



SPEC Kit 294

Managing Digitization Activities

September 2006

Rebecca L. Mugridge

Head of Cataloging and Metadata Services

Pennsylvania State University



ASSOCIATION OF RESEARCH LIBRARIES

Series Editor: Lee Anne George

SPEC Kits are published by the

Association of Research Libraries
21 Dupont Circle, NW, Suite 800
Washington, DC 20036-1118
P (202) 296-2296 F (202) 872-0884
[http://www.arl.org/spec/
pubs@arl.org](http://www.arl.org/spec/pubs@arl.org)

ISSN 0160 3582

ISBN 1-59407-710-X

Copyright © 2006

This compilation is copyrighted by the Association of Research Libraries. ARL grants blanket permission to reproduce and distribute copies of this work for nonprofit, educational, or library purposes, provided that copies are distributed at or below cost and that ARL, the source, and copyright notice are included on each copy. This permission is in addition to rights of reproduction granted under Sections 107, 108, and other provisions of the US Copyright Act.

The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (R1997) Permanence
of Paper for Publications and Documents in Libraries and Archives.



SURVEY RESULTS

Executive Summary.....	11
Survey Questions and Responses	17
Responding Institutions.....	58

REPRESENTATIVE DOCUMENTS

Organization Charts

Brown University	
Brown University Library	62
University of Florida	
Digital Library Center—Organizational Chart	63
University of Louisville	
University Libraries Organizational Chart	64
University of Michigan	
University Library. Information Technology. Technical and Access Services.....	65
Digital Library Production Service	66
Michigan State University	
Michigan State University Libraries Organization.....	67
Smithsonian Institution	
Smithsonian Institution Libraries.....	68
University of Virginia	
Content Management Services	69
Digital Library Production Services	70

Mission Statements

Brown University
 Center for Digital Initiatives 72

University at Buffalo, SUNY
 About UBdigit..... 73

University of Georgia
 Digital Library of Georgia. Mission..... 74

University of Illinois at Urbana-Champaign
 Digital Services and Development Unit..... 75

University of Iowa
 Digital Library Services. Mission 76

University of Michigan
 Digital Library Production Service. About DLPS 77

Michigan State University
 Digital & Multimedia Center 78

Syracuse University
 Digital Imaging Services Center 79

University of Virginia
 Digital Library Production Services (DLPS) 80

Position Descriptions

University of Chicago
 Preservation Reformatting Librarian 82
 Digital Conversion Manager 84

University of Connecticut
 Preservation Librarian 87

University of Connecticut
 Digital Reformatting Librarian 90

University of Kansas
 Scholarly Digital Initiatives Librarian..... 93
 Digital Imaging Librarian..... 95

University of Kansas
 Coordinator of Digital Content Development 97

Ohio University
 Digital Initiatives Coordinator..... 98

University of Virginia
 Head, Digital Library Production Services 103

Policy and Procedure

University of Alabama
 Digital Projects and Programs: A Comprehensive Strategy 106

University at Buffalo, SUNY
 Digitization Standards: Non-transmissive Materials 108
 Digitization Standards: Transparent Materials 110
 UBdigit Digitization Decision Tree 112

University of Connecticut
 Policy Statement on Cost Recovery for Grant-funded Partnership Projects with Digital
 Collections at the UConn Libraries 113
 Policy Statement on Description & Access for Digital Collections at the UConn Libraries... 114
 University of Connecticut Libraries Digital Collections Partnership Proposal 115
 Guidelines for Submitting Progress Reports on Digital Collections Projects at the Univeristy
 of Connecticut Libraries 116

Purdue University
 Digital Initiatives. About Us..... 119

Syracuse University
 Syracuse University Library Digital Library Project Proposal 127

University of Virginia
 DLPS Text Digitizing Services 129

Selection Criteria

University of Alabama
 Project Evaluation Criteria 132

Brown University
 Digital Projects Selection Checklist 133

University at Buffalo, SUNY
 UBdigit Collection Development Statement 136

University of Connecticut
 " Idea Sheet " & " Selection Worksheet " for Library Staff 137

Cornell University
 Selecting Traditional Library Materials for Digitization 141

Duke University
 Selection Criteria 144

University of Georgia
 Collection Development Policy. Digitized Collections 145

University of Iowa
 Digital Library Services. Project Selection Criteria..... 147

Syracuse University	
Syracuse University Library. Digital Library Project Proposal—Step 1. Criteria Checklist ..	148
University of Virginia	
Framework for Digitization. Guidelines for digital collection building	150

Assessment

Johns Hopkins University	
Usability Services.....	152
Syracuse University	
Digital Imaging Services Center. Customer Survey	154

SELECTED RESOURCES

Books.....	159
Journal Articles.....	160
Web Sites about Digitization Activities	161



SURVEY RESULTS

EXECUTIVE SUMMARY

Introduction

Increasingly, academic and research libraries are becoming involved in reformatting materials from their collections to create digital content and are providing access to that content through metadata. As the management of digital projects and initiatives is a relatively new endeavor for most libraries, there is a significant impact on libraries' budgets, organizational structures, and staffing.

Digitization activities require different models for funding, collection development (to provide broad access to otherwise inaccessible materials), acquisitions (the material being digitized is already part of the collection), cataloging (metadata standards may differ depending on the material being digitized), preservation (migration of formats between software platforms and file formats is critical), and systems office support (for a suite of software instead of just the integrated library system). Staff skill sets are different, as are supporting equipment and computer hardware and software.

This SPEC survey was designed to identify the purposes of ARL member libraries' digitization efforts, the organizational structures these libraries use to manage digital initiatives, whether and how staff have been reassigned to support digitization activities, where funding to sustain digital activities originated and how that funding is allocated, how priorities are determined, whether libraries are outsourcing any digitization work, and how the success of libraries' digital activities has been

assessed. The focus of the survey was on the digitization of existing library materials, rather than the creation of born-digital objects.

Background on Digitization Activities

This survey was distributed to the 123 ARL member libraries in February 2006. Sixty-eight libraries (55%) responded to the survey, of which all but two (97%) reported having engaged in digitization activities. Only one respondent reported having begun digitization activities prior to 1992; five other pioneers followed in 1992. From 1994 through 1998 there was a steady increase in the number of libraries beginning digital initiatives; 30 joined the pioneers at the rate of three to six a year. There was a spike of activity at the turn of the millennium that reached a high in 2000, when nine libraries began digital projects. Subsequently, new start-ups have slowed, with only an additional one to five libraries beginning digitization activities each year.

The primary factor that influenced the start up of digitization activities was the availability of grant funding (39 responses or 59%). Other factors that influenced the commencement of these activities were the addition of new staff with related skills (50%), staff receiving training (44%), the decision to use digitization as a preservation option (42%), and the availability of gift monies (29%). An additional factor that motivated many survey respondents was the need to improve access to library resources. Others commented that participat-

ing in digitization activities was a strategic goal of the library.

In addition to being one of the instigating factors in many libraries' decision to begin digitizing library materials, improving access to the library's collection was cited by all of the respondents as an ongoing purpose behind these efforts. Other purposes that were highly ranked by respondents are support for research (85%), preservation (71%), and support for classroom teaching (70%). For a smaller number (24 or 36%), the purpose of their efforts is to support distance learning. Several respondents reported that promoting the library and its collections was also a reason to participate in digitization activity.

Only four libraries reported that their digitization activities are solely ongoing functions; the great majority (60 or 91%) reported that their digitization efforts are a combination of ongoing library functions and discrete, finite projects.

Staffing

The survey asked whether staff efforts for selecting material, digitizing material, creating metadata, and administering digitization activities are centralized in one unit or distributed across the library. The majority of the responding libraries distribute some or all digitization activities across various library units; only five appear to have a totally centralized organizational structure. Material selection is distributed across the library organization at 50 of the responding institutions (76%) and centralized at 10 (15%); six respondents (9%) report both structures. Material digitization is decentralized at 37 institutions (57%), centralized at 20 (31%), and eight respondents (12%) report both structures. Metadata creation is distributed at 45 institutions (68%) and centralized at 12 (18%), while nine (14%) report both structures. Administration is more evenly divided, with 29 respondents (45%) indicating that it is centralized and 30 (46%) that it is distributed; six (9%) report both structures.

Centralized units that manage digitization activities are, in the majority of cases, specifically

designated digitization units with names such as "Digital Initiatives Program" or "Digital Library Center." In other cases, the centralized unit is the special collections library or department (13%), or the preservation department (9%). In most cases, the head of the centralized unit reports to a high-level library administrative officer such as an assistant or associate library director (38%), or reports directly to the library director (30%).

Survey respondents were asked to indicate the names of units in which specific digitization activities (material selection, material digitization, metadata creation, and administration) take place. The units with primary responsibility for material selection are, unsurprisingly, collection development and special collections. Material digitization occurs in preservation and special collections units, as well as in units designated specifically to support digital initiatives. Even in those libraries that have a unit designated to support digitization activities, material digitization often occurs in other units in addition to that unit. Metadata creation is also widely distributed, although cataloging, metadata, and technical services units were indicated by two-thirds of the respondents. Other units responsible for metadata creation are digitization, special collections, and other public services units. Surprisingly, the word *metadata* only appeared in five of the unit names reported. Administration is the most centralized of the functions and is the least likely to be distributed over a second or third unit. The digital library program was most often mentioned as the administrative unit, followed by archives / special collections, systems, preservation, and library administration.

Decisions about the allocation of staff support for digitization efforts are likewise widely distributed across the library. They are made most often by the heads of the centralized units (64%) or a digitization team, committee or working group (55%). Heads of cataloging, collection development officers, and bibliographers also share this responsibility at a number of libraries. In only two cases do the library business office staff have this responsibility.

In addition, respondents indicated that high-level library administrators and the heads of special collections and other units also help make these decisions (23 and 12 responses, respectively).

In order to address staffing needs for digitization activities, all of the responding libraries redefined some existing positions *to add responsibility for* digitization activities, primarily for selection (80%), but also for metadata creation (66%) and digitization (63%). More than half redefined existing positions *to be dedicated to* digitization and metadata creation. Seventy-seven percent also created new positions to be dedicated to digitization activities, most often for digitization (93%) and metadata creation (67%). Staff who select material for digitization and create metadata for the new items most often had their positions redefined to add this responsibility to their other duties. Digitization staff positions also were most often redefined, though a significant number were newly created (72%).

Survey respondents were asked to estimate the number and full-time equivalence (FTE) of librarians, other professionals, support staff, student assistants, and other staff who participate in digital activities. Forty-eight respondents reported a total of 277 staff who are involved in selecting material for digitization. Librarians make up the largest portion of that group (188 or 68%). The number of librarians per institution ranges from one to 14, but at the majority the number is three or fewer. Only 28 of the 188 librarians work full-time on digitization; the remainder spend only a small portion of their time on this activity. Support staff is the second largest category. Thirteen respondents reported a total of 36 support staff, ten of whom work full-time selecting material for digitization. The number ranges from one to six per institution, but the majority have three or fewer. Of the 22 other professionals involved in selection at 11 reporting institutions, three are full-time. One respondent has four staff in this category, but half of the remainder have only one. Only five respondents report using student assistants for selection and, not surprisingly, all of the 16 work part-time. One respondent

reported that faculty also make selection decisions, but that activity is a very small percentage of their time.

Of the 501 staff who reportedly digitize material, the largest categories are student assistants (256 or 51%) and support staff (110 or 22%). There are almost an equal number of librarians (67) and other professionals (60) who participate, too. Five libraries involve a few others, including interns, volunteers, and a programmer. Support staff are most likely to be employed full-time with digitizing material (43 or 39%). Only 16 librarians (24%) and 17 other professionals (28%) do this task full-time. Three libraries report a small number of full-time student assistants.

Survey respondents reported 327 staff who create metadata. These are most often librarians (124 or 38%) but there are also a large number of students who assist (103 or 32%). While the number of metadata librarians ranges from one to 13 per institution, at most libraries the number is one or two. The number of student assistants ranges from one to 16, but only a few have more than five. Only 23 librarians have this as a full-time responsibility; none of the students do. Twenty-seven respondents report a total of 70 support staff who also create metadata, 14 of them full-time. At 17 libraries other professionals create metadata, although only four of these 28 are full-time. Two libraries report using interns for this work, too.

Budget

Slightly more than half of the respondents reported that they have no dedicated budget for digitization activities. Only 19 (30%) reported that there was a dedicated budget for both start-up and ongoing costs for digitization activities. Six (9.5%) reported a dedicated budget for start-up costs but not for ongoing costs. The 19 reported start-up budgets range fairly widely, from a minimum of \$5,000 to a maximum of over \$366,000 with a mean of \$97,027. The ongoing budgets vary even more widely, from a minimum of \$5,000 to a maximum of over \$1,000,000 with a mean of \$303,916.

The sources of funding for start-up costs are most often the library's budget (85%), grants (57%), and one-time supplemental funds (40%). Less than a third of the respondents received funds from gifts, the parent institution, or information technology units. A few respondents received income from fees, consortial money, and development funds. Sources of funding for ongoing costs for digitization activities are mostly the library's budget (97%), grants (49%), and gifts (33%). Less than a quarter of respondents receive support from one-time supplemental funds, the parent institution, or information technology units. Some receive income from contract scanning, from hosting journals, and other fee-based services. As digitization activities move from start-up to ongoing status they increasingly rely on the library's permanent budgets, gifts, and information technology funds. Another noteworthy trend is the reliance on fee-based service income to support ongoing costs for digitization efforts.

Budget allocations for digitization activities differ somewhat from start-up to ongoing operations, as is to be expected. Hardware and software acquisition and staff are the major expenses during start-up, followed by vendor fees. Ongoing operations shift a higher percentage of their budgets to staffing and benefits, vendor fees, and hardware and software maintenance; they decrease the percentage for hardware and software acquisition. Only a few respondents expend any funds on promotion or assessment of digitization activities and then only a small amount.

The survey asked how operational costs are covered when there is not a dedicated budget for digitization activities. Most of the respondents reported that all or part of the expenses are absorbed by the library's operations budget; several also rely on gifts and grants. One respondent replied, "Creatively." Some libraries allocate and manage funds on a project-by-project basis. Funds are distributed through unit budgets. This is to be expected as the majority of responding libraries' digitization activities are managed in a distributed fashion,

and as was noted above, much of the ongoing costs are staffing and benefits.

Over the past five years, the majority of respondents have seen expenditures for staff, hardware, software, and vendor fees increase. Expenditures for hardware and software maintenance, promotion, and assessment have remained more level. Only a few report any decrease in expenditures in any category.

Material Selection

A wide variety of materials are being selected for digitization. The most popular include still images and photographs, archival material, manuscripts, rare books, monographs (complete volumes), audio recordings, and moving images and videos. Fewer than half, but still a substantial number, of the respondents digitize parts of monographs, complete issues of journals, and journal articles. Other materials selected for digitization range from art works to university photographs and include maps, newspapers, 3D objects, slides, prints, and theses and dissertations. It is noteworthy that the materials most likely to be digitized (still images and photographs, archival materials, manuscripts, and rare books) are those for which access would be extremely limited without digitization.

An item's subject matter is the top criterion for selection for digitization, followed closely by whether it is part of a collection that's being digitized, and its rarity or uniqueness. Items that fit the criteria of a cooperative digitization project, or are in suitable physical condition or format are also likely candidates. Other respondents select items based on requests from users, faculty or student needs, a high demand for or use of the material, or its research value, among other criteria.

Material Digitization

Sixty percent of respondents reported that they outsource some or all parts of digitization production work. A wide variety of vendors were identified, including OCLC Preservation Services, TechBooks, Apex CoVantage, Backstage Library Works, and

iArchives, along with 31 others. The high number of vendors likely indicates that the widely dispersed survey respondents are using local vendors.

Metadata Creation

The most widely used metadata standards in digitization projects are Dublin Core (92%), MARC (84%), XML (75%), and EAD (69%). Fewer than half of the respondents, but still a substantial number, use TEI (45%), METS (38%), VRA Core (31%), and MODS (25%); 25% report using a range of other standards.

Assessment

How libraries evaluate the success of their digitization efforts varies according to whether they are assessing material selection, material digitization, or metadata creation. Material selection is most often assessed through user feedback, testing, and surveys, but also through usage data. Material digitization is most often assessed through quality control inspections, but also through user surveys and

feedback, and usage statistics. Benchmarking, best practices, and meeting project deadlines also serve as assessment tools to assess material digitization. Metadata creation is most often evaluated based on quality assurance reviews and inspections. Best practices and user surveys and feedback are also used.

Conclusion

Comments throughout the survey indicate that many libraries are in a period of transition as they attempt to determine the best organization, staffing, and budgeting models for their particular digitization operations. Small-scale operations are ramping up for more substantial activity. Collaborative projects are common. Digitization activities increase the availability and access to information for everyone, not just an institution's local users. As libraries continue to pursue digitization activities, it's important to share what is learned in order to benefit from each other's experiences and develop a collective knowledge of best practices.