


[quick links](#)courses > 200708 fall lis-615-01 collection management > **preservation and digitization of historical document collections** **200708 Fall LIS-615-01 Collection Management****Preservation and Digitization of Historical Document Collections**

last edited by James Gwynn on Saturday, 10/20/2007 2:31 PM

Introduction:

This bibliography will be focused on historical rather than historic documents; it is crucial to make the distinction between the two. The former generally refers to documents that are history-making in and of themselves (e.g. the Magna Carta or the Declaration of Independence), while the latter describes more mundane documents, such as photos, records, manuscripts, and even audio and video presentations, which are a part of history or are somehow related to the study of history.

Most of the current research on digitization schemes and projects focuses on variations on the following issues:

- The selection and prioritizing of what items to digitize: There seems to be a consensus that not all items will ever be digitized, and that not all items are worth digitizing to begin with.
- Whether digitization will be to facilitate use and access, for archival and [preservation](#) purposes, or for both (and whether digitization is an appropriate archival format at all): Most agree that digitization is a more appropriate tool for facilitating access than for archival purposes, although there is some support for moving forward with digitization as an archival format as well.
- The mechanics of digitization (equipment, file formats, etc.): Discussions of equipment generally center around speed, reduction of labor costs, and -- perhaps most importantly -- minimizing damage to the original items. Discussions of file formats involve comparisons based on compression of images (JPEG vs. TIFF), context (text-only vs. PDF), and resolution for archival purposes (the higher the better).
- Ethics and legal issues, particularly [copyright-related](#): What sort of materials can various types of libraries digitize legally, and how much public access can they provide to these materials? And with photos, what level of manipulation is acceptable?
- Initial and ongoing funding: There is also agreement that digitization projects are time-consuming and expensive. However, grant money is available, and it is possible to begin a project with relatively little investment, and then to build on it.

The following resources are recommended for a better understanding of current issues related to library and archival digitization and preservation projects:

Blogs:

Minow, M. (2007). *Library law blog*. Retrieved October 6, 2007 from <http://blog.librarylaw.com/librarylaw/>

Recommended by the Special Libraries Association (2007) blog and by respected blogger Jessamyn West of Librarian.net (West, 2004), this ALA member blog focuses on legal issues pertaining to libraries and librarians, and features numerous posts specifically related to digitization projects.

Jordan, M. (2007). *Digitizationblog*. Retrieved October 4, 2007 from <http://digitizationblog.interoperating.info/>

Listed as a resource by the National Library of Australia (n.d.) and by Metronet (2006), a network of public libraries in the Minneapolis area, this blog by Mark Jordan, head of Library Systems at Simon Fraser University focuses on news, trends, and developments in library digitization projects.

Books:

Feather, J. (Ed.) (2004). *Managing preservation for libraries and archives*. Burlington, VT: Ashgate.

Edited by John Feather of Londonborough University, this book covers all aspects of library and archival preservation, from preservation of digital information to digitization and ethics. Of particular interest are chapters two ("The malleability of fire: preserving digital information" by Colin Webb) and three ("Selection for digital preservation: dilemmas and issues" by Majlis Bremer-Laamanen and Jani Stenvall), but all eight chapters have material which would be of value to archivists and librarians contemplating digitization projects.

Starmer (2005) found it to be "a good overview of the issues facing preservation today and where the field is headed" (p.143), and Macevi (2005) calls it "a valuable addition to the collection on various aspects of preservation for libraries, archives, museums and other relevant institutions as well as to universities and

privacy

To see who can read what you write, [click here](#).

page

[edit](#)
[new](#)
[delete](#)
[history](#)
[print \(w/ comments\)](#)

page stats

views: 154
 edits: 62
 contributors: 1
 comments: 2

page contributors

james
 gwynn

search

[search](#)

site navigation

[acquisitions in the academic library av materials in school libraries-upchurch bookstores, libraries, and the "new breed" of booksales: how they're affecting each other's success in distributing books and media censorship and school](#)

professional education institutions offering study programmes or courses within this area." (par. 6).

For more discussions of this title see [this bibliography](#) on Preservation and the Public Library, as well as [this one](#) on Digital Presevation and the Internet.

Hughes, L. (2004). *Digitizing collections : strategic issues for the information manager*. London: Facet.

Reviewed in numerous scholarly and trade journals, including *Online* and *Program: electronic library and information systems*, this overview of digitization projects and strategies offers a more management-centered view of the process. The book is divided into two parts, the first covering "Strategic decision making" and the second "Digitizing collections". Individual chapters touch on costs and benefits, material selection, legal issues, planning and funding, project management, and material types.

Wiley (2005) describes the title as a "good outline" designed for managers who need to understand various aspects of the project but not all the details" (p. 62). Johnson (2005) also notes the management focus and praises the overall organization of the book and the chapter on collaboration in specific, while offering criticisms of Hughes' literature review, and lack of specifics on archival vs. access formats; his overall view, though is that the book is "a compulsory read for any student of digitisation" (p. 285). Like Johnson, Riley (2005) questions certain technical assumptions and oversights, but concludes that "manager ought to come away from this book with the tools to effectively determine when an institution should choose to digitize" (p. 66).

Koelling, J. (2004). *Digital imaging: a practical approach*. Walnut Creek, CA: Altamira Press.

This title compliments Hughes nicely, as it is concerned specifically with mechanics and formats for digitizing images and visual resources, and includes tips and tricks about scanner calibration, selecting the correct hardware for the correct project, and deciding on various archival and use formats. There is also some discussion of copyright and ethical issues.

Grant (2004) praises its "practical approach" (p. 443), and its extensive coverage of technical specifications and image enhancement. Lawson (2005) is similarly enthusiastic about the book's clear language and "stunning visual evidence of why digital imaging is of value to historians and researchers" (p. 266).

Case Studies:

Ballard, T., & Donald, E. (2007). A digitization and multimedia project at Quinnipiac University, Hamden, Connecticut, USA. *New Library World* 108, 445-452.

This case study of a book digitization project at a Connecticut university library focuses on the mechanics of the project (proper scanner types for fragile books, formats, etc.) as well as the fact that the project was begun with no grant money whatsoever. Books were scanned in both image and OCR text formats, and then converted to HTML for web access. The project initially centered on public domain texts on the Irish Famine, and led to a collaboration with a county library in Ireland, which provided access to some primary historical sources that were also digitized. The collaboration resulted in a successful website (<http://www.thegreathunger.org>) that receives over 45,000 visits per month.

Bond, T. (2006). Sustaining a digital collection after the grants: the Early Washington Maps Project. *OCLC Systems & Services* 22(1), 56-66.

This case study of a Washington State University project to digitize historical maps and integrate them into the OPAC emphasizes both the processes involved in effectively digitizing map collections, and in sustaining the project once the original grant funding runs out. The study involves an analysis and projection of ongoing costs for the project, and also touches on archival formats, backup methods, and handling of reference requests for scanning and printing. Access to the maps has been provided at <http://www.wsulibs.wsu.edu/holland/masc/xmaps.html>.

Downer, S., Medina, S., Nicol, B., & Trehub, A. (2005). AlabamaMosaic: sharing Alabama history online. *Library Hi-Tech*, 23(2), 233-251.

This case study of a digitization project of the Network of Alabama Academic Libraries and others focuses on (1) planning the project, (2) completing working groups, technical issues, and digital production centers, and (3) the creation and unveiling of the website. The project was a collaborative venture, involving the collections of numerous academic and public libraries, and was the result of a three-year federal grant to create a statewide digital collection. The study is heavy on planning and technical aspects. The results of the project are available online at <http://www.alabamamosaic.org/>

Galloway, E. (2004). Imaging Pittsburgh: creating a shared gateway to digital image collections of the Pittsburgh region. *First Monday*, 9(5). Retrieved September 18, 2004 from http://www.firstmonday.org/issues/issue9_5/galloway/index.html

This is a case study of an ongoing initiative among three Pittsburgh area institutions to digitize and provide online access to over 7,000 historical images from over twenty different collections. The initial project involved creating a single web gateway to integrate the numerous collections. The paper discusses selection criteria, metadata issues, workflow, website creation, and applications of the service for research and education. The

[media centers censorship in the public library censorship issues in the school library collecting multimedia resources in school libraries: beyond books collection evaluation in public libraries collection preservation in the school library collections management community's role in collection development cooperation between libraries copyright issues for school media specialists copyright issues in public libraries copyright issues in public libraries digital preservation and the internet digital preservation of electronic media digital preservation of electronic media electronic resource](#)

website may be seen at <http://digital.library.pitt.edu/pittsburgh/>

Gemmil, L., & O'Neal, A. (2005). Ohio Memory Online Scrapbook: creating a statewide digital library. *Library Hi Tech* 23(2), 172-186.

Similar in scope to the AlabamaMosaic project, this case study discusses a collaborative, inter-library project in Ohio, led by the Ohio Historical Society, and focuses on creation of content, development of tools, and community outreach. The website is located at <http://www.ohiomemory.org/>

Schlumpf, K., & Zschernitz, R. (2007). Weaving the past into the present by digitizing local history. *Computers in Libraries* 27(10), 10-15.

This a case study of a suburban Illinois public library system that coordinates a regional local history digital archive. The project began in 1998 with a grant from the Illinois State Library and was online by 1999. The goal was to increase access, with protection (not necessarily archival preservation) a secondary issue. There is discussion of material types, server platforms, and even a guide to setting up a well-equipped digitization station.

Database:

Library of Congress. (2006, October 30). Prints and Photographs Online Catalog (Prints and Photographs Reading Room, Library of Congress). Retrieved October 6, 2007 from <http://www.loc.gov/rr/print/catalog.html>

Cited by numerous public and academic library sites (Yeo, 2007; San Diego Public Library, n.d.) around the country and the world, this is the online interface for the Library of Congress' photographic collection, a model for library and archival digitization projects, with over one million digital images. It is useful both as a resource for digitized documents and as a reference point for the study of digitization projects themselves.

Scholarly and Professional Journals:

Arthur, K., Byrne, S., Long, E., Montori, C. & Nadler, J. (2004), *Recognizing Digitization as a Preservation Reformatting Method*. Washington, DC, Association of Research Libraries.

This paper considers the issues surrounding the controversial use of digitization as a preservation format. Its appendices are particularly useful, providing comparisons between various archival formats (microfilm vs. paper facsimile vs. digital facsimile) and standards and best practices with respect to file formats, metadata, and web access. The paper takes the position that the digital file as archival medium is an idea whose time has come.

Besek, J. (2003). *Copyright Issues Relevant to the Creation of a Digital Archive: A Preliminary Assessment*. Washington, D.C.: Council on Library and Information Resources.

This paper is an overview of copyright law with respect to libraries and digital archives, specifically focused on the Library of Congress. Most issues are discussed in general terms, but there is information on fair use, the DMCA, and relevant exceptions for libraries.

Hafner, K. (2007, March 10). History, digitized (and abridged). *The New York Times*. Retrieved September 24, 2007 from <http://www.nytimes.com/2007/03/10/business/yourmoney/11archive.html?pagewanted=1&ei=5124&en=ac8d2f50c8dfc12d&ex=1331355600&partner=permalink&exprod=permalink>

This article discusses the current state on digitization projects in libraries and archives nationwide, including the National Archives and the Library of Congress, and touches on funding issues, the "invisibility" of information not available online, copyright, and the selection and prioritizing of materials to be digitized.

Liu, Y. (2004). Best practices, standards and techniques for digitizing library materials: a snapshot of library digitization projects in the USA. *Online Information Review* 28(5), 338-345.

This survey of digitization projects, primarily among academic libraries in the United States, attempts to determine standards for digitization projects and determine types of materials and issues surrounding digitization. It also discusses special challenges at the Internet Archive, a project that involves archiving of websites ("The Wayback Machine") and other digital material, including audio and video.

Lopatin, L. (2006). Library digitization projects, issues and guidelines: a survey of the literature. *Library Hi Tech* 24(2), 273-289.

As the title suggests, this is primarily a bibliography and literature review, which provides a summary of useful literature off the subject of library digitization projects, most of it published between 2000 and 2005.

Maroso, A. (2005). Educating future digitizers: the Illinois Digitization Institute's Basics and Beyond digitization training program. *Library Hi Tech* 23(2), 187-204.

This article describes an training program for individuals involved in digitization projects. Three levels of training were offered: (1) a one-day workshop offering an overview with no hands on training, (2) a three-week

[sharing in academic libraries electronic serial collections electronic sources gifts to libraries graphic novel collection managing ebook collections marketing and the public library marketing to young adults thru high school & public libraries open access preservation and digitization of historical document collections preservation in the public library ready reference and web site index collections selecting electronic resources for public library collections selection and acquisition of electronic resources for youth the ebook industry virtual libraries in public schools virtual](#)

online course with multiple homework assignments and an online bulletin board, and (3) a hybrid online course with an on-site hands on training component. The article concludes that training is sorely needed, and offers suggestions for enhanced content in future training programs: specifically, more emphasis on non-print and non-photographic materials, and a broader discussion of the longevity of digital materials.

Westney, L. (2007). Intrinsic value and the permanent record: the preservation conundrum. *OCLC Systems & Services* 23(1), 5-12.

This literature review/viewpoint paper attempts to analyze selection criteria for digitization projects in terms of the "intrinsic value" of the materials selected. It discusses prioritizing based on the importance of the item. The discussions of the context and the integrity of digital surrogates are particularly interesting. The author concludes that these surrogates may aid research but are inadequate substitutes for the original items, and run the risk of becoming useless bits of data without continual human involvement.

Websites:

(Note: the websites referenced with the individual case studies section above may also be of interest.)

Library of Congress. (2007, September 25). Standards at the Library of Congress. Retrieved October 4, 2007 from <http://www.loc.gov/standards/>

A list of links to Library of Congress standard data formats, including metadata and XML standards useful to archivists working on digital preservation projects. It is recommended by numerous authoritative sources, including an ALA Collection Development Syllabus (Waters, 2005) and the technical support documentation for Adobe's Acrobat PDF software (Adobe Systems Incorporated, n.d.).

References:

Adobe Systems Incorporated (n.d.) Set the document language. Retrieved October 7, 2007 from http://help.adobe.com/en_US/Acrobat/8.0/Professional/WS58a04a822e3e50102bd615109794195ff-7cee.html

Cujic, M. (2004). Managing preservation for libraries and archives. *Journal of Documentation* 61(3), 445-447.

Grant, K. (2004). Digital imaging: a practical approach. *International Journal of Information Management* 24, 443-444.

Johnson, P. (2005). Digitizing collections : strategic issues for the information manager. *Program: Electronic Library & Information Systems* 38(4), 284-285.

Knapp, A. (2007). New challenges facing academic librarians today: electronic journals, archival digitization, document delivery, etc. *Journal of Scholarly Publishing* 38(2), 112-116.

Lawson, K. (2006). Digital imaging: a practical approach. *Serials Librarian* 49, 265-266.

Macevi, E. (2005). Managing preservation for libraries and archives: current practice and future developments. *Information Research* 10(3). Retrieved October 6, 2007 from <http://informationr.net/ir/reviews/revs169.html>

Maxwell, L. (2006). New challenges facing academic librarians today: electronic journals, archival digitization, document delivery, etc. *College & Research Libraries* 67(5), 484.

Metronet (2006). Metronet professional issues - digitization. Retrieved October 7, 2007 from <http://www.metrolibraries.net/pro/digitization.html>

National Library of Australia. (n.d.).PADI - Organisations and Websites. Retrieved October 6, 2007 from <http://www.nla.gov.au/padi/format/org.html>

Riley, J. (2005). Digitizing collections: strategic issues for the information manager. *Library Resources & Technical Services* 49(1), 65-66.

San Diego Public Library (n.d.) Reading resources | Public library. Retrieved October 6, 2007 from <http://www.sandiego.gov/public-library/news-events/reading.shtml>

Special Libraries Association (2007). *SLA Blog*. Retrieved October 6, 2007 from http://slablogger.typepad.com/sla_blog/

Starmer, M.(2005). Managing preservation for libraries and archives. *Library Resources & Technical Services* 49(2), 141-143.

Waters, D. (2005). ALA | 2005 - Syllabus for Serials Collection Management and Acquisitions Teaching and Training. Retrieved October 7, 2007 from <http://www.ala.org/ala/alctcontent/alctpubsbucket/alctresources/alctserialsb/syllabiforserial/syllabusserialscollection.htm>

West, J. (2004, November 26). Library law blog. *Librarian.net*. Retrived October 6, 2007 from <http://www.librarian.net/stax/954/library-law-blog/>

[libraries in public schools visual resource collections weeding and its complexities](#)

toolbox

[page list](#)
[export site](#)

Wiley, D. (2005) Recommended reading on digitization, PDFs, information architecture, and the digital divide. *Online* 29(2), 61-63.

Yeo, S. (2006). SULAIR: Research quick start guides: PWR students. Retrieved October 6, 2007 from <http://www-sul.stanford.edu/guides/pwrimages.html>

> [Comments \(2\)](#)



Inspire Thinking Online.™
Teams LX - 2.7.13.6403
Copyright © 2003-2007, Learning Objects, Inc.