# Knowledge for the Ages: Preserving Rare Collections through Digitization and Content-Based Access

## Constantia Constantinou

## Contributors: Shafeek Fazal; Joseph Williams[2]

Maritime College, The State University of New York

**Abstract.** The mission of the SUNY Maritime College Library Archives is to preserve, protect and secure long-term access to its rare collections. The SUNY Maritime College Library aims to promote and provide access to rare collections by providing access to researchers, faculty, students and the public throughout the world. The rare historical collections housed in the SUNY Maritime College Library serve a variety of purposes, providing legal, historic and evidentiary value to individuals and society as a whole. The paper, "Knowledge for the Ages: Preserving Rare Collections through Digitization and Content-Based Access" describes the methodology for recruiting and acquiring rare historical documents, the technology standards, equipment and techniques used for digitizing the collections, the metadata and cataloging standards for making the historical collections searchable and accessible to researchers across the world as well as issues of storage and access.

**Keywords:** Rare Collections; acquiring historical collections, preservation; content-based access; digitization; metadata;

## 1. Introduction

The study "Knowledge for the Ages: Preserving Rare Collections through Digitization and Content-Based Access" demonstrates the work of the Stephen B. Luce Library of SUNY Maritime College in its efforts to preserve and promote its rare collections by providing access to researchers throughout the world.

Maritime College is one of the 64 colleges and universities of the State University of New York. A four-year college is located at the historic Fort

Schuyler in New York. SUNY Maritime College[1] offers undergraduate and graduate degrees in science, engineering, maritime transportation, business, and the humanities. The Stephen B. Luce Library of SUNY Maritime College is the oldest academic maritime library in the United States. The Luce Library is a highly specialized library in maritime collections; its holdings represent a well rounded collection of over 100,000 volumes with special strength in maritime engineering, naval architecture, marine transportation, oceanography, transportation economics and management.

## 2. Special Collections

The Stephen B. Luce Library[2] is nationally and internationally known for its rare collections and archival documents in maritime history dating as early as 1700's. The scope of these collections include journals and proceedings of professional maritime organizations, such as the Marine Society of the City of New York; genealogical collections of journals reports, daybooks and personal letters, company documents of shipping companies and cruise liners, as well as detailing the history of ships and maritime activities during World War II. Most important are the collections pertaining to genealogical records of mariners and passenger manifests of ships arriving at the New York City harbor early 19th century. New York City harbor was one of the busiest ports in the United States during the major wave of immigrants arriving from Europe.

This study focuses on one the library's major collections, the Sailor Snug Harbor collection. The study describes the Library's efforts to acquire, treat, preserve and make the collection accessible to the scholarly community in the United States.

## 3. Sailor Snug Harbor – Background

Sailors' Snug Harbor was founded in 1801 upon the death of Captain Robert Richard Randall. Captain Randall created a family fortune as an entrepreneur, privateer, and patriot in the United States. His bequest established an institution "for the purpose of maintaining and supporting aged, decrepit and worn-out sailors." This institution became firmly established on Staten Island with a site that overlooked Upper New York Bay and the Kill Van Kull on Staten Island. On this site, the architect Minard Lafever designed a complex in the Greek Revival-style that would accommodate 200 seamen. Since that time, the complex expanded until by the early twentieth century there were more than 1,000 residents, or inmates as they were called, living at Sailors' Snug Harbor.

The men who resided at Sailors' Snug Harbor came from a wide variety of backgrounds both ethnic and racial. During the mid-nineteen century, approximately one-third of the sailor's were foreign born and throughout its history, Sailors' Snug Harbor had African American residents beginning in 1842. Despite their varied backgrounds, the sailors of Snug Harbor had a common heritage. All had lived hard lives

as common seamen which often resulted in permanent physical injury or other problems such as alcoholism.  Each one of these men and their records, offer a unique glimpse of the history of maritime sailors and their culture, health care and medical treatment administered on the premises, turning these records scholarly gems for historians, sociologist and anthropologists.

## 4.   Sailor Snug Harbor – Collection

The collection consisting of approximately 27,600 pages of documents belonging to these mariners are as diverse as their background.  Genealogical materials, such as birth certificates, licenses, and citizenship papers are prevalent, but also extant are medical records, photographs, and scrapbooks. Each one of these items has unique historical significance in detailing the social history of American mariners.

In 1976 the Stephen B. Luce Library of Maritime College, SUNY became the sole owner of the information archives of Sailors' Snug Harbor.  The collection, covering the period 1797 to 1969 consists of documents and records representing a fairly complete picture during this time period.  These materials further document the activities of the Harbor and their residents.   These materials are wide in scope ranging from the business records of the institution to registries to scrapbooks.  Of these records, there is a subset consisting of the official and personal records of the inmates.  These materials date from the late nineteenth century until the mid-twentieth century.   These records are of historical significance because they consist of citizenship certificates, professional licenses, medical records, photographs, and scrapbooks.

### 4.1 Sailor Snug Harbor – Research Value

The Sailors Snug Harbor collection inspires high interest because of their uniqueness and how they complement other Maritime collections.

The library has been receiving many requests for access to the Sailors' Snug Harbor archives by researchers in the fields of Maritime and New York City history, and genealogy.  Researchers in the humanities who undertake research that documents the social history of the Maritime world benefit greatly by the access that digitization provides.  The collection remains unique because aside from the collections at other maritime institutions there are few fully digitized maritime collections extant.

### 4.2 Sailor Snug Harbor – Digital Project

Recognizing the value of the Sailors' Snug Harbor collection, the Stephen B. Luce Library staff decided to pursue the avenue of digitization to increase access and awareness and also to preserve the documents by minimizing handling of the physical collection.  In 2006, the Library laid the groundwork for the Sailors' Snug Harbor Digital Collection by acquiring a small grant to digitize a sampling (approximately 500 pages of documents) of the Sailors'Snug Harbor collection to illustrate the founding of the institution, the daily activities, and the subsequent disposition of the property, including photographs and newspaper clippings on the history of the institution.  This grant was funded to

purchase equipment, software, and hire staff to digitize 500 pages of documents. All documents are uploaded to CONTENTdm management system and are fully searchable with appropriated metadata.

The digitized portion of the collection served as a complement to the entire collection, increasing awareness to researchers and the public in general of the type of original documents that exist in the Sailors' Snug Harbor archival collection [3].

In addition, the value of this collection to New York State's history has been recognized by the New York Heritage organization which has added the Sailors' Snug Harbor digital collection to its repositories of collections on New York State's people, and places [4].    In an effort to expand access and improve on searching the Sailors' Snug Harbor Digital collection, the Library collaborates with other New York library organizations to upload metadata and thumbnail images for all digital records into the heavily accessed OCLC Worldcat catalog.

## 5.   Methodology and Standards- Digital Project

As funding continues to become available through grants and private donations the library digitization team, continues to create digital images for each of the 9,200 personal records, the Governors' journals which consist of 50 bound volumes, and 1,000 pages of records pertaining to the physical grounds and properties of the Sailors' Snug Harbor. Using the specifications outlined in the BCR's CDP Digital Imaging Best Practices Version 2.0, the team either scan in or digitally photograph each document. For the most part, the entire collection could be scanned in safely using the Epson 10000XL flatbed scanner. The Epson 10000XL has been used with great success in many large scale historical digitization projects.  However, approximately 15% of the Sailors' Snug Harbor documents have been deemed too fragile for scanning, including most of the bound Governor's journals. For fragile documents a DSLR (digital single-lens reflex) Canon EOS camera is used for digital imaging.  Carefully, case-by-case decisions are made about the digitization of these fragile documents.

Records are scanned or photographed in color to show the shade and condition of the document. Three images are created for each record, a master image, an access image, and a thumbnail image. Master images are in Tagged Image File Format (TIFF) with maximum resolutions of approximately 1200 dpi for typewritten documents and approximately 2000 dpi for handwritten documents. Handwritten documents, about 10% of the collection, will be scanned or photographed at a higher resolution to allow for easier reading. Also, two copies of the master images will be saved, one copy on a network hard drive and another backup copy on DVD-R.  The network drive is specified as two terabyte network access storage (NAS) drive.  The drive is easily shared on a network and the original images stored and made accessible on the Library's public computers.  This will facilitate the need for anyone requesting higher resolution images.   Access images are saved in JPEG format with a resolution ranging

between 150 - 300 dpi, depending on text size and matching required image file size for comfortable viewing. Thumbnail images, in JPEG format with an approximate resolution of 72 dpi, are created to use as links to the larger access images.  All image editing is done using Adobe Photoshop software.

The metadata and cataloging process is done almost simultaneously with the digital imaging. The metadata specialist, aided by the digital archivist and two library school interns, oversee this portion of the project. Metadata is created for each digital file following the latest standards of the Dublin Core Metadata Initiative and using Library of Congress Subject Headings vocabulary.  Each typed document is converted to a text file using Optical Character Recognition (OCR) software – ABBYY Finereader.  This permits searching the full text of the digital images.

## 6.   Storage of and Access to Content

As the project progress and more content is becoming digital, the 30,000 pages of digitized documents will eventually join our collection's 500 Snug Harbor images already housed on the METRO CONTENTdm server as part of the Digital Metro New York initiative [5].   Specific images can easily be searched through the CONTENTdm search interface. Digital Metro New York houses significant digital collections for member libraries from the New York City area. The Stephen B. Luce Library collections are also accessible through *NewYork Heritage.org* [6], a research portal for hundreds of digital collections about New York history.  In addition, the digital images are   uploaded to SUNY's Dspace repository [7].

The major public entry point for the Sailors' Snug Harbor Collection is a comprehensive web site devoted to the collection and hosted on the library's web server[8].  The site is designed by a web design consultant and includes a detailed description of the history of Sailors' Snug Harbor, the Sailors' Snug Harbor finding aid, a narrative outlining the documents, as well as, a list of all the documents in the collection. Most importantly, the site is prominently features a searchable database of names, for both the inmates and governor. Names are linked to the image of its respective digitized record stored on the CONTENTdm server. Persistent links to "search the digital collection" and to "browse all digitized documents" are also available on all pages of the site and these links, when clicked, are redirect users to the CONTENTdm search interface. The web site features historic images of Snug Harbor from the Snug Harbor Cultural Center and our own collection and a link to the Snug Harbor Cultural Center.

The main venue for access to the Sailors' Snug Harbor Digital collection is an open access web site available 24/7.  Another venue for access is the heavily used worldwide online public access catalog, OCLC Worldcat. All metadata records with thumbnail images are becoming available into the Worldcat

catalog. Higher resolution images are also becoming available for download onsite via the NAS network drive that stores all original images.

## 7. Conclusion

In the Library world, digitization is now recognized as the most effective method to preserve and promote access to rare collections. Undertaking a digital project poses several challenges – from seeking funding to having skilled team members to choosing the right equipment to following image and metadata standards to ensuring adequate access. The Sailors' Snug Harbor digitization project was a success as it assures these maritime treasures are well preserved, but most importantly, through the use of modern digital technology these treasures are as open and accessible as possible. With the success of the Sailors' Snug Harbor digital project, the Luce Library is now venturing into digitizing other rare collections, including the Marine Society of the City of New York records and the Maritime College institutional archives.

## References

1. SUNY Maritime College. http://www.sunymaritime.edu/

2. SUNY Maritime College Library. http://www.sunymaritime.edu/stephenblucelibrary/

3. Sailor Snug Harbor digital Collection.

http://cdm128401.cdmhost.com/cdm4/browse.php?CISOROOT=%2FMC

4. New York Heritage. http://www.newyorkheritage.org/browsecollections.php

5. Metro Digital Projects. http://www.metro.org/digital-collections/

6. New York Heritage. http://www.newyorkheritage.org/browsecollections.php

7. SUNY Dspace. http://dspace.sunyconnect.suny.edu/

8. Sailor Snug Harbor website at SUNY Maritime
   http://www.sunymaritime.edu/stephenblucelibrary/sailorssnugharbor/index.html