Copyright Notice

This paper was presented at EDUCAUSE 2000 in Nashville, October 10-13, 2000. It is the intellectual property of the author(s). Permission to print or disseminate all or part of this material is granted provided that the copies are not made or distributed for commercial advantage and the title and author(s) of the paper appear. To copy or disseminate otherwise, or to republish in any form, requires written permission from the author(s).
Southeastern Native American Documents, 1763-1842

By
Robert E. Henneberger
Nan McMurry
University of Georgia
Athens
Georgia
James B. Lloyd
University of Tennessee
Knoxville
Tennessee

In 1999 the University of Georgia Libraries, the University of Tennessee at Knoxville Library, the Frank H. McClung Museum, and the Tennessee State Library and Archives were awarded a one-year National Leadership Grant from the Institute of Museum and Library Services to digitize 1,000 original documents and visual images relating to the Native American population of the Southeastern United States. These documents and images were selected from the most significant holdings from each collection and range in date from 1763 to 1842. The original documents reside in many separate manuscript collections within participating institutions, but as digital entities they are being brought together into a single electronic collection. The final product will be a database of facsimile images and transcribed texts, individually cataloged and full-text searchable, mounted on GALILEO (Georgia Library Learning Online).
Southeastern Native American Documents, 1763-1842

Project Description

In 1999 the University of Georgia Libraries, the University of Tennessee at Knoxville Library, the Frank H. McClung Museum and the Tennessee State Library and Archives were awarded a one-year National Leadership Grant from the Institute of Museum and Library Services to digitize 1,000 original documents and visual images relating to the Native American population of the Southeastern United States. These 1,000 documents and images were selected from the most significant holdings from each collection and range in date from 1763 to 1842. Although the primary focus of these collections is the Cherokee tribe, other tribes are represented, such as the Seminole and Creek. These documents have been identified over the years by the librarians and scholars who have worked with them, and they include official documents such as treaties, laws and military orders, personal letters written by both Native American and white authors, and the first newspaper published in a native language, the Cherokee Phoenix. Individually, most documents are relatively brief, from one to ten pages in length, but when combined they form a rich corpus capable of supporting any level of research or educational outreach.

Collectively, the Southeastern Native American holdings of the participating institutions are the most comprehensive in existence. The original documents reside in many separate manuscript collections within the participating institutions, but as digital entities they are being brought together into a single electronic collection. Work on this project began in November, 1999, and the final product will be a database of facsimile images and transcribed texts, individually cataloged and full-text searchable, mounted on GALILEO (GeorgiA Library Learning Online), the World Wide Web-based server sponsored by the Board of Regents of the University System of Georgia, available to the public at large at http://www.galileo.peachnet.edu.

We have requested funding to continue the project for a second year, incorporating another 1,000 documents and including two new partners, the Tennessee State Museum and the Museum of the Eastern Band of Cherokee. The limitations of traditional print collections described below (see Need for Digital Access) are one of the driving forces behind this second proposal, for in gathering materials for the current project we have identified additional documents that far exceed the number originally projected and extend back in time earlier than the original starting date of 1763. Doubling the size of the digital collection, extending the starting date back to 1730, and including the holdings of two additional institutions would bring us significantly closer to the long-term goal of creating a comprehensive electronic repository of primary source materials on Native American history in the Southeast and offering a model for the digitization of similar materials in other parts of the country. It is our intention to utilize the practices and work flows developed through this grant to continue developing the Southeastern Native American digital collection and to include additional partner institutions that have expressed an interest.

Significance and Scope of the Documents

“Southeastern Native American Documents, 1763-1842” has well-defined limits of geography and time period, but its content is of national significance. The importance of native tribes in the early history of America has been universally acknowledged, but primary sources remain scarce, especially for students below the college level and for the general public. The documents selected for this project serve to communicate to the modern reader how Native Americans viewed the European settlers as fellow
human beings, both positively and negatively, from the first contacts to the point when they were forcibly removed from their lands. Conversely, these documents also reveal how Europeans perceived Native Americans during the same time span. The interaction between European settlers and the Native Americans was the first step on America’s long and continuing journey towards cultural diversity. These documents will also serve to provide insight to the everyday life and social structure of Southeastern Native American societies through seldom seen source documents. The depth and scope of the resulting digital collection of source documents cannot be found anywhere else.

Documents and images selected for the first year of the project include formal treaties between British/American officials and tribal representatives, letters recording the negotiations that formed the background to the treaties, Native American law codes, and other sources that testify to the richness and continued viability of Native American culture even as it was encroached upon and eroded by European settlement. If we receive a second year of funding similar materials will be added from the University of Georgia’s Hargrett Manuscript and Rare Book Library, the University of Tennessee Library’s Penelope Johnson Allen collection and the Tennessee State Library’s Cherokee collection. The McClung Museum’s Tellico Reservoir Archaeology Project collection, on which digitization began in the first phase, will receive greater emphasis in the second, as it represents the comprehensive archaeological record of Cherokee town and house sites in the eighteenth and nineteenth centuries. The maps, photographs and report tables in this collection provide information on town and house structure and histories of individual Cherokee settlements. This material is of great research value, but is difficult to use in its traditional formats due to their diversity and vulnerability to deterioration. Documents from the two new partners in the second phase of the project, the Tennessee State Museum and the Museum of the Eastern Band of Cherokee, include the papers of William Holland Thomas, the individual who is almost single-handedly responsible for the continued residence of the Eastern Band in North Carolina.

**Need for Digital Access**

The early history of Native Americans in the Southeast, as in most of America, is a story of voices falling silent. At times the silencing was sudden and violent, as in the forced migration known as the Trail of Tears that tore so many natives from their homes forever. It also proceeded more subtly, as when native languages lost their last speakers. The silence can persist even in carefully preserved historical documents when wide public access to them is lacking.

The organization of many manuscript collections, while faithful to the principle of provenance, is often at odds with the needs of scholars, who pursue subject themes that can range across many collections. The Hargrett Rare Book and Manuscript Library of the University of Georgia, for example, houses a rich store of materials documenting the early history of Native Americans in Georgia, yet no single collection or even a group of collections is the source of this wealth. Because they are dispersed throughout the holdings of this library as well as others, documents concerning Native Americans are as scattered and silent as the native authors who produced them. Traditional collection descriptions and finding aids, when they exist at all, are of limited value in identifying subjects covered at the item level, leaving scholars to rely on the expertise of archival staff members, whose collective knowledge varies over time as they take other positions or retire.

Even when identified, many of these documents are available only in their original form, and access to them is possible only through prior arrangement with the staff of the collection in which they reside, and only under strictly controlled circumstances due to the age, value and condition of these materials. Given these necessary restrictions, it is not surprising that the primary users of these original documents are advanced scholars. Special collections in state institutions and museums are generally open to all, and classes of school children and other non-university groups do make frequent visits, but usually only for tours rather than for actual use of collections.
Digital technology can end the silence of Native American voices by creating a new, virtual collection of Native American documents and revolutionizing access to these materials. Users will not have to travel to the particular university or museum to view them; vulnerable originals will not be handled; and multiple users, including entire classes of school children, will be able to examine the same document simultaneously.

The research potential of this collection is immense, but it also holds great promise for education at the secondary and elementary levels as well as interest for the public. Students below the college level seldom receive adequate exposure to primary source material, particularly documents that portray Native Americans as complex individuals in a dynamic society. Many items have obvious dramatic appeal, but they can also function as bridges to other materials in which Native American voices speak for themselves anew.

Public interest in history, especially the history of specific regions, ethnic groups, local communities and families, is also widespread, but has been hampered by the lack of easily accessible source material. Recent years have seen a great increase in the number of Native Americans and African-Americans searching for their roots, but traditional genealogical sources are not as helpful to groups whose ancestors suffered subjugation and dispersal. Collections such as “Southeastern Native American Documents, 1763-1842” can fill this void with a generous supply of specific detail. One letter describing slaves allegedly "stolen" by Native Americans, for example, includes the names of all the slaves and their owners, as well as a physical description of each slave.

Because “Southeastern Native American Documents, 1763-1842” will be a public database in Georgia’s GALILEO system (please see http://www.galileo.peachnet.edu), available to all World Wide Web users, we anticipate that the audience for original documents concerning Native Americans will expand from a small cadre of scholars to a wide range of interested viewers of all ages and backgrounds.

Ownership, Privacy and Copyright

With an ending date of 1842, it is doubtful that free distribution of this collection will infringe on the privacy of any contemporary individual. In the period covered by this project, however, it is difficult to tell what is and what is not, official correspondence. This is important because the writers of private letters retain the copyright until cleared by current law, while the writers of official ones do not. The individuals represented in this collection almost all held official posts of one kind or another, so we have to determine on a case-by-case basis whether the documents in question seem to be preponderantly personal.

Preparation and Preservation

About three-quarters of the documents in this collection are handwritten. Although the paper has yellowed and the ink has bled through in some of the oldest documents, their overall condition is quite good, since they are generally on rag rather than pulp paper. Some have been encapsulated, and most are in acid-free storage files and boxes. Individual items are pulled from their respective collections and assembled as one unit for this project. Those items not already in containers are placed in the proper containers. All of this work takes place within the manuscripts area of each participating organization. Scanning is performed on a flatbed scanner or an overhead scanner, whichever better minimizes harm to the documents. Transcriptions are prepared using the scanned images to avoid any unnecessary handling of the originals. After scanning is complete, selected portions of the University of Georgia collection will be microfilmed for long-term archival storage.
Work Process and Standards

Every page of every document is scanned at 300dpi, in color with a color depth of 24 bits. After many experiments with handwritten manuscripts, we have determined that using this level of resolution and scanning in color most faithfully reproduces the original. In addition, a 24-bit color image scanned at 300 dpi generally has a better readability in manuscript materials than a grayscale image of the same dpi.

The writing on the vast majority of the documents is quite large, and there are no fine details in the handwritten pieces. Large format scanners can scan originals up to 11 X 17 inches (most documents in this project are smaller than this). Pages exceeding 11 X 17 inches are scanned in sections and pieced together using Adobe PhotoShop.

The original scans are LZW compressed TIFF (Tagged Image File Format) format files, averaging 12 to 20 megabytes in size. The LZW compression is a lossless compression, and the only image manipulation of the TIFF images is rotation and cropping. The high-resolution TIFF images will be retained by the respective institutions for archival purposes.

From the TIFF images, we produce 300dpi DjVu images for use by the public. DjVu is a file format developed by AT&T and now supported by Lizardtech (http://www.lizardtech.com). A sample of images in this format is available at: http://www.libs.uga.edu/darchive/hargrett/bits/native/creek/creek.html. Even though the images will be very large, the DjVu viewer allows the user to zoom in or out on each image so that the whole image can “fit” on a computer screen, or a small portion can be “zoomed” into to see detail. Because access to the information on the images will be through the full text searching of their transcriptions, it will not be necessary to provide thumbnails of the page images. We use Adobe PhotoShop for all image manipulation and the DjVu software package for conversion from TIFF to DjVu format.

Every document is transcribed into an SGML (Standard Generalized Markup Language) document with pointers to the DjVu images of the original documents. SGML markup not only makes it possible for every word of text to be searchable, but also provides for more sophisticated searching strategies as described below.

Each item is also cataloged as a Dublin Core, CORC record. Cataloging of each item is done according to accepted standards and practices, and the resulting records reside in the OCLC CORC master file, accessible for searching to users at all participating OCLC institutions. Bibliographic elements from the cataloging records are included in the headers of the SGML documents so that they may be searched along with words in the document texts.

All the documents will be mounted together as a single fully searchable text database on Georgia’s statewide GALILEO system. Although the commercial databases on GALILEO are not available to locations outside of Georgia, several non-commercial databases are universally available, such as the Georgia Government Documents database produced by the University of Georgia Libraries (also accessible at http://www.galileo.peachnet.edu). “Southeastern Native American Documents, 1763-1842” will be a universally available database.

As a fully searchable text database, this collection will be searchable by any word or phrase found in any of the documents, including the information found in the headers of those documents, such as author, title, and subject headings. By using SGML, searching can also be done by data type. For example, a user could search for a surname only within correspondence, or for a place name only within treaties. Search results will be available as both DjVu images of the original handwritten documents and the SGML transcriptions.

A test database containing many of the documents digitized in the first phase of “Southeastern Native American Documents, 1763-1842” is in place now. Following a final review, this database will go into production through Galileo before the expiration of the funding period, with more documents being added as transcription and SGML markup are completed. The availability of this database will be widely announced through existing channels established for other GALILEO products as well as in national
forums such as the NINCH (National Initiative for a Networked Cultural Heritage) and Imagelib electronic lists.

**Sustainability**

The original documents used to create “Southeastern Native American Documents, 1763-1842” will of course be retained in the special collections where they now reside. In addition to digitizing these materials, we have also preserved most of them on microfilm for long term archival retention. Long-term retention of the TIFF images and SGML files created in this project will be on CD-ROMs. As with all of our data retained on CD-ROM, we will sample 20% of our CD-ROM collection each year and re-record any questionable CDs. Also, as new, more permanent storage technology becomes available, we will migrate data to that media. As regards the database, all data on the GALILEO servers is backed up daily by an automated tape system.

**Evaluation**

“Southeastern Native American Documents, 1763-1842” will be subject to the extensive review and evaluation process required for all GALILEO databases. Each new database is first tested and evaluated by project staff and the GALILEO Director of Virtual Library Development for design and functionality. It is then announced as a demonstration database to the GALILEO Reference Subcommittee composed of reference librarians from university, school and public libraries throughout the state. This group reviews the prototype database with their varied constituencies in mind, consulting other professionals as appropriate and making recommendations for adjustments and improvements. At this point the internal review is complete, and the database is announced and made available to all GALILEO users.

After a new database goes into production there are additional avenues for ongoing evaluation. Every major GALILEO screen includes a “leave a comment” button which any GALILEO user can employ to communicate with GALILEO central-office staff. Feedback concerning “Southeastern Native American Documents, 1763-1842” will be routed back to project staff at the University of Georgia and its partner institutions for their consideration.

In addition to this direct method of evaluation, GALILEO also collects use statistics and conducts annual user surveys. This information is available to all GALILEO users at http://www.peachnet.edu/galileo/evaluate.html and will be used by project staff to monitor the amount and type of use received by “Southeastern Native American Documents, 1763-1842” database.