The National Park Service's Heritage Documentation Programs Draft a Collections Management Strategy for the 21st Century

Martin Perschler and Jennifer Hall; National Park Service; Washington, DC/USA

Abstract

The National Park Service's Heritage Documentation Programs, in coordination with the Prints and Photographs Division of the Library of Congress, have drafted a long-term strategy for managing its shared architectural records collection of more than 350,000 drawings, photographs, and historical reports. The strategy involves a gradual migration from paper and film to digital for service copies, a proposed eventual phase-out of the traditional booklet-in-binder format in place since the 1930s, the transfer of all paper and film documentation to remote storage, and a more robust information system for managing the collection and supplying text and images for the Library of Congress's Built in America website.

A decade or so ago, anyone observing the day-to-day operations of the National Park Service's heritage documentation programs (the Historic American Buildings Survey, the Historic American Engineering Record, and the Historic American Landscapes Survey, hereafter referred to collectively as HDP) would have concluded that the protocols in place for everything from documenting historic sites to making that information available to the public were so much a part of the history of the nation's oldest federal historic preservation program that they would outlive even its youngest recruits. Field teams traveled to historic buildings and sites, and, with tape measures and other tools in hand, set to work recording all the dimensions required for representing the site faithfully in measured architectural drawings just as the first field teams had done in 1933. Photographers, meanwhile, hauled their large-format view cameras from one site to the next, while historians rummaged through libraries and interviewed inhabitants so as to provide as complete a description and history of the site as time and the resources allowed.

Back in the office, the architects translated their measurements into detailed drawings, the photographers processed their sheet film, and the historians wrote their histories. They then passed their work on to the support staff at HDP, who typed report covers and index cards, placed the contact prints in photo mount cards, and then combined all the items into booklets. They then handed the completed booklets to the collections managers, who made duplicate, or service, copies of the drawings and any color transparencies, logged the booklets into a database then called the Automated Inventory, and physically transported all the materials to the Library of Congress's Prints and Photographs Division, where they were inserted into three-ring binders and then unceremoniously shelved for the benefit of future generations.

Then, beginning a decade or so ago, a number of things happened.

First, in 1994, the Library of Congress launched the National Digital Library, an ambitious, multi-year effort to digitize a number of historical collections and make them available without restriction on the Web. A Kellogg Foundation grant, awarded a few years later to help the Library identify collections appropriate for online use in the classroom, resulted in a report from the Center for Children and Technology that singled out the HDP collection as one of the best suited to meet K-12 educational requirements. This sudden realization of a vast, untapped audience for our documentation, along with a generous donation from the Shell Oil Foundation, put the collection on the fast track to becoming one of the first collections to be digitized under the American Memory-Built in America masthead.

Then, in 2001, one of HDP's longtime vendors for making the contact prints filed for bankruptcy. A year or so later, the supplier of photo mount cards stopped supplying them, and HDP's color transparency duplicator stopped duplicating them because of decreased demand for that and other traditional photographic services. Large manufacturers of everything from photographic paper to slide projectors to view cameras likewise announced that they were discontinuing traditional items and going digital or, like HDP's first vendor, filing for bankruptcy.

Just as the Library began its digitization effort, new documentation stopped flowing from the office to the Library, resulting in a backlog of approximately 5,800 booklets containing nearly 3,200 measured drawings, 40,000 pages of historical data, and more than 44,000 negatives with prints. Each time a vendor agreement collapsed, HDP responded in knee-jerk fashion by scrambling around for suppliers who had not abandoned the old ways of doing things so that HDP did not have to abandon its old way of doing things.

As a result, the costs to the program soared. The price of contact prints in the Washington, DC, area increased more than 300 percent, prompting HDP to look as far as New York City for suppliers of those goods and services. Prices for photo mount cards and color transparency duplicates also rose. Instead of taking a critical look at its output and how to produce that output, HDP opted in the short term to pay a premium rather than deal with the changes taking place in the photographic and related industries.

Then, one more thing happened: The Library of Congress asked HDP to take over the task of digitizing its drawings, photographs, and histories for the Built in America website.

Looking back, the launch of Built in America on the Library-side, the gradual disappearance of traditional goods and

services on the supply-side, and HDP's paralysis in the midst of it all, were all signs telling HDP that it needed to start thinking seriously about making major changes in the way it collected, processed, and disseminated information. By working through these developments separately, HDP was able to understand their connections and, based on that understanding, draft a reasoned, long-term collections management strategy.

This strategy actually began modestly as a short-term strategy for eliminating the backlog. At the time, preparing the booklets for the Library involved countless hours spent adding or updating each record in the HDP's Automated Inventory, then retyping the same title and address information onto cover stock, and then retyping the information again onto an index card, or onto multiple cards for cross-referencing purposes. The first step in eliminating the backlog was to eliminate the Automated Inventory and the outmoded and labor-intensive process for preparing documentation.

The most effective way of addressing the backlog was by adopting a "key once" approach and automating as much of the document preparation process as possible. HDP replaced the Automated Inventory with a relational, more robust networked collections management database that auto-generated booklet covers, index cards, and box labels either individually or in batches, along with any packing lists or reports required for transmitting the booklets to the Library.

Thankfully, that investment paid off. HDP's small collections team of three succeeded in eliminating that documentation backlog in less than five years. The new system paid other dividends, too. Anyone on staff suddenly had the power to generate lists, enter and edit data, print covers, index cards, and reports, and conduct their own collection searches instead of asking the collections staff to do these things for them. Whereas the Built in America website had replaced the Prints and Photographs Division's card catalog as the public's preferred entry point into the HDP collection online, HDP's new database replaced Built in America as the preferred point of entry for much of the HDP staff. When HDP realized it was auto-generating index cards for a card catalog at the Library that hardly anyone was using, and that the HDP staff was generating them from a system from whence came the data for Built in America, HDP began to question why it was generating index cards at all.

With the backlog gone, HDP turned to the matter of assuming responsibility for digital imaging. The first task was to figure out how, in an era of flat budgets, HDP could afford to do more with less. Calculating the material replacement costs of a typical documentation package (including the booklet) turned out to be a helpful exercise, since it revealed that the service copies made for each drawing and color transparency accounted for more than 50 percent of the costs of producing the entire package on average. Before Built in America, the Library had relied on those service copies for fulfilling patron requests for reproductions. Once patrons were able to download digital images of the documentation immediately and at no cost, requests for traditional reprographic goods and services declined. A simple comparison of the costs of a vellum service copy of a drawing and a scan of that same drawing

that met the Library's specifications for digital images showed how HDP might recover the costs of digitization. A similar comparison of the costs at the time of a duplicate color transparency and a scan of the original confirmed that a move to digital service copies and away from paper and film was the way to go.

This cost exercise also helped separate those items that were essential for documenting and describing a site from those produced out of habit or for the convenience of the researcher at the Library. Since the launch of Built in America, usage of the binders at the Library had declined so significantly that the Library staff decided to start moving the binders offsite. The removal of those binders not only meant that the Web was now the preferred go-to place for immediate access to the full collection, it also led HDP to question the value of continuing the age-old practice of generating all those extra things required for the presentation of the documentation in booklets set in binders but in no way considered irreplaceable or essential to documenting a site.

The exercise also helped put the crisis on the supply-side in perspective, specifically concerning the long-term availability of traditional photographic goods and services. [That concern, along with the planned removal of the collection out of Washington, DC, entirely to a state-of-the-art, bar-code based storage facility at Fort Meade in suburban Maryland may be enough to persuade HDP and the Library to abandon the old format and limit future output to full-sized drawings, historical and descriptive data, film negatives and positives, and digital surrogates.]

Out of these developments emerged a long-term collections management strategy for HDP that involves a gradual migration from paper and film to digital for service copies, a proposed eventual phase-out of the booklet-in-binder format, the transfer of all documentation to remote storage, and a more robust system for managing the collection and supplying data for the Built in America website.

Although the new collections management database had helped address the backlog, automate certain tasks, eliminate others, and mine the collection in ways not thought possible a decade ago, the addition of digital images to the mix required a switch to an item-level system. The next generation of the database had to do everything the other one did plus track the movement of digital images between directories, across a network, and onto the Library while at the same time allow for the search and retrieval of images, batch harvesting of metadata, renaming of files, and editing of image file header tags according to the Library's technical specifications. It also had to be efficient and user friendly.

Conceptually speaking, this move to an item-level system marked the beginning of the end for the booklet and binder format of the 1930s. Whereas HDP used to track drawing sets as a single unit (one drawing set could contain 10 individual drawing sheets, for example, which were all grouped together in a booklet), it now tracks the movement of digital images of individual drawings through a parallel, electronic universe of standards and procedures for accessioning, quality-control, and transmittal to the Library.

The timing also seemed right for a leap to item-level subject cataloging, one outcome of which was an increase in the program's inventory of un-cataloged legacy documentation from roughly 41,000 booklets of grouped items to more than 350,000 individual items.

HDP is working overtime to automate the item cataloging process so that it might make an appreciable dent in those un-cataloged items. To that end, on the photography side, the Library and HDP added more than 150,000 descriptive photo captions to the collections management database in an effort to make the photos searchable in Built in America, and the program now enters new photo captions directly into the database instead of typing the captions in word-processing applications and printing them out on paper for the booklets.

When it came to mapping out a strategy for creating the digital images, once HDP realized it was creating drawings and histories on computers and then printing them out only to have them scanned so that they could be made digital again, it decided the more efficient route for creating the images for Built in America was via image-file converters and other applications that helped transform those AutoCAD and word-processing files from their native formats into TIFFs, the image file format required by the Library. Now that those CAD and word-processing files are central to HDP's digital-image-making process, it had to establish protocols for accessioning, processing, and archiving those files, as well as plan for software and hardware upgrades and other future computing needs on the processing end.

As HDP's architects have learned from their various experiments in measuring historic sites using remote laser scanning, technology and architectural documentation are not always a perfect fit. None of the developers of Built in America, it seems, anticipated a move to item-level cataloging, or to image-file converters or other applications instead of scanners, or even the possibility of a move away from the booklet format. They designed the system based on what they had in front of them and on the assumption that the flow of paper and film to the Library would continue indefinitely. Needless to say, their technical specifications and system architecture reflect those assumptions. HDP had a difficult time, for example, finding image-file converters that would produce digital images of drawings matching the Library's scan-based specifications; thankfully, the most recent version of Adobe Acrobat Professional has solved that problem. And even though the TIFF file specification for digital images works well for drawings and photos, stand-alone TIFF images of a multi-page report that are neither searchable nor printable in batches, while satisfactory digital surrogates, are not very useful to researchers.

Furthermore, as HDP's dependence on databases and other applications and technologies grows, it must adjust the way it does things in order to meet the expectations of the software, or it will have to customize the software. In most cases, the adjustments have paid off, but they can be costly, and they have left HDP nearly as vulnerable to external forces and changes in the availability of goods and services as silver-based large-format photography has.

Even though HDP has set aside the issue of digital large-format photography for the time being, it must address the important issue of digital field photography, namely, those useful, detailed reference photographs taken by architects and historians while working at a site but that are not considered formal documentation. HDP has no plan for preserving those records currently other than to print them out in black and white on archival bond, burn them onto CDs, and drop them into a folder with other field notes. In the meantime, HDP continues to expand and refine its collections management strategy in response to internal demands and external changes in the industries that produce the things its needs in order to carry out its mission.

Since completing the first, full transmittal cycle in fiscal year 2005 involving digital images of drawings and historical reports generated and processed by HDP and then sent to the Library, a number of important things have happened. HDP has—

- migrated to a relational database for tracking and cataloging down to a single photograph, report, drawing, and digital image (that is to say, down to the item level);
- trained its photographers and other staff to enter new photo captions directly into its database;
- working with its GIS experts, implemented geospatial data standards so that it can export and share spatial and attribute data quickly and efficiently;
- compiled guidance for creating PDFs and then TIFFs from AutoCAD (in the case of drawings) and word-processing applications (in the case of historical reports); and
- taken on the task of creating digital images of color transparencies.

In turn, the Library has—

- dropped the index cards, duplicate transparencies, and full-size vellum copies of measured drawings from its list of requirements;
- launched an effort to re-key an additional 100,000 photo captions; and
- launched a new version of the Built in America website that makes it possible for researchers to search on those photo captions, on media dimensions, and the names of photographers, delineators, and historians (where available).

The next items on the agenda for HDP and the Library to discuss are the future of contact prints for the notebooks and PDFs as an alternative to TIFFs for the historical reports.

Since 2004, HDP has focused on internal production, specifically on the production of service copies, with the intention of cutting costs and reducing the amount of time required to get digital images of the documentation to the public via the Built in America website. HDP staff now supply PDFs of drawings and historical reports so that the collections team can generate the TIFFs required by the Library. In the near future, it will be asking for PDFs from independent contractors and others outside HDP who produce drawings and historical reports for the collection.

Martin Perschler earned his PhD in the history of architecture from the University of Virginia (1999). Jennifer Hall earned her MS in historic preservation from the University of Pennsylvania (2002). Both work in the cultural resources division of the National Park Service in Washington, DC.