CSS Alabama *Digital Collection*: A Special Collections Digitization Project

Andrea Watson, with P. Toby Graham

Abstract

In today's high-tech information environment, archivists and special collections librarians can use technology to enable the informational contents of rare and unique materials to transcend their physical boundaries. To explore this possibility, a team in the William Stanley Hoole Special Collections Library at the University of Alabama sought to create access via the Internet to a small set of the Hoole Library's materials relevant to the famous Confederate raider, the CSS *Alabama*. The creation of the *CSS* Alabama *Digital Collection* served as a learning experience for the staff in the digital imaging of a wide variety of special collections materials and resulted in an electronic learning resource for students of all ages. To date, the website has logged over three thousand hits and has evoked highly positive responses from users, proving that the World Wide Web is an effective access medium for special collections materials.

Introduction

In the past, the scarce or unique character of materials housed in archives and special collections libraries has rooted them in terra firma. In today's high-tech information environment, however, archivists and special collections librarians can use technology to enable the informational contents of these materials to transcend their physical boundaries. To explore this possibility, a team in the William Stanley Hoole Special Collections Library at the University of Alabama sought to create access via the Internet to a small set of the Hoole Library's materials relevant to the famous Confederate raider, the CSS *Alabama*. This pilot digitization project to create the *CSS* Alabama *Digital Collection* is the University of Alabama Libraries' first attempt in developing an original electronic information resource from rare and unique items among its holdings.

Let us consider for a moment the traditional approach to research requiring the examination of manuscripts and other primary sources, scarce or

unique texts, images, and maps, in sum—the stuff of which special collections are made. First, a scholar has to locate materials that relate to his topic. Published reference tools are useful, as are footnotes and bibliographies in scholarly publications; and sometimes word-of-mouth along the scholarly grapevine provides valuable leads to materials in far-flung libraries and archives. Seasoned researchers also understand that because these special materials are usually irreplaceable, they cannot be allowed to circulate as ordinary library materials do. Consequently, scholars must conduct their investigations not only on-site but according to institutional rules and within prescribed hours. This requires a carefully planned trip, taking all these factors into account while attending to sundry personal logistics.

Once inside the library, the careful scholar who wishes to do an exhaustive study of a given topic will often find it necessary to examine items in many different formats, such as print, manuscript, microform, sound recording, etc. using any devices necessary to access the information therein. These complexities weary even the most determined and skilled researchers. Clearly, electronic access to special collections materials can overcome these physical impediments to research.

Genesis of the Project

In the summer of 1995 the Alabama Power Foundation made a contribution to the University of Alabama's Capital Campaign Fund in the form of a grant in the amount of $44,500 to the University of Alabama Libraries. The terms of the grant stated that it should be used for the specific purpose of exploring the potential of information technology to provide electronic access to library materials. The money was shared among three library interest groups, and from it the *CSS* Alabama *Digital Collection* project received $4,254 for equipment, plus the salary of a graduate assistant working twenty hours per week for two years.

In July 1995 James Kuhlman, associate dean of libraries for access services, and Hoole Special Collections Library staff met to define the objectives for the project. First of all, it should serve as a learning experience for staff in the digital imaging of special collections materials. Because of the current emphasis in the state of Alabama on improving primary and secondary education, the group felt that the subject matter of the project should be accessible and appealing to grade school and high school students as well as those more advanced. In addition to creating an electronic learning resource for students of all ages, the variety of materials to be digitized should promote public awareness of special collections libraries and the many types of materials they hold. Responses to this pilot project should be used to evaluate the World Wide Web as an access medium for digital collections and provide

guidance for further efforts toward the creation of a virtual library. Finally, widening access to rare and unique special collections materials should be the project's ultimate and ongoing goal.

As a topic that would encompass a variety of material formats and have the necessary wide appeal, reference librarian Andrea Watson suggested the Confederate raider, the CSS *Alabama.* Reference records over the past four years showed that queries about the ship have come from school children, history buffs, and professional modelers throughout the United States, Great Britain, and Canada. Materials relevant to the ship's history receive considerable on-site use, and requests to borrow them, which must be denied, are painfully frequent. Interest in the ship had been rekindled some months earlier by the appearance in the December 1994 *National Geographic* of a feature article describing diving explorations of it off the coast of Cherbourg, France.[1] Watson knew that young students were fascinated with the ship, having recently answered a reference query from a schoolboy, William Owen Perry, of Birkenhead, England, where the ship was built. More than anything, Watson's work with this particular foreign student had convinced her of the need for a web resource of CSS *Alabama* information, and she was able to make a strong case for it. Other members of the Hoole Library staff were equally well aware of the interest in the ship and the library's extensive holdings on the topic, including plans and color images, that could make a visually exciting website. The CSS *Alabama* became the immediate and unanimous choice of the group as the subject of the project.

Clearly, few ships in recent history have captured the imaginations of so many for so long as has the CSS *Alabama.* Built in secrecy for the Confederacy in the Birkenhead shipyards of John Laird Sons and Company, the *Alabama* became the subject of controversy even as her keel was laid. The Union did not take kindly to this expression of British sympathy for the cotton-producing South, and much diplomatic subterfuge was required to complete and launch "290," the *Alabama's* first *nom de guerre.* Subsequently she would be called *Barcelona* and *Enrica* until her real identity as the *Alabama* became known.

Bark-rigged and steam-powered, the 220-foot cruiser *Alabama* was afloat on the high seas by the summer of 1862. Sailing as far west as Galveston, Texas, and as far east as Viet Nam, the *Alabama* harried Yankee traders and took scores of prizes until she was finally sunk by the USS *Kearsarge* off the French coast near Cherbourg in June 1864. In all, she destroyed nearly sixty American merchantmen and released several others on ransom bonds. The exploits of the *Alabama* and other Confederate raiders resulted in a crippling of the American merchant marine that took decades to overcome.

[1] Max Guerout, "The Wreck of the *Alabama*: Avenging Angel of the Confederacy," *National Geographic* 186 (December 1994): 67–83.

The *Alabama*'s brief but brilliant career has been well documented for over a century. Accounts of the *Alabama*'s construction and actions appeared in contemporaneous news sources such as *Harper's Weekly*. The memoirs of her famous captain Raphael Semmes and the reminiscences of her officers John McIntosh Kell and John Low are firsthand chronicles of the *Alabama* at sea. Official documents of the Confederacy provide further historical data. Professional historians have ruminated at length on the *Alabama* and her exploits in scholarly articles and monographs. Guerout and others have created works of more popular appeal. The sheer beauty of the *Alabama*'s bark rigging inspired artists of her era and continues to fascinate modelers today.

This persistent interest in the CSS *Alabama* has generated a body of knowledge in various formats that resides in discrete areas within the Hoole Special Collections Library. Many of these sources are published books, but there is additionally a particularly significant group of unique manuscript materials. Other relevant items are drawn from the newspaper, map, sheet music, image, and artifact collections. As objects they constitute a microcosm of special collections materials, and project planners believed that the scanning and digitization of their many formats would prove technically instructive as well as serving as a contribution to scholarship. Planners agreed that their priorities for materials to be included in the *CSS* Alabama *Digital Collection* would be first, unique and rare documents from manuscript holdings, images of the ship, its plans and personnel, relevant items from contemporaneous published sources; and second, as many copyright-free monographs as time and energy might permit.

A description of core materials in the Hoole Library for the *CSS* Alabama *Digital Collection* follows.[2] Of signal importance among these are the William Stanley Hoole Papers, which include seven folders containing source material for Hoole's edition of the diary of Edward Anderson, one of the *Alabama*'s officers.[3] There is also a prize artifact—a leather-bound album containing fifty *cartes de visite* of the CSS *Alabama* crew, friends, and associates made in London or Liverpool during the summer of 1864 and presented to Anderson in Liverpool on July 22, 1864. Other folders contain a transcript of Lieutenant (later Captain) John Low's *Reminiscences* and newspaper stories about the cruise of the *Alabama*, more clippings about Low from a Liverpool paper of 1904, the *Cape Argus*, 1933, the *Illustrated London News*, 1863, and others. There are photographs of Captain Raphael Semmes and Low, and Low's pistols, coat, and grave monument. A final folder contains typescripts of a

[2] The authors are indebted to technical archivist Clark Center for the thorough survey of CSS *Alabama* materials in the Hoole Library's manuscript collections.

[3] Edward Clifford Anderson, *Confederate Foreign Agent: The European Diary of Major Edward C. Anderson*, edited with a prologue and an epilogue by W. Stanley Hoole (University, Ala.: Confederate Publishing Co., 1976).

Union agent's report on activity in the Laird shipyards and of other spy reports on "290."

Another invaluable set of source documents are Low's manuscript logs kept on the CSS *Alabama* and the CSS *Tuscaloosa.* This collection also includes photographs of the masts and rigging of the CSS *Alabama* and of the shipyard model used by Laird.

Miscellaneous materials relating to the CSS *Alabama* from the Hoole Library's collections are two pieces of sheet music, three colored engravings of the ship at sea, news items and political cartoons from *Harper's Weekly* during 1862-1864, and a full-scale copy of the plans of the ship. Because the team wished to experiment with digitizing as many different image and textual formats as possible, they chose to include a broad sample of these materials as well as items from the manuscript holdings in the *CSS* Alabama *Digital Collection.*

Although several copyright-free monographs relating to the ship also reside in the library's collections, digitization of these proved to be beyond the scope of this pilot project primarily because of the failure of optical character recognition software (see discussion under "Technical Considerations" below).

Armed with an exhaustive list of the Hoole Library's holdings on the CSS *Alabama* and charged with digitizing as great a variety of materials formats as possible, P. Toby Graham, a doctoral candidate in the University of Alabama's School of Library and Information Studies who also holds a master's degree in American history, designed and constructed the *CSS* Alabama *Digital Collection,* calling upon his principal advisers, technical archivist Clark Center and reference librarian Andrea Watson for help as needed. The collection is coded in HyperText Markup Language (HTML) and is available over the World Wide Web at <http://www.lib.ua.edu/hoole/cssala/main.htm>.

The website has three principal parts: an image gallery, a documents page, and an innovative "virtual journey" image map. The image gallery provides a list of subject categories, *e.g.*, "The *Alabama* at sea" and "The Battle with the *Kearsarge.*" Each category in turn has an individual page that includes a list of image files accompanied by small, low-resolution thumbnail images. Because image files are often large and slow to transfer over the Web, the thumbnail images help the user make intelligent choices about which full-size images are worth the wait.

The documents page lists primary source material that has been either entered as computer-readable text or scanned as an image. Computer-readable text is preferable from the standpoint of file size, legibility, and searchability. Team members felt, however, that images of primary documents serve an educational end for young students who have never seen a primary source document: by illustrating their physical appearance they convey a sense of

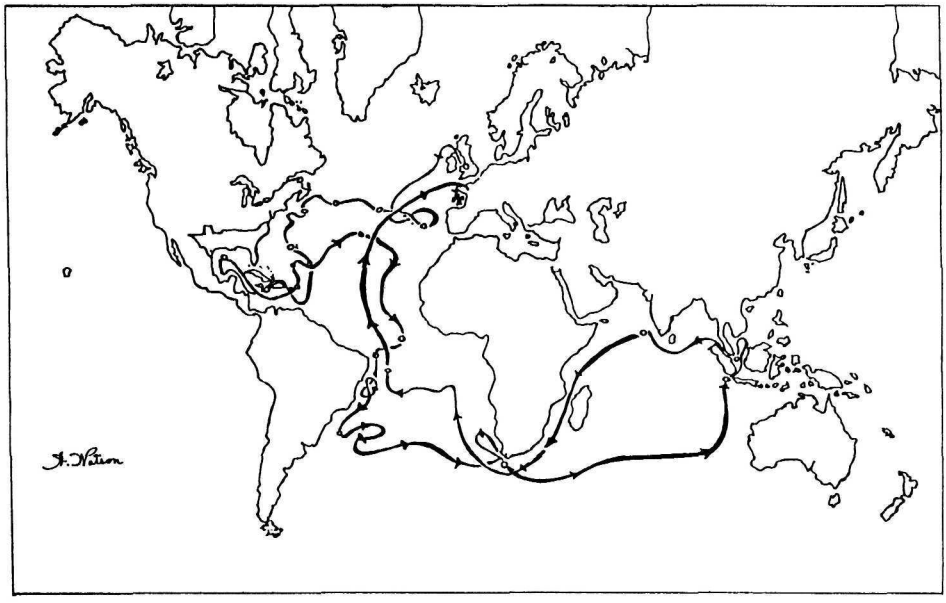

FIGURE 1: The Cruise of the CSS *Alabama* 1862–1864 preliminary pen-and-ink drawing used for generation of clickable map.

their significance as unique objects that are invaluable to serious scholars. For this reason, document images of the log book of John Low, a unique primary source, are included, with textual transcripts accompanying the images.

Developing the "Virtual Journey"

The most unusual feature of the CSS *Alabama* project is its "virtual journey" page, an exciting experimental method of access that the team hoped would prove appealing as well as geographically instructive, especially to younger users. Because the collection lends itself readily to geographical division, Graham and his advisers decided to depart from the traditional method of access to embedded information via an HTML menu and employ a clickable map. Users sit before an on-screen color image of a world map marked with the circuitous route of the CSS *Alabama*, from its launch in 1862 at Birkenhead to its destruction two years later off the French coast (see Figure 1). The route is punctuated with red dots marking the vicinities of the ship's various engagements. Users can navigate the route, clicking on the dots to reveal linked log entries, newspaper reports, historical accounts, and illustrations that correspond to events that occurred in the area. For example, by clicking on the Birkenhead dot the user can choose to examine Union spy reports, the construction plans for the *Alabama*, or an image of the Laird

shipyards. This is a cogent illustration of the multiplicity of material formats and creative perspectives that can document a single event.

Because the clickable map is one of the more interesting features of the Web page, it is appropriate to give a more detailed account of how it was developed. In his peregrinations along the World Wide Web, archivist Clark Center saw clickable maps used as gateways to geographical information. This, along with finding maps of the cruise of the CSS *Alabama* among the manuscript collections chosen as the basis of the *CSS* Alabama *Digital Collection*, inspired him to suggest the creation of the "virtual journey" as an adventurous approach to organizing information about the ship's exploits. Website designer Graham and team adviser Watson were immediately taken with the idea. Unfortunately, those maps already in the library's collections were unsuitable for the website; they were either too detailed with copious geographical annotations or had superimposed images in critical locations. Watson, who pursues graphic design as a hobby, decided to take on the challenge of creating an original graphic of the cruise of the CSS *Alabama.*

For the sake of authenticity, her first step was to find and reproduce an accurate world coast line for the Civil War era. The Hoole Library's rare book collection yielded one usable item, *Johnson's New Illustrated (Steel Plate) Family Atlas* (New York: Johnson and Ward, 1864). This contained a map of the world spread over two facing pages. These were photocopied, pieced, and the coast line traced, omitting all extraneous markings on the original map. Next, using a number of maps of the ship's travels from various sources as rough guides, Watson drew the ship's voyage on the map. She and Graham then decided where to place the red dots that would serve as links to the documents relating the ship's activities in certain vicinities. Because the CSS *Alabama* took more than sixty prizes, several of these events were grouped on a single dot. The ship's launch at Birkenhead and her sinking at Cherbourg were accorded individual dots. When the pen-and-ink drawing was finished, Watson scanned it as a sharp black and white tiff image. She then used Photoshop software to clean up the outlines, create captions, and add color.

Once the graphic was in its final form, building an image map from it was "a fairly simple process" according to Graham. The first step was to convert the image from a tiff file to gif, a file type in common use on the Web. Then, using an image map editor, the HTML author created a map file, selecting the coordinates around each red dot along the route of the *Alabama* as a hot-link. A central Web page couples the map file with the image file. The result is the *CSS* Alabama *Digital Collection*'s "virtual journey" that provides geographically based access to thirteen images and forty-four text files associated with the Confederate raider's cruise.

Technical Considerations

Generous grant funding from Alabama Power allowed the purchase of a state-of-the-art digital imaging work station. The CPU, a Dell Pentium 100 with 32 MB of RAM and a 500-MB SCSI hard disk drive, is equipped with a graphics driver capable of 24-bit color and a 17-inch monitor. Text and images were digitized using a Hewlett-Packard ScanJet flatbed scanner with a sheet feeder. The overall cost of the work station was approximately $4,500. Graham found the work station to be fast and reliable. The monitor and graphics driver produce high quality displays that delight the eye.

Graham used Corel Photo-Paint to manipulate scanned images with great ease and success. Team member Watson, who created the original graphic for the clickable map, preferred Adobe Photoshop software to sharpen and color her original pen-and-ink drawing. Both packages have similar functionalities, so personal preference was clearly the operative factor in their choices.

The *CSS* Alabama *Digital Collection* project certainly proved instructive to its participants as they confronted a number of technical issues. Decisions regarding the conversion of documents to machine-readable text, the synergy between preservation and access, speed of information transmission, image resolution and size, and the possibilities and limitations of the World Wide Web had to be carefully considered and addressed.

Keeping file sizes to a reasonable minimum was a guiding principle, since large files dramatically reduce access speed on the Internet. To that end, all seventy-five scanned images are jpg condensed files. A few gif files do appear, however, because Graham saw no need to compress stock symbols such as backgrounds, lines, and bullets, and the HTML author used to select coordinates around each red dot required that the map be in gif format. The average image file size is approximately 36K. All told, the website contains 252 image and text files totaling some 6.5 MB.

The major disappointment in the course of the project was the failure of optical character recognition (OCR) software when applied to special collections materials. Initially project planners had hoped to use OCR to digitize at least a few copyright-free monographs from the Hoole Library's rare book and Alabama collections in order to give the website a small full-text delivery capability. A low-end OCR package was totally inadequate when tested on the typefaces of Civil War-era imprints. Grant monies allowed for the purchase of a high-end, "teachable" OCR package, but it too proved unequal to the task of converting the uncommon and infinitely flawed typefaces of the nineteenth century (Figure 2). Once team members saw this, they realized that text from older items would have to be keyed-in manually. This seriously curtailed the number of full-text items that could be added to the collection. Of the forty-seven books about the CSS *Alabama* in the Hoole Library's col-

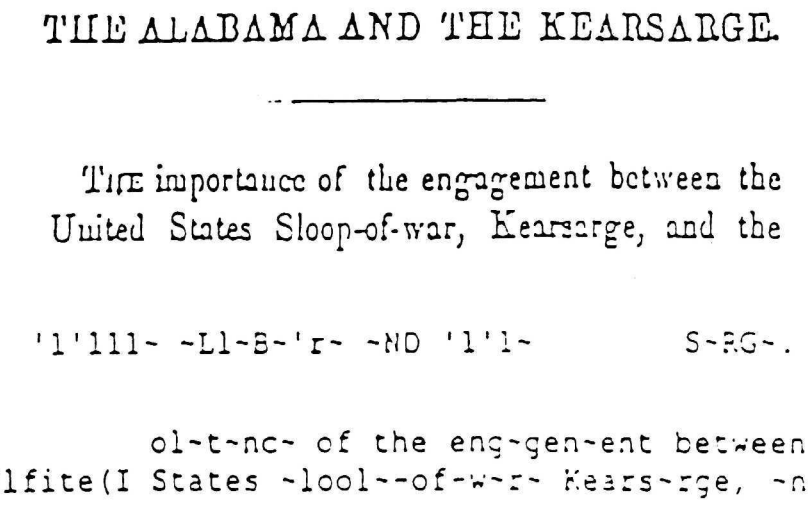

FIGURE 2. The bottom three lines illustrate an OCR attempt at reading nineteenth-century text as found in *The Alabama and the Kearsarge*, by Frederick Milnes Edge (London: W. Ridgway, 1864).

lections, twenty-three are English-language monographs published prior to 1920 and now free of copyright restrictions. These obvious candidates for digitization total approximately five thousand pages. The failure of the OCR software packages at our disposal meant that these books had to be excluded from the *CSS* Alabama *Digital Collection.* Further investigation of OCR versus other, more accurate means of creating full-text digital files,[4] must be pursued to determine the feasibility of developing full-text delivery components in future websites.

The Hoole Special Collections Library is guided by a mission that includes both preservation and access. In practice Hoole faculty and staff take their responsibility for the preservation of the physical vehicles of information stored in their several collections very seriously. Some among us, truth be told, side with skeptics who are still not satisfied that electronic media are durable and flexible enough to be considered a sound preservation choice. This accounts for the definition of the project's goal of providing access to special materials rather than using digital reformatting as a preservation technique. Yet while its digitized form does not replace the original object, it does preserve it indirectly by functioning as a surrogate that can save the object from handling. The digital imaging process may also prompt encapsulation of fragile originals prior to scanning. As access to digital images increases the demand for photographic copies, negatives and prints will be created. In the end, digitization, though it does not solve the problem of physical storage,

[4] William Saffady has published a set of tables with instructions that, accurately completed, will give a conservative estimate of the costs of digitizing a given set of materials (*Library Technology Reports*, May/June 1995).

results in preservation of the physical object as well as access to its intellectual content.[5]

This emphasis on access rather than preservation had another significant advantage; it freed the designer from having to adhere to the 600-dpi (dots per inch) resolution standard that is the accepted minimum for preserving images. The large files created at this level of resolution, even when compressed, make a web resource sluggish and inefficient. Scaling images in the *CSS* Alabama *Digital Collection* to the 300-dpi resolution standard for access keeps the website's response time at an acceptable rate.

Conclusions

Despite the problems involved, the participants were able to accomplish the stated objectives of the project and complete the website within two semesters. We were fortunate to have grant funding in support of the project, as well as helpful resources readily available in other units within the University of Alabama (for example, a free training course in HTML offered in the University's Seebeck Computer Center). Preliminary responses to the *CSS* Alabama *Digital Collection* have been extremely positive. Particularly gratifying has been the continuation of the Birkenhead—Tuscaloosa, Alabama, correspondence. In June 1996 William Owen Perry wrote to Andrea Watson, "Thank you also for your very useful information about the *CSS* Alabama *Digital Collection* on the Internet. Yesterday I had the opportunity of using the college's Internet system with the help of one of my tutors. I am delighted with the amount of interesting information which I found there. You asked about my reaction to it. I think the concept of your collection is brilliant, the amount and variety of information is excellent. The time it took for the graphics to load was a matter of minutes and the tutor accessed it with no difficulty." Perry did go on to say that his one criticism of the website was that the images were blurred and some were incomplete. This points up a home truth that all web designers must live with: no matter how much care one takes to generate high-quality images, one has no control over the capabilities of the receiver's equipment.

The magic of the Web lies not only in the access to information it provides, but also in the connections it can make among discrete texts and im-

[5] For a discussion of issues on preservation and access in a digital environment, the following works are useful: Paul Conway, *Preservation in the Digital World* (Washington, D.C.: Commission on Preservation and Access, 1996); Kristen L. Garlock and Sherry Pointek, *Building the Service-Based Library Web Site* (Chicago: American Library Association, 1996); Ross Harvey, *Preservation in Libraries: Principles, Strategies, and Practices for Librarians* (London: Bowker-Saur, 1993); Michael Lesk, *Image Formats for Preservation and Access: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access* (Washington, D.C.: Commission on Preservation and Access, 1990); Picture Elements, Inc., *Guidelines for Electronic Preservation of Visual Materials: Study for the Library of Congress Preservation Directorate* (Washington, D.C.: Picture Elements, Inc., 1995).

ages. Participants in the *CSS* Alabama Digital Collection project believe that its greatest success lies in using this unique Web capability to create a highly interactive and flexible educational experience. By offering a variety of materials in digitized form, it accommodates diverse learning styles. A letter from a young midshipman is infinitely more meaningful when it is linked to his photograph, to an image of his ship, and to other descriptions of his experiences than when it stands alone. By taking full advantage of the strengths of HTML, most notably in the "virtual journey" page concept, the collection encourages user interaction through visual appeal. Compared to massive projects such as Yale University's "Project Open Book," which plans to digitize ten thousand volumes from preservation microfilm, or the Library of Congress's multiplicitous "American Memory Project," the *CSS* Alabama *Digital Collection* is quite small. Nevertheless, by focusing on an interesting topic for which the Hoole Library holds a strong and varied set of documentation, it succeeded in meeting its goals of providing wide access to rare and unique information resources through technology and increasing awareness of the library's holdings. Additionally, the experience gained by team members during the creation of the *CSS* Alabama *Digital Collection* will inform choices and procedures for future digitization projects in the Hoole Special Collections Library.