AFRICAN-AMERICAN DOCUMENTARY
RESOURCES ON THE WORLD WIDE WEB:
A SURVEY AND ANALYSIS

BY ELAINE L. WESTBROOKS

ABSTRACT: Numerous institutions have launched historical digital collections on the World Wide Web (WWW). This article describes, analyzes, and critiques 20 historical African-American digital collections created by archival institutions, academic institutions, public libraries, and U.S. government agencies. In addition, it explores issues that are an important part of historical digital collections, such as preservation, integrity, and selection criteria, as well as trends in collection content, institutional policy, technology, Web-site organization, and remote reference. Finally, this article assesses the value of individual digital collections as well as the overall value of digitization.

Introduction

Currently, there is a scarcity of printed African-American documentary resources in the United States. There is also little information about archives, historical societies, museums, and repositories whose primary goal and purpose are to collect and organize these resources. This suggests that either African-American history is poorly documented or that documented African-American history is not valuable to researchers. Neither is the case. African-American historical documents are available but not easily accessible. The trend to digitize historical collections is making historical documents more accessible; however, there is little research about digital collections of African-American history on the WWW. Though unnoticed by researchers, archivists, and librarians, dozens of institutions have digitized primary African-American documents. This paper addresses this information gap and will show how African-American historical documents are being made accessible to researchers. It will also explore the purpose of digitization.

An analysis of a sample of African-American digital collections is valuable for a number of reasons: First, to measure the digital representations of African-American historical documents from various institutional perspectives. Second, to expose similarities and differences among four institutional categories (archival institutions, academic institutions, public libraries, and U.S. government agencies), as well as the simi-
larities and differences among digital collections within the given institutional categories. Third, to measure "representations of how basic archival functions and policies are being incurred or ignored in the networked environment." Fourth, to show what formats, standards, and procedures are prevalent on the WWW. Once problems associated with Web sites and digitization projects are identified, better Web sites containing better collections will be available for researchers.

Methods

The exact number of institutions with on-line African-American historical collections is unknown. However, for the purposes of this analysis, two tools were used to retrieve as many collections as possible on the WWW. The first tool, Digital Librarian: a librarian's choice of the best of the Web, <http://www.digital-librarian.com/>, contains indexed Web pages that are updated daily by librarian Margaret Vail Anderson. Digital Librarian offers a good place to begin an exploration of which African-American collections are available on the WWW. According to Vail Anderson, "There were over 1,200 libraries of various sorts linking to Digital Librarian, as well as a number of non-profit organizations, businesses, and individuals." In addition, Digital Librarian is divided into nearly one hundred subject categories that include over 6,000 hyperlinks to on-line resources. The remaining Web sites were found via keyword searches ("African American" and "Digital") on the WWW using the second tool, AltaVista.com, <http://www.altavista.com/>. AltaVista was selected because of its efficiency and comprehensiveness. Collectively, AltaVista and Digital Librarian retrieved a total of 47 Web sites.

Selection of Web sites for inclusion in this analysis is based on three criteria: First, that the digital collection primarily consist of African-American historical documents on the WWW. Second, that the publishers of digital collections be archival, academic, public library, or U.S. governmental institutions. Third, that the digital collection consist of 20 or more different digital files, which may include SGML/HTML encoded text as well as any combination of digitized photographs, maps, video, and audio. Of 47 Web sites initially listed or retrieved by Digital Librarian and AltaVista, 20 digital collections satisfied all three criteria. It is these digital collections that are the focus of this evaluation and analysis.

The digital collections selected for review represent four broad institutional categories. The first category, Stand-Alone Archives, Museums, and Historical Societies, includes Black Archives of Mid-America (Black Archives), Kendall Whaling Museum (KWM), and Ohio Historical Society (OHS). The second category, Academic Libraries, includes ten institutions: Duke University, Howard University Moorland-Spingarn Research Center (MSRC), Michigan State University (MSU), Louisiana State University (LSU), University of Michigan, University of North Carolina (UNC) Chapel Hill, University of Pennsylvania, University of Virginia (UVA), Virginia Commonwealth University (VCU), and Wright State University (WSU). The third category, Public Libraries, includes four institutions: Kansas City Public Library (KCPL), Public Library of Charlotte and Mecklenburg County (PLCMC), the Schomburg Center for Research in Black Culture (Schomburg Center), and New Orleans Public Library
(NOPL). The final category, Governmental Institutions, includes Library of Congress (LC), National Archives and Records Administration (NARA), and National Park Service (NPS). See Appendix B for a short description and citation of these digital collections.

This article reviews and analyzes the digital collections of African-American documents created and organized by these four types of institutions according to five broad dimensions: Collection and Content; Institutional Information and Policy; Technology and Technical Notes; Design and Organization; and Information and Reference (see Appendix A). Evaluative criteria have been established to answer such questions as: Who is the audience? Is preservation a goal? Are there selection criteria? Each broad dimension is divided into smaller subdimensions that address important issues. As a result, there are 20 subdimensions. These subdimensions, though subjective and somewhat simplistic, cover a range of issues that give sufficient data to determine what each institution is doing with its Web pages and digital collections. First, Collection and Content evaluates the quantity, depth, and quality of collections as well as document integrity and collection selection criteria. The second dimension, Institutional Information and Policy, brings together important issues concerning institutional policy such as mission statements, intellectual property, and funding sources that impact the digitization process. Technology and Technical Notes identifies technology and equipment used. The Design and Organization dimension examines how well collections are presented to users. The final dimension, Information and Reference, assesses the ability of an institution to remotely gather information from its users. What follows is a digital collection content analysis and a Web-site evaluation focusing primarily on collection content as well as on the documentation that accompanies collections.

The remainder of the article is organized according to these evaluative dimensions. Summaries of similarities and differences among institutional categories in the way they digitize African-American historical documents are presented, along with summaries that highlight the similarities and differences among and within individual collections. All conclusions are based on the extent to which each collection measures up to the evaluative criteria.

Collection and Content

Quantity of Items/Documents

Thirty percent of institutions list the number of documents in their digital collections and only 12 percent indicate the percentage digitized of a given collection. In cases where numbers are not listed, documents were counted or estimated. This information is not readily available from Web sites such as Historical African-American Autographs from the Ramos Collection at KCPL whose collection was completed over two years ago. Collections consisting of several formats (audio, encoded text, images, and video) fail to indicate the quantity of each given format. The size of projects ranges from 20 photographic images mounted by MSU to over 6,000 audio, video, and textual documents mounted by UVA. Public libraries have the smallest collections, averaging approximately three hundred digital documents followed by stand-alone archives, his-
torical societies, and museums whose average is nearly four hundred. Academic institutions and government archives have the largest digital collections, averaging over nine hundred digital documents. The disparity can be attributed largely to the amount of funding an institution is able to procure. Most information centers are financially constrained by budgetary problems due to increasing serial prices, technology investments, and funding reductions. However, Congress and the National Digital Library Program, by mandate, ensure that governmental institutions (NARA, LC, and NPS) will continue to digitize their collections. Such a funding mandate sets governmental institutions apart from all others in the analysis. Of the government agencies, only LC lists numbers and percentages of items that are digitized within a collection as well as numbers to account for documents that will be digitized in the future.

Format

Of the four types of formats (images, encoded text, audio, video), images are used by 95 percent of institutions. Perhaps the visual appeal of images, their ability to enhance textual material, and the minimal viewing requirements account for the overwhelming use of images. Images are essential for on-line exhibitions and collections that must attract users, much like brochures and advertisements attract visitors. Furthermore, images are less complex than video or audio files and require less software. All Web sites use images, with the exception of VCU, whose oral history project is limited to HTML encoded text. SGML and HTML encoded texts are desirable because they are searchable. Forty-five percent of Web sites include text encoded with SGML. Sixty-six percent of government Web sites and 60 percent of the academic institutions use SGML encoded text. SGML encoded text is used by one public library, the Schomburg Center, and one archival institution, the OHS. Thirty percent of all institutions have audio files of music or oral histories/interviews, while only 15 percent have digital videos within their historical collections. The University of Pennsylvania and UVA include all formats, making them the only true multimedia collections.

With the popularity of digital collections, information professionals have the opportunity to inform users how historical documents in digital form differ from historical documents on paper. Users conducting research on the WWW must be reminded that digital images are nothing more than surrogates of original documents. Information professionals tend to take the concept of surrogacy for granted; however, it is worthy of more attention. Institutions also fail to inform users that digital reformatting eliminates artifactual information embodied in books, such as binding, color, size, and page material. As Peter Graham states in his article “New Roles for Special Collections on the Network,” in the networked environment, the user has access only to the information, not to “the physical design and construction of the book [which] provides evidence as to the way readers were intended to perceive it.” The term “reformatted” implies a trade-off. Users have access to informational content, but not to all of the physical features of the original document. Although the majority of users have no use for such artifactual evidence, information professionals digitizing historical documents could attempt to address the impact that digital reformatting has on the way materials are received and perceived on-line. On the other hand, if access continues to be the
main purpose of digitization, information professionals may not necessarily want to
dress the issues raised by Peter Graham on the collection Web site.

Collection Background and Context

Few Web sites include background information and context that are sufficient or
dquate for scholarly research. In the article “Open All Night: Using the Internet to
Improve Access to Archives,” Thomas Ruller contends that “Information made accessible on the Internet requires adequate context for the end user to understand its source, context, and authenticity .... This is particularly important for archival materials that are dependent upon their context in order to be understood.” What is the purpose of posting historical photographs or manuscripts on the WWW if users have no context? Stand-alone archives and governmental agencies included the most background information and context. As much documentation as possible should be added to collection Web sites. For example, users can benefit from knowing how the paper version was acquired, its provenance, extent, and relation to other collections inside or outside the institution. The Paul Laurence Dunbar Digital Text Collection at WSU provides the most in-depth background information of all sampled collections. This digital text collection is one of two Web sites in the sample devoted exclusively to the work of one individual. The narrow focus may account for the considerable depth of background information and context and the relationships among documents are much easier to recognize. The other digitization projects with a narrow focus and detailed background information are the Marian Anderson Collection of Photographs, 1898–1992 at the University of Pennsylvania and the Historical African-American Autographs from the Ramos Collection at KCPL. Yet, at least one broader project provides proper context as well. North American Slave Narratives, Beginnings to 1920 at UNC Chapel Hill is broad and focuses on several individuals, yet the creators managed to include valuable information about the project. In fact, in the introduction to Slave Narratives, the project coordinators include in-depth essays on the value of the project, the literary context for slave and ex-slave narratives, the historical context of slavery, and the importance of the project to the nation.

Range, Accuracy, and Depth

The range, accuracy, and depth of documents are determined by the presence of
dated images, currency of the Web pages, errors, bad links, and the interrelatedness of
documents. Range, accuracy, and depth are factors that help users to determine the
reliability, authenticity, and integrity of a collection. Fifty percent of sampled Web
sites are rated “excellent.” The government agencies and archival institutions fare extremely well in this subdimension. The range and depth of the Schomburg Center is also excellent due to the completeness, interrelatedness, and accuracy of documents. Academic institutions produced collections that were either limited and inaccurate (poor), or extensive and accurate (excellent). Cultural Heritage Initiatives for Community Outreach (CHICO) by the University of Michigan School of Information contains inaccurate information and heavily depends on external links that are no longer
accessible on the WWW. With the exception of the Schomburg Center, public libraries are below average. Charlotte-Mecklenburg Public Library and NOPL posted images lacking bibliographical information and coherent organization. Users are not given an opportunity to determine the source of photographs on display. In addition, images are accompanied by captions that do not include bibliographic information, nor do they indicate the physical location of the analog version of images. Images should be posted with informative captions and users should be able to manipulate them (zoom in, enlarge). These features make digital collections more useful and interesting to users.

Reliability, Authenticity, and Integrity

In light of the fact that historical and primary documents are the focus of this analysis, reliability, authenticity, and integrity are important factors. Before digital collections are discussed and evaluated, a definition of “reliability” is in order. According to Luciana Duranti, “reliability refers to the authority and trustworthiness of a record as evidence, the ability to stand for the facts they are about.”17 Duranti adds, “Degree of completeness and degree of control of the procedure of creation are the only two factors that determine the reliability of a record.” The majority of documents within the digital collections in this analysis lack reliability because the authority and trustworthiness of documents are questionable. University of Virginia’s The Jackson Davis Collection of African-American Educational Photographs is an exception because it posts “image tracking information” that includes the size of the scanned negatives, dates of image creation, titles of the compact disc (CD) storing digital information, and dates the CDs are burned. This information helps to document the procedures used to create, preserve, and locate digital files. As a result, users know more about the creation of documents they are viewing.

No sampled institution captured what Duranti calls “completeness.” Completeness is an elusive goal of digital libraries today and will continue to be so in the future. Authenticity is also desirable and just as elusive. A record is authentic “when it is the document that it claims to be.”18 Ensuring authenticity of digital collections means guaranteeing that documents have not been manipulated or falsified after their creation. Verifying the authenticity of digital collections is difficult. As David Bearman and Jennifer Trant have stated, “when scholars encountered original sources many physical clues assisted in establishing their authenticity. If sources are studied in the surrogate, all the questions concerning the authenticity of the original are overlaid with additional questions about the methods of representation.”19 Although institutions can embed digital watermarks into images, information professionals “are without widely understood and employed methods of assessing and establishing authenticity of digital sources.”20 No Web sites list information regarding the authentication of their collections, nor is there evidence that institutions are attempting to establish authenticity.

Related to authenticity is integrity. According to Clifford Lynch, “As this Internet becomes an increasingly critical part of . . . the historical and cultural record of our society, there is a resurgence in the interest and concern about the integrity of electronic information.”21 Integrity comes in several forms. Informational, bibliographical, and structural integrity of digital collections are essential. Informational integrity re-
fers to the consistency and constant content of an information package. As Clifford Lynch states, "When one retrieves a work from one year to the next, one should get the same content." In the dynamic networked environment where content is constantly updated, such integrity is difficult to guarantee. Bibliographical integrity is equally important. Bibliographical integrity is the consistency of reference through citation and retrieval. Every digital document should have one and only one bibliographical description. Most collections are successful in capturing either bibliographic or informational integrity, but not both. Physical, or structural, integrity refers to the degree to which digital surrogates retain physical features such as pagination, color, and size. In other words, how well does the digital version resemble the paper version?

Academic institutions are more aware of the importance of reliability, authenticity, and integrity than are other institutions. Although the high quality of academic collections suggests that reliability, authenticity, and integrity are important to academic institutions, most did not provide enough to support scholarly research. Only 35 percent have collections of sufficient depth for academic research. These are: the University of Pennsylvania, OHS, the Schomburg Center, UVA, UNC Chapel Hill, NARA, and LC. These institutions come fairly close to satisfying all criteria for integrity, authenticity, and evidence. Public libraries, on the other hand, are not successful in capturing integrity. For example, Historical African-American Autographs from the Ramos Collection at KCPL, posts images of the autograph in isolation as well as an image of the letter that contains autographs. Unfortunately, KCPL did not include more context or information about the letters. The date, subject, location, and/or accession number of the letters should be included in the on-line collection as well as a description of each letter. Collection designers appeared to have captured the physical integrity of the letters, yet there is little bibliographical integrity.

One way to be sure that digital collections will be consistently cited is to show users how to cite them. Schomburg and UVA are the only libraries that have Web pages that show users how to cite their digital collections. In fact, the Schomburg Center devotes an entire Web page to the citation of its digital collection. University of Virginia includes citation examples on a section of its "Use, Reproduction, and Publication of Materials" Web page. Providing examples helps users to feel more comfortable with the bibliographical integrity of on-line documents. Although many of the projects in the sample are able to achieve some degree of reliability, authenticity, and integrity, it is difficult to capture all three factors.

Selection Criteria

The final subdimension within Collection and Content is selection criteria. Many institutions like OHS do not include statements on selection criteria, but imply their use. The statement, "The African-American Experience in Ohio, 1850–1920 brought together from a number of individual sources specifically for this project," implies some minimal selection criteria. LC and NARA are the only institutions that include clear statements of selection criteria. Schomburg and PLCMC address selection criteria in their introductory essays, but statements lack detail. Selection criteria are important in light of the large costs involved in such digitization projects. In fact, selecting
Digitization is one of the earliest and most important steps taken in the formation of a digitization project. There is a growing literature on criteria for digital selection. It is perplexing that most institutions do not consider selection criteria important enough to post on their collection Web sites.

Institutional Information and Policy

Mission Statement

Mission statements are found in 50 percent of digital collection Web sites. Only 20 percent of academic institutions list mission statements. In many cases, archives and special collection departments do not publicly post mission statements independent of the library in which they are physically located. Independent and “freestanding” institutions that are not deeply embedded within university/public library systems are more likely to post their mission. This explains why Black Archives, OHS, and KWM have highly visible mission statements. In other words, the organizational structure and purpose might explain why some institutions (archival and governmental) are more likely to have their statements posted than are others (academic and public). The National Archives, LC, and NPS take mission statements seriously. In the case of NARA, there are separate Web pages with formal mission statements submitted and signed with the approval of the institution’s highest administrative officer. LC not only posts the mission statements on American Memory Web pages, but also posts goal statements from the National Digital Library Program.

Preservation

For most institutions preservation is not a goal. Even though institutions such as Black Archives, LC, UNC Chapel Hill, and the Schomburg Center claim to digitally preserve their historical documents, it appears that they really are not doing so. For example, the Black Archives of Mid-America claims, “this digital image database is a tool to prolong the natural life of fragile materials by reducing the physical wear and tear they receive during handling.” There are several problems with this statement. Digitization is a passive form of preservation that perpetuates a document’s existence if it is not damaged during digitization and if it is put away for safekeeping. Digitization can bring about outcomes that can degrade rather than preserve documents. The process of digitization can put additional stress on original documents because they are exposed to digitization tools such as scanners as well as prescanning processes such as debinding, which can be damaging. In addition, digitized surrogates can generate more interest in the original documents, whereby users are more likely to use the original documents, which ultimately increases wear and tear. The goal to never rely solely on the original paper document may perpetuate its existence, yet it cannot be considered digital preservation. The preservation of paper-based materials through digitization and the preservation of digital media are two different discussions. The potential for a digital version of a document to serve as a preservation copy is highly
questionable. According to Caroline Armes, "Traditionally, preservation of content has focused on creating a facsimile, as faithful a copy of the original as feasible, on a long-lasting medium."

Digital libraries are creating and storing within a digital infrastructure that is far from being a long-lasting medium. Black Archives and some other institutions are confusing the digital preservation with traditional preservation. Instead of digitally preserving historical documents, institutions are using "network surrogates which can enhance the preservation of artifacts . . ." It is not a coincidence that LC, Black Archives, and the Schomburg Center mention preservation as a goal and also provide selection criteria information. This fact suggests that preservation is related to the process of selecting materials for digitization.

Collaboration with External Institutions

Collaboration is noted when an institution shares resources (personnel, equipment, and documents) with an external institution. Forty-five percent of all institutions collaborated with external institutions that are academic, governmental, and commercial in nature. Some form of collaboration was practiced by all governmental institutions, 66 percent of archival institutions, and 50 percent of academic institutions. Internal collaboration, on the other hand, is by and large omnipresent. That is, building a digital library requires the input of experts from a variety of departments, centers, and divisions within an institution. University of Virginia is a perfect example of an institution that depends on both internal and external collaboration. The Uncle Tom's Cabin & American Culture multimedia collection is a collaboration among university departments, such as the Institute for Advanced Technology in the Humanities, and outside institutions such as the Harriet Beecher Stowe Center in Hartford, Connecticut. Another highly collaborative effort exists between the Amistad Research Center and LSU. Amistad provides the content (monographs and photographs) from its American Missionary Association collection, while the LSU digital library system provides the equipment and infrastructure to host the on-line collection.

Governmental institutions heavily relied on each other as well as on external government agencies. LC and NARA shared a substantial amount of information. In addition, the National Park Service cooperated with three government agencies such as the Historic Landmarks Commission to conduct research and borrow artifacts for digitization. Although public libraries and governmental institutions are collaborating with external institutions, the impetus for collaboration varies. Public libraries are probably motivated to collaborate to save money, while governmental institutions probably collaborate to gain access to content.

For some institutions collaboration embodies community outreach and volunteerism. For example, PLCMC mobilized the community through its digitization project. This grassroots project began with a series of community fund-raisers to launch the local history digitization project. In the process, members of the African-American community donated photographs and volunteered to make the project successful. VCU also uses its oral history digital project as an opportunity for the university to build a more positive relationship with the African-American community in Richmond, Virginia. As more institutions build digital collections, interinstitutional and intraintstitutional
collaboration will increase. It makes good sense for any institution to initiate cooperative projects to maximize resources and save money.

Intellectual Property, Privacy, and Copyright

For legal, ethical, and financial reasons, intellectual property and copyright statements are found in 85 percent of Web sites sampled. Needless to say, information professionals are highly cognizant of these issues, and the study results are proof. There is a great deal of variety in the placement, form, and content of these statements, but their purpose is clear. The Moorland-Spingarn Research Center at Howard University is one of the more proactive institutions. Copyright statements are explicit and highly visible in red text on the index of the Prints and Photographs Department as well as on every page linked to the index. In its statement on “Conditions for Use,” the Special Collections Department of UVA includes excerpts from United States copyright law. The National Park Service and two public libraries do not post intellectual property statements. Institutions should indicate the copyright status of a collection even if all materials are in public domain. If there is no information to address copyright, then users may assume that any use of downloaded images is permissible.

Funding

The type of funding (other than self-funding) received to create digital projects is listed on 55 percent of Web sites. Because it is important to acknowledge funding sources, presumably Web sites that do not post information are not being funded. Fifty-four percent of Web sites receiving funding were academic institutions while only 50 percent of public libraries received special funding or grants. The majority of funding comes from sources such as the Ford and Rockefeller Foundations. The National Endowment for the Humanities (NEH) exclusively funded academic institutions such as the University of Pennsylvania and UNC Chapel Hill. The Institute for Museum and Library Services (IMLS), an independent agency of the federal government, provided University of Virginia with over $240,000 for the The Jackson Davis Collection of African-American Educational Photographs and awarded the Amistad Research Center and LSU $186,000 for its American Missionary Association collection. The Church in the Southern Black Community from UNC Chapel Hill and The African-American Experience in Ohio, 1850–1920 from OHS were each granted over $72,000 from the LC/Ameritech National Digital Library Competition. The two public libraries that were funded received the funding from state library and education agencies. Consortia have been a source of funding for academic institutions. For example, VCU received funding from the library consortium to which it belongs: the Virtual Library of Virginia.

Much of the information concerning grants and funding is not readily accessible to users. Institutions should devote more attention to acknowledging their funding sources. University of Virginia is the only institution to provide a hyperlink of the grant proposal to its digital collection homepage. Grant proposals provide a great deal of suc-
cinct information about the grantor, the grantee, and staffing, as well as about the digital collection.

**Technology and Technical Notes**

**Image Quality, Image Display, and Plugins**

The quality of images varies from institution to institution. The highest quality images are found in academic library collections. The University of Pennsylvania and UVA used 600 dots per inch (dpi) black and white Tag Image File Format (TIFF) images. Black Archives and the Schomburg Center use 200 dpi black and white images as well as thumbnail images at 100 dpi. The Schomburg Center states that the quality of documents is compromised to ensure short download times for quick and easy access for all users. In addition, poor quality images are often found in collections that do not outsource scanning, such as *The Black Archives Digital Image Database*. When scanning is done in-house, equipment quality and the expertise of the personnel tend to be lower, which leads to poor images.

Plugins for display and/or audio/video playback are necessary for any digitization project that includes video and audio files. None of the collections include images that require viewers other than those associated with Netscape and Internet Explorer. Duke University, University of Michigan, UVA, and the University of Pennsylvania are the only four Web sites that require RealAudio. The University of Pennsylvania, UVA, and PLCMC also require the MPEG player QuickTime for video playback. Although these special features are great enhancements, downloading software can be time-consuming. LC, UVA, PLCMC, and the Schomburg Center provide information about using plugins effectively to help users optimize their viewing experiences. In fact, the Schomburg Center includes a “Compatibility Note” suggesting that, “These pages are best viewed with a table- and frame-compliant browser, such as Netscape version 2 or higher.” Although none of the American Memory Web sites in this analysis require plugins, LC posted an informative Web page entitled, “Viewing and Listening to American Memory Collections.” This Web page includes hyperlinks to Web sites that permit users to download plugins. The *Uncle Tom’s Cabin & American Culture* Web site at UVA recommends different viewers for users who own older modems. In addition, UVA’s plugin Web page explains why .ra files were more suitable for older modems than were .wav files. By devoting more attention on their Web sites to plugins, institutions such as the University of Pennsylvania, University of Michigan, Duke University, and WSU help users receive the full benefits of on-line collections.

**Equipment**

Forty-five percent of Web sites list equipment used for digitization. Academic institutions are least likely to provide this type of information. The University of Pennsylvania and UVA, which have their own electronic text centers, include detailed information about the scanners and digital cameras used in their digitization projects. A
number of Web sites such as Black Archives and the Schomburg Center list the graphics and Web publishing software, hardware, and peripheral devices (scanners, printers, and digital cameras). Institutions that outsource digitization do not post vendor information. As much vendor, software, and hardware information as possible should be provided. The information may not be valuable to researchers, but it is valuable to information professionals who are looking for guidance in selecting competent vendors and equipment for new digital libraries.

**Design and Organization**

**Organization, Accessibility, and User Friendliness**

In many respects, WWW browsing and the Internet culture have shortened the attention span of most Web users. This is exacerbated by the fact that the vast majority of information on the WWW is not meaningful or useful. Consequently Web-site organization, accessibility, and user friendliness are important. In this subdimension, navigation tools, colors, spacing, typography, accessibility, and organization of information are examined. Many Web sites, such as OHS, are able to strike a nice balance without inundating the Web page with oversized images and moving text. In addition, OHS allows users to navigate within a collection of materials as well as among different collections. Users are provided with many avenues to achieve goals. One can search, browse, and navigate to any part (serials, newspapers, photographs, manuscripts, pamphlets) of *The African-American Experience in Ohio, 1850–1920* collection. The left side of the Web page is a permanent and seamless frame that is always accessible to users. Although the font size is small, this Web site is well organized. The archival institutions (OHS, KWM, Black Archives) fared extremely well in this subdimension: all are rated either excellent or good. As a whole, Duke University, UVA, the Schomburg Center, the University of Pennsylvania, KWM, and OHS are the only Web sites that are well organized. Most Web sites in the analysis are restrictive, inundated with images, poorly organized, and not intuitive. NOPL, MSRC, and NPS Web sites are the most poorly constructed because the text is difficult to read, users must scroll endlessly, there is no visible structure to the homepage, and navigation is quite difficult. This problem is prevalent throughout most institutional and individual Web sites on the WWW today. Most of these problems could be resolved if Web-site standards emerged that could guarantee well-organized, user-friendly Web sites.

**General Institutional Information**

General institutional information is included on all Web sites. General information helps users gain virtual and physical access to collections. Many Web sites include operation hours, staff directories, and institutional histories; on the other hand, user statistics and directions are not often displayed. The only collection that does not include much general information is the University of Michigan’s CHICO Web site, whose primary goal is outreach and access to kindergarten through twelfth grade (K–12).
CHICO, however, is a virtual research center with no physical counterpart. UNC Chapel Hill provides more information about its digital collections than any other institution. It provides information on the status of each electronic text at each step of the digitization process: OCR, encoding, availability on the WWW, and assignment of OCLC number. Also included are the names of personnel who scanned and encoded each document along with a completion date. Curriculum vitae of all digital project staff are available. More important is the detailed information about the methodology, standards, archiving, and cataloging available for users. The University of Pennsylvania is the only institution that includes collection development and acquisitions policies.

Access statistics are another type of information being made available to users. If users want to know how many people visit a particular collection on-line, UVA and Duke University provide statistical user feedback. The user is given the choice to view the access statistics (numerically or graphically), which are compiled in a sophisticated manner. Similarly, Duke University has a Web page entitled “WWW Access Statistics,” but it is less detailed than UVA's Web page. The relevance of this type of information to the average user is unclear.

Virtual tours, floor plans, and tutorials can also make archives as well as their on-line historical documents more accessible. If archivists intend to reach wider audiences, they should provide unsophisticated researchers with orientation information. A tutorial to explain what archives are and directions on how to use them is useful. A number of institutions have done as much. In fact, Yale University and the University of California at Berkeley's archives have mounted on-line tutorials to “orient new researchers to techniques of searching for primary source material and to details about doing research.” The CHICO project at University of Michigan is the only institution in this analysis to provide instructional information about its on-line collection for its audience (K–12 students). Adults as well as school-age students often need assistance when it comes to primary or historical documents. Web sites must include a minimal level of information (hours, rules, regulations, and policies) but also more instructional information to help users conduct research with primary documents. Access statistics and detailed staff information may not be as useful as directions and methodology. The goal is to make archives more accessible for all types of users.

**Finding Aids**

Only 30 percent of Web sites incorporate finding aids into their digital collections. None of the public libraries or governmental Web sites have finding aids in their digital collections. Finding aids are primarily present in digital collections that focus on textual information such as oral histories. *North American Slave Narratives, Beginnings to 1920* at UNC Chapel Hill and the oral histories at VCU are text only and are the only two Web sites that use both HTML and SGML encoded finding aids. All Web sites that include finding aids have search and retrieval interfaces, which make the documents easily accessible. The depth of description varies from institution to institution.
Search and Retrieval Interface

Sixty-five percent of evaluated Web sites provide on-line search and retrieval interfaces within their Web sites. The majority of search and retrieval interfaces are quite sophisticated. Many institutions create interfaces that use combinations of browsing and searching. Other Web sites such as Black Archives provide users with a choice of using a simple or advanced search. OHS’s search and retrieval interface is permanently situated on the left within a frame that makes it accessible to users wherever they travel within the Web site. Wright State University uses the LibNet search engine to search for electronic resources such as databases and URLs. KWM uses the “external” search engine Excite to enable users to search the museum’s catalog. National Archives has one of the more elaborate search and retrieval systems known as NAIL (NARA Archival Information Locator), which is a database that contains information about a wide variety of NARA’s holdings across the country. Although NAIL holds a limited portion of NARA’s vast holdings, it is sophisticated enough to locate manuscripts as well as films, videos, maps, and charts. In the Digital Classroom, an educational component of NARA, items are selected for K–12 users to retrieve information.

Many institutions have one search function for retrieving digital and analog materials. Consequently, searchers are not aware that they may encounter on-line finding aids or paper manuscripts. This variety further complicates searching, which is often difficult for users to understand. Easily browsed and hyperlinked indices are available at 10 Web sites. For example, the Paul Laurence Dunbar Digital Text Collection uses an alphabetically arranged index to help users locate poems, books, and librettos. Even when search and retrieval interfaces are well designed, users are likely to have difficulty applying Boolean logic. Consequently, help pages are important. Over 70 percent of the collections have Web pages to help users search their Web sites.

Audience

In some cases, the target or intended audience can be easy to ascertain, but in others the audience may not be clear. Whether or not an institution identifies its intended audience on the Web site is very important. Only 30 percent of Web sites post information about the audience. Institutions whose focus is on K–12 are much more likely to post this type of information. If we exclude the CHICO project from University of Michigan (whose audience is clearly K–12), then only 40 percent of academic institutions consider it necessary to inform users about the target audience. These institutions assume that audiences for digital documents are the same as those who use analog documents. This assumption is probably false because on-line information almost always attracts a new and diverse population of users who would not otherwise use archival materials or visit archives to conduct research. Overwhelmingly, LC and NARA include more information about their target audience than do any other institutions. It is worth noting that historically, LC has denied K–12 students access to their 22 reading rooms. In response to this condition, the “Periodic Report” from the National Digital Library Program states, “the reading rooms will continue to be reserved for researchers over high school age [while] many of the library’s core historical American
collections will be available on the Internet to everyone.\textsuperscript{45} Similarly, NARA’s Digital Classroom reaches out to school-age children by granting access to primary resources that were traditionally inaccessible on paper.

**Metadata**

Metadata are organized data that describe information in digital and analog form. As digital information increases, so does the need for metadata. Even the simplest digital collection cannot be successfully completed without the creation of some form of metadata (administrative, technical, preservation, use, descriptive). Fifty percent of all institutions are embedding either TEI headers\textsuperscript{46} or Dublin Core tags into their digital files.\textsuperscript{47} Descriptive metadata are primarily being used by academic and governmental institutions. The only public library that posts information about its application of TEI-lite is the Schomburg Center.

Embedding digital information with metadata is the first step. Organizing or cataloging that information is the next. Metadata collected for *The Church in the Southern Black Community* at UNC Chapel Hill is extensive. In addition, the Status Report Web page lists OCLC numbers of the digital texts that tell us that the item is cataloged. Unlike UNC Chapel Hill, most institutions fail to catalog their digital collections. If digital collections are not cataloged, they do not become part of national databases. Consequently, users, catalogers, and vendors do not know they exist. As the WWW grows, it is even more important for information professionals to gain bibliographic control over digital information, and metadata are key to that end.

**Information and Reference**

**User Feedback and Remote Reference**

At a basic level all Web sites provide some sort of reference in the form of hyperlinks to various libraries, museums, organizations, and archives. However, the capability for user feedback or reference is determined by the presence of an E-mail address (hyperlinked or not). Twenty percent of institutions do not provide one. Is the omission of this information deliberate? It is possible that MSRC, Black Archives, and KCPL cannot handle the onslaught of inquiries that may accompany E-mail addresses. At the same time, it is not so clear why institutions would go to such lengths to digitize collections and not allow users the opportunity to gain access to reference staff as they gain on-line access to collections. On the other hand, there are several institutions such as LC and KWM that encourage users to send comments, contact staff, or sign guest books. Given that many institutions may not be staffed to handle more reference questions than they already have, it seems that archivists would welcome the opportunity to log and track the quantity and quality of questions coming in as a result of the on-line collection.

In evaluating Web sites, it was not possible to ascertain if institutions have implemented tools to create usage logs that generate monthly reports. According to Thomas
Ruller, usage logs "provide a rough indicator as to what types of materials are used by visitors." The number and length of visits to a Web page and the domain type of the visitor (.edu, .gov, .com) provide valuable information. Ruller goes on to say that if institutions are not making "an effort to measure . . . the level of use of the resource, [then they will not be able to calculate] the extent to which the resource meets customer needs." Are archivists underestimating the potential benefits of providing remote reference services? Feedback can be valuable and may help staff make important decisions about the long-term maintenance of their digital collections.

**Conclusion**

Can we respond to Murle Kenerson's question, "Is digitization a true revolution in the effort to ensure more access to and preservation of African-American collections, allowing what has been for too long hidden resources to at last be revealed to all seeking the rich vein of tradition . . . ?" Although these digital collections show a great deal of promise, the answer to the question is "No." Institutions analyzed and evaluated for this paper have sought to add a newer dimension to their Web sites in the form of a digitization project. Granting access to primary resources over the WWW is the goal of these institutions. There is some similarity among projects, but there is even more diversity. Each institution has created proprietary projects according to its financial restrictions, audience, personnel, and technology. Consequently, we have seen unique digitization projects from 20 institutions. Such variety can be attributed to the lack of standards in the area of digital libraries, Web design, network architecture, software, and hardware.

There are standards in the areas of cataloging, encoding, and imaging. USMARC and MARC-AMC are accepted national standards for the representation and communication of bibliographic information in machine readable form. International Standard Archival Description (ISAD [G]) and Encoded Archival Description (EAD) are emerging standards used in constructing finding aids. The open international protocol (ISO 8879) is used for maximum computer interoperability. In terms of imaging, TIFF is the uncompressed image standard while Joint Photograph Experts Group (JPEG) and Graphic Interchange Format (GIF) are the image compression standards observed in these digitization projects. Academic institutions are already accepting these standards, but there are other procedures and components that are essential to digital libraries that are not standardized.

Web-page design and user interfaces lack standardization. When entering any given digital collection, one has no idea what to expect. One may be confronted with a host of formats, some may be converted on the fly, some may be enhanced, and some images may or may not be enlarged. Many Web sites have search and retrieval interfaces, but they are all quite different. Finding aids are not standardized either in the analog or digital environment. The multitude of interfaces a user might encounter is endless. Although LC and UVA are excellent models, there is no prototypical digital collection that may even begin to serve as a comprehensive guide for institutions.

The Digital Library Initiative, a division of LC, lists 10 challenges that must be met if effective digital libraries are to be created in the twenty-first century. The list of
challenges addresses many of the problems that digital libraries repeatedly face. The fact that LC has not managed to meet the majority of them is problematic. In and of themselves, the challenges are descriptive, but do not address social, political, and economic issues that make it impossible for most institutions to create effective digital libraries. Information professionals must be confident that LC will set standards that will positively impact digital libraries. Currently, standards are lacking and institutions will continue to create unfriendly interfaces, low quality images, and digital documents that lack reliability, authenticity, and integrity.

Lack of standards and the unique challenges facing digital libraries lead one to ask, “Why digitize?” The answer is access. Now that users have access to collections documenting African-American history and culture, what is the quality of the materials that they are accessing? After evaluating digitization projects from 20 institutions, it is not clear why investments were made because the quality of the materials is only average. If the target audience is scholars and researchers, then these digital collections are unsuccessful and extremely wasteful. In 1992, Michelson and Rothenberg argued that networking and access to digital sources will change all dimensions of the scholarly work process, including identifying sources, interpreting and analyzing data, disseminating research findings, and teaching. According to the results of this analysis, Michelson and Rothenberg’s prediction will not easily become reality. Only a small number of these projects could actually support scholarly research. Kindergarten-through-twelfth-grade users stand to gain the most from these projects because traditionally they have had limited access to archives and primary resources. Also, teachers can provide guidance and valuable context for students who are not likely to seek such information. In addition, for K–12, digitization projects are more successful because they tend to be less interested in the informational, physical, or bibliographical integrity of documents than are researchers. If the audience is broad—much like that of a public library—then these digitization projects are nothing more than glamorized exhibitions serving as institutional advertisements.

The idea of digitizing portions of archival collections can be powerful, if implemented properly. LC was the only institution evaluated in this analysis that implemented user studies and ran a five-year pilot project (1990–1995) on the American Memory historical collection. If archivists are not conducting pilot studies or collecting information from and about the users, then they are working from nothing more than assumptions. How can archivists and their institutions justify expenses invested in digitization projects that are based more on assumptions than on facts? Obviously, many institutions are tolerant and eager to experiment and develop digital libraries. In the future, users’ attitudes and behaviors towards on-line historical collections must be studied. The final (and minimal) goal for all institutions should be to have well-planned digital collections that are practical, meaningful, and user centered.

Even though most digital collections contained plenty of detailed information, more relevant information is needed. As David Wallace contends, “There is a clear need for more information regarding the archives themselves such as their enabling legislation, operational responsibilities and collecting missions.” Furthermore, he adds, “The availability of institutional policy documents, appraisal reports, preservation and conservation reports, and the like could ignite greater professionalization and intellectual
dynamism within the profession. This is essential given the paucity of formal archival education programs."

In addition, archivists should be utilizing the Web to change the derogatory views that many have about archives that store “old stuff.” Richard Cox goes further stating, “Archivists need to consider how . . . the web provide[s] an opportunity for creating a greater understanding of archival value.” Currently, however, there is little evidence in the digital collections that indicate the role archivists are playing in digitization projects. The fact that selection criteria are barely mentioned testifies to this. It is not clear what archivists are doing, what their role is, or what their responsibilities are.

In the networked environment, archivists should apply the same archival theories and have the same responsibilities as they do in the analog environment. It is also important for archivists to take the opportunity to invent new procedures and abandon unsuccessful ones. Instead, archivists and other information professionals are following procedures adopted before the WWW, which are inappropriate and costly in the networked environment. Unfortunately, archivists and users maintain the idea that the archivist’s primary job is to house documents. This custodial function has often been stressed over acquisitions, reference, appraisal, outreach, and description. Consequently, many archivists are under the assumption that if they collect and organize it, researchers will come. This must be the case if repositories are willing to invest thousands of dollars without conducting pilot projects or user studies. The most common procedure is to digitize documents, organize them, and wait for users to come. In the digital environment, we cannot afford to make such assumptions.

There are other fundamental weaknesses in these collections, too. Preservation of digital libraries is not much of an issue. Although most institutions are cognizant of the problems that will arise with the obsolescence of technology, they are not doing anything that would assure users that digital collections will exist for decades to come. It is unlikely that institutions in this analysis are prepared to take on the provision, organization, and preservation of digital information with the same long-term commitment they have had for printed materials. A more national, coordinated, and planned approach is needed to foster access and research, and to ensure that preservation goals are met in the digital environment.

It is clear that archivists have more questions than answers. They are asking: What are the long-term and short-term goals of digital collections? What factors are critical to the success of digitization? How can we measure the benefits of digital collections? In the final analysis, information professionals and others who have a vested interest in digital libraries will have to be prepared to explain why archivists should digitize. Digital collections of historical African-American documents, in and of themselves, do not adequately speak to this question.
## Evaluative Criteria

<table>
<thead>
<tr>
<th>DIMENSIONS</th>
<th>SUBDIMENSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Collection and Content</td>
<td>1 Quantity of items/documents</td>
</tr>
<tr>
<td></td>
<td>2 Format</td>
</tr>
<tr>
<td></td>
<td>3 Collection background and context</td>
</tr>
<tr>
<td></td>
<td>4 Range, accuracy, and depth</td>
</tr>
<tr>
<td></td>
<td>5 Reliability, authenticity, and integrity</td>
</tr>
<tr>
<td></td>
<td>6 Selection criteria</td>
</tr>
<tr>
<td>II</td>
<td>Institutional Information and Policy</td>
</tr>
<tr>
<td></td>
<td>7 Mission statement</td>
</tr>
<tr>
<td></td>
<td>8 Preservation</td>
</tr>
<tr>
<td></td>
<td>9 Collaboration with external institutions</td>
</tr>
<tr>
<td></td>
<td>10 Intellectual property, privacy, and copyright</td>
</tr>
<tr>
<td></td>
<td>11 Funding</td>
</tr>
<tr>
<td>III</td>
<td>Technology and Technical Notes</td>
</tr>
<tr>
<td></td>
<td>12 Image quality, image display, and plugins</td>
</tr>
<tr>
<td></td>
<td>13 Equipment</td>
</tr>
<tr>
<td>IV</td>
<td>Design and Organization</td>
</tr>
<tr>
<td></td>
<td>14 Organization, accessibility, and user friendliness</td>
</tr>
<tr>
<td></td>
<td>15 General institutional information</td>
</tr>
<tr>
<td></td>
<td>16 Finding aids</td>
</tr>
<tr>
<td></td>
<td>17 Search and retrieval interface</td>
</tr>
<tr>
<td></td>
<td>18 Audience</td>
</tr>
<tr>
<td></td>
<td>19 Metadata</td>
</tr>
<tr>
<td>IV</td>
<td>Information and Reference</td>
</tr>
<tr>
<td></td>
<td>20 User feedback and remote reference</td>
</tr>
</tbody>
</table>
Appendix B

Evaluated Digital Collections

Stand-Alone Archival Institutions:
Archives, Museums, and Historical Societies

1. Black Archives of Mid-America
   

2. The Kendall Whaling Museum
   

3. The Ohio Historical Society
   

Academic Institutions

4. Duke University: The Digital Scriptorium Rare Book, Manuscript, and Special Collections Library
   
   *African-American Women On-line Archival Collections*, <http://scriptorium.lib.duke.edu/collections/african-american-women.html>, features scanned pages and texts of the writings of the memoirs of Elizabeth Johnson Harris (1867–1942), an 1857 letter from Vilet Lester, a slave on a North Carolina plantation, and several letters from other slaves on the estate of David Campbell, the governor of Virginia.
5. Howard University, Moorland-Spingarn Research Center (MSRC)

*Prints and Photographs Department*, <http://www.founders.howard.edu/moorland-spingarn/PRINTS.htm>, is a small image gallery of 27 black and white images that depict African-American life and culture from the mid-eighteenth century through the early part of the twentieth century.

6. Michigan State University Archives & Historical Collections

*The African-American presence at MSU: Pioneers, Groundbreakers, and Leaders, 1900–1970*, <http://www.msu.edu/unit/msuarhc/frame_exhibits.htm>, is an exhibit of 40 black and white photographs that honor some of the pioneers of the university and give recognition to their achievements.

7. Louisiana State University Digital Library in collaboration with the Amistad Center of Tulane University


8. University of Michigan, the School of Information, Cultural Heritage Initiatives for Community Outreach (CHICO)

*Harlem 1900–1940: An African-American Community*, <http://www.si.umich.edu/CHICO/Harlem/>, is a multimedia collection designed for the K–12-grade audience. It consists of approximately three hundred lesson plans, images, and texts that have been collected by the educational programs unit of the Schomburg Center for Research in Black Culture. Within this collection is a database devoted to the writers, artists, and musicians of the Harlem Renaissance.

9. University of North Carolina at Chapel Hill Libraries: Documenting the American South

*North American Slave Narratives, Beginnings to 1920*, <http://metalab.unc.edu/docsouth/neh/neh.html>, is a collection that documents the individual and collective story of the African-American struggle for freedom and human rights in the eighteenth, nineteenth, and early twentieth centuries. *North American Slave Narratives* is one division of the collection of electronic texts that are part of the Documenting the American South Project at UNC. This division includes 29 texts of some of the most famous narrators such as Frederick Douglass and Nat Turner.
The Church in the Southern Black Community, <http://docsouth/church/index.html>, contains over 19,000 pages from approximately one hundred works, including autobiographies, sermons, church reports, religious periodicals, and denominational histories, tracing the experience of Southern African-Americans and the transformation of Protestant Christianity into the central institution of black community life.

10. The University of Pennsylvania Library, Annenberg Rare Book & Manuscript Library

Marian Anderson Collection of Photographs, 1898–1992, <http://www.library.upenn.edu/special/photos/anderson/>, contains over 4,500 photographs of Marian Anderson, her friends, colleagues, and admirers. Users can conduct searches as well as view documents in the browse mode. A register of the Marian Anderson papers is also available on-line.

11. Special Collections at University of Virginia Library


Uncle Tom’s Cabin & American Culture, <http://jefferson.village.virginia.edu/utc/sitemap.html>, is a full-text-searchable multimedia archive of texts, images, songs, three-dimensional objects, and video clips on a wide variety of issues and themes that are associated with Harriet Beecher Stowe’s novel Uncle Tom’s Cabin. All information is accessible in interpretive, browse, and search modes.


The Church Hill Oral History Collection, <http://www.library.vcu.edu/jbc/speccoll/vbha/church/church.html>, contains interviews with 35 individuals, all of who were then current or former residents of Richmond’s Church Hill neighborhood.

13. Wright State University Libraries Special Collections and Archives

*Paul Laurence Dunbar Digital Text Collection*, [http://www.libraries.wright.edu/dunbar/], is a tribute to Dayton poet and novelist Paul Laurence Dunbar. This collection provides access to over two hundred poems published at the turn of the century. In addition, there are a libretto and audio clip of some of his most famous poems.

**Public Libraries**

14. Kansas City Public Library


15. The Public Library of Charlotte and Mecklenburg County (PLCMC)

*An African-American Album, Volume I*, [http://www.cmstory.org/african/album/volume1/default.htm]. The first part of the on-line exhibition of close to a dozen photographs and five audio clips that document African-American life in Charlotte and Mecklenburg County from the late nineteenth century until the 1940s.

*The Black Experience in Charlotte and Mecklenberg County, Volume II*, [http://www.cmstory.org/aaa2/default.htm]. The second part of the on-line exhibition of over 45 photographs and 20 audio clips that document African-American life in Charlotte and Mecklenburg County from the 1940s to the present.

16. The Schomburg Center for Research in Black Culture

*African American Women Writers of the 19th Century*, [http://digital.nypl.org/schomburg/writers_aa19/], is a digital collection of some 52 published works by nineteenth-century black women writers. This searchable collection provides access to the thought, perspectives, and creative abilities of black women as captured in books and pamphlets published prior to 1920.

*Images of African Americans from the 19th century*, [http://digital.nypl.org/schomburg/images_aa19/], is a pictorial database of over five hundred illustrations and photographs that document the life and history of peoples of African descent worldwide.
17. New Orleans Public Library


*African Americans in New Orleans: Les Gens de Couleur Libres,* <http://nutrias.org/~nopl/exhibits/fmc/fmc.htm>, is an exhibit designed to provide firsthand examples of the role that free people of color played in antebellum New Orleans. It uses original documents from the city archives along with materials from other Louisiana Division collections.

*African Americans in New Orleans: Family History Sources,* <http://nutrias.org/~nopl/exhibits/bhm98/black98.htm>, is an exhibit that displays some of the sources available in the Louisiana Division and in the city archives for genealogical research into African-American ancestry.

**Government**

18. Library of Congress American Memory

*The African American Odyssey: A Quest for Full Citizenship,* <http://lcweb2.loc.gov/ammem/aahome.html>, showcases the incomparable African-American collections of LC. Displays more than 240 items, including books, government documents, manuscripts, maps, musical scores, plays, films, and recordings.

*African American Perspectives: Pamphlets from the Daniel A. P. Murray Collection, 1818–1907,* <http://lcweb2.loc.gov/ammem/aap/aaphome.html>, has 351 titles in the collection including sermons on racial pride and political activism; annual reports of charitable, educational, and political organizations; and college catalogs and graduation orations. Also included are biographies, slave narratives, speeches by members of Congress, legal documents, poetry, playbills, dramas, and librettos. Several of the pamphlets are illustrated with portraits of authors.

*Creative Americans: Portraits by Carl Van Vechten, 1932–1964,* <http://lcweb2.loc.gov/ammem/vvhome.html>, consists of 1,395 photographs taken by American photographer Carl Van Vechten (1880–1964) between 1932 and 1964. The bulk of the collection consists of portrait photographs of celebrities, including many figures from the Harlem Renaissance. A much smaller portion of the collection is an assortment of American landscapes.
19. National Archives and Records Administration (NARA), The Digital Classroom: Primary Sources, Activities, and Training for Educators and students.

*The Amistad Case*, <http://www.nara.gov/education/teaching/amistad/home.html>, contains documents related to the circuit court and Supreme Court cases involving the *Amistad*. It includes over 19 pages from five primary documents about the *Amistad* slave revolt case. In addition, teaching activities are included to assist teachers and school-age students in using such documents on the WWW.

*The Fight for Equal Rights: Black Soldiers in the Civil War*, <http://www.nara.gov/education/teaching/usct/home.html>, includes teaching activities, historical documents, and photographs that explore the issues of emancipation and military service. It includes 13 pages of documents and one teaching activity.

*Beyond the Playing Field: Jackie Robinson, Civil Rights Advocate*, <http://www.nara.gov/education/teaching/robinson/robmain.html>, is an exhibition of manuscripts and photographs that trace Robinson’s career as a civil rights leader. There are nine documents and three lesson plans available that highlight Jackie Robinson’s legacy.

20. National Park Service

*Historic Places of the Civil Rights Movement: We Shall Overcome*, <http://www.cr.nps.gov/nr/travel/civilrights/>, is a collection of photographs and documentation on the 41 properties throughout 20 states and the District of Columbia that have been nominated by the states and listed in the National Register of Historic Places.

**ABOUT THE AUTHOR:** Elaine L. Westbrooks was a digital research librarian and cataloger at the University of Pittsburgh and is now a metadata librarian at Cornell University. She is a member of numerous professional organizations, including the Association for Library Collections & Technical Services, the Association of College and Research Libraries, and the Black Caucus of the American Library Association. The author would like to express her gratitude to Elizabeth Yakel for her assistance and valuable comments on this article.
NOTES

1. “Digitization” is the process of converting analog images, text, audio, and visual information into digital information with the aid of Optical Character Recognition (OCR) software or with the aid of hardware such as scanners and digital cameras.

2. The terms “digital collection” and “digital projects” are interchangeable. Both terms refer to sets of digital images meaningfully linked or related.

3. In most cases, digital collections may receive attention when they are awarded grants from institutions such as the Institute of Museum and Library Services (IMLS). This was the case with the American Missionary Association digital collection of the Amistad Research Center at Tulane University. However, little scholarly research has been published about the digital collections, their use, or value to researchers.


9. With the exception of University of Michigan (CHICO), Tulane University, Howard University’s Moreland-Spingarn Research Collection, and University of North Carolina, all of the academic institutions in this analysis launched digitization projects from numerous departments such as special collections and archives. The University of Pennsylvania and University of Virginia benefitted from the assistance of electronic text centers that focus on digital projects.


12. “SGML” (Standard Generalized Markup Language) is a set of rules for defining and expressing the logical structure of documents, thereby enabling software products to control the searching retrieval, and structured display of those documents. The rules are applied in the form of markup (tags) that can be embedded in an electronic document to identify and establish relationships among structural parts. See the Library of Congress, Encoded Archival Description Official Web Site, <http://www.loc.gov/ead/eadback.html> (5 May 1999).

13. “Information professionals” are any personnel whose responsibility is to manage, organize, preserve, or maintain information. This includes librarians such as bibliographers, systems librarians, catalogers, preservationists, and archivists.


18. Duranti, 7.
20. Bearman and Trant.
22. Lynch.
24. “Selection criteria” refers to the list of minimal requirements that items or collections must meet in order to be reformatted, acquired, or preserved. The concept of selection criteria has been applied to the field of preservation to identify endangered materials that are worth the effort and cost of long-term preservation. Some refer to this as "preservation selection criteria." The concept of "digital selection criteria" includes the identification of materials that are worth the effort and cost of long-term digital preservation, or digitization. In some cases, there may be no distinction between digital selection criteria and preservation selection criteria. According to Hazen et al., selection for digitization is a complicated process. Selection entails making judgments about the following factors: the intellectual and physical nature of the source materials; the number and location of current and potential users; the current and potential nature of use; the format and nature of the proposed digital product and how it will be described, delivered, and achieved; how the proposed product relates to other digitization efforts; and projections of costs in relation to benefits. Dan Hazen, Jeffrey Horrell, and Jan Merrill-Oldham, “Selecting Research Collections for Digitization” (Council on Library and Information Resources: August 1998) <http://www.clir.org/pubs/reports/hazen/pub74.html> (19 January 1999).
28. In “Digital Imaging: Issues for Preservation and Access,” Meg Bellinger defines “digital preservation” as the reformattting or copying of information from unstable originals to media and hardware with proven and verifiable standards for longevity. Reformattting entails a quality reproduction that is adequate to reproduce the intellectual content of the original; the media to which the information is transferred must be stable and accessible now and in the future. Information born out of the electronic environment must also be reformatted and copied to media and hardware with proven and verifiable standards for longevity. On the other hand, the preservation of paper-based materials entails either the stabilization of the original artifact and the subsequent control of its environment, the creation of a surrogate to reduce use of the original and, thereby, perpetuate its existence, or when the original is unstable, the transfer of the intellectual content to another more stable medium to ensure availability in the new medium. This distinction can be found in the collected work, Digital Image Access & Retrieval: Proceedings of the 1996 Clinic on Library Applications of Data Processing, 33rd ed., ed. P. Bryan Heidorn and Beth Sandore (Urbana, Illinois: Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 1997): 157–163.
34. This grant information was not part of the Amistad Research Center Web site; however, a link to the IMLS homepage is provided. IMLS, “Recipients of National Leadership Grants for Libraries 1999,” <http://www.imls.gov> (20 April 2000).
38. The “quality” of an image depends on factors other than the dpi or resolution of the digital image. The dynamic range, monitor/printer resolution, image capture (equipment), and compression ratio are all factors that determine image quality. See Howard Besser and Jennifer Trant, “Introduction to Imaging: Issues in Constructing an Image Database” (Santa Monica, CA: The Getty Art History Information Program, 1995) <http://www.getty.edu/gri/standard/introimages/index.html> (27 February 1999).
40. According to the MPEGorg Web site, MPEG (Moving Picture Experts Group) refers to the family of standards used for coding audiovisual information (e.g., movies, video, music) in a digital, compressed format. The major advantage of MPEG over video and audio coding formats is that MPEG files are much smaller for the same quality due to its sophisticated compression techniques. “MPEG Pointers and Resources,” <http://www.mpeg.org/MPEG/> (28 April 2000).
46. “TEI” (Text Encoded Initiative) is the encoding structure and vocabulary developed for electronic texts in the humanities in an international cooperative project. The TEI header provides a detailed framework for documenting the electronic text, its source, the encoding system, and its revision history. University of North Carolina at Chapel Hill Library, “Digitizing the Collection,” <http://metalab.unc.edu/docsouth/digitizingnarr.html> (6 May 1999).
48. Ruller, 167.
49. Ruller.
51. Kenerson, 5.
52. A summary of the 10 challenges to building an effective digital library are as follows:
   1. Develop improved technology for digitizing analog materials.
   2. Design search and retrieval tools that compensate for incomplete description.
   3. Design tools that facilitate the enhancement of descriptive information by incorporating users.
   4. Establish standards to facilitate the assembly of distributed digital libraries.
   5. Address legal concerns associated with access, copying, and dissemination of all materials.
6. Integrate access to both digital and physical materials.
7. Develop approaches that can present heterogeneous resources coherently.
8. Make the National Digital Library useful to different communities for different purposes.
9. Provide more efficient and more flexible tools for transforming digital content to suit end users’ needs.
10. Develop economic modes for the support of National Digital Library.


56. It would certainly be a mistake to create all digital collections for scholarly research. If scholars are the intended audience as archivists have claimed, then historical digital collections must have physical, bibliographical, and informational integrity. These essential features ensure evidential value and authenticity of historical documents.

59. Cox, 32.
60. Ruller addresses many of these issues in his article, “Open All Night,” 168–170.