NATIONAL ARCHIVES

OFFICE *of* INSPECTOR GENERAL

Date : July 3, 2014

Reply to
Attn of : Office of Inspector General (OIG)

Subject : Audit Report No. 14-12: Audit of Selected Aspects of NARA's Digitization Program

To : David S. Ferriero, Archivist of the United States

Enclosed for your information is the subject report. The report contains 11 recommendations. In your response to the report, you concurred with each of the recommendations. Your comments can be found in full in Appendix C of this report.

In accordance with the Supplement to NARA Directive 1201, Audits of NARA Programs and Operations, section S7.m., your written response in the form of a management action plan to the recommendations of this report should be forwarded to our office within 45 days.

As with all OIG products, we will determine what information is publically posted on our website from this audit report. Should you or management have any redaction suggestions based on FOIA exemptions, please submit them to my counsel within one week from the date of this letter. Should we receive no response from you or management by this timeframe, we will interpret that as confirmation NARA does not desire any redactions to the posted report.

Should you have any questions concerning the report and recommendations, please contact me at (301) 837-3000.

JAMES SPRINGS
Acting Inspector General

Enclosure: OIG Report No. 14-12

**Audit of Selected Aspects of
NARA's Digitization Program**

OIG Audit Report No. 14-12

July 03, 2014

Table of Contents

Executive Summary

The National Archives and Records Administration (NARA) Office of Inspector General (OIG) completed an audit of selected aspects of NARA's Digitization Program. NARA's 2009 Strategic Plan states "In a democracy, the records of the Government belong to the citizens, and providing access to them is a vital obligation." NARA has established "public access" as its core purpose, with an objective to make all records available to the public in digital form to ensure that anyone can explore, discover, and learn from NARA's holdings. Similar digitization objectives have been a part of NARA's Strategic Plan for the past eight years.

NARA's Digitization Strategy and Policy have not been updated in over five years, despite significant management changes and NARA's increased reliance upon its Digitization Program to support the agency's overall mission and strategic goals related to public access. Further, although NARA's Digitization Strategy states a combination of approaches will be used in digitizing and making holdings available online, management has yet to fully and adequately implement all of the identified approaches. Finally, although NARA has reported progress in its efforts to digitize traditional records, this reported progress does not fully and transparently reflect the quantity of records in which NARA has made accessible online for public use.

Without an updated Digitization Strategy and Policy in place, there is an increased risk NARA's Digitization Program will not fully and efficiently align with and meet its objectives related to NARA's strategic digitization goals. Further, nearly six years after NARA identified and initiated its digitization approaches, the agency's Digitization Program is at risk of stagnated or declined progress at a time when NARA's mission is ever more closely tied to the Digitization Program's successful outcome. In addition, NARA's progress reporting submitted to Congress and the public inflates the quantity of traditional records NARA's Digitization Program has made accessible anywhere, anytime.

This report makes 11 recommendations to strengthen the management, accountability, and oversight of NARA's Digitization Program. In addition, we intend to conduct follow-on reviews of digitization areas identified during this audit.[1]

---

[1] See Appendix B for a list of future digitization audits.

Background

The National Archives and Records Administration (NARA), as an agency, is responsible for preserving and providing access to the U.S. Government's collection of documents recording the important events in American History. The archival holdings number more than 10 billion pages of unique documents. Many holdings are handwritten, and include formats such as maps, charts, aerial and still photographs, artifacts, and motion picture, sound, and video recordings.

Over the past decade, NARA has consistently identified "access" as a key initiative of the agency. A component of this initiative includes the digitization of NARA's traditional holdings. NARA's 2000 Strategic Plan recognized the American public "increasingly expects immediate electronic access to information at no cost." Further, NARA's 2006 Strategic Plan incorporated its digitization goals, stating NARA "will provide prompt, easy, and secure access to our holdings anytime, anywhere." In addition, NARA's Open Government Plan established a *flagship initiative* to "develop online services to meet our 21st Century needs," which includes focus on "a strategic approach to digitization."

In May 2008 NARA issued its Strategy for Digitizing Archival Materials for Public Access. This Strategy addressed NARA's efforts to digitize and make available historical documents to the public online. Further, during this timeframe, NARA established performance targets to measure its digitization progress, the first being the goal of having 1 percent of traditional holdings available online by 2012. NARA recognized efforts to digitize historical documents and make them available to the public online are of paramount importance for the agency, but are also a massive, complex, and expensive undertaking.

The Office of Innovation currently manages NARA's Digitization Program. The office, led by NARA's Chief Innovation Officer, oversees the agency's Open Government and Digital Government Strategy efforts, the online public catalog, Digitization Strategy, and Digitization Partnerships. In addition, the Office of Innovation runs and coordinates NARA's internal digitization labs.

Objectives, Scope, Methodology

The overall objective of this audit was to evaluate NARA's Digitization Program, and determine whether management controls are adequate to address agency and customer needs and ensure greater access to NARA holdings. Specifically, we assessed the implementation of NARA's policies and controls related to its digitization efforts. Our audit steps focused on NARA's digitization approach, management, partnerships, and progress.

To accomplish our objective, we reviewed NARA's Strategy for Digitizing Archival Materials; NARA Directive 816, "Digitization Activities for Enhanced Access," March 2004; Digitization Standards at the National Archives and Records Administration, April 2013; Technical Guidelines for Digitizing Archival Materials for Electronic Access, June 2004; OMB M-10-06, "Open Government Directive," December 2009; OMB Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies, February 2002; as well as Executive Orders and other requirements applicable to NARA's Digitization Program.

We interviewed NARA personnel from the Office of Innovation and Research Services responsible for managing, overseeing, and implementing NARA's Digitization Program. We reviewed NARA's current and previous Strategic Plans, Performance and Accountability Reports, Open Government Plans, and Annual Performance Plans; Partnership Agreements, Working Group Charter, Performance Measurement and Reporting System inputs, NARA's intranet and public-facing webpages, and other applicable documents and information related to NARA's Digitization Program.

Our audit work was performed at Archives II between February 2013 and April 2014 (with a seven month postponement due to other audit priorities). We conducted this performance audit in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Audit Results

1. NARA's Digitization Strategy and Policy

NARA's Digitization Strategy and Policy have not been updated in over five years, despite significant management changes and NARA's increased reliance upon its Digitization Program to support the agency's overall mission and strategic goals related to public access. This occurred due to a lack of clear accountability and prioritization to the strategic planning and management efforts of NARA's Digitization Program. Without an updated strategy and policy in place, there is an increased risk NARA's Digitization Program will not fully and efficiently align with and meet its objectives related to NARA's strategic goal of providing "prompt, easy, and secure access to [its] holdings anywhere, anytime."

Over the years, NARA's Strategic Plan and updates have consistently identified objectives related to providing online access to NARA's traditional records holdings. These objectives were echoed in a variety of reports and plans developed by NARA for the public, Congress, and President. In NARA's 2000 Strategic Plan, entitled "Ready Access to Essential Evidence," the agency recognizes "users increasingly expect immediate electronic access to information at no cost." The Strategic Plan states the agency is increasing partnerships with external entities to take advantage of new technological means to bring NARA holdings to people no matter where their location. As described below, future updates to NARA's Strategic Plan incorporate these digitization assertions into the agency's overall mission and strategic objectives and goals.

NARA's Strategic Plan, 2006-2016, "Preserving the Past to Protect the Future," establishes a strategic goal that corresponds to the public access component of NARA's mission statement, stating "we will provide prompt, easy, and secure access to our holdings anywhere, anytime." This strategic goal—as it relates to NARA's Digitization Program—includes making digital copies of selected non-electronic records available online.

NARA's 2009 Strategic Plan revision states "despite changes to our Government, NARA's core mission remains—we preserve, process, and provide access to the records of our Government. In a democracy, the records of the Government belong to its citizens, and providing access to them is a vital obligation." The 2009 Strategic Plan reiterates continued growth in the public expectation of easy online access to NARA's

holdings; adding NARA "must refocus attention and resources on making as many records as possible accessible to the public."

Similarly, NARA's most recent Strategic Plan (issued in March 2014) states the agency's mission is "to provide public access to Federal Government records in NARA's custody and control." The 2014 Strategic Plan's vision statement includes collaborating with other Federal agencies, the private sector, and the public to deliver records when, where, and how they are needed. Within the 2014 Strategic Plan, NARA includes the strategic goal "Make Access Happen," establishing "public access" as NARA's core purpose. This strategic goal—as it relates to NARA's Digitization Program—includes the following objective and initiatives:

- Make all records available to the public in digital form to ensure that anyone can explore, discover, and learn from NARA holdings;

- Digitize all analog archival records to make them available online; and

- Accelerate processing of analog and digital records to quickly make NARA records available to the public.

Along with NARA's Strategic Plans over the years, NARA has also included aspects of its Digitization Program in the agency's response to the President's Open Government Directive. In NARA's initial Open Government Plan, dated 25 June 2010, the agency states it intends to move NARA toward increased online participation and collaboration with the public, in part, by approaching digitization strategically as well as transparently with the ultimate goal of providing greater access to NARA holdings online. Further, the 2010 Plan states "because the American public expects online access to records, digitization has become a fundamental part of [NARA's] business." In addition, the 2010 Plan states efforts to digitize historical documents and make them available to the public online are of paramount importance for NARA, but are also a massive, complex, and expensive undertaking.

Similarly, NARA's updated Open Government Plan, 2012-2014 states NARA provides innovative digital access to the records of the Federal Government and develops strategies that support greater online access to these records. Further, NARA's updated Open Government Plan's Flagship Initiative—as it relates to NARA's Digitization Program—includes updating NARA's digitization strategy to expand efforts to digitize records. In addition, the updated plan cites NARA's strategic goal to provide prompt, easy, and secure access to its holdings anywhere, anytime—stating this goal, in part, will enable NARA to fulfill its mission to "serve American democracy by safeguarding and

preserving the records of our Government, ensuring that the people can discover, use, and learn from this documentary heritage."

Current Status of NARA's Digitization Strategy and Policy

In May 2008, NARA issued its "Strategy for Digitizing Archival Materials for Public Access" (Digitization Strategy) in order to address NARA's efforts to digitize and make available historical documents to the public online. The Digitization Strategy states NARA expects the strategy to undergo revision given the fast pace of technical change and the broadening of NARA's own experience. Accordingly, the Digitization Strategy states NARA will review and update the document at least annually. However, despite major changes and events related to NARA's Digitization Program—including those detailed below—NARA has not updated its Digitization Strategy since its original issuance in 2008.

Similarly, NARA Directive 816, "Digitizing Activities for Enhanced Access"— establishing the criteria and guidance for digitizing activities and for the approval and implementation of digitization projects—has not been revised since its issuance in March 2004. Directive 816 establishes roles and responsibilities for NARA's Digitization Program, as well as digitization project coordination procedures and controls. However, as further described below, NARA underwent a significant "Transformation" starting in 2010, and formed the Office of Innovation in October 2012. These changes impacted many of the Digitization Program roles and responsibilities established in the March 2004 Directive. The following paragraphs provide examples of major changes and events impacting NARA's Digitization Program in the time since its Digitization Strategy and Directive were last issued.

NARA Transformation. In 2010, NARA began a multi-year internal transformation process to better align the agency with its customers, and increase NARA's efficiency and effectiveness. This process was described as a "complete transformation of the agency." Further, the transformation includes a new organizational structure intended to eliminate stovepipes and duplicate processes. This new organizational structure resulted in the establishment of the Office of Innovation and the Chief Innovation Officer, responsible for overseeing NARA's digitization strategy and efforts.

Digitization Services Reorganization. In August 2013, the Office of Innovation underwent an additional reorganization within its Digitization Services Branch. The intent was to allow more flexibility for a wide range of digital access projects and to focus on collaborating with other offices to ensure digital copies created in the labs are provided to the public through online access.

Digitization Strategies. Although NARA's Performance and Accountability Reports and Annual Performance Plans have been updated over the past four years to include revised approaches to NARA's major digitization initiatives, these revisions were not made to NARA's Digitization Strategy or policy.

Digitization Partnerships. Many of NARA's digitization partnership agreements[2]— which account for the vast majority of the traditional records digitized—have surpassed their initial five year terms. These agreements include stipulations surrounding the number of years until NARA is granted full rights to the records digitized by partners and the metadata created. Such stipulations have not been revised to reflect lessons learned from the time NARA first entered into such agreements.

Online Catalog/Public Access System. Major changes have taken place in terms of the systems NARA uses to catalog and provide online public access to its traditional records. At the time NARA's Digitization Strategy and Policy were last issued, the agency used the Archival Research Catalog (ARC) system to provide online access to the descriptions of records held by the National Archives. This functionality was later replaced by NARA's Description and Authorities Service and Online Public Access systems.

Digitization Standards. Over the years, NARA has redefined its standards related to records digitization. Most recently, in February 2013, the Office of Innovation initiated a project to develop a set of standards to be used by all entities digitizing NARA records in an official capacity. Although such standards impact a number of NARA's digitization approaches, the agency's Digitization Strategy and Policy have not been updated to reflect these changes.

NARA's FY 2012 Annual Performance Plan—under the goal of providing online access to archival holdings—included the objective to develop a new Digitization Strategy. However, despite this objective—and the major changes and occurrences impacting NARA's Digitization Program over the past five years—NARA has not updated its Digitization Strategy and Policy since 2008 and 2004, respectively.

Management of the Digitization Program

NARA's 2010 Open Government Plan states moving forward, it will continue to implement its Digitization Strategy as well as additional efforts, including convening an

---

[2] NARA establishes partnership agreements with organizations from a variety of sectors (private, public, non-profit, educational, Government) to digitize NARA's holdings. These formal agreements stipulate the terms and conditions of the partnership.

agency-level digitization working group to develop strategies on best approaches for streamlining and leveraging current processes. Further, NARA's Open Government Plan states "the goal will be to increase access to more digital copies of holdings online. This working group will develop a partnership digitization registry as well as an in-house digitization registry to track all progress being made in this area, with the goal to make these registries available online to the public."

NARA established this Digitization Working Group on September 27, 2010. The Charter for this working group included a deliverable—in addition to those mentioned in the Open Government Plan above—to draft a new digitization policy for partnership agreements and digitization projects. However, prior to completing this and other deliverables, the Digitization Working Group was disbanded in July 2012.

The management of NARA's Digitization Program currently falls under the Office of Innovation. This office is responsible for creating "innovative ways to serve [NARA] customers and to increase access to and delivery of records through all forms of media." Further, the Office of Innovation oversees NARA's Digitization Strategy and is led by the Chief Innovation Officer.

During a February 2013 interview with Office of Innovation personnel, the Chief Innovation Officer acknowledged the Digitization Strategy should have been updated on an annual basis. She stated she was dealing with multiple competing priorities within the Office of Innovation. However, the Chief Innovation Officer stated she was in the process of establishing a Digitization Governance Board to replace the disbanded Digitization Working Group. The Chief Innovation Officer stated the Digitization Governance Board's first task would be to develop an updated Digitization Strategy. During a more recent follow-up interview, the Chief Innovation Officer stated she was still in the process of completing the Charter to establish the Digitization Governance Board.

Despite identifying the expectation of the public to have greater online access to its traditional holdings more than 15 years ago—and an ever increasing focus on such efforts in subsequent Strategic and Open Government plans—NARA has not developed its digitization policy and strategy to reflect an effort the Open Government Plan calls "of paramount importance." In NARA's most recent Open Government Plan, the Archivist states "we are improving online access to our records and we will revise our strategy to digitize records to provide the online access [NARA customers] expect." This is further reflected in the plan's flagship initiative, which includes updating NARA's digitization strategy to expand efforts to digitize records. However, as noted earlier, similar objectives related to updating NARA's Digitization Strategy and Policy were not

accomplished in the past. Therefore, increased focus and accountability are needed to ensure these objectives are achieved going forward.

Conclusion

Although NARA has consistently placed high priority on digitization efforts in its Strategic Plan and Open Government Plan, this has not resulted in the update of NARA's Digitization Strategy or Policy Directive in a timely manner. Both NARA and the Office of Innovation have undergone reorganizations without policies and strategies in place that accurately reflected NARA's Digitization Program. Without an updated Digitization Strategy and Policy in place, there is an increased risk NARA's Digitization Program will not fully align with and meet its objectives related to NARA's strategic goal of providing "prompt, easy, and secure access to [its] holdings anywhere, anytime."

Recommendations

We recommend the Chief Innovation Officer:

1. Establish a timeframe for creating the Digitization Governance Board.
2. Establish a timeframe for completing an updated Digitization Strategy and Policy.
3. Develop a process to track progress of Digitization Governance Board efforts in reviewing and updating partnership agreements, establishing Digitization Program roles and responsibilities, and drafting a digitization plan that establishes criteria for prioritizing which holdings NARA makes available online for public use.

Management Response

Management concurred with the recommendations.

2. Digitization Strategy Approaches

Although NARA's Digitization Strategy states a combination of approaches will be used in digitizing and making holdings available online, management has yet to fully and adequately implement all of the identified approaches. This condition exists due to a lack of adequate performance accountability and transparency, as well as management not properly identifying and obtaining the resources necessary to facilitate these initiatives in a timely manner. As a result, nearly six years after identifying and initiating these digitization approaches, NARA's Digitization Program is at risk of stagnated or declined progress at a time when the agency's mission is ever more closely tied to the Digitization Program's successful outcome.

NARA's most recent Digitization Strategy, dated May 2008, identifies the following five approaches for digitizing and making holdings available online:

1. *Digitization Partnerships*. NARA will establish partnerships with organizations from a variety of sectors (private, public, non-profit, educational, Government) to digitize and make available holdings.

2. *Gather Previously Digitized Material*. NARA will gather and make available on the web archival materials that have already been digitized in the course of performing agency functions, but for one reason or another are not yet available online.

3. *In-house Digitization Efforts*. NARA will conduct digitizing projects on its own with materials that are not appropriate for partnerships. For example, "treasure vault," or at-risk material that only NARA can handle, or high-interest materials for which no partner can be found.

4. *Preservation Reformatting*. NARA will pursue digitization of archival materials as part of its preservation reformatting approach.

5. *Online Catalog*. NARA will continue to make its online catalog a hub for discovering NARA's digital images.

More recently, starting in FY 2010, NARA's Performance and Accountability Reports and Annual Performance Plans identify four major strategies NARA states it is engaging in to increase the amount of archival material the agency makes available online. These strategies are similar to the first three approaches identified in the May 2008 Digitization Strategy. The FY 2014 Annual Performance Plan removes Preservation Reformatting and Online Catalog as major digitization approaches, and adds a strategy for making electronic records available online. Of the digitization approaches mentioned above, Office of Innovation personnel stated Digitization Partnerships account for 97 percent of

digitization results at NARA, with the remaining 3 percent attributable to in-house/preservation projects. These individual digitization strategies and their ongoing progress are discussed below.

Digitization Partnerships

NARA's Digitization Partnerships—which first began in 2006—account for the vast majority of the records digitized as part of NARA's Digitization Program. However, Office of Innovation and Research Services personnel stated the annual number of images digitized by Digitization Partnerships is on the decline. This is due in part to simpler projects being done first—such as those involving microfiche scanning—which partners were able to digitize at a more efficient rate. The following graph illustrates the recent decline in the annual number of images digitized by partners based on the data provided by Office of Innovation personnel.

Annual Images Digitized by Partners
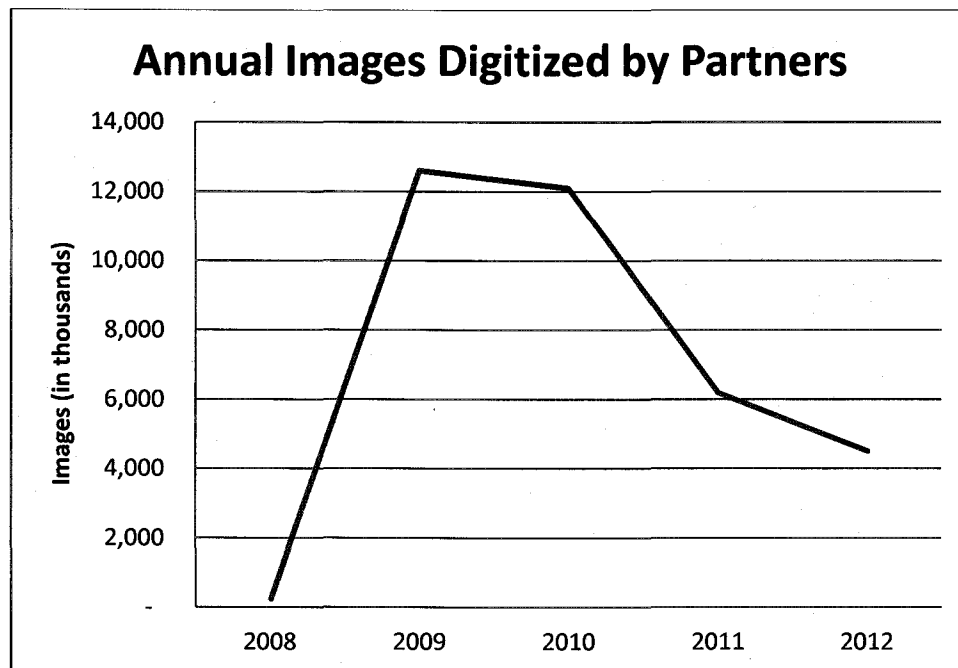
Progress has also recently declined in terms of digitized traditional records made available online per year. According to the data reported in NARA's Performance Measurement and Reporting System, NARA made an additional 19,341 cubic feet of traditional records available online in FY 2012; however, only 3,399 cubic feet were added the following year in FY 2013—representing a decline of 82 percent. At the end of FY 2013, NARA reported a combined total of 56,921 cubic feet of traditional records available online since tracking began in FY 2008—this represents 1.2 percent of the 4.6 million cubic feet in total traditional records holdings at NARA in FY 2013. However, this 1.2 percent figure is made up mostly of "embargoed" partnership records, which in

order to access requires physically visiting a NARA facility or purchasing a membership to the partnership website hosting the embargoed records.

Language within each of NARA's digitization partnership agreements state for a period of five years after the donation of digitized materials by the partnership, NARA will not sell or otherwise make available for downloading the entire contents of digitized materials or a major file segment thereof. Furthermore, NARA does not begin reducing the partnership's five-year embargo period on the digitized records until January 1st of the year following the digitization project's completion. Therefore, if a partnership completes a digitization project in January of a given year, the embargo actually lasts the remaining 11 months of that year plus an additional 5 year embargo period going forward. As a result, even though records have been digitized through partnerships, these records are not accessible online through NARA's public access system until over 5 years later[3].

Although NARA's Performance and Accountability Reports state NARA continually strives to make its holdings accessible to the public as soon as possible, this does not appear to be reflective in the embargo process NARA has implemented. Office of Innovation personnel stated internal discussions have taken place to reduce the embargo periods—this was also noted in the meeting minutes of the September 12, 2012, Executive Leadership Team meeting. However, such embargo period reductions have yet to take place, despite the end of the initial five-year terms of the agreements and their ongoing one-year renewals.

In addition, NARA has not updated its initial five year-term digitization partnership agreements to incorporate lessons learned or more favorable terms despite a number of the agreements dating back to 2007. As noted previously, NARA established a Digitization Working Group in September 2010, which was tasked in part to draft new digitization policy for partnership agreements and explore ways to maximize the benefit from partnership agreements. However, the Digitization Working Group was disbanded in July 2012 without updating NARA's digitization partnership policy or making improvements to the initial digitization partnership agreements.

Further, NARA's public facing webpage "Digitization at the National Archives," citing the 2006-2016 Strategic Plan, "Preserving the Past to Protect the Future," states the agency will work to digitize selected records, including those most requested by researchers. The Chief Innovation Officer stated NARA's partners are in the business of knowing what researchers want; therefore the records partnerships digitize meet this

---

[3] Office of Strategy and Communication personnel stated NARA's public access system provides links to embargoed items hosted on the Partnership websites at the broader series-level description. When these links are selected a notice appears informing the user they are leaving NARA's website and the agency is not responsible for costs incurred while using the Partnership website.

requisite. According to Research Services personnel, the records digitized by partnerships—which focus primarily on genealogy and account for the vast majority of traditional records digitized at NARA—do not necessarily reflect those most requested by NARA's researchers, as NARA's customer base is not exclusively focused on genealogy.

Gather Previously Digitized Material

As with NARA's 2008 Digitization Strategy, NARA's FY 2014 Annual Performance Plan also lists "gathering existing digital copies of traditional archival material and make them available online" as one of the major strategies used to increase the amount of archival material NARA makes available online. The FY 2014 Performance Plan states NARA has digitized a large number of high interest documents for exhibits, many of which have been described and placed online.

NARA's Digital Public Access Branch Chief[4]—who was responsible for the project when it began—stated this was NARA's initial effort in 2007 to get records online and included in the agency's research catalog. The Branch Chief stated early on she had conducted a "digital harvest" in which she canvassed the agency to determine what records had already been digitized. However, according to the Branch Chief, she soon discovered the previously digitized records were of varying levels of quality, both in terms of image resolution and metadata completeness. As a result some of the previously digitized images could not be used.

Although NARA had developed Technical Guidelines for Digitizing Archival Materials for Electronic Access in June 2004, the digitized records gathered, as noted above, did not consistently meet these standards. More recently, in February 2013, the Office of Innovation's Digitization Planning Branch initiated a project to "create an easy-to-understand, widely-available, practical, flexible, and policy-backed set of standardized digitization products to be used by all entities digitizing NARA records in an official capacity." Further, this agency-wide approach was intended to help NARA "make the most of its digitization efforts by defining appropriate technical specifications for a variety of intended end uses, and by creating consistent products around which more efficient production workflows and tools could be built." However, the Digitization Planning Branch Chief announced this Digitization Standards Project was put on hold in May 2013 "due to the departure of key staff and priority changes."

The "Digitization at the National Archives" section of NARA's public facing website includes a webpage entitled "NARA Existing Digital Copies to be Gathered and Made Available Online." This page lists the respective titles, repositories, and digital file

---

[4] The Digital Public Access Branch Chief responsible for this project left NARA in January 2014.

quantities of those records identified—before and after August 2007—for digitization and online availability. Further, hyperlinks are included for records on the list that have now been loaded into NARA's public access system. However, based on the information contained on NARA's webpage, only 35 of the 127 titles and repositories identified on list have been made available online. Further, this represents 11,541 of the 490,775 digital files listed—or approximately 2 percent of the total files identified to be gathered and made available online.

When asked about this list, the Chief Innovation Officer stated it should probably be removed from NARA's Digitization website, as her staff was no longer pursuing these efforts in this capacity. Although NARA's FY 2014 Annual Performance Plan continues to list "gathering existing digital copies of traditional archival material" as one of its major strategies, Office of Innovation personnel stated following the difficulty experienced in 2007, this strategy now only involves individual small scale digitization projects from the field. The Chief Innovation Officer stated there is a person on her staff who works to get "one off documents" into NARA's online access system by helping custodial units write up descriptions of the image. Further, the Chief Innovation Officer stated the digitized records coming from the field in this manner represent a very small amount of the records in the digitized population—however; it is often the "really good stuff."

In-House Digitization Efforts

NARA's 2009 Strategic Plan states "for the broadest possible access, we will make digital copies of selected non-electronic records available online. We will identify these holdings in a digitization plan that sets priorities for putting these holdings online." Similarly, NARA's 2010 Open Government Plan states NARA "will convene an agency-level Digitization Working Group to develop strategies on best approaches to streamlining and leveraging current processes. The goal will be to increase access to more digital copies of holdings online." Further, the plan states the Digitization Working Group will develop a registry to track all in-house digitization progress, with the goal to make these registries available online to the public.

Although NARA has developed a registry for records digitized by partnerships, NARA does not currently have a similar process for identifying and tracking records digitized in-house. Further, NARA has not established a functioning process for proposing in-house digitization projects. Although NARA's "Digitization at the National Archives" includes a link to an "In-House Project Submission Form," the webpage states the form cannot be found. In addition, according to Office of Innovation and Research Services personnel, NARA has not fully developed a digitization plan setting priorities for putting holdings online.

NARA's FY 2014 Annual Performance Plan, under its strategic goal to "provide prompt, easy, and secure access to our holdings anywhere, anytime," lists the following milestones relating to digitizing of traditional records:

- NARA Digitization Working Group to identify improvements to digitization at NARA established (FY 2010)

- Digitization of 1940 Census completed (FY 2011)

- Digitization and online availability of Elena Kagan records at the Clinton Library completed (FY 2011)

- High-use materials from Roosevelt, Kennedy, Nixon, Ford, and Clinton Libraries digitized and made available (FY 2012)

As evident by the milestones reported, NARA has conducted a number of large in-house digitization projects since 2010. For instance, the 1940 Census involved the creation of 3.8 million digitized images. However, according to Office of Innovation personnel, the progress of in-house projects is not calculated separately from the efforts conducted by NARA's digitization partnerships. This is reflective in NARA's FY 2012 Performance and Accountability Report, which states "NARA engages in internal digitization projects and projects with partners to increase the numbers of archival holdings accessible online. For several years we have worked to reach a 2012 goal of achieving 1 percent of archival holdings accessible online." The metric used for measuring this in the report is the "percent of traditional archival holdings available online," which includes both in-house efforts and partnerships.

As noted above, the 1940 Census was one of NARA's major digitization milestone accomplishments. According to data requested from Office of Innovation personnel, the 3.8 million digital images were created from 4,645 rolls of microfilm and represent 69.7 cubic feet of records—to put this in perspective, partnership efforts resulted in approximately 19,270 cubic feet of digitized records that same reporting year. Office of Innovation personnel stated internal and preservation efforts represent 3 percent of the total digitized traditional records made available online. Based on this information, as of FY 2013, NARA's internal digitization efforts have resulted in approximately 0.037 percent of its total traditional records becoming available online.

However, there are also digitization activities taking place at NARA that do not end up online, and therefore are not reflected in this measurement. One example includes much of the digitization efforts in Office of Innovation's Special Media Labs. The archival

records digitized in the Special Media Labs are primarily textual microfilm, still pictures, audio/video, and motion picture. According to Office of Innovation personnel, much of this special media is restricted due to copyright, classification, and legislative concerns and therefore cannot be made fully accessible to the public online. According to Office of Innovation personnel, they have not yet developed a process to measure and track the number of records digitized in the labs in comparison to the special media that ultimately is made accessible online to the public.

An additional concern related to the Special Media Lab in-house digitization efforts pertains to the data storage needed to perform the digitization activities. Concerns related to a shortage of storage space were highlighted in a 28 October 2013 briefing to the Architectural Review Board, indicating only 24 terabytes out of an allotted 363 terabytes of free space remained for in-house special media digitization. The Chief Innovation Officer stated she has been trying to obtain additional storage space from Information Services for the last two years without success. The Chief Innovation Officer stated if circumstances do not change, storage shortages will ultimately hamper their in-house digitization efforts.

Lastly, although each of NARA's Annual Performance Plans since FY 2007 has stated NARA will "explore innovative ways to increase our own capacity to digitize selections of our holdings," Office of Innovation personnel were unable to provide examples of such innovations or provide measurements of increased progress over that timeframe. Further, personnel from Research Services stated NARA currently does not have near the capacity necessary to digitize its holdings in-house.

Online Access

The last of the major digitization strategies identified by NARA involves making digital records available online through NARA's public access systems. On December 27, 2010, NARA announced the public rollout of the prototype of the Online Public Access (OPA) system. OPA is the online public portal to NARA's records and information pertaining to those records. As stated in the FY 2013 Performance and Accountability Report—through partnerships and collaborative efforts—NARA continues to increase the number of digital records available to the public through the online public access portal of its nationwide holdings. Office of Innovation personnel stated there are currently 2 million digital objects in OPA, of which 250,000 were created through digitization partnerships. The remaining 1.75 million represent digital objects that were "born digital" or digitized in-house. Although NARA reports receiving more than 90 million digital objects from its partners—due to the five-year partnership embargoes mentioned earlier—only 250,000 are currently accessible through NARA's online portal.

NARA's Performance and Accountability Report states "we continually strive to make our holdings accessible to the public as soon as possible;" however, this is not reflected in NARA's efforts to make digitized holdings available following the end of partnership embargoes. NARA has been aware for the past five years that on January 1, 2014 partnership embargoes covering 12.6 million digitized records would come to an end, allowing NARA full and unrestricted rights to the digitized records. However, NARA's pilot OPA system was not prepared to handle these records. Despite efforts by the Chief Innovation Officer to address the shortfalls of the pilot system in 2011, it was not until February 2013 that NARA began collecting requirements to re-engineer and increase the scalability of OPA in order to ensure the 12.6 million digitized records would be available online on January 1, 2014. According to the Chief Innovation Officer the production version of OPA is not expected to be completed until the summer of 2014, after which time NARA will begin making the 12.6 million records available online. However, due to the recent departure of NARA's Chief Information Officer—who was largely involved in the project—the Chief Innovation Officer stated she was less certain of the April 2014 rollout.

Conclusion

NARA's Digitization Strategy, Performance and Accountability Report, and Annual Performance Plan lists a number of major strategies NARA is engaged in to increase the amount of archival material the agency makes available online. Although NARA has been implementing these strategies for over five years, consistent progress has not been achieved. The partnership strategy accounting for 97 percent of the agency's digitized holdings is on the decline and the process to ensure the digitization of records most requested by researchers has not been fully established. Further, in-house digitization efforts are not tracked and reported separately from partnership efforts, reducing the accountability and transparency of the success of the projects that resulted in the remaining 3 percent of the digitized traditional records. In addition, the system NARA uses to make these digitized records accessible to the public is not able to scale to provide access to records that have already been digitized. As a result—nearly eight years after establishing NARA's mission to provide prompt, easy, and secure access to its holdings anywhere, anytime—NARA's digitization strategies have yet to achieve consistent, measurable results. Based upon these findings and an assessment of risk, the OIG plans to conduct follow-on audits pertaining to NARA's Digitization Program and Strategy as identified in Appendix B.

Recommendations

We recommend the Chief Innovation Officer:

4. Track and report progress on each of NARA's digitization strategies.
5. Establish performance goals specific to NARA's in-house digitization efforts.
6. Set a public timeline for when partnership records no longer embargoed become available through NARA's public access system.
7. Coordinate with Research Services to develop a process to ensure records most requested by researchers are digitized.
8. Update "Digitization at the National Archives" webpage to ensure accessibility to forms and accuracy of information reported.

Management Response

Management concurred with the recommendations.

3. Digitization Progress and Reporting

Although NARA has reported progress in its efforts to digitize traditional records, this reported progress does not fully and transparently reflect the records in which NARA has made accessible to the public online. This condition exists because Office of Innovation personnel changed the way it measured and reported progress in order to achieve a previously established performance measurement goal. As a result, NARA's progress reporting submitted to Congress and the public inflates the quantity of traditional records NARA's Digitization Program has made accessible anywhere, anytime.

NARA's Open Government Plans issued in 2010 and 2012 emphasize the importance of transparency while embracing innovation and developing best practices. NARA's 2010 Open Government plan states "although [NARA is] in the business of open government, we have the opportunity to do more to strengthen transparency, participation, and collaboration." Similarly, the 2012 Open Government Plan states NARA will seek to strengthen the principles of government and transparency by informing the public of the agency's progress. Within the context of this audit, the plan's flagship initiative includes seeking to update NARA's digitization strategy and expand the agency's efforts to digitize records.

NARA's Strategic Plans, Performance and Accountability Reports, and organizational goals identify and promote the digitization and access of NARA's archival holdings as a priority of the agency. This digitization goal was first quantified in NARA's FY 2008 Performance and Accountability Report with an objective to have 1 percent of archival holdings available online by 2012. NARA states:

> Here, we directly measure the extent to which our holdings have been put online. These records include both electronic records and non-electronic records that have been digitized, either by us or by our partners. By online we mean that the public can retrieve copies over the Internet without an interaction with NARA staff, and the archival context of these records is available in our online catalog. Not counted here are the electronic holdings that we will provide upon request and are therefore not, strictly speaking, online.

More recently, NARA established a long range performance target of 1.4 percent of traditional holdings available online by 2016.

Digitization Progress Measurement

In order to track NARA's digitization progress, the FY 2008 Performance and Accountability Report included an objective to develop a measurement methodology for the number of archival holdings accessible online. According to the FY 2008 Performance and Accountability Report, NARA "developed a methodology to measure the number of archival holdings accessible online" that same year. Office of Innovation personnel described this methodology, which includes tallying the total cubic feet of the traditional records digitized and comparing that to the total cubic feet of NARA's traditional holdings.

In the first two years this measurement methodology was put into place, the following information was reported in NARA's FY 2009 Performance and Accountability Report and Performance Measurement and Reporting System:

	FY 2008	FY 2009
Traditional Records - Total Cubic Feet	3,728,601	3,937,316
Traditional Records Available Online (Cu. Ft.)	1,406	1,634
Percent of Traditional Records Available Online	0.038%	0.042%

The FY 2009 Performance and Accountability Report states an objective for that year was to meet 20 percent of the 2012 (1 percent) target for archival holdings accessible online—this would represent 0.20 percent of traditional records being made available online. With only 0.042 percent of records reported being available online in FY 2009, this objective was not met.

In addition, NARA's FY 2010 Annual Performance Plan (issued on 8 March 2010), states "to date, more than 60 million records are available online through our partners." However, according to Office of Innovation personnel, partners digitized the following number of images[5] during the first three years:

Year Digitized by Partner	Number of Images
2008 (no longer embargoed after 2013)	245,000
2009 (no longer embargoed after 2014)	12,600,000
2010 (no longer embargoed after 2015)	12,100,000
Total	**24,945,000**

Therefore, when NARA reported 60 million records available online through its partners in March 2010, fewer than half had actually been digitized. Four years later, in NARA's

---

[5] The Chief Innovation Officer explained that Digital Records can be made of multiple images; therefore the number of images reported may actually reflect even fewer records.

FY 2014 Annual Performance Plan, the agency was still reporting this 60 million record figure.

Further—as noted previously, and shown in the chart above—digital images of NARA's records created through Digitization Partnerships are embargoed for over five years before NARA has unrestricted rights to place them online. Initially, NARA did not include these embargoed digitized records in the performance measurement calculations in FY 2008 and FY 2009, as they were not available online through NARA's public access system—only through physically visiting a NARA facility or by purchasing a membership to a partnership website.

Office of Innovation personnel stated the factors used to measure online availability of traditional records have always remained consistent. However, starting in FY 2010—as reflected in NARA's annual performance and accountability reporting—NARA began including partnership embargoed digitized records in its performance measurement calculations. Therefore, the 24,945,000 partnership embargoed images created between 2008 and 2010—which represents approximately 23,015 cubic feet of records—were added to the reported traditional records available online in the FY 2010 Performance and Accountability Report. As highlighted in the chart below, by adding the partnership embargoed records from FY 2010 and the previous two years to the number of traditional records NARA reported as being available online in FY 2010, NARA was able to report an increase from 0.042 percent to 0.610 percent in traditional records available online from the previous year. This represents a 15 fold increase over the previous year.

Year	Traditional Records Available Online (Cubic Feet)	Percent of Traditional Records Available Online
2008	1,406	0.038%
2009	1,634	0.042%
2010	24,649	0.610%
2011	34,181	0.805%
2012	53,522	1.194%
2013	56,921	1.232%

As a result of changing its reporting methodology to include partnership embargoed records, NARA reported in its FY 2012 Performance and Accountability Report that the agency had exceeded its goal of 1 percent of traditional records online. However, the report failed to clarify only a small fraction of these traditional records could be accessed through NARA's public access system.

As noted earlier in this report, Digitization Partnerships account for 97 percent of digitization efforts at NARA, with the remaining 3 percent attributable to in-house

efforts. This corresponds with data collected regarding in-house digitization efforts in comparison to Performance Measurement and Reporting System inputs. Therefore, had NARA remained consistent in reporting traditional records available online only once the records were available through NARA's public access system and no longer under the partnership embargoes, the agency would have reported approximately 0.04 percent of traditional records available online in the FY 2012 Performance and Accountability Report—well below the agency's 1 percent objective for that year. The Chief Innovation Officer also questioned the value of this performance measurement when it includes embargoed records, stating it provides no incentive for the agency to work toward making the records available through NARA's public access system.

This lack of incentive is reflective of the Digitization Program's current status. As noted previously, although NARA was aware for the past 5 years the embargo of over 12.6 million partnership digitized images would end on 1 January 2014, the agency did not scale its public access system in time to handle this known influx of records in order to make them available online to the public. However, this is not apparent in NARA's annual performance and accountability reporting, as these records were already reported as being available online back in FY 2010 when NARA changed its calculation methodology.

Conclusion

NARA's updated Open Government Plan highlights the importance of transparency and gives focus to NARA's strategic approach to digitization. Further, the plan states NARA "will seek to strengthen the principles of open government and transparency by...informing the public of our progress." However, the information NARA reports pertaining to the online accessibility of its traditional holdings does not fully embrace these initiatives. Specifically, NARA has used an inconsistent methodology to report on its digitization progress. This resulted in the Office of Innovation prematurely reporting digitization and accessibility accomplishments prior to being achieved.

Management Comments to the Finding

In the Archivist's response to the audit report, he stated "the Office of Innovation did not prematurely report on digitization and accessibility accomplishments prior to being achieved." The Archivist stated the increase in 2010 was not due to suddenly including records hosted by digitization partners, but instead "came from a deliberate change in measurement." The Archivist's complete comments are included in Appendix C.

Audit Response

As highlighted throughout this report, NARA has emphasized the importance of transparency in its Open Government Plans. During the audit, the Office of Innovation personnel responsible for reporting NARA's digitization efforts consistently stated embargoed partnership digitized records were not initially included in calculating records available online. As the 2012 goal of making one percent of traditional archival holdings available online approached, NARA management made the decision to include digitized records externally hosted on partnership fee-based websites. This management decision resulted in a 15 fold increase in online digitized records reported from the previous year. However, no mention of the changed methodology was included in NARA's Performance and Accountability Report or in the notes to NARA's internal Performance Measurement and Reporting System.

NARA's Chief Innovation Officer stated the revised methodology provides no incentive for the agency to work toward making the records available through NARA's public access system. NARA's *Principles for Partnerships to Digitize Archival Materials* states access to the digital copies ultimately should be readily accessible and free. In the context of NARA's most recent Strategic Plan, it is our recommendation that NARA consistently and transparently measure its efforts to "make all records available to the public in digital form to ensure that anyone can explore, discover, and learn from NARA holdings."

Recommendations

We recommend the Chief Innovation Officer:

9. Develop a reporting methodology that tracks total traditional records made available through NARA's public access system.
10. Distinguish in reporting what quantity of records is available online through NARA, and what quantity is accessible online through partnership websites.
11. Ensure future Annual Performance Plans accurately reflect current data regarding traditional record availability.

Management Response

Management concurred with the recommendations.

Appendix A – Acronyms and Abbreviations

ARC	Archival Research Catalog
CINO	Chief Innovation Officer
FY	Fiscal Year
NARA	National Archives and Records Administration
OIG	Office of Inspector General
OMB	Office of Management and Budget
OPA	Online Public Access system

Appendix B – Future Digitization Audits

Based upon the findings of this report and an assessment of risk, we plan to conduct future audits in the following areas of NARA's Digitization Program:

- Partnership Agreements and Digitization Projects
- Development of the Production Version of NARA's Online Public Access System
- Digitization Services Storage and Transfer Capabilities
- NARA's In-House Digitization Efforts
- Prioritization of Records made available Online
- Authenticity of Digitized Records

Appendix C - Management's Response to the Report

NATIONAL ARCHIVES

Date: JUN 2 0 2014
To: James Springs, Acting Inspector General
From: David Ferriero, Archivist of the United States
Subject: OIG Revised Draft Audit 14-12, Audit of Selected Aspects of NARA's Digitization Program

Thank you for the opportunity to provide comments on this revised draft report. We appreciate your willingness to work with us to clarify language in this report.

We concur with all 11 recommendations in this audit, and we will address them further in our action plan.

It is important to note that the Office of Innovation did not prematurely report on digitization and accessibility accomplishments prior to being achieved. The increase in 2010 is not due to suddenly including records hosted by our partners. These had been included all along. The increase came from a deliberate change in measurement. Previously, we counted a series as available online (through NARA or our partners) only when all of its items had been individually described in ARC. In 2010, recognizing the burden associated with item level description of the increasing volume of partner-digitized records, we shifted to series-level descriptions. That change resulted in more records being counted that had been available online all along but for which we did not have a description in ARC.

In addition, the reported percentages of traditional records available online are still accurate according to the definition first provided in the FY 2008 Performance and Accountability Report, "By online we mean that the public can retrieve copies over the Internet without an interaction with NARA staff, and the archival context of these records is available in our online catalog."

If you have any questions, please contact Susan Ashtianie at susan.ashtianie@nara.gov or by phone at 301-837-1490.

DAVID S. FERRIERO
Archivist of the United States

NATIONAL ARCHIVES and
RECORDS ADMINISTRATION

8601 ADELPHI ROAD
COLLEGE PARK MD 20740-6001
www.archives.gov

Appendix D - Report Distribution List

Archivist of the United States
Deputy Archivist of the United States
Chief Operating Officer
Deputy Chief Operating Officer
Chief Innovation Officer
Executive for Research Services
Chief Information Officer
Director, Performance and Accountability