Abstract

Purpose – Aims to provide an overview of methods for digitizing maps, the creation and preservation of a digital collection, interoperable database construction, and an examination of the real costs associated with maintaining a digital collection.

Design/methodology/approach – This paper is an examination of the tools, metadata choices, work with Open Archives Initiative (OAI) research projects, and sustainability costs, associated with a digital project of 946 dynamic historic cartographic images.

Findings – The Early Washington Maps project demonstrates how multiple tools may be integrated and interoperability between databases achieved. Realistic costs associated with digital collections and practical reference techniques are described.

Practical implications – This is an extremely common-sense discussion of digital library issues for library professionals planning or enhancing digital collections.

Originality/value – The paper describes ideas for enhancing access to digital collections during and after a grant award.

Keywords Maps, Collections management, Electronic media, Open systems, Databases

Paper type General review

Introduction

The Early Washington Maps project began in 2000 with funding provided by a Library Services and Technology (LSTA) grant from the State Library of Washington. This collaborative project between Washington State University (WSU) and the University of Washington (UW) Libraries included scanning, providing descriptive metadata, and creating a web-accessible database for 300 copyright-free maps selected from the map collections of the two universities. During this initial grant, WSU and UW librarians selected a broad range of maps from early territorial surveys to unique manuscript maps, such as two of the earliest maps of the region drawn by Jean Pierre De Smet of the upper Columbia (c. 1841) and Edmond S. Meany’s historical map of Washington (1857). In 2001, the Washington State University Libraries partnered with the Tacoma Public Library in a second successful proposal to the Washington State Library to double the size of the Early Washington Maps project to 700 maps. The collection has grown to include 946 maps[1].

Part of the initial grant was to test various scanning options including the use of commercial vendors and in-house imaging methods. The University of Washington selected DigiColor of Seattle[2]. Initially, DigiColor scanned the University of Washington’s maps while they were still encased in mylar. Any slight speck of dust on
the mylar, however, caused thin rainbow-colored lines to appear in the final scan. These distortions were not immediately visible and require extensive staff time to slowly examine each image. At Washington State University, given the University’s rural location (and the lack of nearby commercial vendors), several in-house scanning methods were tested. After experiments with using a Nikon D1 digital camera and scanning color transparencies on a Nikon Super Coolscan 8000, the best method was to scan maps in pieces on a Microtek 9600XL and then use Adobe Photoshop to stitch the individual pieces together (O’English et al., 2002). Since the scanning surface of a Microtek 9600XL is roughly 12 × 17 inches, some of the very largest maps in Manuscripts, Archives and Special Collections (MASC) were not scanned. As a result, WSU MASC staff looked for maps to include in the project beyond those stored in the oversize map cases, and in the process found a hidden treasure trove of maps among the book and manuscript collections. Publications by the major railroads to attract tourists and sell railroad lands proved a rich source of interesting and unusual maps (see Figure 1).

Locating these maps bound in books by searching the library OPAC, however, proved extremely frustrating. The most successful method to locate these hidden maps was to select key subject areas and then physically browse the book vault. During the process, it became clear that it would greatly increase access to these hidden treasures if there was a way to enhance the MARC records in the WSU library OPAC so that users who did not have physical access to MASC’s closed stacks could retrieve more information to locate these maps. By using the Innovative Interfaces (III) Imaging and Document Management product, WSU MASC staff created links between III

Figure 1.
A charming map included in an 1896 Northern Pacific Railroad publication
INNOPAC MARC catalog records and the Early Washington Maps Collection (Bond and Cornish, 2002). Embedding the maps images in the MARC records provided an ideal opportunity to enhance the entire MARC record. As a result, WSU MASC staff enhanced all of the catalog records for any printed items (books, pamphlets, brochures) or manuscript collections from which maps were scanned for the project. These enhancements included subject headings for maps and the addition of a corporate name heading (MARC 710) “Historical Map Collection” used in all map catalog records for items held in MASC. This heading, “Historical Map Collection”, virtually unites all cataloged maps whether the items are physically located in map cases, the book vault, or manuscript collections (see Figures 2-4).

In preparing for the Early Washington Maps project in 2000, librarians at WSU and UW surveyed other web-accessible map sites. Most notable at the time was a site created by the Library of Congress’s American Memory project devoted to Civil War Maps, which included dynamic maps available in the MrSID format[3]. This collection provided the best model, in that it included the functionality of MrSID files without the need to download any plug-ins. Unlike the Library of Congress, however, both Washington State University and the University of Washington Libraries had already begun using DiMeMa’s CONTENTdm to host digital collections. After the receipt of the first LSTA grant for the Early Washington Maps project, DiMeMa enhanced CONTENTdm so that it would accept the MrSID file format. To take full advantage of the MrSID format, Alan Cornish at WSU customized the MrSID Image Server package so that it would function alongside CONTENTdm. After users enter a query and select a map from their results, a link labeled “View Map Image (Zoom)” appears in the item.

**Figure 2.**
Maps bound in an engineer’s report. Clicking on the map images leads to the MrSID image viewer from the Early Washington Maps Project (see Figure 4)
Figure 3. Enhanced MARC fields, such as the 710, collocate maps bound in books with WSU’s Historical Maps Collection.

Figure 4. Users may return to the OPAC or click on this dynamic MrSID image to study details.
metadata, just below the title. If the researcher selects this link, a MrSID viewer opens, which enables the detailed viewing and manipulation, such as zooming in on a specified area or panning, of high-resolution MrSID images in the online collection without requiring the downloading of a plug-in or the MrSID file (see Figure 4). Furthermore, the MrSID viewer includes the same navigation banner and color scheme so that it matches the rest of the site. Additionally, from this viewer users are able to download a copy of the MrSID format full-resolution image to their computer and to further work with the image with programs available through the LizardTech website and GIS software packages, such as ESRI’s ArcView software[4]. One of the Perl scripts making up the MrSID Viewer program, show.plx, was modified so that basic metadata appears with the map image, including the title and date of the map, the identifier, and the size of the MrSID file.

There are a number of searching options on the Early Washington Maps site, including keyword searching, a drop-down list of predefined topics, browse the entire collection, and an advanced search. A USGS Topos index links to a graphic of Washington State by county. Selecting any of the counties will generate a query and retrieve the appropriate map. A similar index was made for a group of San Juan Island T-Sheets. WSU MASC staff also prepared a timeline of notable events in Washington State History and, where appropriate, included a map icon which when selected brings up a germane map for the given topic. For example, the entry for the year 1811 indicates that David Thompson maps the Columbia from its headwaters. Clicking on the map icon after this entry results in the display of Thompson’s “Map of the North-West Territory of the Province of Canada from Actual Survey during the years 1792 to 1812”.

**Collection access**

After completing the first phase of the Early Washington Maps project, the site was posted on the Washington State University Libraries website with links added to various pages, such as the Libraries’ listing of electronic databases and the WSU MASC website. The University of Washington Libraries provided similar links. During the second phase of the project with the Tacoma Public Library, the programming was modified in the Tacoma Public Library’s place name database so that when search displays are presented, a link is provided that reads “Search the Early Washington Maps database”. If the researcher selects that option, the same query is passed directly into the Early Washington Maps database (see Figure 5)[5].

In addition to efforts at regional outreach, WSU’s System Librarian, Alan Cornish, worked with programmers involved with the Gateway to Cultural Heritage Materials hosted by the University of Illinois at Urbana-Champaign (UIUC) to share metadata from the Early Washington Maps project. UIUC’s Gateway to Cultural Heritage site is based on the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH)[6].

Since CONTENTdm (the application used in the Early Washington Maps project) allows collection dissemination through the Open Archives Initiative-Protocol for Metadata Harvesting (OAI-PMH), once the metadata contained in the CONTENTdm Early Washington Maps project was opened, programmers involved in the Gateway to Cultural Heritage Materials project could harvest the metadata and integrate it with
collection information from other institutions into a subject repository. The harvested information includes descriptive information and pointers to the original map images and metadata residing on a WSU Libraries server (Wykoff et al., 2005). This work demonstrates the interoperability of the Early Washington Maps site, as well as the UIUC’s Gateway to Cultural Heritage Materials, which has in turn been harvested by other OAI projects. By sharing the descriptive metadata from the Early Washington Maps project, users searching other sites, such as the Gateway to Cultural Heritage Materials, can locate maps from WSU, UW, and the Tacoma Public Library (see Figure 6)[7].

In addition to sharing data with OAI projects, WSU MASC staff registered the Early Washington Maps site with OCLC and exported the metadata from the collection into WorldCat. Joanne Gullo, a database specialist at OCLC, prepared the data and loaded several test records into WorldCat. After consulting with staff at the University of Washington and Tacoma Public Library, Ms Gullo incorporated suggestions and then loaded the complete database. One final method of outreach is the inclusion of Early Washington Maps metadata in the Western Waters Digital Library project. With funding from the Institute of Museum and Library Services (IMLS), the Greater Western Library Alliance, a consortium of Western academic libraries, created the Western Waters Digital Library (WWDL) utilizing CONTENTdm’s multi-site server product, which harvests metadata from participating repositories and creates a single index. Once a user searching the WWDL site selects a result from the Early Washington Maps project, the user is taken to the WSU server to view the item (see Figure 7)[8].
Sustainability

One of the most difficult questions in any grant proposal relates to sustaining the collection after the grant award. Having completed most of the work on the Early Washington Maps database in 2002, some practical costs associated with maintaining the collection can now be determined. Though there is a growing literature on the topic of building and sustaining good digital collections, there are few published examples of the real costs associated with maintaining a digital collection[9].

Methodology

To arrive at an annual price required to sustain the Early Washington maps project, the expenses of the major components of providing access to library digital collections were totaled. These expenses include the annual cost of the support license to DiMeMa, the maker of CONTENTdm, the annual cost for hardware, expendable digital back-up tapes, and a realistic estimate of the annual time require by our Systems Librarian who oversees WSU’s inactive CONTENTdm collections and generates use reports. To arrive at a per-image cost, these costs were divided by the total number of items in our digital collections (currently 46,273). Once this total cost per image was calculated, this per-image cost was multiplied by 946, which is the total number of images in the Early Washington Maps project (see Figure 8).
The WSU Libraries pays DiMeMa $3,980 annually for support and a license for 64,000 images/digital objects. At the time of writing of this article, WSU has a total count of 46,273 items in library digital collections. These items are primarily image files, but also include texts, sound recordings, and video. Looking at Figure 8, when $3,980 is divided by 46,273 digital objects, the results is an annual cost of 0.086 cents per image. Since the Early Washington Maps project contains 946 images, the portion of the annual support to DiMeMa for the entire collection is $81.37 (i.e. 946 images in the collection, multiplied by 0.086).

Other fixed costs include a $6,000 server and a $4,000 Hewlett Packard auto-loader tape back-up system. The server and the auto-loader each have a four-year life span. The auto loader functions with a set of 16 digital tapes at $40 per tape. Individual tapes need to be replaced every six months to a year. To determine the annual cost for the server and auto loader costs, the price of each ($6,000 and $4,000) divided by four is $2,500. Replacing 16 digital tapes each year at $40 per tape equals $640. A per-image cost for the hardware and back-up tapes to support all of WSU CONTENTdm digital objects is $3,140 divided by 46,273, which equals 0.067. If the 946 images in the maps project is then multiplied by 0.067, an annual cost of $64.19 is the result. The average support provided by a WSU Systems Librarian is two hours per month (this estimate is appropriate for inactive collections only). Two hours per month at $30 per hour (including benefits) equals an annual cost of $720, which when divided by 46,273 and then multiplied by 946 equals $14.39.
The annual cost to support the Early Washington Maps (as a portion of all digital collections), including license payments, hardware, and staff time, therefore equals $160.28.

**Providing reference service**

Helping users with large maps can be a challenge. Maps are large and awkward to handle and, as often happens, the map requested is at the bottom of the oversize drawer. With the added visibility of the website, there was a concern that this would lead to an increase in demand for the original maps. WSU MASC staff experience has been (along with many of WSU’s other digital collections) that users with very few exceptions prefer to work with the digital images from the collections, rather than the originals. Indeed, the ability to quickly page through screens of thumbnail images allows one to rapidly survey a collection. Users also don’t have to go through the pesky registration requirements of working in a special collections reading room, can avoid waiting while library staff retrieve the items from the basement vault, and have the opportunity to work with collections beyond MASC’s 8.30 am to 5.00 pm weekday schedule. Given the remote location of Washington State University in rural Eastern Washington, visiting MASC in person can be a challenge.

Since the majority of reference requests arrive by e-mail and are from users outside of the Washington State University community, all of MASC’s policies, procedures, and the price for reproductions, have been posted on the MASC website. Quick access to these routinely consulted documents greatly expedites common requests. The greatest challenge in satisfying online reference requests for reproductions of maps in
Early Washington Maps collection is managing the very large image files, many of which exceed 500 MB. For requests of scans, MASC staff give the user the choice to have the images reduced in size (when necessary) to fit on a CD or maintain the full resolution by burning to a DVD. To facilitate the making of prints, library staff established an FTP link to the campus graphics lab to transfer these large files.

Conclusion

Waters (2004) notes that one of “persisting problems with digitizing project in libraries, and particularly those in the humanities, is that they rarely build on, enhance, or otherwise connect with work across institutions”. Having avoided this problem, the Early Washington Maps collection, with its regional focus, inter-institutional collaboration, and unusual content, is to date the Washington State University Libraries’ most used digital collection[10]. One of the values of the site is that it brings together, virtually, maps physically housed at multiple institutions around Washington State. As mentioned above, even at Washington State University, the Early Washington Maps site unites online maps drawn from different sections of special collections including rare books, manuscript collections, and traditional map cases. In a digital environment, the ability to connect the various methods of intellectual control over the maps, such as enhancing MARC catalog records with embedded images and metadata, and linking graphic indexes and historical times to their corresponding maps, provides researchers with new avenues to discovery. The interoperability of the metadata in the Early Washington Maps project has been demonstrated by collaboration with OAI sites and OCLC’s WorldCat. Along with many other libraries that have created digital collections, the WSU library is struggling with long-term sustainability. What was initially a research project to test various methods of scanning and provide enhanced access to historical maps is now a preservation project. If all of the data was suddenly lost, WSU, UW and the Tacoma Public library would not have to funds to recreate it. With the full-resolution TIFF and MrSID files saved on servers that are regularly backed-up, as well as burned to DVD, short-term preservation is ensured. Future work on the collection will include utilizing preservation metadata schema, such as METS, and collaborating with other preservation projects, and slowing adding new items to the collection.

Notes

7. Gateway to Cultural Heritage Materials (available at: http://mergal.grainger.uiuc.edu/cgi/bib/bib/bib-idx). Additional sites that have harvested metadata from the Washington State
University include Denmark’s Electronic Research Library (available at: http://preprints.cvt.dk/cgi-bin/egw_sindap/1240/screen.tcl/name = find-a&service = prep&lang = eng) and the University of Michigan’s Digital Library OAIster project (available at http://oaister.umdl.umich.edu/o/oaister/) (accessed April 2005).


References

