Organizing for Digitization at Oregon State University: a Case Study and Comparison with ARL Libraries

By Michael Boock, Assistant Professor/Head of Digital Access Services, Oregon State University Libraries, 121 The Valley Library, Corvallis, OR 97331, Michael.boock@oregonstate.edu

Abstract:

This article presents a case study of how Oregon State University Libraries (OSUL) organized to accomplish digitization activities. Digitization activities are broken down into 6 major categories: management, copyright, digital imaging, metadata, hardware/software/web design, and selection. The OSUL departments responsible for tasks within each of these sets of responsibilities are identified. The OSUL experience of incorporating digitization responsibilities within the existing organizational structure is compared with the results of a previously conducted survey of ARL libraries.

Keywords: Digital Collections; Digital Imaging; Digital Libraries – Organization; Digitization; Libraries – Organization

Introduction

Nearly all large academic libraries digitize print-based materials and provide access to those materials over the World Wide Web.[1] In the last 10 to 15 years, libraries have taken on digitization activities in addition to their other responsibilities. This article discusses the library departments and/or cross-departmental digital project groups that are responsible for digitization tasks at Oregon State University Libraries (OSUL) and Association of Research Libraries (ARL) to give others a sense of the ways libraries might assign digitization responsibilities in order to accomplish digitization activities. For the purposes of this paper, digitization refers to the scanning and description of analog objects such as books and photographs for online accessibility.

A 2005 survey of ARL libraries by the author and Vondracek found that ARL digitization efforts most often begin within the departments that contain the materials to be digitized; for example, libraries' special collections and archives departments.[2] Over time, these libraries most often move selected digitization responsibilities to entirely new digital library departments rather than incorporate the responsibilities within existing departments. Three years have passed since that survey was conducted. It would be interesting to see if this has changed over the intervening years through a follow-up survey of ARL libraries. OSUL has incorporated digitization responsibilities within existing departments. This article contrasts the two approaches and describes the advantages and disadvantages of each approach.

Literature Review

Aside from the ARL survey, few articles or case studies look in-depth at library-wide organizational changes and changes in roles and responsibilities in relation to digitization activities. The literature focuses on discussions of the technical aspects of digitization, such as scanning, metadata creation, standardization, and digitization costs, but does not describe who is doing the work. Other research focuses on the skills and expertise needed by the new 'digital librarian' rather than the organizational structures or specific roles and responsibilities necessary to accomplish digitization. Most case studies continue to document how digital programs emerged or on the digital collections themselves.

Greenstein and Thorin's report describes some of the issues for libraries to consider in the formation of digital programs.[3] Their study of Digital Library Federation libraries discusses the full range of digital library efforts. They include six case studies of larger institutions that describe how they are organized to accomplish digitization, the emergence of the libraries' digital programs, and information about how funding was secured. The case studies indicate that most reporting libraries contain digitization responsibilities within new departments or units rather than transitioning the work to existing units.

Hurlbert and Dujnic describe technical services department involvement in several digitization projects at Carnegie-Mellon and the new procedures and workflows they have developed within the department "to accommodate the various project materials that enter and leave the department."[4] Catalogers are involved in the application of metadata to digitized objects and for ensuring that there are links from the catalog to the digitized

versions. Catalogers are represented on digital project teams and help determine the workflow and procedures necessary to accomplish all aspects of metadata creation for the digitized objects.

Kennedy describes the effects of digitization efforts on ARL preservation departments.[5] She describes the transition of preservation department photocopying and other reformatting activities to include scanning operations and finds that digitization efforts have led to a 10% increase in the total number of books that are processed at a sampling of ARL preservation departments.

In addition to the ARL survey, Vondracek and the author, in separate articles, each discuss how traditional library responsibilities can be evolved to handle digitization responsibilities. Their articles also suggest that blending of responsibilities across public and technical service departments can benefit libraries. For example, reference librarians could be more engaged in metadata assignment and catalogers involved with user interface development.[6]

In an earlier article, the author describes where specific responsibilities were assigned for digitization projects at Cleveland State University, a medium-sized academic library.[7] Responsibilities were assigned to the departments traditionally responsible for specific functions. The necessity of inter-departmental collaboration in completing digitization projects is also emphasized.

The dearth of literature on organizing to accomplish digitization may be due to the relatively recent emergence of digitization programs. There is also a general lack of literature on library organizational structures. The lack of information about how libraries organize to accomplish digitization activities is the reason for this article. OSUL's experience can hopefully serve as one model among many for libraries investigating options for assigning digitization responsibilities within their organizations.

**Distribution of Digitization Responsibilities**

OSUL's earliest foray into digitization occurred in 1994. At this time, as in many libraries' early digitization efforts, the originating department handled most digitization responsibilities. In OSUL's case, and in most other libraries, the originating department is special collections or archives. OSUL gradually began redefining its organizational structure and identifying what departments should have digitization responsibilities as the number of digitization projects increased. Finding the best organizational structure for managing digitization projects was not a straightforward process and is an ongoing effort at OSUL. Whereas ARL libraries have established new digital library departments that report outside of the traditional library reporting structure to handle aspects of digitization, OSUL has largely incorporated activities within existing library departments.

At OSUL, reference librarians, archives staff and special collections staff do most of the selection of content to be digitized. Since 2003 most scanning and all metadata activities are housed within a digital production unit within the technical services department. Three staff positions were reassigned, two from a serials cataloging unit and one from a monographs cataloging unit, to become part of the digital production unit. Hardware,

software and web design work is done in a systems department. Copyright

responsibilities are dispersed throughout the organization. Table 1 illustrates how

digitization responsibilities are dispersed throughout the OSUL organization. In the table,

one box indicates some responsibility for a category of responsibility, two boxes

indicates that the department has primary responsibility or that that department and

another share the primary responsibility. Three boxes indicate that the department is

entirely responsible for that category of responsibility.

Table 1: Simplified Org Chart Showing Where Digitization Responsibilities Reside at OSUL
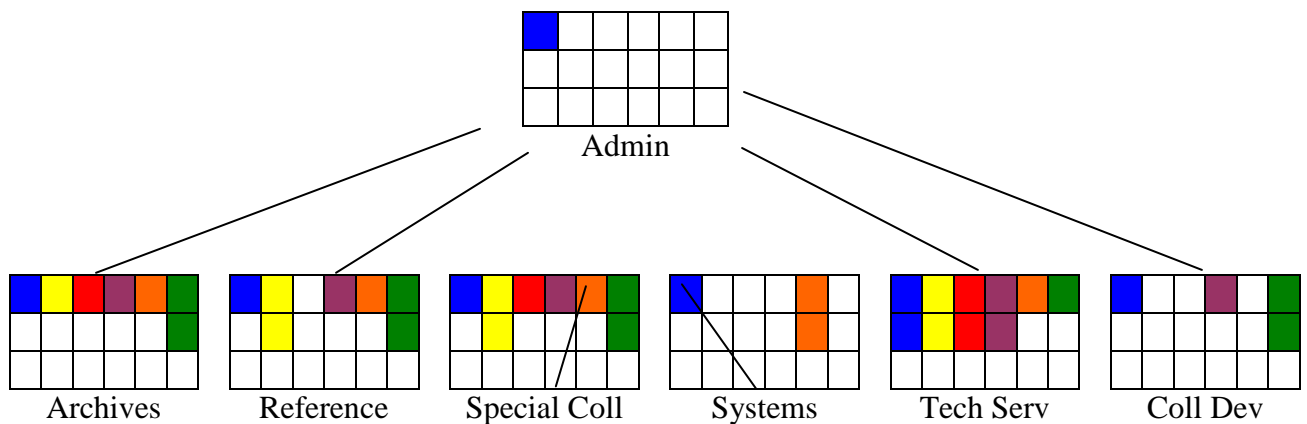
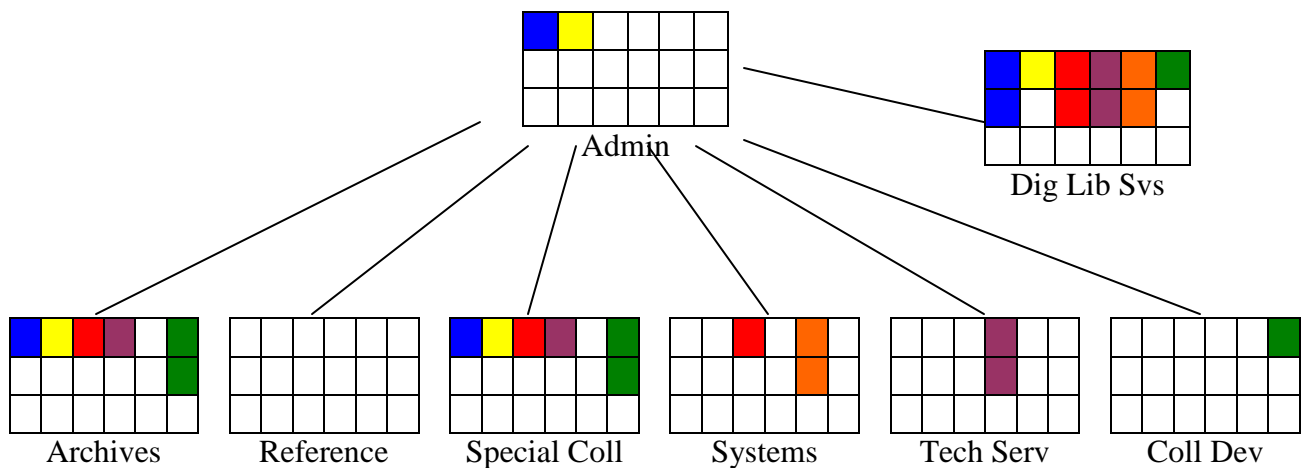|	Management (Blue, col 1)	Copyright (Yellow, col 2)	Digital Imaging (Red, col 3)	Metadata (Violet, col 4)	Hardware/Software/Web Design (Orange, col 5)	Selection (Green, col 6)
Admin	■					
Archives	■	■	■	■	■	■■
Reference	■	■■		■	■	■■
Special Coll	■	■■	■	■		■■
Systems	■				■■	
Tech Serv	■■	■■	■■	■■	■	
Coll Dev	■			■		■■

Management: Blue (column 1)
Copyright: Yellow (column 2)
Digital Imaging: Red (column 3)
Metadata: Violet (column 4)
Hardware/Software/Web Design: Orange (column 5)
Selection: Green (column 6)

At ARL libraries, at least as of the 2005 survey, archives and special collections

departments do most of the content selection for digitization, although many ARL

libraries are currently engaged in "mass digitization" projects with Google, the Open

Content Alliance and Microsoft.[8] Libraries engaged in these efforts do little or no

selection of content appropriate for digitization; rather, everything is selected.[9] Scanning activities at ARL libraries most often happen within newly formed digital library units that report outside of the traditional library reporting structure. As of the 2005 survey there was still little digitization being outsourced although this has no doubt increased at in the last three years. Metadata activities at ARL libraries also most often take place in newly created digital library units, although technical services, archives and special collections departments often have some involvement at these libraries. As at OSUL, systems departments do hardware, software and web design work. Copyright responsibilities, as at OSUL, are widely dispersed at ARL libraries. Table 2 illustrates how digitization responsibilities are dispersed throughout ARL organizations according to the results of the 2005 survey.

Table 2: Simplified Org Chart Showing Where Digitization Responsibilities Reside at ARL Libraries

Admin — Dig Lib Svs

Archives   Reference   Special Coll   Systems   Tech Serv   Coll Dev

Management: Blue (column 1)
Copyright: Yellow (column 2)
Digital Imaging: Red (column 3)
Metadata: Violet (column 4)
Hardware/Software/Web Design: Orange (column 5)
Selection: Green (column 6)

In the tables above and throughout the rest of this paper, digitization responsibilities are distributed into 6 categories of functional responsibility: Management, Copyright, Digital Imaging, Metadata, Hardware/Software/Web Design, and Selection (Appendix A). The departments and position levels that are responsible for these digitization activities at OSUL and ARL libraries are discussed and compared.

In the earlier survey of ARL libraries, respondents were asked to select the department that is responsible for particular digitization functions from these options: administration, archives, collections, reference, special collections, systems, and technical services. Few libraries use these traditional department names any longer, and OSUL is no exception. To compare the results of the ARL survey with the OSUL experience, this article maps OSUL department names to those same generic department names noted above.

Management

Management responsibilities relating to digitization include approving and prioritizing digitization projects, coordinating efforts within the library and with other collaborative partners, securing funding, and promoting the final products. In 2003, OSUL formed a digital library task force charged with overseeing digital library activities, including digitization projects. This task force later merged with the OSUL management team because the membership and responsibilities of the two groups overlapped significantly.

At OSUL, the management team along with the university librarian continues to give final approval for digitization projects to move forward. The group developed a project

proposal form to create a mechanism for library staff to propose projects and to facilitate

proposal review. The management team makes project funding decisions.

Project coordination responsibilities are largely dependent on which department initiated

the project. The heads of the archives and special collections departments coordinate, or

assign coordination within their departments, for digitization of materials from their

collections. These departments work closely with the digital production unit to ensure

that scanning and metadata work is completed in a timely manner and with the systems

department to create collection web sites. The head of technical services tracks most

projects originating from strategies in the OSUL strategic plan, including projects

proposed by reference librarians, and arranges for digital production unit involvement in

all digitization projects.

In addition to the digital library task force, a digital library project group was also formed

in 2003 to direct the more technical aspects of digital production and to research and

agree on standards. After standards were agreed upon, the group dissolved as the digital

production unit within the technical services department was created and OSUL

incorporated the project group responsibilities within existing departments. Informal,

cross-departmental, project groups continue to be formed to handle questions and make

preparations for specific digitization projects.

Digital collections of archives and special collections materials are marketed and

promoted by those departments, with oversight and additional assistance provided by

staff within administration. Marketing includes writing press releases for distribution by

the university communications office. Reference librarians promote digital collections to the academic units to which they are assigned.

At ARL libraries, management responsibilities most often reside in newly created digital library departments, although some libraries report that project coordination is assigned to the originating department. Coordination of digitization projects in ARL libraries moves to newly formed digital library departments as digitization efforts mature.

For medium-sized and smaller libraries, management team involvement in review of project proposals is recommended. It ensures that department heads have an opportunity to address potential stumbling blocks at an early stage; the head of the systems department ascertains storage space availability; the head of technical services – the department responsible for scanning and metadata – determines staff and equipment availability.

Copyright

Every digitization project that OSUL undertakes requires some investigation of intellectual property ownership, copyright verification, and/or securing or negotiating legal rights to materials through correspondence with publishers, authors and other copyright owners. At OSUL, attorneys provide legal guidance on rights management and contract language, but responsibility for copyright within OSUL has been diffuse and not clearly defined. Copyright responsibilities reside in several departments with no real oversight. As a result, consistent copyright processes, policies and standards have not yet been adopted.

OSUL hired a person with responsibility for many aspects of rights management for materials residing in archives and special collections departments. This position also instructs project staff from other departments in how to track down and acquire legal rights for materials selected for digitization and maintains a copyright wiki for OSUL.

Reference librarians track and secure permissions for materials in collections that they propose for digitization. They write the emails that are sent to copyright holders and oversee the work of student assistants responsible for corresponding with copyright holders. Reference librarians also secure permissions for resources that are digitized as part of the OSUL on-demand digitization process that allows librarians and faculty to request digitization of single documents or a small series of volumes.

Collection development departments, increasingly responsible for licensing electronic resources and handling other issues of scholarly communication, may be an appropriate place for copyright-related work to reside within an organization. Paraprofessional staff that report to collection development at OSUL are increasingly responsible for depositing (in a sense, acquiring) university-owned and other librarian identified digital materials into OSUL digital repositories. During the course of depositing digital materials, paraprofessionals could also be responsible for tracking down copyright permissions and storing that information in electronic resource management systems, databases, spreadsheets or registries.

At ARL libraries, copyright-related responsibilities are also widely dispersed. Often these responsibilities are assigned to administrative staff or line staff in coordination with university legal counsel. At least one ARL library employs an attorney with responsibility for coordinating copyright activities; a solution that makes sense for larger libraries. Libraries would benefit by assigning oversight of copyright policies and activities to a single staffperson.

Digital Imaging

Digital imaging refers to scanning and enabling optical character recognition of print materials for full text searching, digitizing other forms of analog media such as photographs, audio and video resources, and assuring the quality of the digital objects.

At OSUL, containing both digital imaging and metadata assignment activities within a digital production unit in the technical services department enables a streamlined workflow and greater coordination between the two sets of tasks. Students doing scanning also apply basic metadata and submit digitized resources to digital repositories where metadata is enhanced by high-level paraprofessional staff.

Almost all digital imaging work at OSUL is accomplished by student employees. Initially, a mix of paraprofessional staff and other higher-level student assistants performed quality control of scanned materials. Involving paraprofessional staff in this process was valuable because it gave them an understanding of what digitization involved and the overall workflow. After initial involvement in quality control and with concomitant training, staff with this responsibility have been transitioned to assigning

metadata. As the quality of student scanning and optical character recognition work has improved, OSUL has decided to no longer check the quality of all items scanned by student assistants.

ARL libraries indicate that digital imaging at their institutions most often takes place in newly created department-level digital library departments that report outside of the traditional library reporting structure and sit at the department level. At many other libraries, special collections departments, where digitization projects began, maintain this role. Other ARL libraries report that digital imaging is done in preservation departments.

Metadata

Metadata responsibilities relating to digitization include selection of metadata structures and schemas (e.g. Dublin Core, MODS, METS), creation of data dictionaries that describe the fields to be used, selection and creation of controlled vocabularies, metadata assignment, batch loading and manipulation, and metadata quality control.

At OSUL, a digital production unit within the technical services department is responsible for all metadata related responsibilities since its inception in 2003. Previously, data dictionary and controlled vocabulary work was either not done or was accomplished by cross-departmental project groups. Prior to technical services involvement, metadata often did not adhere to library standards. Metadata schema selection, development and usage, field property determination, and data dictionary creation and maintenance at OSUL are handled by a newly-created digital production librarian position working closely with a metadata working group. The position was

created out of funds formerly allocated to two now-retired paraprofessional staff from other units in technical services. The metadata work group includes professional catalogers and paraprofessional digital production unit staff. The person that proposes a digital project, ordinarily a reference librarian or member of the archives or special collections departments, attends metadata working group meetings to review and suggest changes to the data dictionary and controlled vocabularies to be used for projects they have proposed.

Paraprofessionals at OSUL have the expertise necessary to describe digital resources using a metadata schema that is less complicated than MARC and content standards less rigid than AACR2. OSUL primarily uses Qualified Dublin Core as the descriptive metadata schema and generates METS records for administrative and structural metadata. Since the schema underlies the digital repository systems in place, it is helpful but not necessary for staff to be familiar with Qualified Dublin Core to assign metadata. They only need to understand how to apply metadata to particular fields they see in the digital repository data entry screens. Students assign brief metadata as part of the scanning process. Paraprofessionals review student-assigned metadata for quality, add descriptive elements, assign subject headings, and perform name authority work prior to the publication of digital objects.

 If metadata is already available for resources that are being digitized, it is reused at OSUL. For example, if metadata is already available in the online catalog, students copy and paste the metadata to the digital repository as part of their work. If metadata is available from another source in the format of a spreadsheet or database, attempts are

made to clean up and batch load the metadata into the OSUL digital repository. This work is also done within the digital production unit.

Involving catalogers and paraprofessionals formerly responsible for cataloging monographs and serials in describing digital objects has been central to the success of OSUL digitization efforts. When the digital production unit was formed, metadata creation responsibilities were assigned to paraprofessionals with cataloging experience and a proven ability to do accurate work efficiently. This was particularly important in getting digital collections up and running quickly in the early stages to show that it was possible and to serve as an example to the rest of the department's paraprofessionals of the type of work with which they would increasingly be involved.

Reference librarians assist technical services staff with the identification of relevant controlled vocabularies for collections that they propose for digitization. They sometimes assist with the assignment of subject terms in their subject areas and assign metadata to digital objects to enable metadata harvesting and retrieval in topical and geographic web portals. In effect, the assignment of metadata to objects in digital repositories is the equivalent of selecting the objects to appear in the portals. Reference librarian involvement is recommended because the librarians have collection expertise and knowledge of subject-specific controlled vocabularies and thesauri within their disciplines. They also have an understanding of how the materials will be searched and retrieved.

At ARL libraries, reference librarian involvement with metadata is not widespread. Most ARL libraries indicate that traditional cataloging departments have some responsibility for metadata activities, although in many libraries those responsibilities are shared with newly formed digital library departments. At many ARL libraries, catalogers are called to serve on project teams led by the newly formed digital library departments, or metadata specialists report to both a cataloging department and the digital library department in a matrix reporting structure.

Hardware/Software/Web Design

Hardware and software digitization activities include the installation, maintenance, support and customization of servers and other computer hardware, digital repository software, digital preservation, and reformatting. Not surprisingly, systems departments at both ARL libraries and OSUL handle these activities.

At ARL libraries and OSUL, systems departments also do most web design and interface development, working closely with project groups, digital production librarians, and staff responsible for proposing collection digitization such as special collections, archives and reference librarians. OSUL has a usability team consisting of a web designer and two reference librarians that are charged with ensuring the OSUL websites are intuitive and easy to use. They work closely with a web design team that consists of two other reference librarians and a web designer.

OSUL also outsources some web design and graphics work to a web design firm. Most ARL libraries, at least at the time of the 2005 survey, did not tend to outsource this or any

other work relating to digitization. OSUL has found that hiring a web design firm to help design web pages has resulted in the best looking and most usable web pages on the library site. Preservation metadata is assigned within technical services. All other digital preservation responsibilities at OSUL and ARL libraries – particularly the tracking of format obsolescence and refreshing content -- reside in system departments.

Programming for digitization at OSUL includes manipulation and batch loading of metadata, accomplished in the digital production unit within the technical services department. Other programming activities relating to maintenance and revision of digital repository software is also accomplished within technical services although increasingly that is moving to the systems department.

Selection

As might be expected, at OSUL, reference, archives and special collections departments are largely responsible for selection of analog materials appropriate for conversion to digital format. Reference librarians at OSUL serve as subject selectors for OSUL research collections. Special collections and archives staff also function as subject selectors because of their specialized knowledge of their collections.

At OSUL, reference librarians collaborate with faculty in their assigned academic departments and programs to identify potential content for digitization. This presents opportunities for OSUL and academic departments to select unique content for digitization that benefits the greater research community. Reference librarians also select individual volumes and small series of volumes for digitization as part of the digitization

on demand process in place at OSUL. This process is an effective way to involve reference librarians in identifying priority items for digitization.

Reference librarians are pivotal to the preparation of digitization proposals at OSUL because of their familiarity with the research that is conducted in their colleges and also their broader knowledge of the content of their subject areas. Reference librarians, because of their reference, selection, instruction and other liaison responsibilities, are best placed to address the significance of digital collections to faculty and students. Reference librarians also help design and assess digital collection user interfaces and web pages to ensure optimal search and retrieval, develop context for digital collections by writing narratives, and create bibliographies and links to other resources.

At ARL libraries, according to the results of the 2005 survey, staff in archives and special collections departments do most content selection for digitization. As digitization efforts have continued to increase, it seems likely that over the last three years reference librarians have become increasingly involved in the selection of content for digitization at ARL and other libraries.

Conclusion

In the 2005 survey of ARL libraries, most respondents indicated that they have created new digital library departments with a full range of digitization responsibilities. Other ARL libraries that had more recently created digital library programs indicated that they had retained responsibilities for digitization in archives, special collections and systems departments rather than transition other existing departments such as technical services to

take on scanning and other digitization responsibilities. OSUL has incorporated digitization responsibilities within its existing organizational structure. For OSUL, digitization projects have been most successful when they involve staff across the library organization. This makes effective use of existing library expertise and results in deeper engagement in digitization efforts throughout the organization.

Having an oversight group at the management level to approve projects and establish standards is critical, as is assigning appropriate levels of funding and ensuring access to necessary resources and training. Centralizing the actual digitization of materials and metadata application in a single unit within technical services is cost-effective, improves workflow and improves adherence to standards. It also allows individuals across the organization to focus their efforts in their areas of expertise rather than take on tasks for which they lack requisite expertise. Grant funds provide opportunities for funding new positions, but reassigning the duties of current staff, particularly in technical services and reference departments, allows staff to transition to new positions that are more effective in moving digital initiatives forward.

Support staff in technical services departments recognize that their jobs are drastically changing. They want the training that is necessary for them to be able to work in the new digital environment. Those that don't want to have a role in the digital environment are allowed to continue doing what they've always done, as long as the work is there for them to do.

As reference librarians spend less time at the reference desk and less time collecting traditional print materials, OSUL anticipates that they will be increasingly involved in working with faculty and staff throughout the university, and as a land-grant institution, throughout the state, building collections of born-digital and digitized resources. In order to do so, they will require understanding of metadata and how metadata structures contribute to accessibility. They will need to be able to review content and understand how it can best be organized and described. They will need to understand how information is harvested and take an active role in selection of resources from other repositories for metadata harvesting.

This transition of reference librarians from traditional public services work to that which has traditionally been the provenance of technical services librarians is currently under way at OSUL. Technical services librarians and staff, in addition to having responsibility for scanning and metadata creation, are also responsible for administration of the OSU institutional repository (IR) and other digital repositories, training university staff in depositing materials to the IR, and soliciting resources from other university departments for inclusion in the IR. It is anticipated that in addition to promoting the IR to academic units and campus offices, reference librarians could also take on collection administration, training responsibilities, and assignment of metadata to materials in the IR for retrieval of the content in OSUL topical and geographic natural resources web portals.

<u>Appendix A</u>

Categories of Digitization Responsibilities

<u>Management</u>
1. Giving final approval for digitization efforts to move forward

2. Coordinating the digitization efforts

3. Negotiating with groups outside the library

4. Promoting the collections and public relations

<u>Copyright</u>

5. Investigating whether or not items are in public domain

6. Soliciting permission to digitize items

7. Coordinating payment for permission to digitize items

8. Managing copyright permissions

<u>Digital Imaging</u>

9. Scanning

10. Optical Character Recognition

11. Digitizing multimedia

12. Quality control of digital objects

<u>Metadata</u>

13. Selecting and developing metadata structures (e.g. determination of schema, field

    selection, field labeling, and data dictionary creation)

14. Selecting and creating controlled vocabularies

15. Assigning metadata

16. Converting and batch loading metadata

17. Quality control of metadata

Hardware/Software/Web Design

18. Installing, maintaining and supporting hardware (servers and PCs) and software

    for digitization

19. Programming for digitization

20. Designing web pages and determining web interface usability for digital

    collections

21. Preserving, maintaining and refreshing digital content

Selection

22. Proposing an item or collection for digitization

23. Selecting content

24. Assessing user needs (e.g. determining the audience, how the audience will search

    and retrieve information, what information and context the audience will require)

[1] Michael Boock and Ruth Vondracek, "Organizing for Digitization: A Survey," *portal: Libraries and the Academy* 6(2): 197-217.

[2] Ibid

[3] Daniel Greenstein and Suzanne E. Thorin, *The Digital Biography, Strategies and Tools for the Digital Library*, Digital Library Association/Council on Library and Information Resources (Washington, D.C.: Sept. 2002).

[4] Terry Hurlbert and Linda L. Dujmic, "Cataloging Department Participation in Digital Initiatives," *Technical Services Quarterly*, 25(1): 19-26.

[5] Marie R. Kennedy, "Reformatting Preservation Departments: The Effect of Digitization on Workload and Staff," *College and Research Libraries*, 66(6): 543-51.

[6] Ruth Vondracek, "Going Beyond Selection," *Library Journal netConnect*, Summer 2003, http://www.libraryjournal.com/article/CA304093 Accessed September 2007.

[7] Michael Boock, et al, "Getting Digitization Projects Done in a Medium-Sized Academic Library: a Collaborative Effort Between Technical Services, Systems, Special Collections, and Collection Development," *Technical Services Quarterly* 20(3): 19-31.

[8] Karen Coyle, "Mass Book Digitization," *Journal of Academic Librarianship* 32(6): 641-645.
[9] ibid