

Richmond Architectural Surveys: Choices in Digitization

by Sam Byrd

In 2008, Virginia Commonwealth University (VCU) Libraries received a donation from the City of Richmond, Virginia, that was a perfect candidate for digitization: the Richmond Architectural Survey (RAS) Collection (<http://go.vcu.edu/ras>). The collection consists of over 600 survey forms, photographs, and other data sheets documenting the vernacular architecture of two districts of Richmond (Jackson Ward

and Oregon Hill). The surveys were completed by Richmond city employees in the mid-1970s, with each survey form being devoted to a particular building or city block.

For any given digitization project, decisions have to be made about how the material is to be presented on the web. These decisions are driven by at least two major factors: the original arrangement of the collection, usually dictated by the physical nature of

the material, and the possibilities (or limitations) of the digital asset management software to be used. One of the goals of digitization is the accurate representation of the material in a way that does not distort its nature, while taking advan-

Sam Byrd is digital collections systems librarian for Virginia Commonwealth University Libraries. He can be reached at sbyrd2@vcu.edu.



tage of the possibilities of the digital medium for new expression and recontextualization of physical items as well as enhanced searchability and discoverability.

The collection in question came to us in a single box, containing manila folders housing the material, each folder containing survey forms pertaining to an area of the respective districts, typically designated by street names. The 8½ x 11 inch forms were color-coded: yellow for Jackson Ward, green for Oregon Hill. Each survey form had space for the surveyor to fill in information concerning a specific building (construction type, architectural significance, use, etc.). In addition, many of the survey forms had photographs attached to them with paper clips. Typically, the photograph depicted the building covered in the survey; if the survey covered a whole block, the photograph could depict either a single building or a row of them. Sometimes there were multiple photographs of the same building or block. Several of the survey forms were also accompanied by thin strips of white paper, called “assessors [sic] property cards,” which contained additional building information not covered in the survey forms.

It was clear that there were two primary organizing factors for the collection—the geographical area of Richmond, designated by the color of the survey report, and the manila folders, organized by area subsection. Beyond that, at the item level, the survey report was the primary document, with the photographs and property cards accompanying the survey, although there were also a few standalone property cards.

VCU Libraries uses CONTENTdm for its digital asset management software. With CONTENTdm, digital files can be presented as standalone digital objects, each with its own set of Dublin Core

metadata, or “compound objects,” CONTENTdm’s term for multiple files connected by an XML structure and presented as one entity. Compound objects are typically used to present multipaged objects such as printed monographs, journal issues, or newspaper issues. In CONTENTdm, the compound object structure is such that there is a “top-level” record containing metadata for the entire object, as well as secondary “page-level” metadata pertaining to each individual page, or item, of the object.

Since the items in the RAS col-

**So, we decided to totally
redo the project,
and treat each separate
component of the
collection ... as a separate,
standalone object.**

lection were physically organized around the survey form, each form representing a street address, we decided that a logical way to present the material in CONTENTdm would be to have compound objects for each form, with the additional material at the secondary level being the accompanying photographs and property cards. In search-result sets, one record would be retrieved per address and would be displayed on the initial result screen. Once that record was chosen, its individual components would display to the side in a separate frame. The advantage to this was that all items pertaining to the same address would be kept together in the initial search results.

We did an initial load of the collection in this way, and while it achieved the goal of bringing the objects together, other aspects be-

came apparent:

1. In the initial record display, the photographs were buried. For the thumbnail display of the top level of each compound object, we used the survey form; if the form had a photograph attached to it, the thumbnail of the form included the photograph. The thumbnails were small enough already, but displaying the photos this way shrank their size so much they were almost illegible. Also, there was no way to indicate at that level if a survey form had multiple photographs. The photographs are the most compelling visual aspect of the collection, and putting them at the page level of the compound object meant users had to click twice to get to the actual photo.
2. In some ways, the organization around the survey form was an artificial one, since some photographs or property cards attached to a survey form were actually for more than one address. There was enough variance in the correspondence of survey form to accompanying material that it seemed forced to display them this way.
3. Looking at the metadata harvested from the collection, CONTENTdm’s compound object structure did not easily convey the context for each sublevel of each object. It was not readily apparent what the parent level of the item metadata was. Typically, for monograph-style compound objects, all of the fields applying to the top level are not repeated at the page level; rather, the page-level metadata consists of the page number and possibly the transcription of the text on that page. For this reason, when we recently began harvesting metadata from our digital collections for our beta discovery tool (Primo), we only harvested the top-level metadata for compound objects.

However, for the RAS collection, there was extensive additional page-level metadata pertaining to the photographs and property cards that was being lost upon harvesting.

So, we decided to totally redo the project, and treat each separate component of the collection—survey form, photograph, property card—as a separate, standalone object. The address associated with the object would always be the first part of the title, and then the component term would follow. That way, a title sort would serve to collocate all the components pertaining to a single address. In the search results display, the photographs would be exposed at that top level and would be easier to select. It would be readily discernible if there were multiple photos associated with the same address. The harvested metadata would be easier to understand, and would better serve us in the future in migrating our data to a different system.

Since the advent of digitization, the possibilities involved with converting physical collections and presenting them on the web are always evolving. These possibilities enable us to experiment with new ways to present collections that take advantage of search and display technology, ways that do not necessarily match a more bibliographically oriented, traditional display model. However, those experiments do not always work. In the case of the RAS collection, we ended up favoring a straightforward, structurally simple approach. If the materials in the RAS collection were displayed physically, the survey forms, photographs, and property forms would probably best be presented by preserving the physical grouping of the individual items, with the survey form being the central point of focus, and the photos and forms associated with it physically adjacent or nearby. The presentation of the same material digitally can afford to take advantage of the abil-

ity to supply context in other ways. While working within the constraints of what the digital asset management software can do, our goal was to maximize its potential while keeping the raw digital material (the files and metadata) structurally intact to allow for flexible use in the future.

This was not a major change that took that much time to rectify, or a major shift in our approach to digitization. Rather, it was merely typical of the kinds of adjustments needed to accommodate problems that arise during a project. These kinds of decisions are made all the time in the course of digitization work, and our practice and decisions continually evolve and build upon what we have already done. While it was a relatively minor problem, it highlights the kind of thinking we need to engage in when we approach the digitization of materials that are not as straightforward as single monographs or groups of photographs. **VI**

MALiA
MID ATLANTIC LIBRARY ALLIANCE

Formerly SWING



www.malialibrary.org

The Mid Atlantic Library Alliance, or MALiA, is a consortium of over 150 public, academic, school, and special libraries. Originally founded in southwestern Virginia, the organization has grown to include libraries throughout Virginia as well as member institutions in northeastern Tennessee and North Carolina. MALiA is open to all libraries, regardless of location, and we welcome new members at any time.

-Buying Power

**-Continuing
Education**

-Scholarships

-Only \$150 a year!