

CREATING AND MAINTAINING DIGITAL COLLECTIONS
What's Wrong With this Picture Workshop
March 14-16, 2007
Dr. David Green
Knowledge Culture

RESOURCE PACKET

CONTENTS

1. Selected Cultural Heritage Digitization Bibliography
2. Principles Of Digitization Good Practice
3. Questions to Ask Before Selecting Material for Digitization:
 - a. Intellectual Value of Proposed Material
 - b. How Does Your Material Contribute to the Ecology? Research Resources
4. When Works Pass into the Public Domain (Laura Gasaway)
5. Checklist for Fair Use (Indiana University)
6. Getting Permission (Georgia Harper)
7. Copyright Decision Tree (Lauryn Grant)
8. Physical Properties of Source Material
9. Selection of Materials For Digitizing: A Decision-Making Matrix (CLIR)
10. Cost Factors
11. RLG Worksheet For Estimating Digital Reformatting Costs (first page)
12. List of Key Funders of Digitization projects
13. RLG Assessing Document Attributes and Scanning Requirements
14. Image File Formats; Estimating File Sizes
15. Metadata Types

INTRODUCTION

1. SELECTED CULTURAL HERITAGE DIGITIZATION BIBLIOGRAPHY

1. Why Digitize? The Digital Difference

Anderson, Maxwell. "Cupped Hands in the Stream: Digital Media's Challenge to Collecting Institutions." American Institute for Conservation
<<http://aic.stanford.edu/meetings/archives/2000/anderson.html>>. *Abstract of AIC talk by the director of the Whitney Museum questioning how to approach the preservation of digital artistic work.*

Bearman, David, and Jennifer Trant. "Authenticity Of Digital Resources: Towards A Statement Of Requirements In The Research Process." *D-Lib Magazine*, June 1998.
<www.dlib.org/dlib/june98/06bearman.html>. *Important statement on the centrality of the notion of authenticity when dealing with digital documents.*

Brooke, Bob. "The Brief History of Digital Photography," Bob Brooke's Digital Studio. [no date] www.bobbrooke.com/DigitalStudio/digitalhistory.htm.

Broun, Elizabeth. "The Future of Art at the Smithsonian," *American Art*, Summer/Fall 1994, Vol. 8, No. 3-4.

Coleman, A. D. *The Digital Evolution: Visual Communication in the Electronic Age : Essays, Lectures and Interviews 1967-1998*. Nazraeli Press, 1998. *Smart, often profound review of how computerization and electronics have affected the photographic arts and our life in general.*

Cullen, C. T., P. B. Hirtle, D. Levy, C. A. Lynch, and J. Rothenberg. *Authenticity in a Digital Environment*. Council on Library and Information Resources, 2000.
<www.clir.org/PUBS/reports/pub92/pub92.pdf>. *Report examining the question of what is an authentic digital object and how the authenticity and integrity of digital files can be guaranteed.*

Day, Louis Alvin. *Ethics in Media Communications: Cases and Controversies*. (Third ed.) Wadsworth Pub. 2000. *Includes consideration of the issue of manipulation of digital images.*

Dietz, Steve, Howard Besser, Ann Borda, and Kati Geber, *The Virtual Museum (of Canada): The Next Generation*. Canadian Heritage Information Network, 2004.
<www.chin.gc.ca/English/Members/Next_Generation/pdf.html>.

DigiCamHistory.Com <www.digicamhistory.com>. *Educational site on history of photography with a particular interest in digital cameras, of which there are many illustrations*

Jones-Garmil, Katherine (ed). *The Wired Museum: Emerging Technology & Changing Paradigms*. American Association of Museums, 1997.

Marchand, James. "The Computer as Darkroom and Camera." *Offline*, 37, Jan. 1992.
<palimpsest.stanford.edu/byauth/marchand/compscan.html>. *An early, but good, guide to the history and characteristics of digital photography and digital imaging by a medieval scholar (and photographer) who uses digital images in his research.*

Mitchell, William. *The Reconfigured Eye: Visual Truth In The Post-Photographic Era*. MIT Press, 1992. *The dean of MIT's School of Architecture and Planning examines the technology of digital imaging and its implications for society.*

Smith, Abby. *Why Digitize?* Council on Library and Information Resources, 1999
<www.clir.org/PUBS/reports/pub80-smith/pub80.html>

von Amelunxen, Hubertus, et al (eds). *Photography after Photography: Memory and Representation in the Digital Age*. G+B Arts. Amsterdam, 1996. *Collection of essays on the impact of the digital manipulation of photographs.*

2. Planning: Audiences, Partnerships, Resources, Copyright

Hamma, Kenneth. "Public Domain Art in an Age of Easier Mechanical Reproducibility," *D-Lib Magazine*, November 2005.
<www.dlib.org/dlib/november05/hamma/11hamma.html>

Hamma, Kenneth. "The Role Of Museums In Online Teaching, Learning, and Research." In *Sharing Digital Resources, Web-Wise*, 2004. *First Monday*, (Volume 9, Number 5) — 3 May 2004 <http://firstmonday.org/issues/issue9_5/>

Harper, Georgia. *Crash Course in Copyright*. University of Texas, 2001
<<http://www.utsystem.edu/ogc/intellectualproperty/cprtindx.htm#top>>

Hazen, Dan, Jeffrey Horrell, Jan Merrill-Oldham, *Selecting Research Collections for Digitization*, Council on Library and Information Resources (August 1998),
<www.clir.org/pubs/reports/hazen/pub74.html>

National Initiative for a Networked Cultural Heritage, *The Copyright Town Meetings, 1997-2003* <<http://www.ninch.org/copyright/>>

The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials, by the Humanities Advanced Technology and Information Institute (HATII), University of Glasgow, and the National Initiative for a Networked Cultural Heritage (NINCH), 2002
<http://www.nyu.edu/its/humanities/ninchguide/index.html>

Shapiro Michael, and Brett I. Miller, *A Museum Guide to Copyright & Trademark*. American Association of Museums, 1999.

3. Digitization

Baca, Murtha, (ed). *Introduction To Metadata: Pathways To Digital Information*. Getty Trust
<http://www.getty.edu/research/conducting_research/standards/intrometadata/>

Colorado Digitization Project. *Digitization Resources*.
<http://www.cdpheritage.org/resource/index.html> Includes *Digital Imaging Best Practices, Metadata Best Practices, Digital Audio Best Practices.*

National Information Standards Organization. *A Framework of Guidance for Building Good Digital Collections*. 2nd edition. 2004. <http://www.niso.org/framework/framework2.html>

Sitts, Maxine K. (ed). *Handbook For Digital Projects: A Management Tool for Preservation and Access*, Northeast Document Conservation Center, 2000.
<http://www.nedcc.org/digital/dighome.htm>

Kenney, Anne R, and Oya Y Rieger. *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. Cornell University Library, 2000.
<http://www.library.cornell.edu/preservation/tutorial/contents.html>

Smith, Abby. *Building and Sustaining Digital Collections: Models for Libraries and Museums*. CLIR, 2001. <http://www.clir.org/pubs/reports/pub100/pub100.pdf>.

Zorich, Diane. *Digital Cultural Heritage Initiatives and Their Sustainability Concerns*. CLIR, 2003 <http://www.clir.org/pubs/reports/pub118/pub118.pdf>.

4. Preservation

Besser, Howard. "Digital Longevity." In *Handbook For Digital Projects (above)*. <http://www.nedcc.org/digital/ix.htm>

Bishoff, Liz, and Nancy Allen. *Business Planning for Cultural Heritage Institutions*. CLIR, 2004. <http://www.clir.org/pubs/reports/pub124/pub124.pdf>

Day, Michael, "Preservation Metadata," in G. E. Gorman and Daniel G. Dorner (eds.). *Metadata Applications and Management*. International Yearbook of Library and Information Management, 2003-2004, London: Facet Publishing, 2004, pp. 253-273. Available at <http://www.ukoln.ac.uk/metadata/publications/iylim-2003/> *Good summary narrative on preservation metadata.*

Digital Preservation Coalition (Neil Beagrie & Maggie Jones). *Preservation Management of Digital Materials; A Handbook*. British Library, 2001. Available at <<http://www.dpconline.org/graphics/handbook/>>.

Kahle, Brewster. "Preserving the Internet." *Scientific American* 276, no. 3, March (1997). <http://www.sciam.com/0397issue/0397kahle.html> *Founder of the Internet Archive examines possibilities for preserving what we find on the Web.*

Lavoie, Brian, and Lorcan Dempsey. "Thirteen Ways of Looking at...Digital Preservation." *D-Lib Magazine*, July/August 2004 <http://www.dlib.org/dlib/july04/lavoie/07lavoie.html#2>

Lavoie, Brian, and Richard Gartner. *Preservation Metadata*. DPC Technology Watch Series Report 05-01, September 2005. <http://www.dpconline.org/docs/reports/dpctw05-01.pdf> *Comprehensive but highly readable update on developments in preservation metadata and METS.*

Preserving Access to Digital Information: subject gateway to international digital preservation resources. <<http://www.nla.gov.au/padi/>>

National Initiative for a Networked Cultural Heritage. *The Price of Digitization: New Cost Models for Cultural and Educational Institutions. A Digitization Symposium presented by NINCH and Innodata*. Report by Lorna Hughes. NINCH, 2003 <<http://www.ninch.org/forum/price.report.html>>

Research Libraries Group. *RLG Working Group on Preservation Issues of Metadata*. Research Libraries Group (1998) www.rlg.org/preserv/presmeta.html. *Seminal work on preservation metadata.*

Rothenberg, Jeff. "Ensuring the Longevity of Digital Documents." *Scientific American* 272, no. 1, January (1995): 24-9. *Rothenberg's classic early statement about the choices available for preserving digital material.*

Russell, Kelly. "Digital Preservation: Ensuring Access to Digital Materials Into the Future." CEDARS, www.leeds.ac.uk/cedars/Chapter.htm.

WHY DIGITIZE?

2. PRINCIPLES OF DIGITIZATION GOOD PRACTICE

National Initiative for a Networked Cultural Heritage 2000

1. Optimize interoperability of materials

Digitization projects should enable the optimal interoperability between source materials from different repositories or digitization projects.

2. Enable broadest use

Projects should enable multiple and diverse uses of material by multiple and diverse audiences.

3. Address the need for the preservation of original materials

Projects should incorporate procedures to address the preservation of original materials.

4. Indicate strategy for life-cycle management of digital resources

Projects should plan for the life-cycle management of digital resources, including the initial assessment of resources, selection of materials and digital rights management; the technical questions of digitizing all formats; and the long-term issues of sustainability, user assessment, digital asset management and preservation.

5. Investigate and declare intellectual property rights and ownership

Ownership and rights issues need to be investigated before digitization commences and findings should be reported to users.

6. Articulate intent and declare methodology

All relevant methods, perspectives and assumptions used by project staff should be clarified and made explicit.

PLANNING: Define the Project

3. QUESTIONS TO ASK BEFORE SELECTING MATERIAL FOR DIGITIZATION:

A. Intellectual Value of Proposed Material

- Is the information content of the analog objects high?
- What is the intellectual significance of the analog materials?
- How do the proposed materials compare in terms of importance, authority, uniqueness, and timeliness?
- Is the object you are considering for digitization a good example of the period under examination?
- Is it original?
- Are there alternatives in the collection more capable of illustrating particular points?
- Is it complete or are there any parts missing?
- Is it up to date?
- Is it accurate?
- Would it withstand the test of time or is it only of ephemeral interest today? (Is it, for example, an important testimony of the past, of another culture or our own; of an artistic movement, an author, or a scientific discovery?)

PLANNING: Define the Project

3. QUESTIONS TO ASK BEFORE SELECTING MATERIAL FOR DIGITIZATION:

B. Where Does Your Material Fit In to Overall Ecology?

Research Resources

Select List of Registries, Catalogs and Portals of Digital Projects and Resources

Active as of May 2006

American Memory. The Library of Congress collection of some 9 million items, organized into more than 100 thematic collections that document U.S. history and culture.

<<http://memory.loc.gov/ammem/>>

ARL (Association of Research Libraries, Washington DC) *Digital Initiatives Database* (in collaboration with the University of Illinois Chicago)

<<http://www.arl.org/did/>>

CETH (Rutgers University Center for Electronic Texts in the Humanities) *Directory of Electronic Text Centers*. Links to Text Centers that offer listings and sometimes direct access to electronic texts

<http://tabula.rutgers.edu/ceth/etext_directory/>

DLF/OCLC *Registry of Digital Masters* <

<www.oclc.org/digitalpreservation/why/digitalregistry >

Access via OCLC's First Search, available through libraries

Humbul Humanities Hub (UK gateway of humanities resources with an extensive catalog of humanities resources) <http://www.humbul.ac.uk/>>

IMAGELIB (University of Arizona Library) *The Clearinghouse of Image Databases*.

<<http://elearn.arizona.edu/imagelib/>>

IMLS *Collection Registry for National Leadership Grant Awards* [under construction; currently contains some 150 digital collections] -

<<http://imlsdcc.grainger.uiuc.edu/collections/>>

Labyrinth (Resources for Medieval Studies)

<<http://www.georgetown.edu/labyrinth/labyrinth-home.html>>

Virtual Museum of Canada. <www.virtualmuseum.ca/English/index_flash.html>

Voice of the Shuttle, (Portal on humanities resources maintained by Alan Liu, University of California, Santa Barbara) <<http://vos.ucsb.edu/>>

4. WHEN WORKS PASS INTO THE PUBLIC DOMAIN

Includes material from U.S. Term Extension Act, PL 105-298

<i>DATE OF WORK</i>	<i>PROTECTED FROM</i>	<i>TERM</i>
Created 1-1-78 or after	When work is fixed in tangible medium of expression	Life + 70 years ¹ (or if work of corporate authorship, the shorter of 95 years from publication, or 120 years from creation ²)
Published before 1923	In public domain	None
Published from 1923 - 63	When published with notice ³	28 years + could be renewed for 47 years, now extended by 20 years for a total renewal of 67 years. If not so renewed, now in public domain
Published from 1964 - 77	When published with notice	28 years for first term; now automatic extension of 67 years for second term
Created before 1-1-78 but not published	1-1-78, the effective date of the 1976 Act which eliminated common law copyright	Life + 70 years or 12-31-2002, whichever is greater
Created before 1-1-78, but published between then and 12-31-2002	1-1-78, the effective date of the 1976 Act which eliminated common law copyright	Life + 70 years or 12-31-2047 whichever is greater

1 Term of joint works is measured by life of the longest-lived author.

2 Works for hire, anonymous and pseudonymous works also have this term. 17 U.S.C. § 302(c).

3 Under the 1909 Act, works published without notice went into the public domain upon publication. Works published without notice between 1-1-78 and 3-1-89, effective date of the Berne Convention Implementation Act, retained copyright only if, e.g., registration was made within five years. 17 U.S.C. § 405.

Notes courtesy of Professor Tom Field, Franklin Pierce Law Center

Chart prepared by Laura Gasaway, University of North Carolina, Last updated 9-18-01, <<http://www.unc.edu/~uncclng/public-d.htm>>

CHECKLIST FOR FAIR USE

Please complete and retain a copy of this form in connection with each possible "fair use" of a copyrighted work for your project

Name: _____ Date: _____ Project: _____

Institution: _____ Prepared by: _____

PURPOSE

Favoring Fair Use

- Teaching (including multiple copies for classroom use)
- Research
- Scholarship
- Nonprofit Educational Institution
- Criticism
- Comment
- News reporting
- Transformative or Productive use (changes the work for new utility)
- Restricted access (to students or other appropriate group)
- Parody

Opposing Fair Use

- Commercial activity
- Profiting from the use
- Entertainment
- Bad-faith behavior
- Denying credit to original author

NATURE

Favoring Fair Use

- Published work
- Factual or nonfiction based
- Important to favored educational objectives

Opposing Fair Use

- Unpublished work
- Highly creative work (art, music, speech, films, plays)
- Fiction

AMOUNT

Favoring Fair Use

- Small quantity
- Portion used is not central or significant to entire work
- Amount is appropriate for favored educational purpose

Opposing Fair Use

- Large portion or whole work used
- Portion used is central to work or "heart of the work"

EFFECT

Favoring Fair Use

- User owns lawfully acquired or purchased copy of original work
- One or few copies made
- No significant effect on the market or potential market for copyrighted work
- No similar product marketed by the copyright holder
- Lack of licensing mechanism

Opposing Fair Use

- Could replace sale of copyrighted work
- Significantly impairs market or potential market for copyrighted work or derivative
- Reasonably available licensing mechanism for use of the copyrighted work
- Affordable permission available for using work
- Numerous copies made
- You made it accessible on Web or in other public forum
- Repeated or long-term use

This document is provided as a courtesy of the Copyright Management Center, IUPUI, 530 W. New York St., Indianapolis, IN 46202. For further information and updates please visit <http://www.copyright.iupui.edu/>. This document last updated March 10, 2003.

6. GETTING PERMISSION

<http://www.utsystem.edu/OGC/Intellectualproperty/permisn.htm>

Getting Permission

[Collective Rights Organizations](#) | [Contacting Owner](#)
[Charged Owner](#) | [Authority](#) | [In Writing](#)
[Difficulty Identifying Owner](#) | [Unidentifiable/Unresponsive Owner](#)

Assuming the work you wish to use is protected, your use is not a fair use or otherwise exempt from liability for infringement, and the work has not been licensed for your use online, you need permission. *Now what?*

Collective Rights Organizations **CCC**

There are no foolproof methods to obtain permission, but there are steps likely to yield results. The steps will vary depending on the nature of the work you need to use. If the work is part of a book or a journal article, contact the [Copyright Clearance Center](#) ("CCC") first. The CCC now offers an experimental electronic permission service and a well-established photocopy based academic permission service. Definitely worth a try. Your library or copy center is probably already working with the CCC and should be able to help you. If the work you want to use is registered with the CCC, you can get permission within 24 to 36 hours. Permission during peak times like the beginning of fall semester will take longer.

Foreign Collectives

Lesley Ellen Harris publishes [information about international collective rights agencies](#) at her website. For example, she notes: ["VERDI \(Very Extensive Rights Data Information\)](#) which is financed by the EU is linking together the services of existing multimedia rights clearance systems in Finland, France, Germany, Ireland, the Netherlands and Spain. [CLARA](#), a Web site organized by 5 Norwegian copyright collectives, since November 12, 1998 informs users of their rights and rights clearances of all types of copyright materials including use in multimedia productions.

"On February 26, 1999, the UK-based [Copyright Licensing Agency](#) ("CLA") launched its first digital licensing scheme for print text. It offers a license for the creation, storage and exploitation of digital versions of existing print works in its repertoire. The first electronic licenses were offered to the higher education and pharmaceuticals sector."

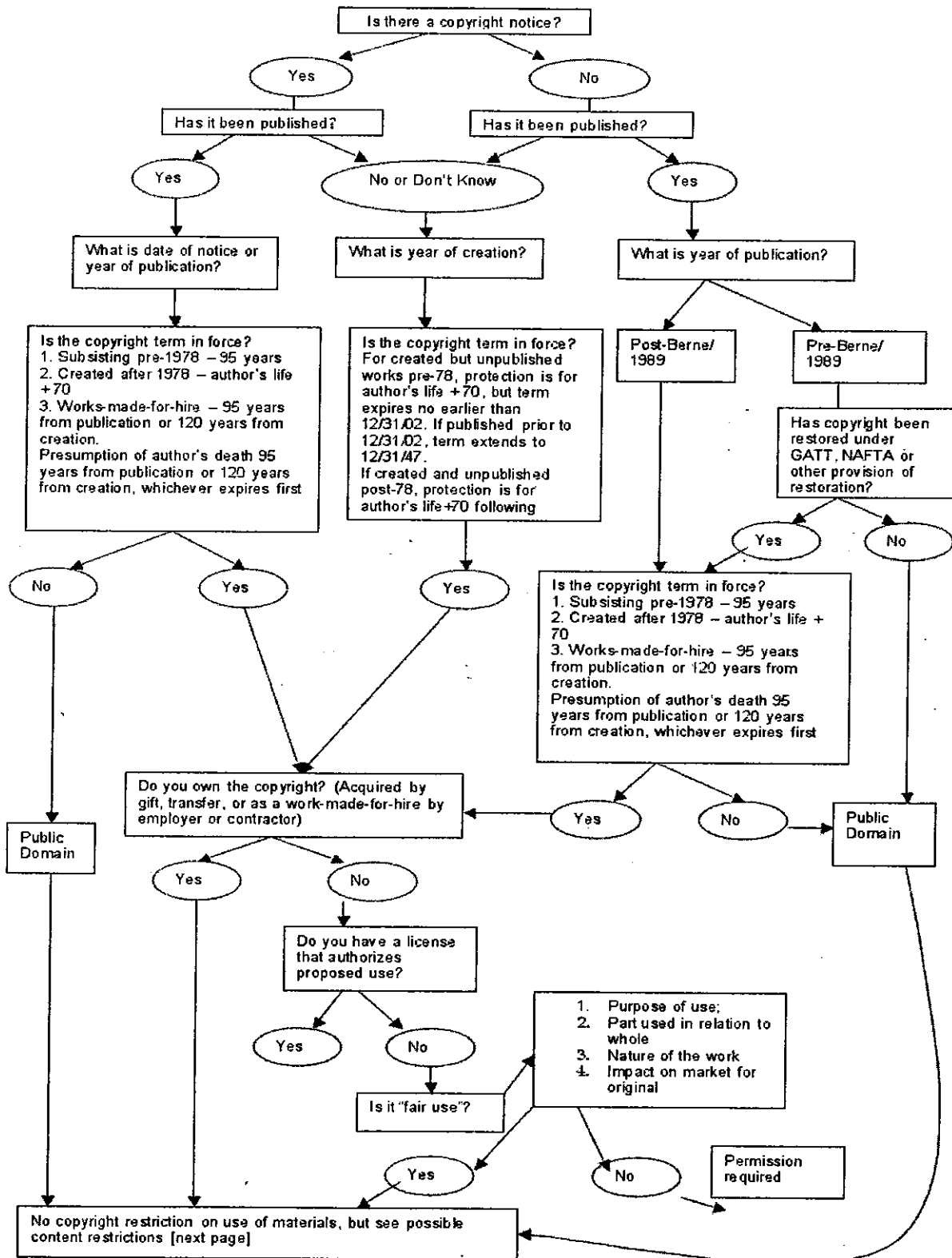
Image Archives

At this time, the professional organizations representing image creators cater to commercial interests and may be unfamiliar with educational needs. There are only a few collections specifically devoted to educators. Until more organizations catering to our needs emerge, these are a sampling of your options.

- [Academic Image Cooperative](#)
- [Allan Kohl's Art Images for College Teaching](#)
- [American Society of Media Photographers](#)
- [Artists Rights Foundation](#)
- [Artists Rights Society](#)
- [Aurora Picture Network International](#)

7. COPYRIGHT DECISION TREE

To determine if the material is protected by copyright, conduct the following inquiry for each piece of intellectual property. Remember that many kinds of materials have multiple layers of copyright, such as a photograph of a work of art, a videotape containing artwork and music, a letter containing quotes from a published book, etc.



PLANNING: Define the Project: Equipment

8. PHYSICAL PROPERTIES OF SOURCE MATERIAL

Type and category of object

Is it a book, manuscript, photograph, sound recording, TV broadcast?

Production process

Printed document, handwritten text, engraving, woodcut, wax cylinder recording, recording in mono or stereo? Is it an original or a reproduction/intermediary?

Date

How old is it? Do you have information about when it was made? If not, can you find out or estimate?

Physical size and dimensions

Is it a book with pages of regular (letter size? Is it uniform? What is its length in cm, inches, duration of an audio-visual tape in hours/minutes/seconds, number of reels of film width; depth; thickness; weight?

Media Type

Paper, leather, wood, magnetic videotape, vinyl record; combination of materials; gold leaf details in manuscripts?

Format

78 rpm disc, wax cylinder, reel-to-reel analog tape recording, DAT (Digital Audio Tape, Betacam SP tape, NTSC or PAL format video recording?

Sensitivity to light

What kind of lighting levels can it be safely exposed to during digitization? For how long?

Color information

Does it contain color information? Does color convey important information in this case? Is it an important element for the understanding and appreciation of the object?

Tonal Range

Does it have a wide tonal range? Is this an important element for the understanding and appreciation of the object/recording?

Noise

Does the audio recording have audio hiss, clicks and pops? Are there background sounds or images, which were captured in the original sound or moving image recording, that are not related to the main material? Is it important to preserve these?

Characteristics of born digital material

File format, resolution or sampling rate, bit-depth or rate, compression method, file size?

Characteristics and structure of informational content

For printed documents: does it include both illustrations and plain text? For sound recordings: what is the duration of the songs recorded and how many are included in the tape? For video recordings: what is the duration of the film recorded?

Structure of the material

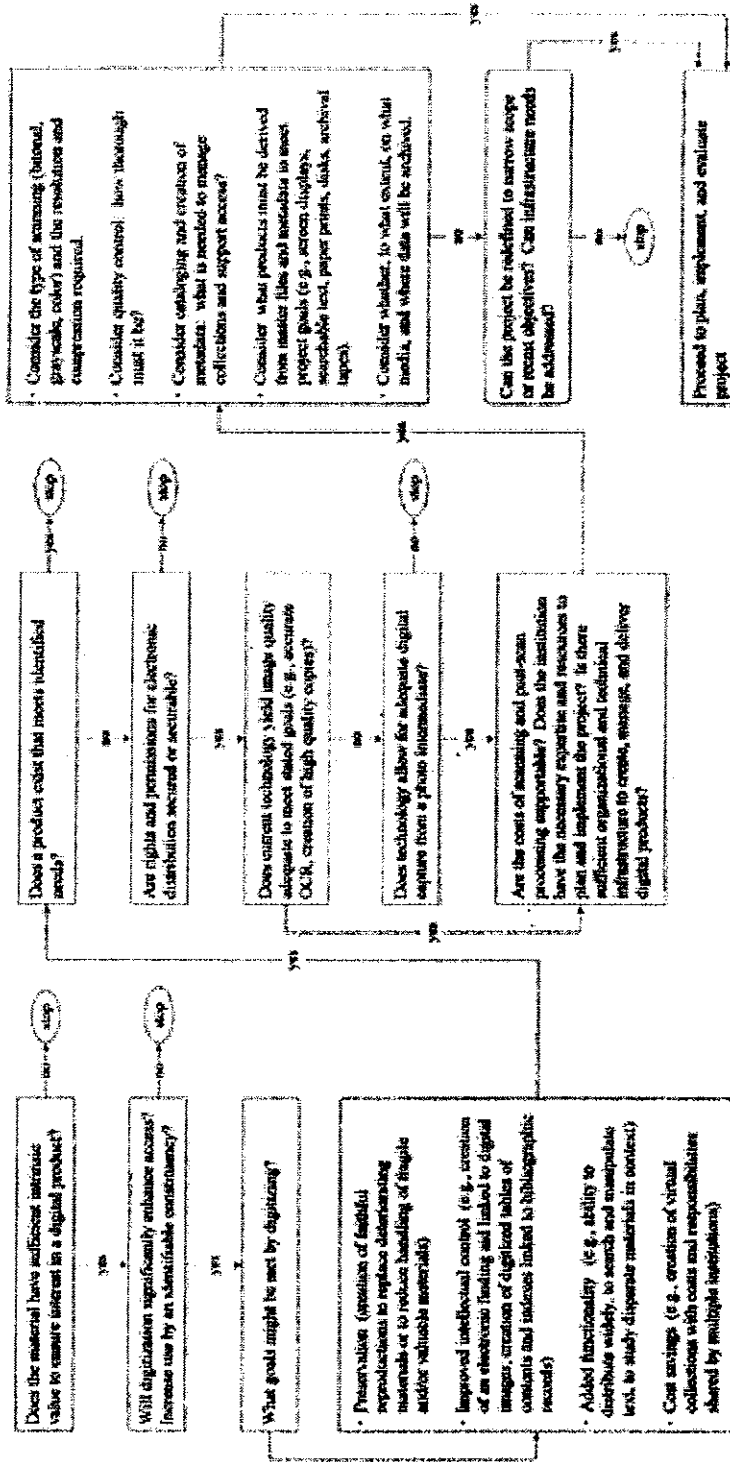
Is the material bound or mounted?

Condition of the material and conservation

What is its state of preservation? Has it been assessed by conservators? Should it be conserved? Does it require any special handling?

PLANNING: Define the Project: Selection of Materials

SELECTION FOR DIGITIZING: A Decision-Making Matrix



©Claremont Colleges, Merrill, Jan Merrill-Oldham, SJ2007
©Harvard University Library

Source: "Selection For Digitizing: A Decision-Making Matrix," from Dan Hazen, Jeffrey Horrell, Jan Merrill-Oldham, *Selecting Research Collections for Digitization*, Council on Library and Information Resources (August 1998), <http://www.clir.org/pubs/reports/hazen/pub74.html>

PLANNING: How Much Will It Cost?

10. COST FACTORS

Per unit costing, within a single media type with uniform capture techniques and settings, can be useful in budgeting

A. Direct Digitization Costs (Usually the least costly component of a project).

1. Selection, Prep and Handling time (from the shelf to point of capture and return as a percentage of total salary costs on a daily basis)
2. Pre-digitization conservation work (only for those items that require it)
3. Capture time (from set-up to naming and saving provided as a percentage of the capture operators total salary costs on a daily basis)
4. Cataloging/Metadata (required for digitization and/or created at capture stage as a percentage of total salary costs)
5. Hardware cost per item
6. Quality Assurance (time as a percentage of salary cost)
7. Software cost per item (both hardware and software costs should be on the basis of the depreciation of equipment or projected replacement cost, rather than the total cost of hardware and software)
8. Hardware maintenance
9. Technical support time (proportion of total salary or contract cost related to capture)
10. Project Management time (proportion of total salary related to capture)
11. Training (directly related to capture)

B. Indirect Digitization Costs

- Additional Project Management salary
- Web Programmer's salary
- Educational Officer's salary (or other additional project staff)
- Cataloging/Metadata (post capture creation % of total salary costs)
- Additional technical support salary
- Additional hardware and software costs
- Consumables (including any storage media such as CDs or DATs)
- Travel and per diem
- Training (non-capture related)
- Storage costs (based on total maintained cost for the gigabytes required)

See RLG's extensive worksheet: *RLG Worksheet For Estimating Digital Reformatting Costs* <<http://www.rlg.org/en/pdfs/RLGWorksheet.pdf>>

PLANNING: How Much Will It Cost?

11. RLG WORKSHEET FOR ESTIMATING DIGITAL REFORMATTING COSTS

www.rlg.org/preserv/RLGWorksheet.pdf

RLG WORKSHEET FOR ESTIMATING DIGITAL REFORMATTING COSTS

This worksheet is a guide to the preparation of a budget for activities involving digitization. It can be used for in-house scanning projects or for those utilizing an outside vendor. The activities are organized in eleven steps:

- A. Select materials
- B. Determine the size of the collection
- C. Prepare documents
- D. Determine imaging requirements (benchmarking)
- E. Determine requirements for and create metadata
- F. Determine imaging costs
- G. Determine text conversion costs
- H. Determine SGML encoding costs
- I. Determine Finding Aid Conversion Costs
- J. Post-process digital files
- K. Estimate additional local costs

The accompanying appendix provides a methodology for determining scanning requirements based on an assessment of document attributes and current scanning technologies. Endnotes follow Summary of Projected Expenses.

A. SELECT MATERIALS

This step of the process includes:

- * Identifying materials
- * Determining legal restrictions (copyright, privacy rights, donor rights, etc.)
- * Investigating the availability of digital and other versions, such as microforms (online search)
- * Eliminating items due to condition or completeness
- * Determining appropriate conversion process (e.g., film then scan, disbind originals, etc).

Estimate staff time required for selection, then the cost based on the level of staff performing each process.

Name/Title: _____
 # of hours _____ at _____ /hour + _____ benefits/hour = _____ <A₁>

Name/Title: _____
 # of hours _____ at _____ /hour + _____ benefits/hour = _____ <A₂>

Name/Title: _____
 # of hours _____ at _____ /hour + _____ benefits/hour = _____ <A>

Name/Title: _____
 # of hours _____ at _____ /hour + _____ benefits/hour = _____ <A>

AND/OR

Name/Title: _____
 % FTE _____ at _____ salary/annual + _____ benefits/annual = _____ <A>

Note: If various levels of staff are used in the process, calculate a blended monetary rate or repeat the formula for each staff member.

TOTAL SELECTION COST = <A> = _____ <I>

PLANNING: Funding

12. LIST OF KEY FUNDERS OF DIGITIZATION PROJECTS:

Andrew Mellon Foundation

Large projects; stress sustainability; stress scholarly resources
<<http://www.mellon.org/awmf.html>>

National Endowment for the Humanities (NEH)

Preservation & Access program
<<http://www.neh.gov/>>

The Getty Grant Program

Cultural heritage; image-based.
<<http://www.getty.edu/grants/>>

Institute of Museum and Library Services (IMLS):

Increasingly important; especially for joint museum-library projects:
Leadership Grants; also see annual WebWise conference
<<http://www.imls.gov/grants/index.htm>>

National Historical Publications and Records Commission (NHPRC)

Smaller grants; archives and preservation
<<http://www.nara.gov/nhprc/>>

PLANNING: Capture Equipment

APPENDIX A: ASSESSING DOCUMENT ATTRIBUTES AND SCANNING REQUIREMENTS

Digital image capture must take into consideration the attributes of the source documents: physical dimensions, physical presentation, level of detail, tonal range, and presence of color. These document attributes will determine the type of scanning approach and devices to be used in digital conversion. For instance, very fine lettering will require high resolution for full informational capture; oversized items (greater than 8.5" x 14") can not be digitized on a flatbed scanner; fragile documents can not be scanned via an automatic sheet feeder.

The physical size and shape of the documents will determine the appropriate scanning device to be used. The first 'sort' is on document dimension; further subdivision is by physical presentation (bound, single-leaf), followed by document classification (e.g., printed text/line art); and then by tonal range and color attributes of the original source document (e.g., a black and white photograph). The final consideration will be the level of detail that must be rendered in the digital file. Capture requirements (resolution, bit depth) can be determined a number of ways. Once this information is gathered, digital conversion requirements and costs can be determined.

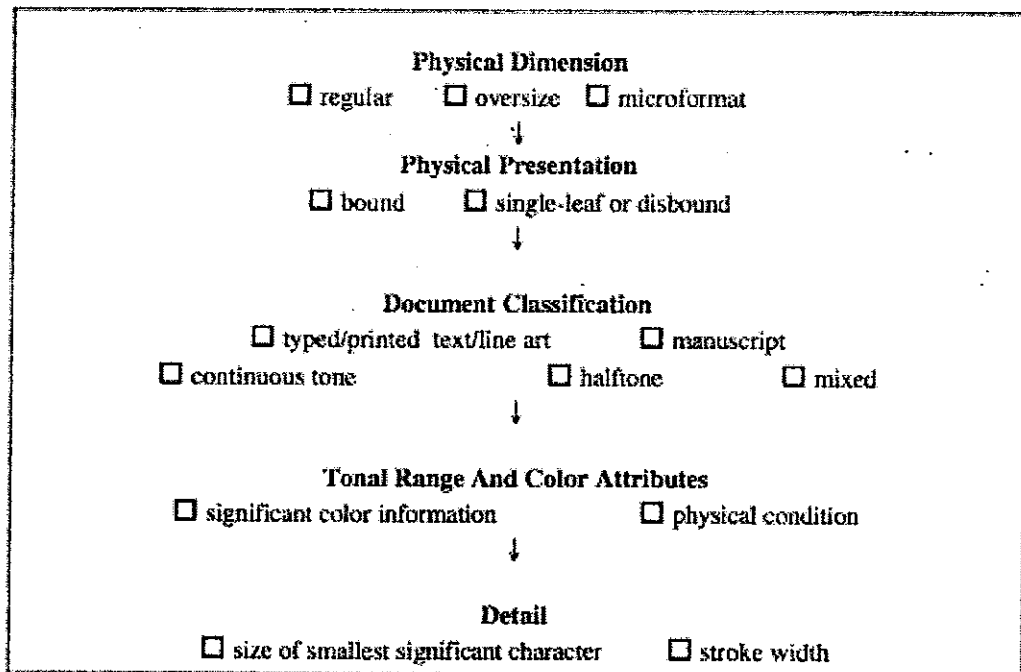


TABLE 1: STEPS IN DETERMINING THE APPROPRIATE SCANNING METHOD

14. FILE FORMATS

Extension	Meaning	Description	Strengths/weaknesses
.tiff, .tif	TIFF (Tagged Image File Format)	Uncompressed file. Originally developed for desktop publishing. 1 to 64 bit depth. Used mostly for high quality imaging and archival storage.	Generally non-compressed, high quality. Large file sizes. Most TIFF readers only read a maximum of 24-bit color. Delivery over web is hampered by file sizes. Although LZW compression can reduce these file sizes by 33% it should not be used for archival masters.
.gif	GIF (Graphics Interchange Format)	This 8-bit file format has support for LZW compression, interlacing and transparency.	Lossless compression. Popular delivery format on web. .png was defined to replace GIF.
.jpg, .jpeg	JPEG (Joint Photographic Experts Group)	Compressed images. 8-24 bit. Variable amount of compression to vary quality and file size.	Lossy compression. Widely used delivery format. Flexible.
MrSid	Multiresolution Seamless Image Database	image-compression technology	Lossy compression. can compress pictures at higher ratios than JPEG; stores multiple resolutions of images in a single file and allows the viewer to select the resolution.
.pcd	ImagePac, PhotoCD	Lossy compression. 24 bit depth. Has 5 layered image resolutions.	Used mainly for delivery of high quality images on CD.
.png	PNG (Portable Network Graphics)	Lossless compression. 24 bit. Replaced GIF due to copyright issues on the LZW compression. Supports interlacing, transparency, gamma.	Some programs cannot read it.
.pdf	PDF (Portable Document Format)	4-64 bit depth. Uncompressed. Used mainly to image documents for delivery.	Need plug-in or adobe application to view.
.pct	PICT	Compressed. Mac standard. Up to 32 bit. (CMYK not used at 32 bit.)	Supported by Macs and a highly limited number of PC applications.

Estimating Approximate File Sizes (in bytes):

Approximate file sizes for material created with a flat-bed scanner can be determined using the following formula:

$$FS = (SH \times SW \times BD \times dpi^2) / 8$$

FS = file size; SH = Source Height (inches); SW = Source Width (inches); BD = bit depth

dpi = dots per inch; /8 because 8 bits = 1 byte, the unit in which file sizes are measured.

CAPTURE: File Formats and File Size

15. METADATA TYPES

TYPE	GOAL	SAMPLE ELEMENTS	SAMPLE IMPLEMENTATIONS
Descriptive Metadata	<ul style="list-style-type: none"> · describing and identifying information resources at the local (system) level to enable searching and retrieving (e.g., searching an image collection to find paintings of animals) · at the Web-level, enables users to discover resources (e.g., search the Web to find digitized collections of poetry). 	<ul style="list-style-type: none"> · unique identifiers (PURL, Handle) · physical attributes (media, dimensions condition) · bibliographic attributes (title, author/creator, language, keywords) 	<ul style="list-style-type: none"> · Dublin Core · MARC · HTML Meta Tags · Controlled vocabularies such as: <i>Art and Architecture Thesaurus, Categories for the Description of Works of Art</i>
Administrative Metadata	<ul style="list-style-type: none"> · facilitates both short-term and long-term management and processing of digital collections · includes technical data on creation and quality control · includes rights management, access control and use requirements · preservation action information 	<p>Technical data such as scanner type and model, resolution, bit depth, color space, file format, compression, light source, owner, copyright date, copying and distribution limitations, license information, preservation activities (refreshing cycles, migration, etc.)</p>	<ul style="list-style-type: none"> · MOA2 Administrative Metadata Elements · National Library of Australia Preservation Metadata for Digital Collections · CEDARS
Structural Metadata	<ul style="list-style-type: none"> · facilitates navigation and presentation of electronic resources · provides information about the internal structure of resources including page, section, chapter numbering, indexes, and table of contents · describes relationship among materials (e.g., photograph B was included in manuscript A) · binds the related files and scripts (e.g., File A is the JPEG format of the archival image File B) 	<p>Structuring tags such as title page, table of contents, chapters, parts, errata, index, sub-object relationship (e.g., photograph from a diary)</p>	<ul style="list-style-type: none"> · SGML · XML · Encoded Archival Description (EAD) · MOA2, Structural Metadata Elements