Marquette University
e-Publications@Marquette

Library Faculty Research and Publications | Library (Raynor Memorial Libraries)

9-1-2009

A Collaborative Workflow for Digitization of Unique Materials

Gretchen Gueguen
East Carolina University

Ann Hanlon
Marquette University

Accepted version. *Journal of Academic Librarianship*, Vol. 35, No. 5 (September 2009): 468–474. DOI. © 2009 Elsevier. Used with permission.
NOTICE: this is the author's version of a work that was accepted for publication in *Journal of Academic Librarianship*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Journal of Academic Librarianship*, VOL 35, ISSUE 5, (September 2009) DOI.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

A Collaborative Workflow for the Digitization of Unique Materials

Gretchen Gueguen
Library, East Carolina University
Greenville, NC

Ann M. Hanlon
Raynor and Memorial Libraries, Marquette University
Milwaukee, WI

Abstract:

Recent literature has emphasized the digitization of unique materials. This paper will examine the experience of the University of Maryland Libraries as it embarked on a program to harness existing workflows for digitization and create more systematic methods for digital capture of unique collections using existing organizational resources.

Introduction

A number of major research universities have undertaken mass digitization of their book collections, including efforts associated with such well-known projects as the Google Book Search Project and the Open Content Alliance.. So it is not surprising that calls for the "mass" digitization of our special collections materials have followed. Indeed, prominent players in the library world, including OCLC and the Council on Library and Information Research (CLIR), have argued for the need to scale up digitization efforts in order to move from project-based digitization to more systematic programs focused on enabling deep research of heretofore hidden or geographically inaccessible (for some) collections.[1] [2] But libraries face the basic challenge of how to scale up in the midst of already strapped budgets and overburdened organizations. Given our existing workforce and workflows, how can we begin to make our unique materials more systematically available online? This paper will examine the experience of one institution, the

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

University of Maryland Libraries, as it made organizational efforts to harness existing workflows and to capture digitization done in the course of responding to patron requests. By examining the way this organization adjusted its existing workflows to put in place more systematic methods for digital capture of unique collections, the authors hope to provide insight into the benefits and pitfalls of one model for scaling up digitization.

Literature Review

Several recent articles have suggested methods to scale up digitization. Much of the focus in this literature has been centered on "mass digitization" projects such as Google Book and the Open Content Alliance. Karen Coyle, in a 2006 overview of such projects, points out that "mass" projects have different qualities from previous "large scale" projects. The uniformity of the book format has made it possible to automate much of the digitization, increasing the scale to that of entire collections. However, Coyle wisely points out that there are two fallacies in the mass digitization model. The first is the assumption that all books are suited to this method, no matter how fragile or uniquely formatted. The second is the assumption that materials time and money will be saved by digitizing materials only once and making the subsequent digital copy universally accessible. The universal accessibility of the digital copy, particularly with regard to fragile materials, would also be a boon to non-book materials. However, the fragile and idiosyncratic nature of special collections and archival materials make automation much more problematic. Human intervention is likely to be necessary at the item level in nearly every case, making it difficult to move beyond "boutique" digitization projects.[3]

This boutique model, and the hurdle it presents to the systematic digitization of special collections materials, is likewise favored by funding models based in grants that focus on digitizing a specific body of materials selected to meet grant guidelines. However well-designed these guidelines might have been for selection, the limits they imply mean that only a portion of any collection can be digitized in this manner within a grant's timeframe and budget. Thus, while grant funding can be an excellent means to establish important digital collections, it cannot be a fundamental

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

part of a digitization program. As Daniel Greenstein and Suzanne E. Thorin (2002) write:

> Many believe that as the digital library becomes library infrastructure, the financial resources needed to maintain it will come from numerous budget lines rather than from one line that is earmarked for digitization. In the adult digital library, electronic resources will be acquired from general collection budgets, and digital preservation activities will be supported with general preservation funds.[4]

Laurie Lopatin (2006) notes that while movement toward sustainable funding has been seen in some quarters (she cites a 2005 survey of libraries in the New York City area in which 51% of respondents reported their budgets were funded internally), many others reported a continual search for new funding to begin and sustain projects. The high profile of the mass digitization projects already noted further muddies the waters.[5] As Nicholas Joint (2008) points out, when Google Books sounds like " 'a 110 million pound scholarly digital library available for free,' administrators may think: Why ever spend another penny on your local library?"[6] While Joint is primarily concerned with scholarly open access projects, the fight to gain recognition for the extraordinary effort put into digital library development remains the same.

Along with inconsistent funding, systematic digitization initiatives may be harmed by a lack of internal organizational support. Boock and Vondracek (2006) conducted a survey of 40 ARL libraries and found that 38 of them (95%) had engaged in digitization. Of these, 84% were found to rely on "cross-departmental project groups" in these efforts. That is, the bulk of those libraries that are making digitization happen are those that are able to leverage the expertise of their larger institution.[7] Although new units and new positions may be created in support of these initiatives, using the existing strengths of the organization appears to be the most viable strategy to adapt to changing needs.

With specific regard to special collections, Ricky Erway and Jennifer Schaffner's 2007 report for OCLC Programs and Research attempts to distill the sentiments and discussions of the "Digitization

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Matters" forum attended by two hundred directors, administrators and curators of special collections in libraries, archives and museums. The report argues "that large quantities of digitized special collections materials will better serve our users," and that we should therefore "optimize procedures primarily for access." The report does not call for librarians to abandon standards and best practices for digitization. But it does call for better decision-making. Erway and Schaffner ask whether this is a viable standard for special collections — does every item we digitize need to be treated as though it cannot or will never be digitized again? Is it possible to digitize for access and assume that the opportunity to digitize for preservation still lies ahead? And as for description of special collections materials, a major impediment to the mass digitization of those materials, might there be room in the item-level world of special collections digitization for group-level description and collection-level decision-making?[8]

Finally, a recent CLIR report on "Reconceiving Research Libraries for the 21st Century" (2008) calls "for more aggressive intervention to better structure and manage the challenges we face." Drawing upon the proceedings of a symposium featuring leaders in the field of digital libraries, the report argues for rethinking what we conventionally consider to be "fringe activities," such as metadata building and digitization, and reprioritizing such activities as core investments that we need to make in order to "make material available to the scholarly community in a systematic way."[9] Shifting our basic orientation in this way is no small task. But the authors of the current case study hope our efforts serve as one example of the processes by which libraries might begin to organize for the systematic digitization of unique materials in special collections and other holdings.

The University of Maryland digitization program sheds some light on how nearly all of these suggestions might be applied. The project systematized digitization by implementing a policy to deposit all digitization done for patron requests into a newly created digital repository. This policy had far-reaching implications. And it provoked many new questions: how would this new policy affect the daily digitization workflow; how could the scope of the digital collections be defined if we were collecting the arbitrary digitization

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

requests of patrons; and how could we adapt our standards and best practices to accommodate this new workflow without overburdening the special collections and digital collections staff affected? The following case study explores how some of these questions were answered and how the University of Maryland, as an organization, adapted workflow and policy to meet the goal of capturing this existing digitization workflow in order to implement more systematic digitization efforts.

Case Study: The University of Maryland's Digitization Workflow

In December 2004, the UM Libraries established the Office of Digital Collections and Research (DCR), to coordinate and plan for digital initiatives, and to develop and manage a central digital repository (using the Fedora architecture) to house digitized objects from across the UM Libraries' departments. The repository would limit the re-scanning of frequently requested material and at the same time repurpose those scans for online digital collections. The repository was to be populated with materials created from patron requests, particularly those generated by the Department of Archives and Manuscripts, along with any materials digitized as part of other digitization projects. As DCR began the task of coordinating efforts to create the repository, the patron scanning workflow in Archives and Manuscripts was growing, particularly due to efforts to document the University's history for its 150th birthday celebration. With the increase in patron and exhibit scanning, joined with the significant time required to program, design, and develop the Fedora-based repository, an urgent need emerged to create at least a stop-gap measure to capture and track the scans being created. In response, a Project Archivist hired to assist with digital image management and the Curator for Historical Manuscripts, working in close consultation with DCR, developed a "scanning database." This was a Microsoft Access database with fields that, when completed by staff and students in the course of scanning materials, would map directly into the repository's newly developed XML metadata schema. Scanning would be done according to specifications published by DCR. A file-naming scheme was added, and a dedicated directory was created for saving new digital images. The database of metadata, along with the directory of images, was to be

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

migrated over when the repository infrastructure was finished. At that point, a web-based administrative interface would be launched, giving staff and students the ability to upload objects and metadata records directly with sophisticated tools for handling metadata and rights management.

Stumbling Blocks

As with any digital initiative, the goals and aims of the repository project changed over time. In some cases this learning process required going back to the drawing board and starting over. But in the case of UM's digitization workflow project, staff continued to add to the scanning database in anticipation of the completed digital repository. Thus, the digitization program already underway had to be robust enough to adapt to changing policies and the repository had to be flexible enough to accommodate legacy data. These issues necessitated answering the following questions: how to create quality digital objects, how to handle the scale of the operation, and how to present this mixture of materials online in a way that made sense to a diverse audience.

Quality Digital Objects

Building a repository while simultaneously populating it led to certain obvious difficulties. First, changing a standard — for example, requiring images to be created with a 24-bit color profile rather than a 48-bit one — meant rendering potentially "unacceptable" thousands of images created up to that point, not to mention thousands of work hours. Second, given the distributed workforce, day-to-day decisions about standards and practices were not easily communicated or implemented. The range of archivists, curators, librarians, and student assistants participating in this project shared an uneven awareness of current digital standards and technology. While many were willing to learn, accurately communicating a message to a diverse and distributed group was a difficult challenge. Finally, the digitization program was not necessarily the top priority of archives and special collections departments dealing with the more immediate pressures of daily patron requests, reference questions, processing backlogs, exhibit building and fundraising.

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Several methods were employed to resolve some of the inherent complications of the project. The first was the creation of in-house documentation and standards. This documentation ranged from statements of mission and responsibility and simple guidelines codifying benchmarks for digital output, to more in-depth explanations of "input referred" scanning techniques and step-by-step instructions for typical practices in which staff might need to engage, such as resizing an image, using a histogram to evaluate target aim points, or sharpening an image.

These efforts were supported by a series of workshops and trainings organized to increase personal communication among staff in Archives and Manuscripts and DCR. In addition, quality control procedures were devised to balance responsibilities among the staff. Curators would be responsible for regularly reviewing the metadata records created by graduate assistants to insure against items piling up. Curators would ensure a consistent level of quality control by checking to see that item records were completely filled out and the information was basically correct. DCR staff were to follow up by checking a statistically significant portion of these records for style and consistency in metadata and image quality. This system played into the strengths of those involved: the curatorial staff's ability to verify the correctness of the information, the DCR staff's familiarity with technical standards. Despite an initial reluctance to interfere with existing scanning operations in the Archives and Manuscripts departments, DCR eventually stepped in to fill the role of trainer and project manager.

The creation of documentation and standards provided much needed limitations for the image creation process. The standards removed the necessity for individual decision-making about digitizing items, synthesizing the wide range of possible color profiles, resolutions, and post-processes into a more manageable range of "if-then" scenarios. By choosing a standard that would be acceptable in most cases (such as a relatively high spatial resolution, or an RGB color profile), context became irrelevant and the workflow process was streamlined. In addition, the organization agreed early on to accept into the repository legacy data that did not meet the current standards. This meant that the Library had to accept the possibility that images might be rescanned in the future

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

if a higher quality version was needed. However, we were guided by the growing realization in the field that access trumps preservation and "digitize once" may not be a foolproof plan.

Scale and Presentation

Perhaps the largest hurdle for the project was the scale of the endeavor. Although close to 3,000 images have been entered into the repository through this method as of January 2009, more than 4,000 are still backlogged, many with only preliminary metadata records. That may seem a small number relative to other digital initiatives; but is significant given that these materials were all "captured" from existing work — a repository built in the interstices between meetings, processing, desk work, and the day to day activities of a typical librarian or archivist. Most impressive, these 7,000 images represent an archive that no one person had curated, collected, or planned for; a wildly diverse collection that was, in a sense, found on the doorstep.

If the repository had been finished, and the web-based administrative interface made available, it is possible that many of those images in the backlog would now be online. However, a relatively robust metadata standard, designed so that records could be easily repurposed and shared, added a significant burden to the existing scanning workflow. Added to the robustness of the metadata was the volume of scanning requests, often so dense that there was little time left over for metadata — staff were more inclined to be preparing their next item for scanning, rather than creating a metadata record.

To address the problems of scale, one solution might have been to divide the labor for creating individual metadata records by assigning initial basic descriptive information (a title, a creator if applicable, and a description) to an image when it was scanned. Once the image was digitized, the "stub record" would go into the repository with the image. Then catalogers from the Technical Services department would go through and augment these records with more detail and controlled subject headings. In this way, items would not sit in a backlog far from the public view, but would be available with some basic metadata even before a fuller description could be created. In addition, the curator's knowledge of the

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

collection could be harnessed, but without asking that curator to acquire the skills of a cataloger along with those of an archivist. Although such processes were not part of the original plans for workflow at the UM Libraries, one of the authors has successfully implemented such a workflow for digital collections at East Carolina University.

Aside from scale, issues of collection-scope proved a concern with this project, and ultimately provoked among administrators a desire for stricter guidelines concerning what was to be captured and stored in the repository. Patrons tend to request a predominance of images of sports events and sports figures – certainly a part of the University of Maryland's history, but not the only part that should be highlighted. It might be argued that the unplanned bias of this collection accurately represents the most-used parts of our archive; it could also be argued that the Libraries have a responsibility to provide materials for all forms of research and inquiry, not just those that present themselves most often. While the project had originally been designed in response to the need for an image management system and a hope to avoid the repeated scanning of the most popular requests that might come from restricting content, administrators also argued for the benefit of having the organization spend its time and resources on getting the best materials online first.

Added to the concern about sports-centric content was a concern that the lack of an overall selection focus for the thousands of captured images represented a problematic departure from the way that other digital collections were created at UM. Indeed, the original concept behind DCR was that digitized objects would be created in "collections." As with traditional archival arrangement, these collections of similar material would be presented together for researchers to examine as a group. But items scanned as a result of a patron request belong to no single collection. Presenting this vast sampling of our holdings online and through an interface that would give users some context was a challenge.

It might be argued that the conventional idea of "digital collections" is itself inherently limiting and potentially outdated. Relying on the "first order of information" concept described by Michael Weinberger in Everything is Miscellaneous, the traditional

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

understanding of the "collection" relies on the idea (and even necessity) that things belong in one particular place and one place only.[10] But the realities of digital access make that unnecessary. Moreover, it is unlikely that many users are arriving at our digital libraries through the "front door" and browsing through our carefully crafted collections as we intend. Instead, they are finding individual objects through search engines. As internet searching statistics show, and numerous usability studies and library web analytics confirm, users look for information using search engines. The Pew Internet and American Life project reported in 2008 that the number of individuals using a search engine daily is just under 50% and is above 60% for certain demographics like college graduates.[11]

Given these statistics, it might well be asked: why put digital objects into collections at all? In answer, it can be noted that, even if most users find content on the web through search engines they might still find useful information in the relationships between objects that collections can provide. Taking that notion further, it might be argued that objects may be part of many different "collections" based on their diverse qualities. For example, a 19th century work on agriculture published at the University might belong to collections on the history of agriculture, the history of the University, the bookshelf of a noted agrarian, or a number of other topics. So the problem with the UM project, then, was not that the digitized material collected in response to user requests fit into no collection, but that with items selected from across the institution's holdings possible collections were too numerous to define. With more materials added every day, the difficulty was in trying to logically group items when there was no idea if what was added in the next day, week, month or year would change the scope of the online materials.

In response to the problems with scale and presentation of the materials, a collection development policy of sorts was created in late 2007 requiring that all content fit into one of 18 broad and browse-able subject categories. This policy was developed by a representative team of staff members from Archives and Manuscripts and DCR. The subject categories would not limit the creation of new collections should they arise in response to other needs. But curators were asked to keep these collecting areas in

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

mind when adding digital objects to the repository. When materials fit the guidelines of the policy, a metadata record was to be created and the object added to the online collection. If the item fell outside of the guidelines, it could be simply scanned and deleted.

Abandoning the idea of attaching every digitized object to a unique collection was a move towards what Weinberger has described as the "third order of information," in which materials are not grouped at all, but retain multiple, not pre-determined qualities.[12] These qualities, like the broad subjects, can be searched and aggregated into groups of relevant results and can promote serendipitous discovery. For example, a search for images on "kindergarten" might lead a user back to an "Education" collection, which could potentially lead to many more relevant images.

With the rich assortment of potentially useful subjects, themes and interesting content hidden beneath the repository interface, other discovery methods were discussed that could utilize emergent web 2.0 and data visualization techniques, such as tag clouds and hyperlinked terms in metadata records.

Outcomes

As a result of this approach more than 7,000 images have been created and either ingested into the repository or await ingest in the scanning database. Out of that 7,000, 1,200 items have been selected for inclusion in two thematic collections. The single biggest beneficiary of this approach, in terms of sheer numbers, was the University Archives. That department, which normally receives the most requests for scans of materials, was also in the midst of the publicity campaign for the University's 150th anniversary. A glossy coffee table book and a full-length documentary were two of the major projects undertaken, and both relied heavily on scans of images and documents from the University Archives. With the addition of images that had been scanned and saved prior to the beginning of this project, approximately 2000 images were documented in this manner and were ultimately added to the digital repository to form a still-growing collection called University AlbUM <http://www.lib.umd.edu/digital/album.jsp>.

Another important set of images captured in this manner was a collection of postcards held by the National Trust for Historic

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

Preservation Library Collection housed at University of Maryland. Thousands of postcards documenting historic buildings, destinations, and important architectural styles proved to be popular requests from patrons. In addition to capturing these requests, the librarian in charge of this collection decided to fill in some of the gaps. She set out to systematically digitize the collection and to use the scanning database to capture metadata for future ingest into the digital repository. That effort is now publicly available as the National Trust Library Historic Postcard Collection < http://www.lib.umd.edu/digital/ntlpostcards.jsp>. Although much smaller than the University AlbUM, the online collection represents only a small portion of the digitized postcards, which will be added regularly to the online collection.

Discussion: What Can Be Learned From This Case Study

Librarianship, and certainly curatorship, does not naturally gravitate toward ceding control over any aspect of collections. However, giving up some control over digital selection at the University of Maryland Libraries created a more efficient path to building digital collections by capturing and supplementing an existing workflow. By involving people across the organization and not just those identified as part of the "digital" department, production increased. By distributing the "burdens" (and the satisfactions that come from building a publicly accessible collection), a digital collection was created that was larger and more diverse than one requiring the careful selection of each digitized item. Capturing the existing workflow from patron requests meant building an ostensibly neutral collection. Nothing is ever really without interpretation or bias, of course; and nothing could highlight that fact more clearly than the very pronounced bias toward sports in the University AlbUM collection. But this concept and practice of "neutral collection-building," as opposed to building a collection based on curator selection, enables a collection to capture items that have built-in value to someone other than the curator.

In the end, that collection will reflect the everyday and heavily used holdings rather than the jewels in the crown. The development of digital collections at the University of Maryland became less of a

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

"trophy" service and, instead, began to build toward the critical mass of online, original research content that will enable our digital collections to be a truly valuable part of how research is done in the 21st century.

This neutral collection building requires a different focus of concentration, however. As the Maryland example shows, rethinking our current paradigms for packaging and presenting information is key to the success of initiatives like this one. The inherited museum model of the earlier part of this new century relied on creating "exhibit-style" digital collections that provided large amounts of context to guide users through the carefully shaped narrative of a given collection. The University of Maryland's intention from the beginning of the repository project was to break that mold and focus instead on access to many more images, in many more ways, in order to allow the researcher to build their own context and connections, just as they do in their current research in the library's archives and special collections. Truly providing access at this level requires trying new methods to bridge the gap between repository and user. It seems counter to this line of thinking to insist that this type of undertaking also requires the creation of clear policies about what will and will not be done, but the Maryland initiative might have been buried under a mountain of unreasonable demands if limitations were not developed. These limitations turned out to be advantageous as they offered the opportunity, once again, to thinkabout presentation and collection-building.

Finally, it's worth noting that not only digitization, but problem-solving was distributed in the University of Maryland model. Many of the ideas to solve particular workflow problems — such as stub records, minimizing collections and using a broad vocabulary of subjects — these ideas came about because the "problems" weren't just owned by DCR, but rather by everyone involved in the project. The meeting of minds between archivists and digital collection librarians is a good example of the ways that digitization can benefit from the input and strategic planning of the entire institution.

By focusing on ways to streamline the process of building digital collections, and building upon the existing workflows and expertise of the organization as is possible and effective, digital

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

collection building can become a core function of the library, and digital collections can begin to build to a critical mass, so that researchers can come to the web to conduct systematic original research using digitized primary sources.

Perhaps the overarching challenge in this endeavor is that digitization is still not considered a core function of most libraries' missions. Even though it may be stated in new mission statements, very little has really been done in most libraries to organize around digitization. But in order to open up our collections to new and exciting forms of scholarship, the digitization of our unique materials must become more central to library operations. The model the authors pursued at the University of Maryland Libraries may point toward at least one method for moving digitization to the core of Library operations by tapping into existing resources. It should not be the final step, but it can be the first.

References

1. Ricky Erway and Jennifer Schaffner. Shifting Gears: Gearing Up to Get into the Flow. Online. Report produced by OCLC Programs and Research. 2007. Available: http://www.oclc.org/programs/publications/reports/2007-02.pdf (December 17, 2008).

2. Council on Library and Information Resources (CLIR), No Brief Candle: Reconceiving Research Libraries for the 21st Century. Report produced by the Council on Library and Information Resources, 2008. Available: http://www.clir.org/pubs/reports/pub142/pub142.pdf (February 5, 2009).

3. Karen Coyle, "Mass Digitization of Books," The Journal of Academic Librarianship 32 (November 2006): 641-645.

4. Daniel Greenstein and Susan E. Thorin, The Digital Library: A Biography. Online. Report produced by the Council on Library and Information Resources, 2002. Available: http://www.clir.org/pubs/reports/pub109/contents.html (February 5, 2009)

5. Laurie Lopatin, "Library Digitization Projects, Issues and Guidelines: A Survey of the Literature," Library Hi Tech 24 (2006): 273-289.

6. Nicholas Joint, "It Is Not All Free on the Web: Advocacy for Library Funding in the Digital Age," Library Review 57 (2008): 270-275.

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.

***NOT THE PUBLISHED VERSION;* this is the author's final, peer-reviewed manuscript.** The published version may be accessed by following the link in the citation at the bottom of the page.

7. M. Boock and Ruth Vondracek, "Organizing for Digitization: A Survey," Portal 6 (April 2006): 197-217.

8. Erway and Schaffner, "Shifting Gears"

9. CLIR, "No Brief Candle"

10. David Weinberger, Everything Is Miscellaneous (New York: Times Books, 2007), p. 17-18

11. Deborah Fallows, Pew Internet and American Life Project: Search Engine Use. Online. Available: http://www.pewinternet.org/pdfs/PIP_Search_Aug08.pdf (February 5, 2009).

12. Weinberger, Everything is Miscellaneous, p. 19

The Journal of Academic Librarianship, Vol. 35, No. 5 (September 2009): pg. 468-474. DOI. This article is © Elsevier and permission has been granted for this version to appear in e-Publications@Marquette. Elsevier does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Elsevier.