuscripts, music, moving images, sound recordings, and machine-readable items as unprocessed arrearage. The exclusion of unprocessed print material, maps, pictorial material, and rare books from the Library's reported unprocessed arrearages understates the amount of unprocessed material in arrears. For its reported historical arrearage, the Library made estimates at the time the material was received based on observation and other methods that we did not validate. The Library does not conduct periodic and comprehensive inventories of its collections or material in arrearage; nevertheless, for this evaluation we relied upon the Library's reported amounts as a basis to assess its logistical

²⁸ The Collections & Services Directorate has two sub-directorates—General Collections & Services and Special Collections & Services. In addition to interviewing the Collections & Services Director, the OIG conducted reviews and analysis that included interviews and/or inspections in six of General Collections & Services' eight divisions. Additionally, Special Collections & Services has eight divisions and the OIG conducted detailed work and analyses that included interviews and/or inspections in six divisions. The OIG conducted its inspections at the Library's three Capitol Hill buildings and the Landover Annex; the OIG did not inspect all buildings.

²⁹ The ABA Directorate has nine divisions. In addition to obtaining data from its Director and conducting analyses, the OIG performed work that included interviews and/or inspections in seven ABA divisions.

impact.³⁰ Because we relied upon the Library's arrearage figures and other reports and did not validate their reported figures, we notified the Library that we changed this effort from an audit to an evaluation and used the Council of the Inspectors General on Integrity and Efficiency, *Quality Standards for Inspection and Evaluation*.

We also reviewed prior audits in this area to determine the progress on our findings and recommendations. In 2006, the OIG released an audit report titled, *The Library's Collections Acquisitions Strategy: Effective, but Some Improvements Are Needed*, report no. 2006-PA-104. The audit objective was to determine whether the Library was efficiently and effectively acquiring materials that met researchers' needs and considering the logistical issues of its acquisitions. The audit found that with the exception of large gifts, the Library did not generally consider cataloging, preservation, or storage issues in its acquisition decisions. As a result, the Library was unable to keep up with the inflow of materials, had an overflow of materials in the collections stacks, and required a new storage module in order to properly preserve and secure its collections.

The audit recommended that the Library explore strategies to reduce the quantity of materials it collected. The report explored three strategies for easing acquisitions growth: (1) reducing the number of items acquired by using other libraries and making some adjustments in the way items were acquired; (2) focusing the collections so as to concentrate on

³⁰ On page *iv* of this report we state, "Security controls...help protect the collections from theft. The lack of these controls increases the risk of material being stolen, lost, and forgotten. *Due to the absence of these controls over unprocessed materials, we could not design tests to determine the extent to which the risks exist* (emphasis added)." Another significant factor preventing the design of audit tests is the Library's use of estimates for the arrearage. The absence of an accurate accounting of the arrearage universe prohibits developing valid samples for testing purposes. Additionally, valid audit tests for measuring risk require an understanding of how items in a universe are distributed. The Library's estimates of arrearage did not provide the necessary information. Therefore, we had to rely on other audit methodologies including inspection, inquiry, and observations. Our comment in this report about damage to collections material based on storage conditions are based on observation. We were unable to determine the extent of damage. Such analysis was beyond the scope of this review and would require an expert consultant. We may pursue such analysis in a subsequent review.

more useful items; and (3) improving the human element of acquisitions. Library Services generally agreed with the audit findings. However, it disagreed with the overall audit premise that “the significant logistical and financial resources needed to acquire, process, store, preserve, and protect collection items dictate that the Library reevaluate its acquisitions policies.”

Our fieldwork was conducted from August 2012 through March 2013. Our evaluation covered collections materials acquired as of September 30, 2012. Our evaluation relied on collections data reported by the Library; however, we did not validate the collections data.

We conducted this evaluation in accordance with the Council of the Inspectors General on Integrity and Efficiency, *Quality Standards for Inspection and Evaluation* and Library of Congress Regulation 211-6, *Functions, Authority, and Responsibility of the Inspector General*. Those standards require that we plan and perform the evaluation to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our objectives. We believe that the evidence obtained provides a reasonable basis for our conclusions.

►► FINDINGS

I. The Library has been Meeting Its Top Priority of Continually Obtaining a Universal Collection of Both Analog and Digital Materials

The Library's analog and digital collection has grown almost 22 percent since 2003. Library management commits itself to acquiring collection materials believing that any pause in its acquisition efforts jeopardizes its commitments to its patrons and stakeholders. Specifically, the Librarian has stated that obtaining a universal collection is the Library's top priority. In its pursuit of acquiring a universal body of knowledge, the Library is guided by the collections policy statements which encompass a wide breadth of materials—from foreign to digital collection items.

At Its Current Growth Rate, the Library's Analog and Digital Collections Will Grow to an Estimated 178 Million Items by FY 2020

The Collections Policy Statements, based on the three "Canons of Selection" discussed in the background section of this report, provide the framework for the Library's collection building activities. The Collections Policy Statements are broad and cover a wide range of disciplines and fields of study. The Library updated all collection policy statements in 2008 and reinstated the Collection Development Officer position in 2012 to help develop, coordinate, and execute Library-wide policies and programs related to the development and selection of Library materials.

The Library adds approximately 11,000 analog and digital items to its collections daily. In FY 2012 alone, the Library acquired more than 3.6 million items through purchase, gift, copyright deposit, and exchange and transfer from other government agencies, see Figure 3.

Acquisition of Library Materials by Source						
Source	Items 2007	Items 2008	Items 2009	Items 2010	Items 2011	Items 2012
Gifts	1,619,338	1,511,181	847,945	1,005,407	1,556,198	1,918,974
Appropriated -GENPAC/LAW	712,129	1,171,227	723,390	888,496	1,883,489	711,363
Copyright Deposits	1,077,152	526,508	739,364	814,243	706,583	636,430
Exchange	151,479	145,981	125,228	126,998	122,954	112,104
Cataloging in Publication/PCN	80,373	87,479	83,551	113,877	101,942	104,203
Government Transfers	58,174	77,511	84,146	322,511	72,982	97,237
Appropriated Other	83,574	17,848	27,686	27,323	11,485	15,337
Gift and Trust Funds	13,494	7,326	8,382	164,202	7,987	8,086
Total	3,795,713	3,545,061	2,639,692	3,463,057	4,463,620	3,603,734

Figure 3: Collections material acquired annually by source.
 Source: Library Services Acquisitions and Bibliographic Access Statistical Acquisitions by Source report.

The expansive collections include, but are not limited to, manuscripts, photographs, maps, sheet music, movies, and talking books for the blind and physically handicapped, see Figure 4.

Estimated Analog and Digital Items in the Collections as of Fiscal Year 2012		
Collections Materials	Total	%
Manuscripts ¹	68,118,899	44%
Cataloged Books (Print Collections)	23,276,091	15%
Microforms ²	16,746,497	11%
Photographs (Negatives, Prints, & Slides)	13,640,325	9%
Books - Large Type & Raised Characters	12,638,773	8%
Pieces of Sheet Music	6,589,199	4%
Cartographic Material	5,478,123	4%
Audio Materials	3,420,599	2%
Machine-Readable Material	1,966,354	1%
Moving Images	1,354,126	1%
Other (Broad­sides, Photocopies, Non-Pictorial Material, etc.)	1,349,451	1%
Prints & Drawings	605,547	0.39%
Posters	104,270	0.07%
Talking Books	69,048	0.04%
Total (Items)	155,357,302	100%

Figure 4: Breakdown of the Library’s collections materials, as of September 2012.

Source: Library of Congress Fiscal Year 2012 Annual Report.

¹ Figure includes manuscripts located in several divisions, principally the Manuscript Division, which estimated that as of September 2012, it held 62,204,303 processed collection items in more than 11 thousand separate collections. In addition to these 62 million items, the Manuscript Division also reported in September 2012 holding an arrearage of 11,573,025 items.

² Microform materials include several different formats of materials in a reduced size to preserve and compact large quantities of information into less storage space.

Since 2003, the Library’s collection of analog and digital material has grown by almost 22 percent from 127 million to more than 155 million. Continuing at this rate of acquisition, the Library’s collections will grow to an estimated 178 million items by FY 2020, see Figure 5.

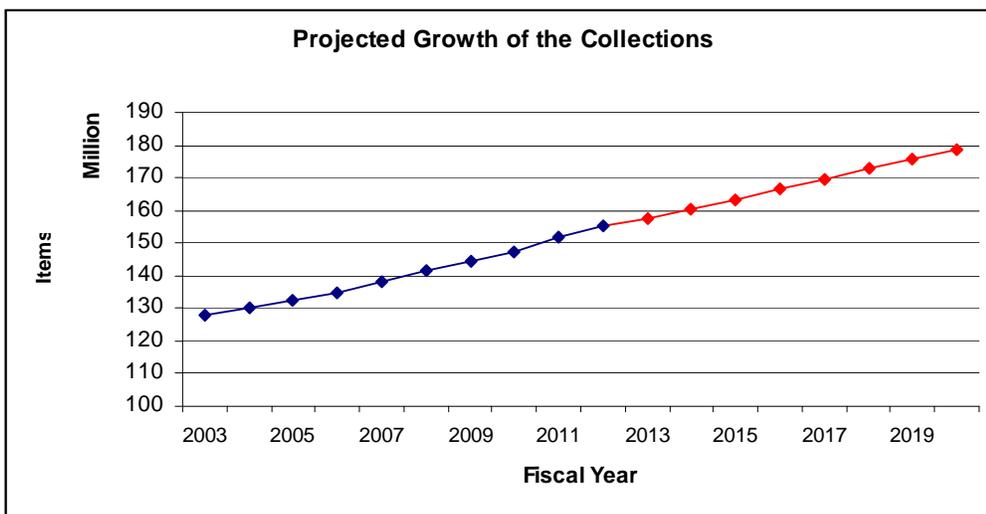


Figure 5: Projected growth of the collections.

Source: OIG projection of the growth of the collections.

Copyright deposits are a major contributor to the Library’s analog collections, with a small amount in digital form. This includes acquiring reprints and new editions of materials that are already part of the Library’s collections. The Library continues to add these reprints and new editions, such as *The Adventures of Tom Sawyer*, because enhancements in the new edition may provide the reader greater insight into the author’s work. Figure 6 illustrates the depth of the Library’s literary works.

Example of Number of Literary Works Acquired by the Library			
Title	Author	Editions/ Reprints	Total Copies
Moby-Dick, or The Whale	Herman Melville	65	84
Little Women	Louisa May Alcott	62	77
Great Gatsby	F. Scott Fitzgerald	15	18
Grapes of Wrath	John Steinbeck	25	34
Adventures of Huckleberry Finn	Mark Twain	90	123
Uncle Tom’s Cabin	Harriet Beecher Stowe	45	61
Adventures of Tom Sawyer	Mark Twain	100	128

Figure 6: New and reprint editions in Library collections for the listed popular literary works, as of 2012.
Source: Library of Congress online catalog.

Foreign Collections Contain Titles from a Number of Countries and Include Newspapers Printed in Almost All Modern Languages

The Library acquires a variety of publications and subjects from all regions, countries, and parts of the world. Half of the Library’s book collections are in languages other than English, representing some 470 languages. Foreign materials are acquired in print (reference works, monographs,³¹ and serials), audio, moving images, microforms, and electronic (databases, Web sites, and electronic serials) formats; however, the majority of foreign acquisitions are serial publications.³² The Library states that it maintains one of the largest collections of newspapers in the world, comprised of important titles from most independent countries and many dependent states that have existed during the past three centuries. The Library also acquires government documents from many parts of the

³¹ A monograph is defined as a publication which is complete in one part or intended to be completed in a finite number of separate parts.

³² Serial publications include newspapers, periodicals, journals, newsletters, etc.

world in numerous languages and in all available and appropriate formats. Over the last three years, the Library has annually acquired an average of 684,824 foreign materials in both analog and digital forms. Figure 7 provides a detailed breakdown.

Foreign Materials Acquired by Format				
	2010	2011	2012	Average
Purchased Books	161,064	156,104	152,936	156,701
Purchased Serials	388,732	369,147	337,464	365,114
Purchased Other Format	38,218	32,346	18,109	29,558
Total Purchased Materials	588,014	557,597	508,509	551,373
Nonpurchase Books	38,671	37,781	36,878	37,777
Nonpurchase Serials	90,880	85,789	75,380	84,016
Nonpurchase Other format	7,317	13,003	14,653	11,658
Total Nonpurchase Materials	136,868	136,573	126,911	133,451
Total Materials Acquired	724,882	694,170	635,420	684,824

Figure 7: Foreign materials acquired in FYs 2010, 2011, and 2012.
 Source: Acquisitions and Bibliographic Access statistical reports on acquisitions by country of publication.

The Library acquires its foreign collections through purchase, international exchange agreements, and gifts. The majority of its foreign works are purchased using Books General Purpose, Acquisitions, and Collections funds (GENPAC).³³ Since 1962 the Library has maintained offices abroad to acquire, catalog, preserve, and distribute collections materials from countries where such items are not readily available through conventional acquisition methods. The coordinated global acquisition efforts are administered by six international buying offices located in India, Egypt, Brazil, Indonesia, Kenya, and Pakistan. The ABA administers the Library’s overseas offices. The Librarian stated in congressional testimony that it has the largest collection of Russian, Chinese, Japanese, and Korean materials outside of those countries. In FY 2012, the Library purchased 508,509 foreign works costing \$11,706,159, see Figure 8.

³³ The GENPAC, administered by ABA, funds the acquisition of serial publications, online content, special format, and domestic and foreign materials of legislative and research value. With the Law Library, ABA manages separate funding for the purchase of law collection materials.

Total Materials Purchased									
Type of Materials	2010			2011			2012		
	Items Acquired	%	Dollar Value	Items Acquired	%	Dollar Value	Items Acquired	%	Dollar Value
Foreign	588,014	54	\$12,395,929	557,597	29	\$10,848,826	508,509	69	\$11,706,159
US	494,019	46	\$9,297,622	1,347,838	71	\$15,754,434	228,326	31	\$9,348,548
Total	1,082,033	100	\$21,693,550	1,905,435	100	\$26,603,260	736,835	100	\$21,054,707

Figure 8: Materials purchased in FYs 2010, 2011, and 2012.
 Source: Acquisitions and Bibliographic Access statistical reports on acquisitions by country of publication.

Digital Collections Have Been Expanding Rapidly

Technological breakthroughs in the mid to late 20th century exponentially widened the Library’s acquisition challenges. In response to the digital information explosion, the Library offered digitized versions of its collection materials online beginning in 1994, including some of its most rare collections that were unavailable anywhere else. The National Digital Library and NDIIPP collections are made up of digitized historical documents, photographs, manuscripts, maps, sound recordings, motion pictures, and books, as well as "born digital" materials such as Web sites. As of February 2013, 4,930 terabytes³⁴ of digital material was available for researchers onsite and/or offsite via the Web, with two exceptions: the Twitter archive (about 65 terabytes as of December 2012), and eDeposit content (approximately 100 gigabytes). In 2005, the Library’s digital collections consisted of 59.9 terabytes. Figure 9 offers a comparison breakdown of the Library’s 2005 and 2012 digital holdings.

³⁴ A terabyte is a multiple of the byte unit digital information. One terabyte is 1,000,000,000,000 bytes. In practical terms one terabyte is equivalent to approximately 143 million pages of stored Microsoft Word documents, 250,000 music files, or roughly 200,000 photographs.

Library's Digital Collections		
Content Type	2005 Cumulative Terabytes	2012 Cumulative Terabytes
Cartographic	3.5	26.0
Mixed (multi-format)	2.5	73.9
Moving Image and Audio (includes Packard Campus)	5.8	3423.3
Still Image/Photographic	11.8	47.8
Textual	10.3	985.0
Web Archives	26.0	374.0
Total	59.9	4930.0

Figure 9: Breakdown of the Library's digital collections.
 Source: Office of Strategic Initiatives.

The National Digital Library, a digital collection of American history and creativity, began as a pilot program in the early 1990s. The pilot set out to digitize some of the Library's historical documents, moving images, sound recordings, and print and photographic media. The collection provides free and open access through the Internet to written and spoken words, sound recordings, still and moving images, prints, maps, and sheet music that document the American experience.

NDIIPP set out to collect and preserve at-risk digital content of cultural and historical importance. NDIIPP seeks to preserve a wide range of born-digital records of public and commercial content and is working to collaboratively establish standards for digital preservation. NDIIPP collaborates with state, archival, and private-sector businesses to achieve a diverse network of partners and develop standards. In 2012, NDIIPP's network of 200 partners collected and preserved 150 billion digital items comprising 29 petabytes (29,696 terabytes) of data. Recent NDIIPP actions include preserving state government information in 35 states, establishing preservation standards for digitized historical content through collaboration with 20 federal agencies, and archiving more than 4,300 Web sites of significance (e.g., Civil War's sesquicentennial, the 2012 Olympics, the U.S. national elections). In a future report, we intend to discuss further the status of the Library's digital collections. See Appendix C for more information on digital preservation and the industry standards that are currently being employed.

II. The Library is Not Able to Timely Process the Inflow of Materials It Acquires—Current Unprocessed Arrearage is Approximately 28 Million and Has Grown Almost 50 Percent Since 2000

The Library's acquisition of material surpasses its ability to efficiently process the items it collects. Consequently, the Library has an arrearage estimated at more than 28 million unprocessed materials³⁵ consisting of manuscripts, maps, print and pictorial materials, and three-dimensional objects.³⁶ These unprocessed materials lack adequate inventory controls to safeguard these assets, making the collections vulnerable to theft or loss, and inaccessible to the general public.

We attempted to gather information on the amount of unprocessed print and pictures, maps, and rare books throughout the Library. We reviewed the ABA and G&S division's annual reports and conducted interviews and walkthroughs to identify unprocessed materials. In FY 2000, the Library reported approximately 19.2 million items of unprocessed materials. The OIG's compilation of all unprocessed materials as of FY 2012 is estimated at 28.3 million items, an increase of about 50 percent since FY 2000. Figure 10 shows our estimate of unprocessed materials on hand in the G&S and ABA divisions as of September 2012.

³⁵ The Library considers an item unprocessed if it has not completed initial processing (including physical preparation, bibliographic access, inventory control, and/or other treatment) required to make it routinely available for use and has not been placed in the location from which it will be served.

³⁶ Three-dimensional objects are materials outside of the formats actively collected by the Library, such as paintings, sculptures, furniture or furnishings, costumes, medals, coins, stamps, badges, emblems, decorations, personal effects, or any other objects or materials, generally out of scope for the Library's collections.

Estimate of Unprocessed Materials in the Library As of September 2012	
General & Special Collections Division	
Prints & Photographs	4,000,000
Manuscript	11,600,000
Geography & Maps	5,100,000
Music	6,100,000
Motion Picture, Broadcasting & Recorded Sound	1,300,000
African & Middle Eastern	Unknown number of periodicals
Asian	Unknown number of materials
European	Known 1,500 monographs
Rare Books	Approximately 300 volumes
Acquisitions & Bibliographic Access Division	
US/Anglo Division	14,000
Germanic & Slavic	23,000
ALawe	16,000
Asian & Middle Eastern	78,000
USGEN & USPL	65,000
Estimate of Unprocessed Materials³⁷	28,297,800

Figure 10: Amount of unprocessed materials in Library Services ABA and G&S divisions.
Source: OIG analysis of unprocessed materials throughout Library Services.³⁸

Manuscript materials make up more than half of the unprocessed materials in the arrearage. The number of unprocessed materials obtained from the divisions is an estimate. For example, we were not able to obtain an exact number of unprocessed periodicals waiting to be microfilmed in the African & Middle Eastern Division; it estimates it had a one to two year backlog of unprocessed periodicals. Figure 11 shows some of the periodicals from the Middle East waiting to be microfilmed.

³⁷ The number of unprocessed materials obtained from divisions is an estimate. “Unknown number” refers to instances where the OIG identified unprocessed materials; however, we could not obtain a reasonable estimate of how much unprocessed material existed. See Objectives, Scope, and Methodology section for limitations in determining the number of unprocessed materials.

³⁸ Currently, the Library reports a lower figure in its estimate of unprocessed materials because it does not include items from Prints & Photographs, Geography & Maps, and those on hand in the Acquisitions & Bibliographic Access Division. Using the Library’s unprocessed arrearage figures, there is a growth rate of almost 52 percent from 2000 to 2012; using that growth rate, we estimate that by 2022 the arrearage would be more than 23.4 million items.

Some Unprocessed Items in the Manuscript Division Date Back Two Decades

The largest amount of the unprocessed arrearage is occurring in the Manuscript Division, which as of September 2012 reported an estimated arrearage of 11,573,025. Library-wide, more than 15 million manuscript materials have backed up resulting in the number of unprocessed manuscripts in arrears growing 71.5 percent since 2000. Figure 12 shows the growth of the unprocessed manuscript materials in arrears, going from 8.9 million in 2000 to more than 15 million in 2012, of which the Manuscript Division portion grew from approximately 5.8 million to 11.6 million. The majority of the Library’s unprocessed manuscripts are stored at the Landover Annex Center (Landover) and date from as early as 1989. The Manuscript Division stores at Landover more than 13,000 unprocessed containers of assorted dimensions and capacities from approximately 160 unprocessed manuscript collections. The containers are stored on industrial shelving, pallets placed in pallet racks, and transport tubs.³⁹



Figure 11: Periodicals from the Middle East waiting to be processed (microfilmed) in the African & Middle Eastern Division.
Source: OIG.

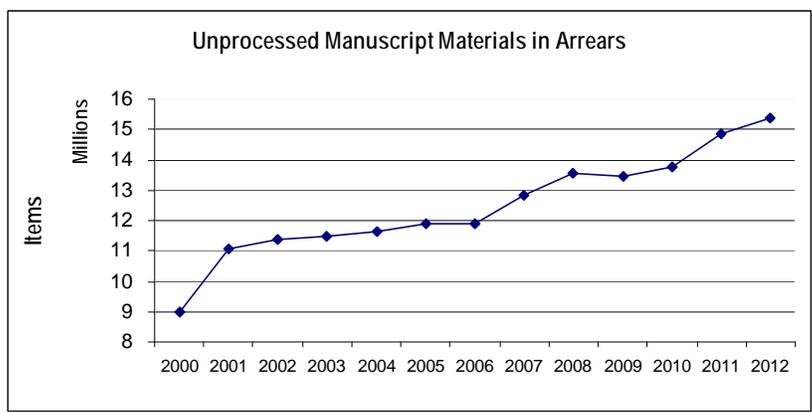


Figure 12: Growth of the unprocessed manuscript materials in arrears.
Source: Library of Congress Annual Reports 2000 through 2012.

³⁹ The Manuscript Division creates a preliminary catalog record in the ILS cataloging module for new collections that are received and accessioned but not processed. The preliminary record includes a Library of Congress Control Number (LCCN), title, main entry, restriction if applicable, and the accession number and shelf location of the material. The preliminary record does not include the estimated number of items in the collection. Materials might stay at this stage (unprocessed) for a few months or years.

According to the Library's Collection Policy Statement, the Library acquires manuscripts for study of the history, law, and civilization of the United States. For example, the Library acquires items such as Supreme Court Justice William J. Brennan's and Dr. Henry Kissinger's papers and those of aviation pioneers Wilbur and Orville Wright. In addition, the Library acquires personal papers of nationally eminent Americans whose activities have significantly influenced American culture or the national policy of the United States government. According to the Library, these papers constitute important source material for the study of the period, throwing light on major movements, developments, controversies, or events. The Library defines personal papers to broadly include correspondence, diaries, notebooks, research notes, logs, scrapbooks, memoranda, journals, and production material for publications, photographs, biographical data, originals, carbons, letterpress copies, microfilm, computer diskettes, and various ephemera. The Library also acquires the papers of individuals who are not nationally eminent, if these papers constitute source material of substantial importance.

The growing unprocessed arrearage also creates a problem for the Library's Integrated Support Services' Facility Services group in planning for long-term space needs. Unprocessed collection materials require interim storage until the items can be processed and moved to permanent storage. In 2012, Facility Services requested that Library Services maintain the level of unprocessed materials so that it would not exceed the planned interim storage space capacity. Library Services agreed to match the outflow rate of materials to the arrival rate as best as possible.⁴⁰ However, in FY 2012 the Library reported that its overall unprocessed arrearage grew 2.6 percent, and the manuscript arrearage alone grew 3.6 percent. We discuss space shortages in the next finding section.

The Significant Number of Arrearage Items Causes Negative Effects—Lack of Access and Accountability

There are significant risks and effects associated with having such a large amount of arrearage—lack of access and accountability.

⁴⁰ The Architect of the Capitol contracted with Wisnewski Blair & Associates to conduct a feasibility study of the Library's storage needs. To assess the Library's current and future storage requirements, certain assumptions were made about the growth in interim collections storage. Library Services assumed no net growth in interim collections storage space.

Access—It is increasingly difficult for the Library to meet one of its critical mission components: to make its collections accessible. Although the Library has some knowledge of the bulk items it receives, if it is not cataloged, then it is not readily searchable or available by patrons without a labor-intensive search. For example, of Geography and Maps⁴¹ 5.5 million cartographic materials, approximately 344,000 items are cataloged and easily searchable in ILS. According to Geography and Maps, cartographic materials acquired before 1970 are not cataloged, except for high value vault items. Geography and Maps primarily catalogs items on demand. A Library constituent cannot search the Library's online catalog system to identify all cartographic materials available. The Library constituent must submit a research request directly to Geography and Maps staff indicating the geographic location of interest. The approximately 5.1 million unprocessed maps are generally filed by the geographic area shown on the map, date produced, and name of publisher/author. Staff subsequently must search map drawers to determine whether they have any cartographic material to satisfy the constituent's request. The Music Division has a similar process in searching through its 2.8 million uncataloged items.

Accountability—Significant backlogs between receiving and cataloging, specifically arrearage and in-process materials, create a vulnerable environment that places collections materials at risk for loss, theft, and misuse. The Library mitigates risks to its collections through bibliographic and inventory controls. Cataloging is the process whereby the Library establishes bibliographic control and an inventory of its material. The lack of these controls decreases the Library's ability to account for and safeguard valuable collections materials. In addition, without these controls, the Library cannot accurately account for what it owns nor can it identify where materials are located. The lack of accountability could result in the Library purchasing duplicate materials.

⁴¹ Geography and Maps maintains a comprehensive cartographic collection of approximately 5.5 million maps, atlases, globes, and other cartographic materials. Cartographic material is collected on a worldwide basis without regard to the time period, language, geographic area, or format. In fiscal year 2012, G&M purchased 7,491 cartographic items at a cost of \$523,030 (approximately \$95,000 of \$523,030 spent were donations). The number of items G&M acquires each year can vary considerably due primarily to government deposits, transfers, purchases, and donations.

For example, the Asian Division could not account for collections material (Balinese lontars)⁴² purchased in 2001. We also discovered newly purchased high-priced lontars unopened/unprocessed for about a year in the Asian Division. These materials are not typical Library materials, which makes them susceptible to theft. The Asian Division's explanation for the recently acquired lontars not being processed was because the division lacked subject matter experts.

During our review, we found that the Asian Division purchased a collection of Balinese lontars between 2001 and 2012, see Figure 13. We sought to determine whether the Asian Division could account for valuable acquisitions and their whereabouts—we attempted to locate 100 Balinese lontars acquired in 2001 and 28 Balinese lontars acquired in 2011.

We were unable to locate 80 of the 100 Balinese lontars acquired in 2001⁴³ and we could not locate 3 of the 28 Balinese lontars acquired in 2011. We also sought to determine whether the Balinese lontars purchased in 2001 were cataloged. We selected 10 lontars from the 2001 purchase order to determine whether there was a bibliographic record in the Library's ILS cataloging module for the lontars. We could not find any bibliographic record for the lontars in the ILS cataloging module.

We were not able to determine a clear cause for the arrearage. It could be a result of the reductions in staff over the years combined with a continual increase in material collected. It could also be a result of a lack of clarity around its canons of collection, thereby enabling each division to collect items that may not be significant or only of interest to a select few, such as the lontars discussed above. Further, we were not able to locate a system-wide strategy that marries its collections efforts to available staff, proper storage capacity, budget, or other factors or a combination of these. The Library recently reinstated the Collections Development Office in order to plan, develop, coordinate, and execute Library-wide policies and



Figure 13: Balinese lontar acquired through the Jakarta Overseas Office.
Source: OIG.

⁴² Lontars are palm leaves with writing on them.

⁴³ Specifically, 20 lontars were identified by consistencies in the box number, title/name of lontar, and number of sheets/pieces. Two additional items matched with the box number and name; however, the number of sheets/pieces did not match what was reported on the purchase order.

programs related to the development and selection of library materials, which encompass holdings in all formats and in all languages. The Collections Development Officer will be responsible for building the collections and maintaining the strength and comprehensiveness of the collections. This office may help address any future arrearage as well as development of a system-wide strategy.

Nevertheless, the Library may also want to consider establishing a national board to assist the Library in making decisions on building its collections and providing the necessary access, similar to that directed by Congress for audiovisual in the National Recording Preservation Act of 2000. In its National Recording Preservation Plan, the Library found that the challenges in audiovisual preservation and access require a well-organized national effort on the part of all stakeholders to achieve success and recommended that the national board consult in the development of a coordinated national acquisition and collections policy for sound recordings. It also tasked the board with promoting partnerships between public institutions, nonprofit organizations, the industry, the collecting community, and companies to accomplish mutual purposes and goals.

The cause for the arrearage could also be the lack of a sufficient number of archivists, preservationists, and catalogers. A significant portion of the Library's arrearage resides in special format divisions, which have sustained a loss of trained archivists and other staff with the specialized expertise necessary for organizing, describing, and cataloging rare and unique materials. The Library Services Preservation Directorate divisions—Binding and Collections Care, Conservation, Preservation Reformatting, and Preservation Research and Testing—are responsible for ensuring long-term, uninterrupted access to the intellectual content of the Library's collections. The preservation of materials requires significant investments in materials, supplies, capital assets, and highly trained specialists. Currently, the Preservation Directorate has an established intern program to help supplement its existing resources, and the Library may want to consider expanding this program to assist in reducing the existing arrearage.

III. Growing Magnitude of the Collections Puts an Overwhelming Demand on Library Storage Space That Comes at a Cost

In addition to processing and accessibility challenges, the magnitude of the Library’s collections is stressing its ability to safely store items it acquires. The Library has exceeded its storage capacity. The long-term strategy to accommodate the current materials and future collections by building more storage modules is not viable at the current time due to lack of funding. Because the Library lacks proper storage, some of America’s Heritage Assets⁴⁴ are damaged and deteriorating at a premature rate. The Library has worked to optimize storage conditions and realize efficiencies in storage configurations by installing modern, high-density shelving where possible. In addition, the Library has begun transferring materials to suitable storage areas in leased facilities. Despite these efforts, the Library’s storage conditions are not currently sustainable.

The Library has more than 24 million catalogued books and bound periodicals (books)⁴⁵ in its collections and is constantly faced with the challenge of properly storing these materials. The Library’s storage facilities have a normal functioning storage capacity of approximately 21 million books, which leaves a current gap of approximately 3 million. Figure 14 shows the total number of books and bound periodicals added to the Library’s collections annually from FYs 2000 through 2012. The average rate of growth for books and bound periodicals was 2.16 percent per year.

Total Books & Bound Periodicals Rate of Growth		
Fiscal Year	Total Books & Bound Periodicals	% of Change
2000	18,899,252	-
2001	19,251,094	1.86%
2002	19,653,492	2.09%
2003	20,064,873	2.09%
2004	20,452,753	1.93%
2005	20,851,111	1.95%
2006	21,363,012	2.46%
2007	21,807,188	2.08%
2008	22,242,339	2.00%
2009	22,881,758	2.87%
2010	23,375,177	2.16%
2011	23,866,960	2.10%
2012	24,417,217	2.31%

Figure 14: Total of books and bound periodicals in the Library's collections annually since 2000.
Source: Library of Congress Annual Reports FY 2000 through 2012.

Assuming that the average current growth rate of books and bound periodicals continues and no future modules are built,

⁴⁴ The Library classifies its collections with historical, cultural, educational, artistic, or national significance as Heritage Assets.

⁴⁵ The Annual Report of the Library Congress, for the fiscal year ending September 30, 2012, reported 23,276,091 cataloged books and 1,141,126 minimal-level cataloging print materials (monographs and serials) in the collections.

the gap increases to over 7 million by 2022. Figure 15 shows the growing gap between the projected growth of the acquisition of books and the Library’s current storage capacity.

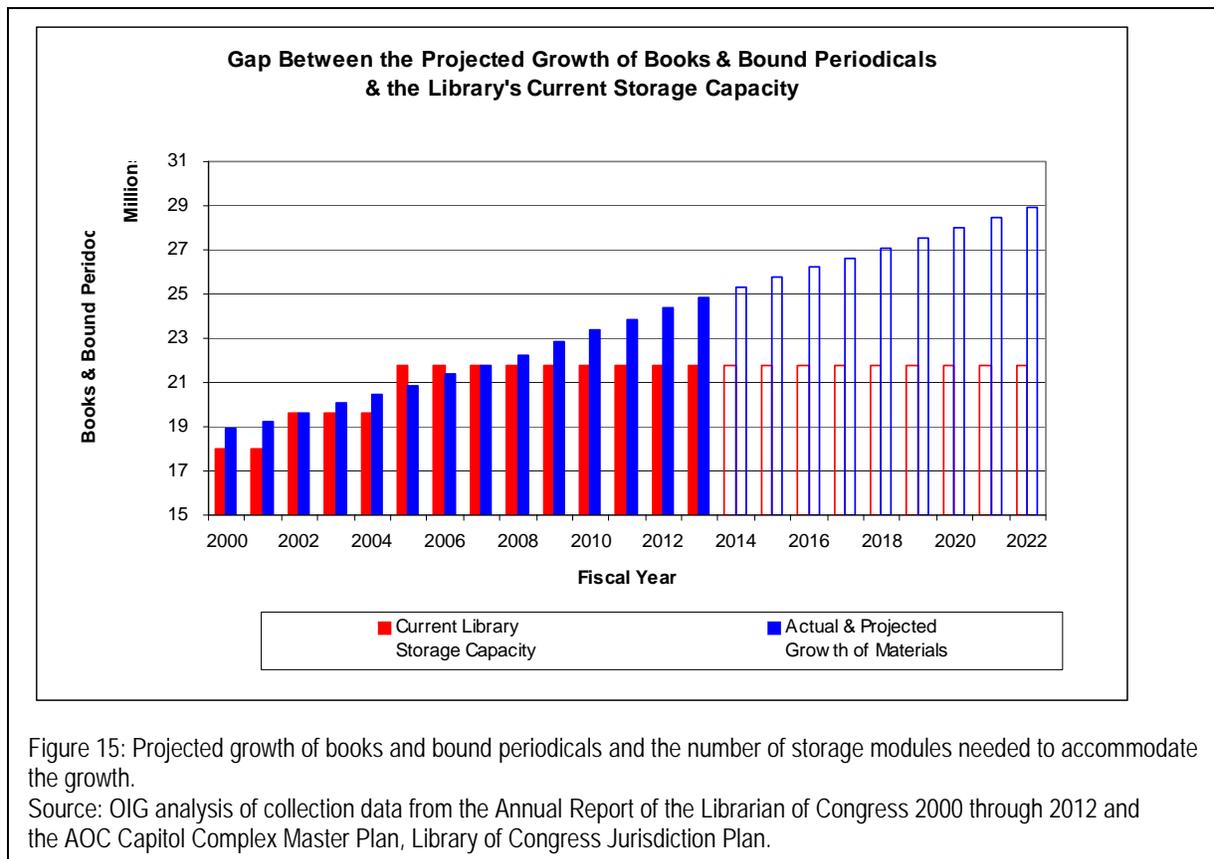


Figure 15: Projected growth of books and bound periodicals and the number of storage modules needed to accommodate the growth.
 Source: OIG analysis of collection data from the Annual Report of the Librarian of Congress 2000 through 2012 and the AOC Capitol Complex Master Plan, Library of Congress Jurisdiction Plan.

If the Library continues to acquire books and bound periodicals at the current rate, we estimate the Library will require three additional modules to accommodate the current and future collections materials through 2020.⁴⁶ This will require the Library to secure funding from Congress for the design and construction of the modules at Ft. Meade, which the Library and Architect of the Capitol (AOC) have not been able to accomplish. The AOC’s master plan for Library facilities at Ft. Meade provided for a total of thirteen storage modules and other facilities to be constructed over time. The

⁴⁶ The Library is currently over storage capacity by approximately 3 million books. If the Library continues to acquire books at the current rate, approximately 450,000 a year, by 2020 it will need space to store approximately 6 million books. Given that a storage module can accommodate 2.2 million books, then the Library will need three additional modules to accommodate the current 3 million overflow books and estimated 3 million books to be acquired. This will sustain the book collections until 2020.

modules were designed with environmental controls to provide an ideal storage environment for the collections materials. As a result, the temperature is controlled at 50 degrees Fahrenheit and 30 percent relative humidity. The materials are projected to last six to eight times longer in this environment than in a typical office environment.

The original construction schedule for Ft. Meade modules provided space to house the collections and for continued expansion. To date, four modules have been built at Ft. Meade. Modules 1 and 2 are configured as repositories for books while modules 3 and 4 are allocated to special format materials. Module 1 has been filled since 2005, and module 2 is now at capacity. Modules 3 and 4 were filled in 2012. All modules have similar environmental controls and security arrangements; however, the interiors have been fitted with different storage container systems to support the different formats.

Module 5 was originally planned to open in 2005 and would have provided 17,000 square feet of storage and held 2.2 million books and bound periodicals. Presently, for the project to proceed, only funding for construction remains to be secured (the design has already been completed). The AOC requested \$16.941 million to construct module 5 in their FY 2011 budget request but did not receive funding.

The plan for all the modules is at least eight years behind schedule. According to the original Master Plan, module 8 would have been constructed in FY 2011.

The Library's Temporary/Leased Storage Facilities Are More Costly Than Building and Not the Most Effective Solution for the Preservation of the Materials

The Library has obtained additional storage space through leasing. The leased space may be more efficient than new construction in the short-run; however, in some cases the Library has to retrofit the leased property to adequately support the collections or pay a premium for an adequately conditioned space. The Library currently has several leased storage facilities, including but not limited to the Landover Center Annex in Maryland (Landover) and a National Archives and Records Administration (NARA) facility in

Valmeyer, Illinois. Both of these leased spaces are not ideal. Specifically:

- The lease of the Landover facility has cost \$51 million to date and is projected to cost another \$36 million – the total cost will be over five times the cost to build a module.* In 1975, the Library leased the Landover facility in Maryland to temporarily bridge a storage gap until the Madison Building opened. Since 1975, the Library has paid over \$51 million for the lease of the Landover facility. Today, the Landover facility continues to provide the Library with 165,000 linear feet of general collections storage space, most of which is not cold storage (the current modules provide about 45,000 linear feet of cold/cool storage). In 2012, the Library entered into a 10-year lease agreement with the General Services Administration to maintain the Landover facility for a projected cost of almost \$36 million, see Figure 16.

Fiscal Year	Annual Cost of Landover
2012	\$ 3,206,878
2013	\$ 3,460,252
2014	\$ 3,491,298
2015	\$ 3,523,276
2016	\$ 3,556,213
2017	\$ 3,590,139
2018	\$ 3,625,082
2019	\$ 3,661,073
2020	\$ 3,698,144
2021	\$ 3,736,328
Total	\$ 35,548,682

Figure 16: Annual lease cost of the Landover Facility over ten years.
 Source: OIG compilation of annual lease agreement payments.

The facility contributes over 14,000 square feet of cold storage⁴⁷ and 4,000 square feet of cool storage⁴⁸ to the Library’s space portfolio. However, the facility is otherwise not suited for the long-term preservation of collections due to its environmental conditions, storage systems, and security provisions.

The most important environmental factors for an item’s longevity are temperature and relative humidity; different formats have different ideal temperatures and humidity set points. The Landover facility was not designed for lower temperature and humidity settings required by many of the Library’s collections, and retrofitting is difficult and expensive. Although the Library is addressing the overflow issue by installing shelving for high-density storage at the Landover warehouse, this facility will only be able to house about 800,000 volumes. As the Library adds on average 447,060 books and periodicals to the collections each year, the shortage of space will continue to grow.

⁴⁷ “Cold” storage refers to an environment as low as -4 degrees Fahrenheit with a relative humidity of 40 percent.

⁴⁸ “Cool” storage refers to an environment of approximately 50 degrees Fahrenheit with a relative humidity between 35-50 percent.

- *The NARA facility has limits in the space it can provide and in access to materials, but is only supposed to be a temporary solution.* In 2012, the Library secured storage space at the NARA storage facility in Valmeyer, IL to house approximately one million items from the Library's collections. According to NARA, 70,000 cubic feet of storage space has been earmarked for the Library. There is no assurance that additional storage space will be allotted to the Library.

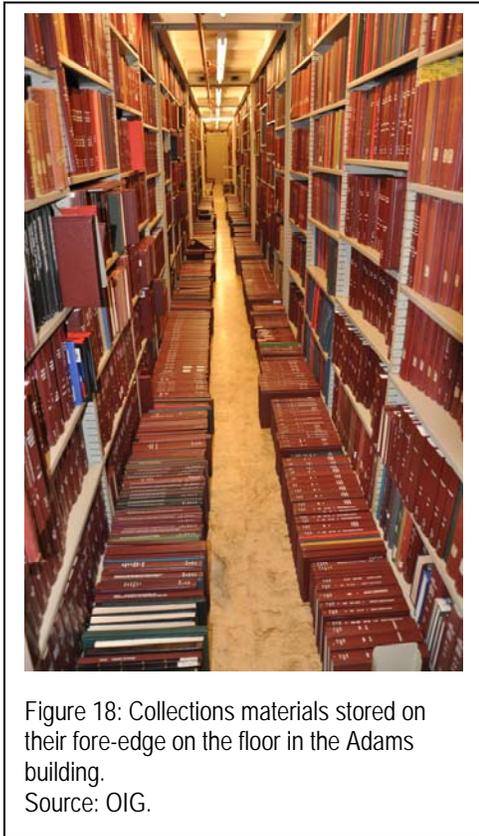
In FY 2012, the Library relocated approximately 35,000 items to the NARA facility. The Library's plan is to transfer 200,000 items annually to the NARA facility. As a result, the number of cubic feet that the Library occupies and costs associated will continue to increase as materials are transferred.

NARA storage charges are based on the total number of cubic feet of storage space that the Library occupies each month. In addition, the Library is responsible for transportation costs and a \$3.50 per cubic foot service charge for the transfer of materials. The retrieval of collections stored in leased space is costly. In addition, NARA charges a fee for the retrieval and reshelving of requested items. The Library is reducing the reshelving expense by sending to the NARA facility only materials with a digital surrogate.⁴⁹ At the time of our evaluation there was no historical data on retrieval rates and cost; material transferred to NARA had not yet been requested. Figure 17 shows the projected annual storage cost of using NARA.

⁴⁹ In its response, the Library stated that it's plans have now been expanded to go beyond the original category of materials sent to the Illinois facility. In addition to the existing issues associated with this off-site storage, the storage of non-digitized material will have associated access costs that may affect an already constrained budget.

Projected Annual Storage Cost Charged by NARA						
Fiscal Year	Annual Amount of Items to NARA	Amount of Items at NARA	Total cu ft of Storage Space Required	Rate Per cu ft	Estimated Cost of Storage Per Month	Estimated Cost of Storage Annually
2013	200,000	235,000	7,121	0.23	\$ 1,638	\$ 19,655
2014	200,000	435,000	13,182	0.24	\$ 3,164	\$ 37,964
2015	200,000	635,000	19,242	0.25	\$ 4,811	\$ 57,727
2016	200,000	835,000	25,303	0.26	\$ 6,579	\$ 78,945
2017	200,000	1,035,000	31,364	0.27	\$ 8,468	\$ 101,618
2018	200,000	1,235,000	37,424	0.28	\$ 10,479	\$ 125,745
2019	200,000	1,435,000	43,485	0.29	\$ 12,611	\$ 151,327
2020	200,000	1,635,000	49,545	0.30	\$ 14,864	\$ 178,364
2021	200,000	1,835,000	55,606	0.31	\$ 17,238	\$ 206,855
						\$ 958,200

Figure 17: Projected annual storage cost of utilizing the NARA storage facility.
 Source: OIG projection of the annual cost of utilizing the NARA facility. This cost does not include transportation costs associated with the transfer and retrieval of material.



Because of the Lack of Proper Storage Space, Some of America’s Heritage Assets Are Damaged and Deteriorating at a Premature Rate

The Library has approximately 3 million items in the Jefferson and Adams buildings that cannot be stored appropriately because there is insufficient space. According to the AOC, the agency that manages the Library’s buildings, the Library’s three Capitol Hill buildings are collectively at 110 percent of their capacity. The Library has implemented a variety of measures to accommodate these materials; however, in doing so, the Library has resorted to storing some books under less than ideal conditions. This includes storing books on the floor, as double and triple shelved, in the dustpans,⁵⁰ or turned down on their fore-edge.⁵¹ Consequently, those books are being damaged and are deteriorating more quickly than should be expected. The Library has over-crowded the decks of the Jefferson Building and stores some books in the aisles/egress pathways of the Jefferson and Adams buildings. Figures 18 and 19 show books that are stored in the aisles and on their fore-edge.

⁵⁰ Dustpan is the very bottom of the shelving section. These books are the most susceptible to water damage, damage by kicking, etc. However, the Library can often get an additional shelf in a section by using the dustpan.

⁵¹ The fore-edge of a book is the trimmed edge that is opposite of the spine. Books are turned on their fore-edge to get an additional shelf in a section. Books on their fore-edge are subject to having the text block pull away from the binding.

In the current state, the collections have become “compressed” due to constrained shelf space and are subject to damage over time. The items must be allotted sufficient space on the shelves without resorting to fore edging and double- or triple-shelving. Fore-edging damages books due to the weight of the books’ pages pulling on the binding. When a book is double or triple-shelved and is removed, its covers, bindings, and internal pages can be damaged.

Further, the Library’s Collections storage areas in all three Capitol complex buildings have experienced excessive humidity levels, which place collections in conditions that fall outside recommended standards. Without a proper storage environment, the Library’s collections may rapidly deteriorate. In warm, moist environments, the materials can become brittle, discolored, weakened, and lose information. In overly dry or rapidly fluctuating environments, the materials can crack, shrink, and warp. Regardless of temperature, environments that are too moist promote the growth of mold.

In addition, the overcrowding of collections materials in the Jefferson and Adams buildings has increased the risk of safety violations. The AOC has expressed concern that overcrowding is restricting access to the fire sprinkler valves and egress pathways. The egress pathways must be open should a fire occur in the bookstacks.

The overcrowding in the decks is also placing a significant amount of weight in the Jefferson bookstacks, which were not built to sustain the excess weight. Consequently, excess materials stored on booktrucks in the aisles must be moved around periodically to prevent compromising the structural integrity of the bookstacks. Figure 20 shows collection materials that are stored on the floor or in carts due to lack of adequate shelving.



Figure 19: Collections materials damaged by water leaks when they were stored on their fore-edge on the floor. Source: Library Services Collections and Services Directorate, Collections, Access, Loan & Management Division.



Figure 20: Collections materials stored on the floor and carts due to the lack of shelf space in the Jefferson building. Source: OIG.

IV. Other National Libraries Face Similar Constraints but Have Addressed Them Differently

The Library is a unique institution but shares a mission and function similar to other national libraries. Because national libraries are all familiar with budget constraints, exploring how these libraries deal with the current fiscal climate is a useful exercise. The way these libraries manage their collections and adjust their collection strategies can highlight some alternative approaches for the Library to consider. Through benchmarking, we sought to determine how other national libraries, tasked with the same role and responsibilities as the Library, were managing with reductions in government funding while still accomplishing their overall mission. We selected two national libraries that have similar mandates to maintain a record of their nation’s cultural and intellectual heritage: the British Library and Library and Archives Canada (Canadian Library). Both of these libraries serve as legal deposit libraries for their nations and receive a major portion of their acquisitions through legal deposit.⁵² Based on the current and future budgetary environment, both the British Library and Canadian Library realized that ‘staying the course’ or ‘going it alone’ were not viable options, and it was no longer achievable or appropriate to collect comprehensively and universally.

For comparative purposes, we have provided their annual appropriations, full-time equivalent personnel (FTEs), and collections/acquisitions in the figures below, along with the percentage changes.

Background Information on Benchmarked Libraries					
	Annual Appropriations (in thousands)				
	FY2010	FY2011	% Change	FY2012	% Change
British Library	£109,464	£105,847	-3%	£101,873	-4%
Library and Archives Canada*	\$140,079	\$129,037	-8%	\$125,373	-3%
Library of Congress	\$643,337	\$628,677	-2%	\$587,344	-7%

Figure 21: Government funding comparison.
 Source: Annual reports and Library of Congress, Office of the Chief Financial Officer.
 *Appropriations are represented in Canadian dollars.

⁵² The UNESCO Guidelines for legal deposit legislation define legal deposit as “a statutory obligation which requires that any organization, commercial or public, and any individual producing any type of documentation in multiple copies, be obliged to deposit one or more copies with a recognized national institution.”

Background Information on Benchmarked Libraries					
	Library Staff (actual FTEs)				
	FY2010	FY2011	% Change	FY2012	% Change
British Library	1,953	1,837	-6%	1,636	-11%
Library and Archives Canada	1,140	1,119	-2%	1,113	-1%
Library of Congress	3,573	3,515	-2%	3,270	-7%

Figure 22: Staffing level comparison.
Source: Annual reports and Library of Congress, Office of the Chief Financial Officer.

Collection Size of Benchmarked Libraries		
	Collection Size (items)	Acquisitions for the Collections (in USD \$ thousands)
	FY2012	FY2012
British Library ⁵³	109,608,428	\$24,366
Library and Archives Canada*	52,642,000	Not Available
Library of Congress	155,357,302	\$27,900

Figure 23: Collection size and acquisition spending comparison.
Source: British Library *Thirty-ninth Annual Report*, and Library and Archives Canada, *State of the Holdings Report*, June 2012.
*Linear meters of government and private textual records not included.

We discussed with these Libraries the following four core areas—collections development strategy, digital preference, partnerships, and storage. We also compared the information collected to that of the Library, and this comparison is summarized in the table below. A more detailed review of our benchmarking can be found in Appendix B of this report.

- Focused/Disciplined Collections Development Strategy***—Both institutions’ collection strategies, although slightly different in approach to each other, acknowledge that a comprehensive collections policy is no longer sustainable.

The British Library’s revised collections development strategy focuses on a smaller number of subject

⁵³ The British Library website: www.bl.uk/aboutus/quickinfo/facts/index.html states, “The [Library’s] collection includes well over 150 million items, in most known languages...” For the purpose of the comparison in Figure 23, we sourced collection holdings (as of March 31, 2012) detailed in the British Library *Thirty-ninth Annual Report*, pg. 33. The same date was used to calculate the FY 2012 British Pound (£) conversion rate = 1.6012500000, www.xe.com.

priorities. The British Library's strategy also focuses primarily on United Kingdom heritage materials and materials of importance to United Kingdom researchers rather than foreign materials. The British Library continues to acquire foreign material and has not specifically changed its acquisition policy towards foreign material; however, it will not expend significant funds toward material that is adequately achieved in the country of origin. Furthermore, the British Library will not acquire in areas well served by other Libraries. The Canadian Library has also adjusted its acquisition strategy to focus on collecting Canadian cultural and intellectual heritage materials and no longer participates in the international exchange program.

- *Digital Preference*—The British Library and the Canadian Library are transitioning to digital deposit as a preferred method for copyright material. Both institutions' acquisition preference has shifted from analog to digital and both intend to acquire increasingly more digital material going forward. The British Library's digital strategy is to connect users rather than acquiring materials. The Canadian Library recognized that the new user environment was decentralized and its Canadian Library could no longer serve as the sole steward of the national documentary heritage. The effect is a collaborative approach with provincial and territorial archives, libraries, and academia. Through these collaboration efforts, the Canadian Library seeks to build complementary collections with partner institutions.
- *Partnerships*—The British Library has placed emphasis on attracting external investments to supplement funding for library initiatives. By developing partnerships with the public and private sector, the British Library has secured an agreement with Google to digitize 250,000 out-of-copyright books from the British Library's collections, at Google's expense. In addition, the British Library has received a financial investment from BrightSolid to fund a project to digitize up to 40 million pages of newspapers from the British Library's collections.

- *Storage*—Both institutions recently received funding for new storage facilities. In 2009, the British Library constructed a fully automated storage building which allowed for discontinuation of three leases. A state-of-the-art building is currently under construction to house the national Newspaper Collection. The British Library has no plans to procure additional storage and is focused on maximizing the use of current storage capacity. The Canadian Library’s holdings are currently housed in a number of owned and leased facilities. The Canadian Library is building a new high-density storage facility that is projected to open in 2014 and plans to move out and close down a number of facilities to bring holdings together.

Figure 24 is a comparison of the libraries we benchmarked to that of the Library of Congress.

National Library Comparison Summary			
Benchmark	Library of Congress	British Library	Library and Archives Canada
Budget Reductions	In FY 2012, received \$587 million appropriations, 9 percent lower than FY 2010.	In FY 2012, received £102 million grant-in-aid, 7 percent lower than FY 2010.	In FY 2012, received \$125 million in total authorities, 11 percent lower than FY 2010.
Acquisitions and Collections Strategy	<p>Collection policy statements are broad, covering virtually every discipline and field of study, with the exception of technical agriculture and clinical medicine.</p> <p>Collections policy statements, based on "canons of selection" created in the mid-20th century.</p> <p>Collection policy statements were last reviewed and updated November 2008.</p>	<p>Not achievable or appropriate to collect everything and will not collect content equally in every subject.</p> <p>Will no longer 'go it alone.'</p> <p>Will not acquire in areas they feel are well served by other Libraries and will work in collaboration with other institutions.</p> <p>The Library should not automatically think of itself as actually holding material; Library's content can be 'connected to' (e.g., licensing, partnerships) but not actually retained.</p> <p>Subject matter based on UK researcher subject demand, UK research donor priorities, and subject matters that are underserved by other partners.</p>	<p>In the current economic environment and with the explosion of information, collecting all the heritage materials being created and comprehensive acquisition and preservation by one institution are unattainable goals.</p> <p>'Staying the course' is no longer an option.</p> <p>Introduced Modernization Initiative and re-examined its evaluation and acquisition strategy. Recently approved a holistic framework based on a new 'Whole-of-Society' approach.</p> <p>No longer participating in the international exchange program.</p>
Legal Deposit	<p>Mandatory deposit requires two copies of every copyrightable work published in the United States.</p> <p>Interim deposit regulations of electronic works published in the United States and available only online establishes that online-only works are exempt from mandatory deposit. Demands may be made only for works published on or after February 24, 2010.</p>	<p>Entitled to one copy of every printed work published in the UK or Republic of Ireland.</p> <p>Under proposed regulations, the Library will be entitled to a print or digital copy (not both).</p> <p>In general, the Library's preference is digital. Anticipates pending regulations will permit a gradual changeover in acquisitions from print to digital.</p>	<p>Required deposit of one or two copies, dependent on print run. Applies to all publishers in Canada, and to all publications in all mediums.</p> <p>In January 2007, the Legal Deposit Regulations were extended to Internet publications.</p> <p>Digital is default choice in the acquisition process.</p>

Figure 24: Benchmarked national libraries comparison.
 Source: Benchmarked libraries' Web sites, interview responses, and reports.

National Library Comparison Summary (continued)			
Benchmark	Library of Congress	British Library	Library and Archives Canada
Storage	<p>Long-term storage plan to construct additional storage modules at Ft. Meade.</p> <p>Holds a number of leased storage facilities.</p> <p>In 2012, secured storage space at a NARA storage facility to house approximately one million items.</p> <p>The Library has exceeded its storage capacity.</p>	<p>A fully automated storage building was recently constructed which allowed the Library to vacate three leases. A second building is currently under construction to house the Newspaper Collection.</p> <p>No plans to procure additional storage. Focus is to maximize the use of the existing storage capacity.</p> <p>Project to have enough storage capacity over the next 20 years.</p>	<p>A new storage facility is projected to open in 2014.</p> <p>Holdings are currently housed in owned and leased facilities. Planning to close down a number of facilities to bring holdings together and allow for additional growth.</p> <p>Project 10-year storage capacity.</p>
Trusted Partnerships	<p>National Agricultural Library and the National Library of Medicine are trusted partners for the collection of technical agriculture and clinical medicine material.</p> <p>World-wide collaborations including (but not limited to) libraries, educational institutions, government agencies, and non-profit organizations.</p>	<p>Developing partnerships with the public and private sector.</p> <p>Working with a number of partners and attracting external investment to enable digitization of parts of the collection.</p>	<p>Collaborative work approach with partners, specifically the pan-Canadian network represented primarily by provincial and territorial archives and libraries and academia.</p>
Digital Preservation	<p>Have well-defined digital preservation and metadata standards in place for content ingestion.</p> <p>NDIIPP has developed and implemented digital preservation standards for all elements of the preservation lifecycle.</p> <p>DuraCloud pilot in 2009. No major use of cloud providers.</p>	<p>Currently revising their digital preservation strategy and working on development of technical solutions and workflows for long term preservation of digital materials.</p> <p>Have well-defined metadata standards for content ingested. Working towards a workflow to ensure consistent management of digital content across the lifecycle.</p> <p>In-house storage servers and a trusted digital organization (not certified).</p> <p>Not using cloud providers, currently revisiting.</p>	<p>Moving to a digital legal deposit model where metadata will be automatically acquired by the Library as material is deposited.</p> <p>Use servers and in-house tape library. Developing a 'trustworthy' repository environment.</p> <p>Canadian federal government is moving to a centralized digital server and storage model. As a result, the Library will not own its servers.</p> <p>Not using cloud providers currently.</p>

Figure 24: Benchmarked national libraries comparison.
 Source: Benchmarked libraries' Web sites, interview responses, and reports.

Major Contributors to This Report:

Kurt Hyde, Assistant Inspector General

John Mech, Senior Lead Auditor

Elizabeth Valentin, Auditor

Christine Cochrane, Auditor

Sarah Sullivan, Management Analyst

► APPENDIX A: CRITERIA FOR SELECTING VIABLE TRUSTED PARTNERS

Assessing an Organizations Performance, Accountability, and Sustainability

At the onset of the evaluation, the Associate Librarian for Library Services requested the OIG's assistance in identifying criteria for selecting viable trusted partners. To satisfy Library Service's request, we compiled criteria from the Trustworthy Repositories Audit & Certification (TRAC): Criteria & Checklist, which is used in assessing whether an institution is administratively responsible, organizationally viable, and financially sustainable as part of the process for certifying whether an organization's digital repository is trustworthy. A task force of Online Computer Library Center's Research Libraries Group (RLG) and the National Archives and Records Administration (NARA) developed TRAC. The checklist was published in 2007 by NARA, RLG, and the Center for Research Libraries.

TRAC articulates a framework of organizational attributes that are used as indicators of a digital repository's comprehensive planning, readiness, ability to address its responsibilities, and trustworthiness. We believe these same attributes can be applied to assess any organization's performance, accountability, and sustainability. The Library may apply the criteria as part of their overall evaluation and assessment of an organization's infrastructure to identify potential trusted partners. Specifically, the criteria outlined in TRAC assess an organization's business plan, readiness, ability to address its responsibilities, and its reliability.

TRAC groups the criteria for assessing organizational performance, accountability, and sustainability into five categories: (1) governance and organizational viability, (2) organizational structure and staffing, (3) procedural accountability and policy framework, (4) financial sustainability, and (5) contracts, licenses, and liabilities. Under each of these categories, TRAC outlines requirements that the organization must demonstrate in order to be certified. We have taken the criteria under each category and rephrased it to apply to collections materials instead of digital information.

Governance and Organizational Viability

1. a mission statement that reflects a commitment to the long-term retention of, management of, and access to collections materials.
2. an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the organization ceases to operate or the governing or funding institution substantially changes its scope.

Organizational Structure and Staffing

1. established duties that it needs to perform and has appointed staff with adequate skills and experience to fulfill these duties.
2. sufficient staff to support all functions and services.
3. an active professional development program that provides staff with skills and development opportunities.

Procedural accountability and policy framework

1. provides clear and explicit documentation of its requirements, decisions, development, and actions to ensure long-term preservation and access to collections materials in its care. This documentation assures constituents, management, producers, and the Library that the organization is meeting its requirements and fully performing its role as a trusted partner.

Financial sustainability

1. short- and long-term business planning processes to sustain the organization over time.
2. processes to review and adjust business plans at least annually.
3. financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.
4. an ongoing commitment to analyze and report on risk, benefit, investment, and expenditure (including assets, licenses, and liabilities).
5. monitors for and bridges gaps in funding.

Contracts, Licenses, and Liabilities

1. manages, preserves, and/or provides access to materials on behalf of another organization, and maintains appropriate contracts or deposit agreements.
2. contracts or deposit agreements specify and transfer all necessary preservation rights, and rights transferred are documented.
3. specified all appropriate aspects of acquisition, maintenance, access, and withdrawal in written agreements with depositors and other relevant parties.
4. tracks and manages intellectual property rights and restrictions on use of materials content as required by deposit agreement, contract, or license.

The National Archives and Records Administration Benchmarking

The National Archives and Records Administration (NARA) is considered the nation's record keeper. The agency is charged with making the Government's *historical* information available to the public. NARA's *Strategy for Digitizing Archival Materials for Public Access, 2007-2016* includes 11 principles for establishing partnerships to increase access to historical government information through the increased availability of information technology products and services. The following are current partners NARA is working with and the principles to these digitized partnerships:

- EMC corp
- Fold3.com (formerly Footnote)
- Genealogical Society of Utah
- Ancestry.com
- Google

NARA Principles for Partnerships to Digitize Archival Materials:

1. Agreements with partners to digitize archival materials will be non-exclusive...
2. ...NARA gains unrestricted rights to the digital copies, including the right to give or sell unrestricted digital copies in whole or part to other entities.
3. Partnerships will support the goals of increased access and enhanced preservation of archival materials...Our objective is to both digitize high-use records, and to continue to develop our own capacity to digitize archival materials.
4. To provide for full access and effective preservation, partners will digitize full series or file segments of records, not just selected documents...NARA may choose to permit the digitization of selected archival materials rather than full series or file segments of records.
5. Public access to publicly owned resources will remain free. Partners may develop and charge for value-added features, but access to the digital copies ultimately should be readily accessible and free...
 - a. Access to the products of the partnership will be free to the public in all NARA's research rooms.
 - b. Partners shall provide NARA without charge a full set of the digital copies produced by the partnership...NARA will have unrestricted ownership of these copies, including the right to make these copies freely available online.
 - c. Partners shall provide NARA without charge a set of metadata generated by the project sufficient to make the digitized copies usable by NARA, and that adheres to NARA's descriptive standards. Ultimately, NARA will have unrestricted ownership of this metadata.

6. NARA will structure partnerships to balance the interests of the American public with the partner's financial investment in the project. There is no single required partnership model.
7. NARA cannot guarantee the authenticity of the digital copies... NARA cannot endorse the authenticity of archival materials hosted on a digitizing partner's Web site. Although NARA will guarantee that our digital copies have not been altered after we take possession, copies for users requiring certification will be made from the original documents or NARA-created microfilm.
8. No partnership agreement to digitize access-restricted materials shall permit the release of these materials before an agreed-upon date or specific contingency, nor shall it delay timely public access.
9. The safety and accessibility of original records will be safeguarded at all times during the digitizing process.
 - a. Archival materials will be handled according to the relevant preservation and security standards at all times.
 - b. NARA makes the final determination whether archival materials are too fragile for digitization through a partnership.
 - c. To minimize handling wear-and-tear, original materials normally will only be digitized once.
 - d. Digitization will take place at a NARA facility or at another facility which has been approved by NARA.
 - e. NARA and partners will seek to minimize the amount of time archival materials will be removed from public access during the digitizing process.
10. NARA will seek to protect and enhance its own institutional interests, while at the same time respecting the interests of our users and our partners.
 - a. NARA makes the final determination regarding whether materials may be digitized or not.
 - b. Any use of the NARA brand must be approved in writing by NARA.
 - c. The partner shall pay all direct costs associated with the digitizing partnerships, to include project management, document identification, collections security...document preparation... metadata collection and quality control, data management, digital conversion, and partner's delivery, marketing, and maintenance...NARA may exercise more leeway when the partner is not-for-profit or if the project is especially important to the mission of the agency. NARA will seek partner assistance in defraying NARA's own delivery, marketing, and maintenance costs.

- d. NARA does not allow for wholesale downloading of a large or complete body of digital content on our Web site...
 - e. The partner may not claim copyright in the digital copy.
11. NARA will publicize and seek written comment on proposed partnerships before they are signed.

►► APPENDIX B: BENCHMARKING OF NATIONAL LIBRARIES

Our benchmarking sought to determine how other national libraries, tasked with the same role and responsibilities as the Library, were managing with reductions in government funding while still accomplishing their overall mission. We selected two national libraries to benchmark which have similar mandates to maintain a record of their nation's cultural and intellectual heritage, the British Library and the Library and Archives Canada (Canadian Library). These two highly respected national libraries are experiencing similar budget constraints as the Library and find themselves in the transitional period of the analog/digital age. Both libraries serve as their national legal deposit library and receive a major portion of their acquisitions through legal deposit.

Budget Reductions

Based on their limited budget outlooks, both the British and Canadian Libraries realize that 'staying the course' or 'going it alone' is not an option and it is no longer achievable or appropriate to collect everything. Both institutions have recently revisited their collection strategies to address these challenges. The abundance of digital material presents new complex challenges for storage and preservation for any institution. Like the Library of Congress, both recognize the immediate need for collaboration in the acquisition of digital material. Both institutions have recently received funding for storage facilities which has enabled the discontinuation of lease-held facilities as well as providing better environmental standards for preservation. Both have long-term storage plans in place with future 10-20 year space capacity.

The British Library is facing a 15 percent cut to their budget over the next four years with more cuts likely to follow. The main source of funding is Grant-in-Aid from the United Kingdom government. Grant-in-Aid provides for acquisitions and funds for other functions serving the Library's core activity, for example staff costs, estates costs, capital expenditure, etc. Government funding is supplemented with other sources of income including philanthropy, commercial revenue (e.g., fees), and partnerships (e.g., research councils).

The Canadian Library's spending power has been reduced by 30 percent in recent years. The Canadian federal government announced a \$9.6 million cut (roughly 10 percent) to the budget by 2014-15.

Acquisitions and Collections Strategy

The role of the British Library is to develop the national published archive through legal deposit, to support United Kingdom research through collecting and connecting to contemporary content, and to support research and culture through developing world-class primary research collections. In the current environment, the British Library realizes that it is not achievable or appropriate to collect everything and will not collect equally in every subject. They will no longer 'go it alone'. They will not acquire in areas they feel are well served by other Libraries and will work in collaboration with other institutions. The

Library will focus on a smaller number of subject priorities that will be more focused than what would be collected in a broader collection strategy. Subject matter will be based on United Kingdom researcher subject demand and subject matter that are underserved by other partners. The British Library continues to collect foreign material and has not specifically changed its acquisition policy towards foreign material; however it will not expend significant funds toward material which is adequately achieved in the country of origin.

The Canadian Library's core responsibilities include: managing the legal deposit for Canadian publications; storing government records of importance, similar to NARA; and acquiring and preserving Canadian heritage material. The Canadian Library recognizes it will never be possible to collect all the heritage materials being created and that comprehensive acquisition and preservation by one institution are unattainable goals. Simply, 'staying the course' is no longer an option. To address these challenges the Canadian Library recently launched a Modernization initiative. Through the initiative, it intends to transform itself from an institution that once gave priority to the acquisition and preservation of analog materials; to an institution that promotes open access to Canada's documentary heritage. This modernization is based on five key principles (1) collaborating with institutions that share complementary mandates, (2) redefining the selection process to ensure it evolves in line with priorities, (3) improving access to current holdings, (4) preserving both digital and analog documentary heritage, and (5) building its capacity to manage and carry out its mandate. The Canadian Library re-examined its evaluation and acquisition strategy and recently approved a holistic framework based on a new 'whole-of-society' approach to take into account how well an acquisition represents the 'whole' of Canadian society, and which organization should house it. In order to focus on its core mandate of collecting Canadian Heritage material, the Library adjusted its acquisition program and is no longer participating in the international exchange program.

Legal Deposit

Like the Library of Congress, the British Library and the Canadian Library function as legal deposit libraries for their nation. The British Library is entitled to one copy of every printed work published in the United Kingdom. At the time of our benchmarking, pending legal deposit regulations state the British Library will be entitled to have print or digital deposits, but not both. This will occur with a mutual agreement between the publishers and the Library. The British Library's *Content Strategy 2013-2015*, discusses a natural shift into digital materials and, in general, the library's legal deposit preference is for digital rather than analog. It anticipates that the regulations will permit a gradual changeover in acquisitions from print to digital, enabling the growth in digital processing to be supported by savings from print processing. However, during this transitional period, the Library will keep analog holding secure as it moves into digital.

The Canadian Library's primary tool for acquisition is through legal deposit and is the means by which a comprehensive national collection is assembled. Legal deposit applies to all publishers in Canada, and to all publications in all mediums. In January 2007, the Legal Deposit Regulations were extended to Internet publications. Digital is Library and Archives Canada's default choice in the acquisition process, and the Canadian Library will acquire increasingly more digital material going forward.

Storage

As of 2012, all collections except newspapers and microfilm are now stored in one of two British Library-owned sites. In 2009, a fully automated storage building was constructed which allowed it to vacate three leases, and a second fully automated building is currently under construction to house the Newspaper Collection. The British Library has no plans to procure additional storage and it is unlikely they will receive further significant government funding in the foreseeable future. The focus is to maximize the use of the storage capacity. The British Library feels they have enough capacity over the next 20 years and that timeline will likely increase with the pending legal deposit legislation and the shift to digital deposits.

The Canadian Library has a new high-density storage facility that is projected to open in 2014. Because of funding cuts there is very little flexibility to invest in new collections storage or in new technology. The Library's holdings are currently housed in a number of owned and leased facilities. Over the coming years, the library plans to move out and close down a number of facilities to bring holdings together and allow for an additional 10 years of growth space.

Trusted Partnerships

The British Library believes the digital era offers the opportunity of reducing the need to retain and preserve some works, which are available by remote access. The Library has shifted its strategy to not automatically think of itself as holding material, but rather 'connecting to' (through licensing/purchase agreements) and not actually physically retaining in its collection. By developing partnerships with the public and private sector the British Library has been successful in attracting external investment to supplement Grant-in-Aid to enable digitization of parts of the collection. This has involved working with a number of partners:

- Google - partnership to digitize 250,000 out-of-copyright books from the Library's collections. Selected by the British Library and digitized by Google. Google will cover all digitization costs.
- BrightSolid - project to digitize up to 40 million pages of Newspapers from the collection. BrightSolid recently invested £40m in the project.

The Canadian Library has adopted a collaborative work approach with partners, specifically the pan-Canadian network represented primarily by provincial and territorial archives and libraries, and academia. Through collaboration efforts, partner institutions seek to build collections that complement each other. The library continuously reviews its holdings to determine what is relevant and what is surplus, duplicate, or could be transferred to institutions where they are more appropriate. It has come to the realization that they must be proactive in order to perform the functions of acquisition, preservation, and resource discovery of digital information. They recognize that the new environment is decentralized and the monopoly as stewards of the national documentary heritage is over.

Digital Preservation

As of 2012, the British Library has the infrastructure to accept/ingest, process, archive, preserve, and access digital collections materials, but recognizes there is a need to improve digital infrastructure to meet the immense challenges of the digital era. Since 2004, the Library has been selectively archiving Web sites with research value that are representative of British social history and cultural heritage. The British Library has in-house storage servers and a trusted digital repository, although not certified. It has found in the past that cloud providers are not cost effective, but is revisiting this as a long-term option for digital storage. They are currently revising their digital preservation strategy and working on development of technical solutions and workflows for long-term preservation of digital materials. They are partners in the European Union co-funded Scalable Preservation Environments (SCAPE)⁵⁴ project and intend to deploy many of the tools and workflows developed in SCAPE alongside their in-house digital library system. The British Library has implemented rigorous characterization practices as part of their ingest processes and have well-defined metadata standards for ingested content. They are working towards a workflow to ensure consistent management of digital content across the lifecycle.

⁵⁴ SCAPE is a European Union-funded project which is directed towards long-term digital preservation of large-scale and heterogeneous collections of digital objects. Its objectives are to develop scalable services for preservation planning and preservation actions on an open source platform.

► APPENDIX C: DIGITAL PRESERVATION

As part of our discussion in addressing all elements of our collection policy the Associate Librarian asked the Office of the Inspector General to identify the following: the gold standard of preservation practices; best practices in identifying trusted partners; and the benefits and risks associated with cloud computing as a solution to our storage issues. In this appendix, we focus on those industry standards which are currently being employed.

Industry Standards for the Format and Preservation of Digital Material

Digital preservation is an ongoing commitment to a digital lifecycle that needs constant consideration and management. The Digital Preservation Coalition's *Digital Preservation Handbook* states, "Digital Preservation [r]efers to a series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation is defined...[and] refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change..." Our research found the term 'gold standard' is allusive concerning digital preservation practices. No 'gold standard' exists, but rather industry standards and best practices. These best practices continue to evolve and vary between institutions and for different format material.

To better understand the challenges and best practices in digital preservation we researched the guidance that is produced and readily referenced by the digital preservation community. The Open Archival Information System (OAIS) Reference Model is widely adopted by the digital community and many institutions reflect its standards in their approach to long-term preservation. The OAIS Reference Model is dedicated to preserving digital information and making it available over the long term. The term 'open' reference model was developed and released in an open public forum. The functional model illustrates how data is ingested, stored and managed for long-term preservation, maintained in the databases, preserved through strategies and planning, and finally, accessed.

Preservation metadata is a key component of nearly all digital preservation processes. By attaching metadata, often referred to as 'data about data,' to an object greatly increases the findability and usability of that data. A commonly used approach to preservation of metadata is PREMIS (Preservation Metadata Implementation Strategies). The PREMIS Editorial Committee is an international working group concerned with developing metadata for use in digital preservation. The Library of Congress serves as the permanent home for the PREMIS data dictionary. The dictionary is a comprehensive description of core preservation metadata supported by recommendations and guidelines for its creation, management, and use. It includes international standards for metadata to support the preservation of digital objects and ensure their long-term usability.

A trustworthy repository to house all the digital data ingested is another piece of digital preservation. The International Organization for Standardization (ISO) is an international standards-setting body that provides recommended practices for assessing the trustworthiness of digital repositories. ISO16363:2012 is often used as a basis for this certification. As discussed in Appendix A, *Criteria from the Trustworthy Repositories Audit & Certification (TRAC): Criteria and Checklist* is also commonly used in the process for certifying whether an organization's digital repository is trustworthy.

In its 2004 collaborative report, *Trusted Digital Repositories: Attributes and Responsibilities*, published by the Research Library Group, the challenges of digital information are described as follows, "By its very nature, digital information can be transitory and difficult to preserve; certainly the traditional methods of preservation are less applicable. The digital landscape—and the digital management landscape—are also quickly evolving with the exponential growth of digital information. No one institution will be able to preserve it all—in fact, it will take many institutions, organizations, businesses, and others to preserve the cultural information that had previously been placed solely in the care of cultural institutions."

Digital Preservation Challenges

The creation of digital information and the capacity to ingest it has grown quickly over the past decade; however, the ability to manage this amount of information has been comparatively slow to develop. There are many challenges presented by digital preservation. Problems include the fragile nature of storage media (i.e., tape, compact discs, and hard drives) and obsolete software. Another major challenge is the speed and amount of digital information being created. For example, the volume of tweets the Library receives each day has grown from 140 million beginning in February 2011 to nearly half a billion tweets each day as of October 2012. The Library of Congress amassed 233 billion tweets from Twitter in 2012, with each tweet containing more than 50 fields of metadata. Digital preservation presents economic challenges as well. The ongoing management of the digital lifecycle includes continued costs for data ingest, long-term storage, and human capital resources to execute the process.

IBM's Raymond Lorie describes in his report, *Long Term Preservation of Digital Information*, "The preservation of digital data for the long term presents a variety of challenges from technical to social and organizational. The technical challenge is to ensure that the information, generated today, can survive long term changes in storage media, devices and data formats."

National Digital Information Infrastructure and Preservation Program

In 2000, the Congress appropriated \$100 million to the Library of Congress to develop a national strategy and network of partners to collect, preserve, and make available significant digital content, especially information that is created only in digital form, for current and future generations. This resulted in the establishment of the National Digital Information Infrastructure and Preservation Program (NDIIPP) with a mission to ensure

access over time to a rich body of digital content through the establishment of a national network of partners committed to selecting, collecting, and preserving at-risk digital content.

In 2003, Congress approved the NDIIPP plan that set forth a strategy for the Library of Congress, in collaboration with other federal and non-federal entities, to identify a national network of libraries and other organizations with responsibilities for collecting digital materials that will provide access to and maintain those materials. The plan was organized around four strategic goals: develop a growing national preservation network; develop a national content collection plan that will collect and preserve at-risk content; build a shared technical platform for networked preservation; and develop recommendations to address copyright issues and to create a legal and regulatory environment that both encourages incentives and eliminates disincentives to preservation.

NDIIPP is based on an understanding that digital stewardship on a national scale depends on public and private communities working together. The program is a network of world-wide partnerships to preserve at-risk digital collections and build a distributed digital preservation infrastructure. As of FY 2012, NDIIPP reported 200 partners, over 47 states and 39 nations to collect and preserve valuable digital content with special attention to public policy, education, and research. Through these partnerships the Library found that each participating institution brings its own resources, interests, and organizational culture to the network and over time the network becomes more complex, but stronger and better able to withstand stresses and strains as a result.

At-Risk Content

The report, *Preserving Our Digital Heritage: The National Digital Information Infrastructure and Preservation Program 2010 Report* states digital content is at risk for loss because the infrastructure for its collection, management, and preservation is inadequate. NDIIPP has identified four categories of risk: technological risks, legal and policy risks, content risks, and organizational risks. These risks are detailed in the Figure 25.

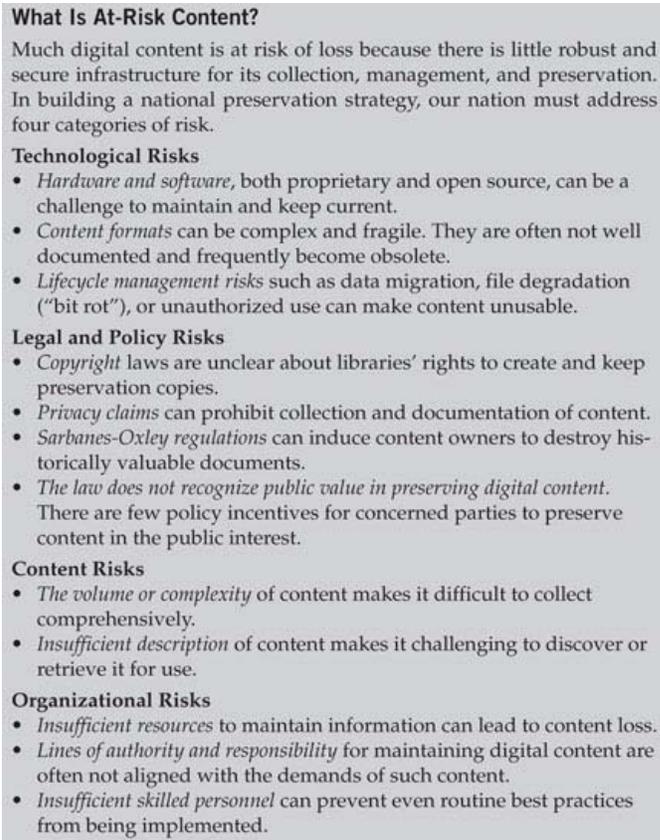


Figure 25: Description of digital preservation risks.

Source: Preserving Our Digital Heritage: *The National Digital Information Infrastructure and Preservation Program 2010 Report*.

Content Preserved and Available Through NDIIPP

Digital content production will continue to outpace the ability of any single institution to collect and preserve a comprehensive collection, requiring an alliance of organizations committed to stewardship. NDIIPP partners and the Library have selected, collected, and preserved content in areas of public policy, education and research, and cultural heritage. The value of digital content grows when it can be shared, and NDIIPP found that a distributed approach to collecting is an effective strategy to achieve breadth and depth of content. Historically significant digital content collected through NDIIPP partnerships are accessible through www.digitalpreservation.gov. The comprehensive list of collections includes:

- American History
- Arts and Culture
- Government, Politics, and Law
- Maps and Geography
- News, Media, Journalism
- Religion and Philosophy
- Science, Mathematics, and Technology
- Social Sciences
- Sports, Recreation, and Travel
- World History and Culture

Noteworthy Accomplishments of NDIIPP

Major partnerships and collaborations established through NDIIPP focused on the collection and preservation of at-risk digital content. Noteworthy accomplishments resulting from NDIIPP include the National Digital Stewardship Alliance (NDSA), Viewshare, BagIt, Lots of Copies Keep Stuff Safe (LOCKSS), and the DuraCloud pilot.

National Digital Stewardship Alliance

The National Digital Stewardship Alliance was initiated in 2010 to sustain partnership efforts and to leverage the expertise and resources of a growing community of preservation organizations. It is a collaborative effort among government agencies, educational institutions, non-profit organizations, and business entities whose mission is to establish, maintain, and advance the capacity to preserve our nation's digital resources for the benefit of present and future generations. Members of the Alliance are committed to managing digital content for current and long-term use. Approaching digital stewardship collaboratively allows the Alliance to coordinate effort, avoid duplicate work, build a community of practice, develop new preservation strategies, flexibly respond to a changing economic landscape, build relationships, and increase capacity to manage content beyond institutional boundaries. The National Digital Stewardship Alliance believes the preservation of digital information is a pervasive challenge and that engaging across different communities strengthens the nation's digital preservation practices and increases the likelihood of preserving content now and into the future. Members encourage the open exchange of ideas, services, and software.

Viewshare

The Library began a pilot project in 2009 to develop an environment that can be used to collect and explore information about digital collections. Viewshare.org is a linked data platform which helps curators assess the completeness and accuracy of metadata by viewing it in a variety of forms (i.e., interactive maps, timelines, lists, and charts). It allows users to integrate data across digital collections.

BagIt

The Library of Congress and the California Digital Library jointly developed a format for transferring digital content called "BagIt." BagIt is an attempt to simplify large-scale data transfers between cultural institutions, streamline the process, and reduce the number of moving parts. A "bag" consists of a base directory, or folder containing the "tag" and a subdirectory that holds the content files. The tag is a simple text-file manifest, like a packing slip, that consists of two elements: an inventory of the content files in the bag and a checksum of each file. A checksum is a way to validate that everything in the bag arrived accurately. Once the content is validated, the receiver notifies the sender, usually by e-mailing a confirmation receipt. In another optional file, users can add metadata about the content, such as a description of the package contents and detailed contact information for the sender. A slightly more sophisticated bag lists uniform resource

locators (URLs) instead of the actual content files. This type of simultaneous multiple transfer greatly reduces the overall data-transfer time.

Lots of Copies Keep Stuff Safe

The LOCKSS Program is an open-source, library-led digital preservation system built on the principle that “lots of copies keep stuff safe.” The LOCKSS system allows librarians at each institution to take custody of and preserve access to the e-content to which they subscribe. LOCKSS libraries acquire digital content in their local LOCKSS Box. Through a LOCKSS distributed network, libraries are cooperating with one another to ensure their preserved content remains authentic and authoritative. When the publisher’s Web site is unavailable for any reason, LOCKSS delivers a copy of the original publication to authorized users in real time, whenever it is needed. Because LOCKSS preserves the original publisher’s copy of each item, it ensures that the most authoritative version persists, unchanged, with full credit to the publisher. The LOCKSS software continually monitors the content in each LOCKSS Box to ensure that it is being properly preserved, by cooperating over the Internet with other LOCKSS boxes to compare each box’s copies of the same content.

DuraCloud

Cloud technologies use remote computers to provide local services through the Internet. The cloud lets an institution provide data storage and access without having to maintain its own dedicated technical infrastructure.

In 2009 NDIIPP launched a one-year pilot program to test the use of cloud technologies to enable perpetual access to digital content. The objectives of the DuraCloud pilot were to explore the use of cloud computing technologies to test the perpetual access to digital content, allow institutions to provide data storage and access without having to maintain dedicated technical infrastructure, and explore strategies to help make content accessible, in addition to providing preservation services.

The New York Public Library and the Biodiversity Heritage Library, both NDIIPP partners, participated in the pilot. DuraCloud provided both storage and access services, including content replication and monitoring services that span multiple cloud-storage providers. The New York Public Library focused on converting images from the TIFF⁵⁵ format to JPEG⁵⁶ 2000 and to serve these images using a powerful JPEG 2000 image engine within DuraCloud. The Biodiversity Heritage Library provides access to historical journal literature in biodiversity. Its pilot focused on replication of digital content to provide protection for valuable biodiversity resources. The library will use the cloud computing capabilities offered by DuraCloud to analyze biodiversity texts to

⁵⁵ TIFF (Tagged Image File Format) is used for storing digital images.

⁵⁶ JPEG (Joint Photographic Experts Group) is a compression technique for digital images.

extract key information such as species-related words. The institution will also deploy a JPEG 2000 image engine via DuraCloud to process and serve digital images.

According to the Library's Office of Strategic Initiatives, the cloud can be a cost effective way to store digital information. The cloud is useful for converting a mass of images at once to upgrade to JPEG 2000, similar to what the New York Public Library did in the pilot. However, there are costs involved in the preparation of digital data for submission and transferring it to the cloud. Currently the Library has no significant collections on the cloud and is exploring opportunities for future cloud storage.

The Seven Sustainability Factors for Digital Preservation

The seven sustainability factors for digital formats have been identified by the Library for the purpose of preserving digital information. These factors reflect the quality and functionality of digital formats and are taken into consideration when making format decisions for preservation. The seven factors are:

1. Disclosure
2. Adoption
3. Transparency
4. Self-documentation
5. External dependencies
6. Impact of patents
7. Technical protection mechanisms

The Preservation Lifecycle

The preservation lifecycle is the overall method in which material is maintained and preserved. Unlike the traditional approach to preservation of analog material, preservation of digital content requires active management throughout its entire lifecycle, see Figure 26. To ensure long-term access to digital materials, a life-cycle management approach is required. Only good life-cycle management can ensure preservation and access to digital materials. There are actions and decisions to be made at each phase of the life cycle that will mitigate the risk of loss of and/or damage to digital materials.

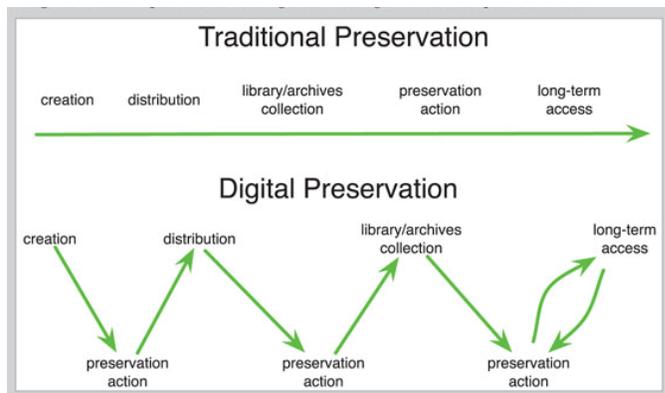


Figure 26: Flow chart of traditional vs. digital preservation.
 Source: Preserving Our Digital Heritage: *The National Digital Information Infrastructure and Preservation Program 2010 Report*.

NDIIPP extensively tested a three-layer architecture model for digital preservation, see Figure 27. The three layers approach supports preservation across time and through various changes in technology. The bottom layer stores and maintains the data. The middle layer provides services for content description and management, and is the layer associated with libraries and archives. The top layer, the access layer, provides services to view and use the content.

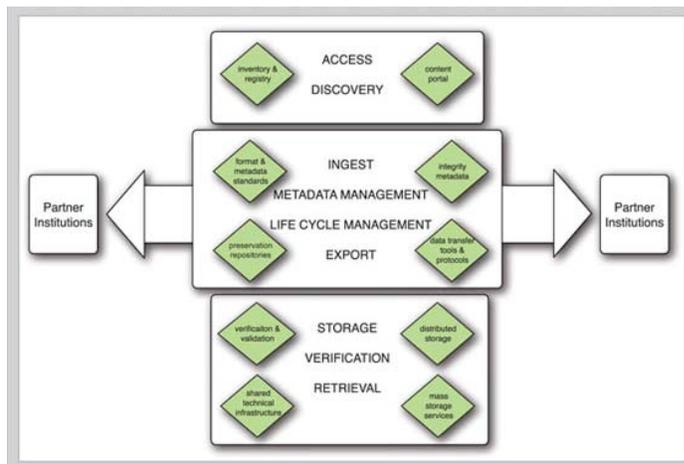


Figure 27: NDIIPP's technical architecture for digital preservation.
 Source: Preserving Our Digital Heritage: *The National Digital Information Infrastructure and Preservation Program 2010 Report*.

Digital Benchmarking

We benchmarked the National Archives and Records Administrations (NARA), a federal agency responsible for safeguarding and preserving the records of the U.S. Government. Similar to the Library’s activities through NDIIPP, NARA implements best practices for the management of digital material. NARA manages large amounts of digital material and is at the forefront of establishing preservation standards in the digital community.

The following is a summary of NARA's best practices:

1. Methodologies for ingesting/receiving, cataloging, preserving, and managing born digital materials—NARA uses the Electronic Records Archives (ERA) to ingest, store, and preserve federal records that were born digital and identified as records of permanent value.
2. Standards established for preserving digitally born materials – Open Archival Information System (OAIS) reference model for systems, PREMIS (Preservation Metadata: Implementation Strategies) model for metadata standards, TRAC (Trustworthy Repositories Audit and Certification: Criteria and Checklist) for repositories also using ISO 16363 (new in 2012) for space data and information transfer systems, not yet finalized but can be used as a basis for certification.
 - a. OAIS—The *CCSDS Recommended Practice for an OAIS Model* defines OAIS as, "...an Archive...that has accepted the responsibility to preserve information and make it available for a [designated community]...The information being maintained has been deemed to need [long term preservation]." The *Trusted Digital Repositories: Attributes and Responsibilities* report states, "the OAIS provides both a functional model—the specific tasks performed by the repository such as storage or access—and a corresponding information model that includes a model for the creation of metadata to support long-term maintenance and access." PREMIS—the international standard for metadata to support the preservation of digital objects and ensure their long-term usability.
 - b. TRAC—discussed in Appendix A.
 - c. ISO 16363:2012—defines a recommended practice for assessing the trustworthiness of digital repositories. It is applicable to the entire range of digital repositories. ISO 16363:2012 can be used as a basis for certification.
3. NARA's best practices and greatest learnings for the management of born-digital material:
 - a. Use standards—OAIS, PREMIS, TRAC and ISO 16363.
 - b. Seek to build interoperable, modular systems and repositories, since both will evolve over time.
 - c. Don't let perfect be the enemy of good. Start by capturing born-digital content and build capacity for increasingly sophisticated preservation actions over time.
 - d. Share information with and learn from other institutions. Many institutions are struggling to build capacity to manage and preserve born-digital content and they can make more progress by collaborating in a rapidly developing field.
 - e. It is necessary to continually evaluate your capabilities and ensure that your policies and practices support a consistent, achievable program.

► APPENDIX D: CURRENT STATUS *LC21: A DIGITAL STRATEGY FOR THE LIBRARY OF CONGRESS*

The study committee that issued, *LC21: A Digital Strategy for the Library of Congress* recommended that the Library put in place mechanisms to systematically address the infrastructure required for it to “collect” digital materials and set new standards for the appropriate formats for digital materials acquired as well as revisiting the Library’s ‘best edition’ statement. We requested a status update from the Office of Strategic Initiatives (OSI) regarding these specific recommendations. Below is a summary of the Deputy Associate Librarian for Strategic Initiatives response.

LC21 Infrastructure Recommendation Status

“The Library has developed and put into production guidelines, procedures and infrastructure for the transfer, receipt, inventorying, processing and preservation of digital content.

OSI has addressed this recommendation on a number of fronts, including software/tools development for the ingest and management of digital materials, reinvestment in server and storage environments for large volume digital materials, development of preservation policies, development of standards (metadata, format and hardware/software), the creation of a disaster recovery site, an additional data center at Culpeper for the ingest and preservation of film and audio, and the implementation of high speed networks.”

Ingest and Management Software/Tools

“OSI formed the Repository Development Group (RDC) [and created] open source reusable tools for content ingest and management...[including] the BagIt specification for the packaging of digital content...”

The RDC has also developed Content Transfer Services (CTS). This suite of tools, put in production at the Library in 2009, supports the secure receipt, inventorying, quality review, and auditing of incoming digital collections as well as collections digitized here at the Library...CTS is increasingly integrated with other Library systems, including the Integrated Library System (ILS), the eCO system in the Copyright Office, and the Delivery Management System for eSerials processing...

The RDC has collaborated with Library Services on the development of requirements for data transfer and repository services which were implemented as features in CTS...”

Reinvestment in Server and Storage Environment

“OSI requested funding from Congress for the Cyclical Investment in Technical Infrastructure (CITI) to invest in both tools for content management and content delivery, and the underlying core infrastructure for the Library’s digital collections. These Cyclical Investments in Technical Infrastructure (CITI) are designed to invest in the Library infrastructure to build resilient flexible and scalable system configurations based on an architectural approach and best practices...OSI proactively monitors the industry’s best practices and technologies and invests in initiatives to streamline and strengthen the infrastructure.

[OSI develops] standards for the underlying hardware and software components, and is establishing processes for updating and using the standards across the Library. The new architecture includes five dedicated ingest servers, which are optimized for receiving digital content from both the network and via portable media. OSI has built increasingly complex workflows for non-technical users in the collecting service units to be able to use these ingest servers to transfer content from portable media such as hard drives ...

In the area of storage management, ITS has built a tiered storage architecture ... This approach provides increased capacity for cost-effective scaling without affecting the way that users or applications interact with the storage. This will include the deployment of software that allows users and applications to keep track of very large numbers of files across the virtual enterprise of storage environments.

The infrastructure also incorporates best practices in development and operations as well as technology components consistent with the Library’s evolving technical architecture program. Best practices include optimization of use of resources in the computer center such as power, cooling and physical equipment.”

Investment in Networks

“ITS...re-architected the Library’s networks and continues investing in the networks components of the infrastructure to expand the availability, resilience, speed, and security of the networks...Some examples of these improvements include increasing the bandwidth within the data centers to 10 gigabytes, increasing local area network backbones to 1 gigabyte, a new Internet POP with more bandwidth and ability to burst as the need arises, a separate Internet presence for the Alternate Computer Facility, and a POP on Internet 2.”

Investment in Content Management

“...ITS is in the midst of transitioning to a technology infrastructure designed with virtualized server and storage components for more cost-effective capacity and services. At the same time, ITS is planning and implementing higher-performance paths for content movement among the Library’s data centers ...

Regarding content management policies for digital collections, they span the gamut from policies governing access, such as Copyright laws and donor agreements (Twitter), to findability access, to policies governing integrity controls, which are influenced by the size of the collections, the value, and the type of medium on which the collections are stored. For example, the Library has developed policies, procedures and tools for improving the findability and usability of descriptive information (metadata) for the Library's digital content collections, both existing and newly ingested. Policies now in place include a minimum set of metadata required for digital content, as well as guidelines for metatags to be applied to content served on the web. In addition, tools and services have been developed for the automatic generation and enrichment of metadata, resulting in more complete and consistent data supporting the Library's online collections..."

LC21 Standard and Format Recommendation Status

"In 2011-2012, five groups of digital format specialists at the Library met and developed a revised and expanded "Best Edition" statement for the preferred formats for all forms of digital and analog content that might be offered to, or received by, the Library for addition to its collections. The revised document includes a new framework that accommodates the Library's need to acquire works in many media, both physical and digital, reflecting current publishing and media production realities...The Best Edition Statement covers Textual Works and Musical Compositions, Still Image Works, Audio Works, Moving Image Works, Computer Software, and Datasets/Databases.

The...eDeposit effort covers electronic-only serials...eSerials [are] the initial content type to be included in mandatory deposit for addition to the Library's Collections, and is the sole format covered by the interim regulation issued by the Copyright Office in 2010."

►► APPENDIX E: CATALOGING PRIORITIES

Priorities

All cataloging will be done in the order of the priority that is assigned to the material as described below. At each stage in the process items of highest priority (lowest priority number) will be processed before any others. Exceptions to established priorities may be made as appropriate, e.g., to provide suitable materials for trainees, and to provide materials for catalogers with special skills or special language abilities. Adequate supervisory controls will be provided to insure that absences of personnel do not delay materials of high priority.

Priority 1

- Titles requested by Members of Congress or their staffs, by agency heads or higher officers of the Executive Branch, by Supreme Court Justices, or by division chiefs or higher officers of the Library of Congress
- Pre- and Post-Publication Cataloging in Publication titles

Priority 2

- First number or volume received of a numbered monographic series, and serials issued annually or less frequently, as well as the first and subsequent volumes of a multipart monograph that are published over a period of time
- All titles destined for the Rare Book and Special Collections Division or rare titles destined for other custodial units
- High-need and/or research value titles
- U.S. Congressional publications
- Major publications in the humanities, social sciences, law, and the physical, natural, and behavioral sciences, both primary and secondary sources
- Substantial publications of topical interest (official and nonofficial) in the following categories (in case of foreign publications, preferably those relating to the country of origin as a whole):
 - All genealogies and compilations of local records of genealogical value
 - Dictionaries: English, English to foreign language, and foreign language to English
 - U.S. Federal, state, and local materials, that are primary sources, and their search tools (government and privately printed):
 - search tools include encyclopedias, manuals, guides, digests, indexes, citators, and loose leaf services
 - Official publications of foreign countries and major subdivisions (states, provinces, capitals, etc.), that are primary sources, and their search tools (government and privately printed)

Priority 3

- Medium-need and/or research value English-language material titles
- Substantial publications in the humanities, social sciences, law, and the physical, natural, and behavioral sciences, both primary and secondary sources, not given higher priority
- Encyclopedias, almanacs, and other works of general reference not selected for reference assignment or given higher priority
- Dictionaries (foreign language only)
- Substantial travel guides
- All U.S. local histories and substantial foreign local histories
- U.S. Federal documents not given higher priority
- U.S. state documents, except primary sources, generally dealing with subjects of national interest
- Official publications of foreign countries and their major subdivisions (states, provinces, capitals, etc.), except primary sources, generally dealing with subjects of broad interests

Priority 4

- Low-need and/or research value English-language material
- Other materials selected for addition to the collections, but not given higher priority (priorities 1-3) including:
 - Children's books
 - College level textbooks
 - Official publications of foreign countries
 - Privately printed works (however, foreign works, especially for developing countries, may, in some cases, require higher priority)
 - State and local government publications
- The following types of materials are rarely given higher priority:
 - Anthologies
 - Applied arts and crafts
 - Secondary level textbooks
 - Popular instructional and devotional publications
 - Popularizations in all subject fields
 - Sports and recreation
 - Unrevised reprints

►► APPENDIX F: MANAGEMENT RESPONSE

 LIBRARY OF CONGRESS
OFFICE OF THE LIBRARIAN

MEMORANDUM

DATE September 30, 2013

TO Karl W. Schornagel
Inspector General

FROM Robert Dizard Jr. 
Deputy Librarian of Congress

SUBJECT Comments on draft OIG Report No. 2013-SP-102: *The Library Collects Extensively but Faces Increasing Challenges in Processing, Controlling, Storing, and Making Accessible All It Collects*

This memorandum provides the Library's comments on the findings of the subject report. The report addresses the Library's acquisitions and the work the Library performs in processing, preserving, and providing access to its collections.

Findings:

I. The Library has been Meeting Its Top Priority of Continually Obtaining a Universal Collection of Both Analog and Digital Materials (pg. 20)

Agree.

This finding goes most directly to the Library's mission. Sustaining a collection that meets the needs of the Congress and the American people, both now and in future generations, is essential and the pre-eminent consideration as we address the report's other findings.

II. The Library is Not Able to Timely Process the Inflow of Materials It Acquires – Current Unprocessed Arrearage is 28 Million and Has Grown Almost 50 Percent Since 2000 (pg. 27)

Agree.

This is primarily a matter of resources. The Library does, however, have a number of initiatives in place to manage the arrearage. These include:

Page 1 of 3

- Increased acceptance of bibliographic descriptions and cataloging data provided by publishers, established partners, and third-party intermediaries;
- Expansion of collection categories that qualify for minimal-level cataloging;
- Development of an entirely new bibliographic system called BibFrame, which will replace the aging MARC record system and provide a more streamlined method of bibliographic description and control of different types of content;
- Recently implemented changes in collections policies relating to acquisition and processing of additional service copies, acquisition of digital alternatives to analog materials; and
- Reestablishment of the Collections Development Officer position (March 2012) and approval for the establishment of a full Collections Development Office (September 2013).

The Library has been and will continue exercising the profession's best practices with regard to how certain types of materials are described and controlled bibliographically.

I offer a few additional comments related to this finding. First, as the tables in the report show, acquisitions fluctuate from year to year. As such, processing times will often vary even with a consistent level of processing resources.

Second, manuscripts, the personal papers of individuals or the records of an organization, are rarely processed at the item level as this may constitute a single sheet of paper. The report notes that more than half of the current arrearage is in manuscripts and does not take into account the fact that many of these collections will be processed, consistent with best practices, at the macro level.

Third, the report provides a number of examples, some photographic, that mischaracterize the impacts of delayed processing. While there are certainly consequences of not having optimum storage facilities or fully processed collections, we do not believe that these examples are commonplace or representative of overall conditions at the Library.

Related to this, the report often cites impacts related to access and security in such a way as to indicate that Library Services has little control over its growing collections. While collection items may not be accessible via online catalogs, for example, they are accessible in person. In fact, as part of the arrearage reduction efforts of a decade ago, it was determined

that, in some instances, special format materials would be advanced for fuller processing and cataloging on an as-needed basis as users request materials not under the desired level of bibliographic control.

Further, bibliographic control is only one component of the total collections security framework. Physical security measures and preservation also provide protection for items at various levels of bibliographic control.

III. Growing Magnitude of the Collections Puts an Overwhelming Demand on Library Storage Space That Comes at a Cost (pg. 34)

Agree.

There is a significant demand for proper collections storage areas to preserve the national collection. The Library has put a priority on funding for Module 5 at Ft. Meade in its budget requests in recent years and we are confident that the Congress will provide resources for this purpose.

IV. Other National Libraries Face Similar Constraints but Have Addressed Them Differently (pg. 41)

Agree.

The report stated that the British Library and Libraries and Archives Canada have dealt with budget constraints differently than the Library of Congress. The report lists four areas of response by these two respected institutions. The Library is acting in three of these areas – digital preference, partnerships, and storage. The Library does not believe that the fourth area, changing collections development policies, is a responsible option for this institution.

The British Library and Libraries and Archives Canada acted in the face of significant budget reductions over a number of years, with the expectation of more reductions coming. While the Library's budget has been reduced, we do not believe that a universal collection policy is no longer sustainable. As mentioned above, such a restriction of collections will fundamentally change the nature of the Library. We need to do the best we can with the resources we now have. The fact that other national libraries are reducing their collections scope makes it all the more important that the Library of Congress continue fulfilling its unique mission.

Thank you for the opportunity to provide comments.