Digitizing Sanborn Fire Insurance Maps for a full color, publicly-accessible collection
http://www.lib.utah.edu/digital/sanborn/

By Kenning Arlitsch
Head of Digital Technologies
J. Willard Marriott Library
University of Utah
kenning.arlitsch@library.utah.edu

Introduction

From 1867-1970 Sanborn Fire Insurance Maps documented the rise of American cities, and their block-by-block detail has assured their popularity and importance to modern-day researchers. Held in the paper and microfilm collections of libraries and historical societies around the country, "the Sanborn Maps are of considerable value … not only for their age but also for their detailed portrayal of street patterns, lot lines, and the placement, size, shape and use of buildings."[1] The maps were so dominant in the insurance industry that *Fortune Magazine* in 1937 declared the Sanborn Map and Publishing Company "the only U.S. fire insurance map maker."[2] Aside from the intended audience of insurance companies, genealogists, demographers, environmentalists, urban planners, historians, and laypersons have all found value in the maps, and it is no surprise that attempts have been made in recent years to digitize them.

Existing Projects

In 1997 the Library of Congress announced a cooperative arrangement to scan one million maps held by the Library and Environmental Data Resources (EDR) Inc.[3], but the effort floundered when contractual differences arose. In 2001 ProQuest Information and Learning unveiled *Digital Sanborn Maps, 1867-1970*, a fee-based product providing access to more than 660,000 maps from 12,000 American towns and cities, scanned from the collection at the Library of Congress.[4] And in 2001, the Ohio Library and Information Network and the Ohio Public Library Network announced a collaborative project to provide 40,000 digital Sanborn Maps for the citizens of Ohio.[5]

While the ProQuest and OhioLINK/OPLIN projects are outstanding for the enormous number of maps they bring to the digital arena, they are limited by two aspects: 1) they restrict access to the maps by requiring fees or, in the case of OhioLINK/OPLIN, citizenship in a specific state; and 2) because the digital files were scanned from black and white microfilm, they fail to reproduce the important color features of the original maps.

Color codes represent the primary Sanborn method for identifying building materials. "Five tints distinguish each building's exterior material – olive for adobe, blue for stone, pink for brick, yellow for wood, and gray for metal."[6] Digital files for both the OhioLINK/OLPIN project and ProQuest's *Digital Sanborn Maps* product were created

from the Chadwyck-Healey collection. Chadwyck-Healey, which has been purchased by ProQuest, had in the early 1990s microfilmed the maps from the Library of Congress collection. Reviewing that microfilm collection in 1993, map librarian Glen Creason of the Los Angeles Public Library lamented that "Despite loud protestations from map librarians, Chadwyck-Healey chose to reproduce the atlases in black and white, and therefore the maps lose a considerable amount of data contained in the colors of the originals."[7] It is difficult to overstate the importance of color reproduction in the Sanborn Fire Insurance Maps.[8] It is, of course, also much more labor-intensive to scan each map from its original, rather than from the microfilm, but the project described in this paper outlines a method that could be duplicated by an institution in each state, thereby dividing the labor and possibly leading to a larger aggregated collection.

University of Utah Digital Sanborn Project

The J. Willard Marriott Library has created a digital collection of nearly 1,300 maps that overcomes issues of restriction and color, making the first (to my knowledge) publicly-accessible full-color digital reproduction of Sanborn Fire Insurance Maps. The procedure could be duplicated, or improved by other libraries.

Copyright Issues

No discussion of digital projects is complete without tiptoeing into the quagmire of intellectual property law. The Sanborn Library, LLC owns the largest and most complete collection (over 1.2 million) of historical Sanborn Maps, and copies from that collection are sold through Environmental Data Resources Inc.[9] Prior to beginning our project, Marriott Library Assistant Director for Special Collections, Walter Jones, contacted the Sanborn Library, LLC and requested permission to scan and post all 3,000 maps of Utah in our collection. Permission was denied, and we limited our scanning to those maps in the public domain, i.e. through 1922. There may be hope for scanning more maps, however. Current law states that materials published from 1923-1963 and whose copyright was not renewed are now in the public domain.[10]

After the digital collection was posted, The Sanborn Library LLC informed us that even the terms "Sanborn" and "Sanborn Maps" are registered trademarks of The Sanborn Library, LLC and that we were required to insert trademark symbols on every occurrence of those words on our website, as well as a statement of that ownership on our main page. We complied.

Scanning Procedures

The Sanborn Maps were photographed with a Leica S1 Pro digital scanning camera, equipped with a 50mm CFi Hasselblad lens. The camera was mounted on a copy stand, and Kaiser Softlite ProVision 6x55W fluorescent 5400K lights, on wheeled stands, were directed at the maps.

The maps were not removed from their plastic Mylar encapsulation, which created some difficulty with light reflection and glare. Numerous test scans were made and lights had to be positioned farther away and higher than for any other project. In the end we were

successful in adequately lighting the maps without glare, and without spending the extra time and effort to remove encapsulation for the nearly 1,300 maps.

Scanning software was SilverFast, operated as a plug-in for Adobe PhotoShop 6.0. Maps were scanned at roughly 5,000 x 4,000 pixels spatial resolution (400 ppi), and saved as uncompressed TIFF format, with file sizes averaging 60Mb. Each scan took approximately 3-4 minutes, and all master files were eventually archived offline to DLT tape.

After scanning, the TIFF files were compressed with MrSID encoding software, from LizardTech, Inc. 20:1 compression was achieved, reducing the approximately 60Mb TIFF files to 3Mb SID files, without noticeable loss of resolution. Users are required to download the free MrSID plug-in software to view the maps. The plug-in initially displays the entire map, and allows the user to zoom in on details.

Digital Collections Database

In January 2000, the Marriott Library selected CONTENTdm [http://contentdm.com] as its digital collection software suite. At the time, CONTENTdm was still a research project of the Center for Information Systems Optimization (CISO) at the University of Washington. Working closely with UW Libraries CISO had developed a product that addressed our needs for capturing, indexing, managing, and displaying images. It was, and still is, a highly flexible system that is adept at managing and displaying a wide variety of materials, including images, text, audio, and video. In 2001 CISO was spun off into a private firm called DiMeMa Inc., which in May 2002 announced a partnership under which OCLC became the sole marketing agent for CONTENTdm. This arrangement allows DiMeMa Inc. to continue improve their software by focusing on product development and user support.

The CONTENTdm software suite includes an Acquisition Station software that runs on Microsoft Windows, and a server software that runs on Windows NT or 2000, Solaris, or Linux servers. The Acquisition Station is used to import digital files for metadata application before being uploaded to the server. Once on the server the files may still be edited through administrative tools that are accessible by connection to the server, either through the Acquisition Station or any Web browser, with authentication.

CONTENTdm uses open source file formats. Metadata fields are based on the Dublin Core framework, with the ability to modify local field names while remaining mapped to DC elements for cross-collection searching. Metadata can be exported as delimited ASCII text or XML, easing concerns about proprietary software and the ability to migrate to other systems.

We have used CONTENTdm to create digital collections of photographs, maps, books, documents, and art prints. Display options for these collections is flexible, ranging from simple thumbnail display to XML wrappers for compound documents, monographs, postcards, and picture cubes to represent the six sides of an object. In the Sanborn Map

collection, search and browse results are displayed as thumbnail images in a customizable template. Clicking on the thumbnail images invokes the MrSID plug-in software.

Workflow

Following the scanning of the maps and conversion to MrSID format, full and part-time staff batch-imported the files into the CONTENTdm Acquisition Station in groups arranged by city. The Template Creator tool was used to enter repetitive metadata; these metadata were automatically applied to each file as they were imported into the Acquisition Station. Unique metadata were then applied to each file in the project spreadsheet, including street names, city, map sheet number, map date, and scan date. Unique identifiers were automatically imported as filenames.

After completing metadata entry for each group, the staff uploaded the files to the CONTENTdm server via built-in FTP. Once on the server, the files were approved by a collection administrator, and then the index was built automatically. At this point the files were live on the Web through CONTENTdm's standard browse and search interfaces. [http://content.lib.utah.edu] Our additional processes included the creation of a project website [http://www.lib.utah.edu/digital/sanborn - [see figure 1] , whose HTML code for browse links and search boxes was automatically generated from CONTENTdm's Query Builder tool.

Display methods

Maps may be browsed by city and year [see figure 2], or may be searched by city, year, street name, and date. Either method retrieves thumbnail images [see figure 3], and clicking on a thumbnail displays the MrSID file in the browser [see figure 4]. Once the SID file is displayed, four levels of zoom are available to view the detail of the map [see figure 5]. Full metadata for each map may be displayed by clicking on the title under the thumbnail image.

Evaluation

Based on server statistics, the Sanborn Fire Insurance Maps website has proven to be our second most popular digital collection. The site is accessible through our list of digital collections http://www.lib.utah.edu/digital/digcol.html [see figure 6], through our library catalog, and through most Internet search engines.

In hindsight we should have taken advantage of CONTENTdm's Full-Resolution Manager. This feature tracks the location of the higher-resolution files, whether they are stored online or off. It also generates display JPEG images automatically, and with the MrSID files stored online, would have given the user the option to pull up the SID file from a link in the metadata, after first viewing the much smaller JPEG file. To clarify this a bit further, the order of viewing could have been [thumbnail—JPEG—SID] rather than [thumbnail—SID] as it stands. It may not be too late to change. We have confirmed with DiMeMa that it is possible to update the collection to this format and will pursue this in the future as resources allow. We will use the [thumbnail-JPEG-SID] format for future map collections.

While the site is small, compared to the other projects mentioned earlier, it still encompasses nearly half the Utah Sanborn Maps held at the University of Utah. More importantly, it makes the color maps available free of charge to any user with an Internet connection, and hopefully will serve as a model for others to follow. In the future we may try to digitize more maps, as copyright allows.

Software Used For This Project

- CONTENTdm Software, DiMeMa Inc. http://contentdm.com
- MrSID, Lizardtech Inc. http://www.lizardtech.com
- Leica Cameras, http://www.leica.com
- SilverFast Software, LaserSoft Imaging, Inc., http://www.silverfast.com/
- Adobe PhotoShop, http://www.adobe.com

Notes

[1] Robert A. Sauder, "The use of Sanborn maps in reconstructing 'geographies of the past': Boston's waterfront from 1867 to 1972," *Journal of Geography*, 79 no. 6 (November 1980): 204-213.

[2] "Map Monopoly," *Fortune* (February 1937): 41-42.

[3] "Library Announces Digital Map Project," *LC Information Bulletin* (December 1997).

[4] About Digital Sanborn Maps, *Digital Sanborn Maps, 1867-1970*, http://sanborn.umi.com/HelpFiles/about.html (2001).

[5] Charly Bauer and Carol Lynn Roddy, "Implementing Digital Sanborn Maps for Ohio: OhioLINK and OPLIN Collaborative Project," *D-Lib Magazine* 7 no. 12 (December 2001).

[6] Kim Keister, "Charts of Change." *Historic Preservation: quarterly of the National Council for Historic Sites and Buildings*, 45 no. 3 (May 1993): 42-49.

[7] Glen Creason, review of *The Sanborn Fire Insurance Maps in the Library of Congress*, by Chadwyck-Healey, *WAML Information Bulletin* 24 no. 2 (March 1993): 97.

[8] Author's note: ProQuest has shown some interest in full-color scans. When we began our Sanborn digital project in 2001 ProQuest contracted with us to provide them digital copies of the first 256 maps we scanned. According to Andrea Eastman-Mullins, Product Manager at ProQuest Information and Learning, they are planning to use those scans to gauge interest in a full-color digital collection.

[9] "Historical Sanborn Maps," *Environmental Data Resources Inc.,* http://www.edrnet.com/sanborn.htm

[10] "How Long Copyright Protection Endures." *U.S. Copyright Office* http://www.copyright.gov/circs/circ1.html#hlc See also the Cornell Institute for Digital Collections website http://cidc.library.cornell.edu/copyright/