

Balancing Boutique-Level Quality and Large-Scale Production: The Impact of “More Product, Less Process” on Digitization in Archives and Special Collections

Introduction

FEW WRITINGS WITHIN THE REALMS of archives and special collections have reverberated throughout the field to the degree of Mark A. Greene and Dennis Meissner’s 2005 article “More Product, Less Process: Revamping Traditional Archival Processing.”¹ Its clarion call to shift archival arrangement and description away from time-consuming, detailed processes toward rapid, minimalist strategies was met with both cheers and groans that continue unabated. In a sign of its ubiquity, the approach advocated by Greene and Meissner is now simply referred to as “MPLP,” and its implementation is the subject of ongoing discussion in numerous venues, including professional journals, conference sessions, blogs, and listservs.

Although its original focus was on arranging and describing collections more efficiently to remove backlogs and expand user access, MPLP subsequently “has cut to the core of Special Collections and Archives” and is being applied to many other areas.² Greene and Meissner’s recent follow-up article, “More Application While Less Appreciation: The Adopters and Antagonists of MPLP” offers a thorough discussion of MPLP’s impact, including its adaptation to preservation, reference service, electronic records, and digitization.³ This applicability across the spectrum of archival responsibilities is not surprising, given MPLP’s underlying philosophy of rethinking how resources should be allocated and managed to improve user access to collections. Although many readers have tended to focus on the compromises involved in accepting Greene and Meissner’s challenge to alter traditional processing

1. Mark A. Greene and Dennis Meissner, “More Product, Less Process: Revamping Traditional Archival Processing,” *The American Archivist* 68.2 (Fall/Winter 2005).

2. Lisa Carter, “It’s the Collections That Are Special,” *In the Library with the Lead Pipe* (blog), July 15, 2011, available online at www.inthelibrarywiththeleadpipe.org/2009/its-the-collections-that-are-special/ [accessed 1 September 2011].

3. Dennis Meissner and Mark A. Greene, “More Application While Less Appreciation: The Adopters and Antagonists of MPLP,” *Journal of Archival Organization* 8.3/4 (July–Dec. 2010).

practices, the authors maintain that their message has always been about a bigger picture: “MPLP, fundamentally, is not about specific processing actions. It is about resource management, whether on a program or enterprise level.”⁴

The strategic management of resources is especially critical to success in digitization, which has emerged as a fundamental duty in archives and special collections over the past fifteen years. An ongoing shift away from resource-intensive digitization processes toward large-scale production models is being driven by both MPLP principles and the increasing need to maximize online access to collections in an environment of shrinking staff and budgetary allocations. This article will examine the influence of MPLP on digitization practices and discuss how a major digitization project involving the John Muir Papers included decision making that reflects elements of the MPLP philosophy. It will also analyze how the minimalist metadata practices advocated by MPLP require careful implementation within the context of evolving user expectations for locating information on the Web. In considering the scope of MPLP’s impact on digitization, archives and special collections repositories can more effectively adopt its principles to improve online access within their unique institutional contexts.

MPLP and Digitization

The emergence of digitization as a core archival duty has required institutions to dedicate significant resources to its support. In the initial phases of integrating digitization into their workflows during the 1990s and early 2000s, institutions were often guided by best practices that advocated resource-intensive processes resulting in what came to be known as “boutique” collections. These digital collections feature high-resolution, full-color images of carefully selected (or “cherry-picked”) items with great aesthetic appeal, accompanied by extensive item-level metadata. These materials typically represent the highlights of an institution’s holdings. The general philosophy behind the boutique approach is to use scanning technology to capture archival items in great detail for preservation purposes and to provide online access that replicates the experience of engaging the original. While these projects are largely successful in meeting these objectives, the amount of time and resources required to support this strategy limits the amount of material that can be digitized.

The conventional wisdom that archival materials should be selected, scanned, and described at high levels of detail is now being widely challenged as the MPLP approach to processing collections has informed a reconsideration of digitization practices. In 2010, Paul Conway observed, “We are at the end of the era of ‘boutique’ digital scanning projects for which the principal goal is experimentation

4. *Ibid.*, 175.

with new technologies and extraordinary attention to the unique properties of each artifact.”⁵ Large-scale strategies that reject the notion of hand-selecting and intensively describing items are increasingly supplanting strictly boutique forms of digitization. Instead, they advocate the bulk digitization of entire collections or series within collections with greater speed and less description.

The relevance of MPLP to digitization is readily apparent in its major tenets:

- Make user access paramount: Get the most material available in a usable form in the briefest time possible.
- Expend the greatest effort on the most deserving or needful materials.
- Establish an acceptable minimum level of work and make it the processing benchmark.
- Embrace flexibility: Do not assume all collections, or all collection components, will be processed to the same level.
- Do not allow preservation anxieties to trump user access and higher managerial values.⁶

All five of these concepts are germane to digitization, but the third one represents the linchpin for moving from boutique practices to large-scale production. Greene and Meissner contend that benchmarks for both image capture and metadata should focus on less-detailed levels of production to make more collections available online. “Given the demand by our researchers for more and more digitized material, and given the proposition—which Greene has presented in several previous venues—that researcher use is the purpose of all archival effort, we should adopt approaches to scanning and metadata that dramatically increase how much we can make accessible for the same or even fewer resources. What is that approach? It is one that abjures item-level metadata.”⁷

Instead of describing each item individually, the MPLP perspective indicates that metadata should describe digitized items collectively according to the collection, series, box, or folder in which they are hierarchically organized. Collective description is often employed in the finding aids that are created when an archival collection is processed; in some cases, online finding aids are being used as the descriptive framework for digitized items. The overarching goal is to provide enough description to get users “in the ballpark” at a collection, series, box, or folder level, then let them take over responsibility for finding the specific items that meet their needs.

5. Paul Conway, “Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas,” *Library Quarterly* 80.1 (Jan. 2010): 79.

6. Greene and Meissner, “More Application While Less Appreciation,” 175–76.

7. *Ibid.*, 195.

In terms of image capture, the MPLP perspective indicates that institutions should scan at the minimum level required to make the digitized item usable to most researchers instead of creating high-resolution images that replicate the original item down to the finest detail. Greene and Meissner do acknowledge that some collections may legitimately merit detailed processing and/or digitization, but this treatment should be reserved for exceptional material that warrants the corresponding investment of time and resources.

This philosophy is reflected in a growing number of large-scale digitization models that are based on the MPLP-oriented belief that it is in the best interests of users to have less detailed access to more content. OCLC has been especially active in facilitating this school of thought. Its 2007 publication *Shifting Gears: Gearing Up to Get Into the Flow* presents a number of provocative arguments for adopting values from the mass digitization of books to the digitization of archival material. Taking a page from the MPLP playbook, Ricky Erway and Jennifer Schaffner argue: “Scaling up digitization of special collections (here defined as non-book collections, such as photographs, manuscripts, pamphlets, minerals, insects, or maps) will compel us to temper our historical emphasis on quality with the recognition that large quantities of digitized special collections materials will better serve our users. This will require us to revisit our procedures and policies.”⁸ Metadata creation is undoubtedly a major part of these processes, as acknowledged by Oya Y. Rieger in her article on large-scale digitization of special collections: “Operating in a high-throughput environment requires that we reconsider what constitutes essential elements of metadata and how they facilitate discovery and access to materials. Metadata creation efforts need to be much more focused and move away from a ‘just-in-case’ mode.”⁹

Four years after the appearance of *Shifting Gears*, OCLC published *Rapid Capture: Faster Throughput in Digitization of Special Collections*.¹⁰ This publication presents nine vignettes that describe how large-scale digitization of archives and special collections is being implemented at institutions using a broad range of techniques and equipment. All of the projects share a common commitment to maximizing the amount of material that is digitized and reflect MPLP principles within specific institutional and collection contexts. Similarly, a session at the 2008 Society of American Archivists annual meeting analyzed large-scale digitization practices

8. Ricky Erway and Jennifer Schaffner, *Shifting Gears: Gearing Up to get into the Flow* (Dublin, Ohio: OCLC Programs & Research, 2007), 2.

9. Oya Y. Reiger, “Enduring Access to Special Collections: Challenges and Opportunities for Large-Scale Digitization Initiatives,” *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 11.1 (Spring 2010): 16.

10. Ricky Erway, *Rapid Capture: Faster Throughput in Digitization of Special Collections* (Dublin, Ohio: OCLC Research, 2011).

based explicitly on MPLP, including how their outcomes compare with those of boutique digitization processes at three universities.¹¹

The digitization of the John Muir Papers at the University of the Pacific Library represents not an either/or comparison of boutique vs. large-scale approaches, but rather an integration of them that embraces the MPLP tenant of adopting rapid, minimalist processes when possible and intensive, detailed processes when merited. Greene and Meissner seem to sanction this kind of strategy in stating: “Rather than falling prey to the supposed dichotomy between ‘Googlization’ (ultra-mass digitization) and ‘boutique’ digitization (most of what we are currently doing—unique items on a very small scale) we should be planted on a middle ground: larger quantities of unique material.”¹² The Muir project also represents an underappreciated aspect of MPLP, which acknowledges that different components of a single collection may necessitate different levels of arrangement, description, or digitization. Analysis of the content and use patterns of separate series within the Muir Papers in conjunction with the resources available for digitization led to varying levels of detail in image capture and metadata practices. As a result, the Muir Papers offer a useful case study in how the influence of MPLP can result in the combination of boutique and large-scale practices.

Digitization of the John Muir Papers

John Muir (1838–1914) is one of the most important figures in California history, and widely recognized as the forefather of the modern environmental movement. A preeminent nature writer, Muir founded the Sierra Club and was instrumental in the development of the National Parks System. The University of the Pacific Library’s Holt-Atherton Special Collections is home of the John Muir Papers, the largest collection of Muir documents in the world. Scholarly and popular interest in Muir continues to increase over time, as indicated in the ongoing flow of books, articles, and films produced about him and related topics. Inevitably, the scholars producing these works need extensive access to this collection, making it an excellent candidate for digitization.

In the 1980s, archivists at University of the Pacific recognized the need to make the Muir Papers more widely accessible and reduce wear and tear on the original documents. A microfilm project was undertaken to capture the Muir Papers and Muir-related documents at multiple institutions in a format that could be repro-

11. The presentations from the session “Less Process, More Pixels: Alternate Approaches to Digitization and Metadata” can viewed online at www.archivists.org/conference/sanfrancisco2008/docs/session701-ranger.pdf; www.archivists.org/conference/sanfrancisco2008/docs/session701-focke.pdf; and www.archivists.org/conference/sanfrancisco2008/docs/session701-belden.pdf [accessed 15 July 2011].

12. Greene and Meissner, “More Application While Less Appreciation,” 196.

duced and distributed to various libraries. This project successfully expanded access to this material, although researchers still had to visit an institution that held a copy of the microfilm and cope with the inherent limitations of this format in terms of legibility and ease of use.

Twenty years after completion of the microfilm project, the Holt-Atherton Special Collections staff began digitizing selected series in the John Muir Papers to establish global 24/7 access over the Web. In the process of digitizing separate series of 242 photographs, 78 journals (3,200 pages), 385 drawings, and 6,500 letters (12,400 pages), the fundamental activities of selection, image capture, and metadata creation progressed from “mostly boutique with some large-scale practices” to “mostly large-scale with some boutique practices.” Grant funding supported most of these projects, and the preparation of a Library Services and Technology Act (LSTA) grant application in partnership with the University of California, Berkeley’s Bancroft Library was a turning point in accepting large-scale standards of “good enough” for both image capture and metadata. This evolution and the underlying strategies within each area are summarized below.

Setup

Before discussing selection, image capture, and metadata creation, it is important to review the setup phase that preceded these activities. The first three series (photographs, journals, and drawings) were all digitized in-house, while the correspondence series was outsourced. As a result, the setup phase varied significantly between the photographs/journals/drawings and the correspondence. For the in-house projects, the same scanning equipment, scanning standards, and metadata practices were employed so that, once these were established for the photographs, they were already in place for the other materials. The journals and drawings were scanned by student employees who had to be trained in the handling of the material and the use of the scanners. This process was not overly burdensome, and the mostly homogenous size of the material allowed for the transfer of the setup infrastructure from one project to another, saving time and resources.

Outsourcing digitization work can result in significant savings in terms of cost, time, and resources, but the initial setup phase in this approach can prove quite time-consuming. Digitizing the Muir correspondence involved a vendor scanning the letters on microfilm, extracting descriptive metadata from the microfilm, and creating full-text transcripts of the letters. While the image capture standards were easy to establish, the metadata processes were much more complex in terms of deciding which elements of the existing metadata would be used, how some metadata would be altered, and how the transcribers would deal with illegible words and other issues. These deliberations required multiple conversations between the project managers

and the vendor, along with some trial and error with samples, but the time invested in the setup phase helped to ensure quality and consistency in the final product.

The setup activities in the correspondence project also included one other time-consuming process. Since the original microfilm set included letters from thirty-five institutions, forms were sent to all of them requesting permission to digitize their letters and offering assurance that the home institution would be acknowledged in a credit line with each letter. Every institution ultimately gave permission, but the process of communicating and following up with all of them took a significant amount of time. While both boutique and large-scale digitization models require investment in project setup, the shift to large-scale methods involves additional emphasis on establishing standards and practices that can be reused repeatedly to facilitate faster production.

Selection

The initial decision to digitize portions of the John Muir Papers was made on the basis of its being the most highly used manuscript collection at the University of the Pacific Library, both in terms of on-site and remote (e-mail and telephone) researchers. Within the collection, the selection strategy focused on choosing the most significant and highly used series as the top priorities. This determination was based on data documenting the number of on-site and remote requests involving the different series within the collection, as well as the Special Collections archivists' curatorial opinions on which series were most important in terms of studying Muir's life and influence.

The first series involved photographs, and a boutique approach was employed in selecting only the 242 photos that include Muir instead of digitizing the thousands of photos that make up this series. This decision was based primarily on patterns of user requests for reproductions over the previous five years and the assumption that there would not be sufficient researcher interest in photos that do not include Muir to merit investment in their digitization. Selection within the subsequent series (journals, drawings, and correspondence) embodied a large-scale approach by including all the materials within the respective series instead of cherry-picking selected items. The research value of the items within these series is fairly consistent, and the items provide valuable context to each other, especially within the correspondence series.

Image Capture

For the in-house digitization of the photographs, journals, and drawings, a boutique approach was used in choosing to scan originals in high-resolution (600 ppi) 24-bit full-color instead of scanning the grayscale microfilm version of these series at a lower resolution. This decision provided greater clarity and level of detail compared to scanning the microfilm, which is especially problematic in its reproduction

quality for finely detailed drawings. The scanning of the journals was also informed by additional preservation and usability concerns: most of the journals are in pencil that is susceptible to smudging and fading, and portions of them can be very difficult to read on the microfilm.

A shift in image capture strategy occurred in planning the digitization of the correspondence. A large-scale approach was used in scanning the microfilm at 400 ppi, 8-bit grayscale instead of scanning originals in higher resolution and full color. The time and resources required to scan the originals would not significantly improve legibility, since most of the letters were in ink that displays clearly on the microfilm. An additional benefit of this approach was the rapid speed and low cost at which the vendor could scan the microfilm, a clear indication of moving toward a large-scale production model.

Metadata

A consistent large-scale approach to metadata was applied in digitizing the photographs, journals, and drawings. Only pre-existing descriptions of the materials were used, with no additional descriptive metadata added in the digitization process. Existing descriptions were deemed sufficient to support use of the material online, and the time and labor required to examine the items and create additional metadata would be prohibitive in terms of available staff. For most of the material, this involved only title/caption and date. A large-scale approach also was employed with the correspondence by using pre-existing descriptions (creator, recipient, date, and location), but this project did also include the boutique practice of creating full-text transcriptions. The vendor was instructed to extract descriptive metadata as it appeared on the microfilm despite inconsistent use of punctuation, abbreviations, and formatting of names. A quality standard of “good enough” prevailed in assuming users could understand the metadata without additional editing for standardization.

The decision to fund the creation of full-text transcriptions was based primarily on the availability of the LSTA grant. Although creating transcriptions is anathema to MPLP, its recommendation to strategically employ available resources was reflected in our willingness to jettison this aspect of the project if the grant application was not approved. Even with the transcriptions in place, the project managers had to accept a less-than-perfect level of accuracy as it was impractical to correct mistakes in over 12,000 transcript pages. Instead, the quality-control review of transcript samples early in the project was used to convey expectations and test for accuracy in collaboration with the vendor. Resisting perfectionist tendencies can be a challenge, but implementing the MPLP philosophy involves avoiding this elusive standard.

In spite of the imperfect nature of the transcripts, they have been met with great enthusiasm by scholars who can now search the text of 6,500 letters in a matter of seconds. While user feedback has not been formally quantified, it has been uniformly positive with the exception of lamentations from authors who completed their works on Muir prior to digitization of the correspondence. So far, no one has said, “I wish you hadn’t spent all that those resources on letter transcriptions so you could have digitized other parts of the collection instead.” This illustrates one of the continuing challenges of measuring success in digitization, as users are generally unaware of what other collections might have been made accessible if resources were directed in a large-scale, rather than boutique, manner. Archivists and special collections librarians, however, should be well aware of how their decisions on a specific project impact the ability to digitize other materials. In the case of the Muir correspondence, the project managers decided that full-text searchability was important enough to explicitly include it in the LSTA grant application even if that meant those funds would not be used to digitize additional materials in the Muir Papers.

The radical improvement in access brought about by the searchability of item-level metadata is a major user benefit that would be impossible with strictly minimalist metadata. Of course, this level of access came at a cost that was paid in full by a grant, but decisions to pursue value-added metadata such as full-text transcriptions should be made by weighing the potential benefits to users in conjunction with the practicalities of resource allocation. The MPLP concept of making minimalist practices the standard and making exceptions when warranted was reflected in the metadata policies employed in the digitization of the Muir Papers.

In general terms, the MPLP-inspired transition to less intensive image capture for archival collections seems to have little user impact beyond the aesthetic experience. Institutions may be willing to compromise on this issue, as their digitization objectives place more emphasis on the scope of the content they make accessible rather than its visual appeal. The impact of the minimalist metadata practices, however, needs to be fully assessed to ensure that this approach does not overly compromise the ability to meet user needs and expectations for discoverability in the online environment.

Metadata Considerations

The implementation of minimalist descriptive practices can clearly increase the speed at which archival collections are digitized. Yet tradeoffs are also likely to be made in how effectively items can be found through the Web search engines that users often employ to locate digitized content. Schaffner offers insights on the connections between metadata practices and users’ search behaviors in *Metadata*

is the Interface: Better Description for Better Discovery of Archives and Special Collection, Synthesized from User Studies. She observes:

“People expect to find archives and special collections on the open Web using the same techniques they use to find other things, and they expect comprehensive results. Invisibility of archives, manuscripts and special collections may well have more to do with the metadata we create than with the interfaces we build. Now that we no longer control discovery, the metadata that we contribute is critical. In so many ways, the metadata is the interface.”¹³

This conclusion is echoed by the findings presented in *The Digital Information Seeker: Report of Findings from Selected OCLC, RIN, and JISC User Behaviour Project*. Its analysis of twelve user studies from the United States and United Kingdom asserts, “High-quality metadata is becoming more important, not less, for discovery and evaluation of appropriate resources... This is a direct consequence of the explosion of resources and formats, and the expansion of results lists.”¹⁴ The importance of metadata is inarguable, yet the definition of “high-quality” metadata requires further articulation. For some institutions, this may mean comprehensive descriptive frameworks, while others may regard quality as related primarily to the precision and succinctness of descriptive practices. Within this context, however, the ability of a given institution’s digital collections platform to make metadata fully available to Google and other search engines for indexing is of paramount importance. Beyond this technical issue of search engine optimization, the role of metadata as the primary interface for discovery, and the related trend of users who parachute into digitized collections from the open Web, should be fully taken into account in adopting large-scale production models and metadata practices.

This perspective is reflected in a recent study at the North Carolina State University Libraries that analyzed how metadata impacts the frequency with which digitized items are located by users via Google.¹⁵ Using Google Analytics, this project documented how often images in a group of 600 digitized photographs with “enhanced” metadata were retrieved in comparison to a comparable set of 600 photographs without enhanced metadata. (Enhanced metadata was created by staff members by adding information such as subject headings, geographic locations, and descriptive titles.) The results revealed that, over a five-month period, images with enhanced

13. Jennifer Schaffner, *Metadata Is the Interface: Better Description for Better Discovery of Archives and Special Collection, Synthesized from User Studies* (Dublin, Ohio: OCLC Research, 2009): 4.

14. Lynn Silipigni Connaway and Timothy J. Dickey, *The Digital Information Seeker: Report of Findings from Selected OCLC, RIN, and JISC User Behaviour Project* (London: Higher Education Funding Council for England, 2010): 49.

15. Joyce Chapman, “Evaluating the Effectiveness of Manual Metadata Enhancements for Digital Images,” NCSU Libraries Intranet, last modified June 3, 2011, <http://go.ncsu.edu/llzhzy>.

metadata were accessed four times more often than those without it. The study also found that 44 percent of traffic to all of the photographs came through Google.

One would expect more extensive metadata to result in increased hits from search engines, but at what cost? In this study, the average time spent creating additional metadata was seven minutes per image. No information was disseminated on how this figure relates to salary expenditures, but capturing this information in terms of monetary and time costs that can be linked to subsequent use statistics is crucial for being able to assess the return on investment of value-added metadata. From the MPLP perspective, there should also be a corresponding consideration of the degree to which creating additional metadata diverts staff resources from digitizing more material with less description.

To support their advocacy for rapid, minimalist approaches to processing archival materials, Greene and Meissner cite user surveys that convincingly indicate most users prefer access to more archival collections with less detailed descriptions than access to fewer collections with more detailed descriptions.¹⁶ While it seems reasonable to assume these views would also hold true for digitized collections, other studies indicate that user preferences are more varied in the online environment. In reviewing the contradictory results of user studies on preferences for metadata levels in digitized archives and special collections, Schaffner concludes: "These conflicting recommendations suggest that minimum description may come as a relief to some users, but others prefer a full description.... Users support concise minimum-level description, which can also be effective for discovery when it is done well."¹⁷

Mixed messages from researchers were also observed by Joshua Ranger in a comparative study where students navigated digitized letters with item-level descriptions as well as digitized letters with folder-level descriptions. The students stated a preference for detailed metadata to support effective browsing and searching, yet they also accepted the legitimacy of large-scale strategies that result in more content with less description. Perhaps the most telling insight of this study revealed, "Interviews suggest a big difference in how long students will spend researching archival documents in person versus the amount of time they'd spend with the same material on-line."¹⁸ This observation suggests the strategy of providing enough metadata to get users "in the ballpark" so they can continue looking for individual items that may be more accepted in the reading room than it is on the Web.

16. Greene and Meissner, "More Application While Less Appreciation," 178–79.

17. Schaffner, *Metadata Is the Interface*, 10.

18. Joshua Ranger, "More Bytes, Less Bite: Cutting Corners in Digitization" (paper presented at the annual meeting of the Society of American Archivists, San Francisco, California, August 26–30, 2008), available online at www.archivists.org/conference/sanfrancisco2008/docs/session701-ranger.pdf [accessed 1 September 2011].

Further research is needed to analyze how researcher behaviors and expectations vary between using collections on site and using them on the Web. In conjunction with additional data on how users search for digitized content on the Web, this information could contribute significantly to determining if the shift to minimalist metadata practices is equally advisable in digitization as it is for the processing of the original archival collections. There is also a need to determine how minimal metadata can be both precise and sufficiently descriptive to support online discoverability. Establishing appropriate standards of metadata quality to simultaneously meet production goals and facilitate accessibility on the open Web will help to avoid a scenario in which more collections become available but users cannot efficiently find them.

Conclusion

The influence of MPLP on digitization and other aspects of archives and special collections will likely increase as institutions seek to maximize access to their unique holdings in an environment of shrinking resources. The corresponding implementation of large-scale digitization will continue to inspire a rethinking of current practices in both image capture and metadata input. As illustrated in the John Muir correspondence project, the digitization of surrogate copies of collections, such as microfilm and photocopies, is especially well suited to the use of large-scale production models. Many institutions microfilmed some of their most important collections during the twentieth century. These microfilm sets offer a treasure trove of archival material waiting to be repurposed for improved user access through digitization.

Large-scale models may also benefit from the emergence of mechanisms for users to contribute descriptive metadata to digitized content. This shift toward user responsibility for creating metadata is consistent with the MPLP expectation that researchers take on greater ownership of the discovery process. Some archives and special collections are currently employing this strategy, with users transcribing everything from historic menus¹⁹ to the weather readings from ships' logs.²⁰ Although issues of accuracy and authority need to be monitored, user-supplied metadata could mitigate to some degree the limitations of minimalist metadata practices and facilitate the adoption of large-scale models.

In moving in this direction, institutions would do well to recognize that, while MPLP recommends a shift to less-detailed benchmarks to support increased levels of throughput, it also acknowledges the need to avoid a rigid "one size fits all" mentality. The Muir Papers project demonstrates how different series within the same collection can legitimately merit varying levels of image capture and description. This approach does not betray MPLP's focus on speed and brevity; rather, it

19. What's On the Menu project, available online at <http://menus.nypl.org/> [accessed 7 March 2012].

20. On Weather, available online at <http://oldweather.org/> [accessed 7 March 2012].

embodies a strategic application of its call to improve user access in a manner that actively looks for opportunities to increase production. As Greene and Meissner themselves acknowledge, "MPLP recommendations are broad strokes that can help archivists make decisions about balancing resources so as to accomplish their larger ends and achieve economies in doing so. Practitioners must shape them into their own institutional contexts."²¹

Many of the lessons learned in digitizing the Muir Papers can be applied to the planning and management of digitization at other institutions. In adopting an MPLP-inspired mindset, repositories should critically examine their existing practices in selection, image capture, and metadata creation in search of adjustments that can increase production without compromising the usability of the material. This involves not only establishing the standard processes to be followed in most cases, but also explicitly allowing for exceptions to be made when appropriate.

In image capture, this includes questioning the need to automatically engage in full-color scanning at high resolutions and considering the benefits of scanning in grayscale or black-and-white at lower resolutions as a standard practice to enable faster levels of production without necessarily inhibiting the use of the images by researchers. For example, textual documents are likely to be fully legible as grayscale images, while exceptions may involve materials such as colored drawings or other visually oriented documents whose usability may be compromised without full-color scanning. A related consideration in image capture involves scanning for access, which can often be achieved at lower resolution without full color, versus scanning for preservation, which typically employs high resolution and full color. While there are some cases, such as the Muir journals' pencil-written text with its susceptibility to smudging and fading, which require capturing the original in the greatest fidelity possible, in many cases archival items are not endangered to a degree that merits across-the-board implementation of this approach. Weighing the various tradeoffs in terms of overarching objectives for content accessibility, aesthetic impact, and preservation concerns with an eye toward increasing speed and production is key in applying MPLP principles to image capture.

The need to assess the impact of various tradeoffs is also true for metadata creation. Generally speaking, repositories should establish standard descriptive practices that incorporate whatever metadata exists in finding aids and other sources, while allowing for flexibility in choosing to expand or enhance existing metadata when warranted. For example, in digitizing the Muir Papers, only pre-existing descriptions of the material were used, with the exception of the correspondence transcripts that were created through grant funding. This involved a measure of

21. Greene and Meissner, "More Application While Less Appreciation," 176.

self-restraint by archivists who knew they could add value to the descriptions from their own knowledge base but refrained in the interests of time and expending their efforts on other worthy projects. This approach was dictated by the planning that guided the project, which involved extensive discussion among the Special Collections staff on the pros and cons of different levels of metadata creation and image capture.

In planning and managing their digitization programs, archives and special collections units will benefit from avoiding the extremes of adopting an inflexible, universal approach to selection, image capture, and metadata creation at one end of the spectrum and an ever-shifting project-by-project approach at the other. By taking part in the continuous self-reflection advocated by MPLP, institutions can seek an appropriate balance in increasing the production of digitized content while ensuring that appropriate levels of quality are met to effectively provide online access to their archives and special collections.



B & L ROOTENBERG

Rare Books & Manuscripts

15th-20th Century

Early Printed & Illustrated Books.

Post Office Box 5049 - Sherman Oaks, CA 91403

Telephone: (818) 788-7765

Telefax: (818) 788-8839

blroot@rootenbergbooks.com

www.rootenbergbooks.com

