WIU Libraries Digitization Policies Recommendations

Scope

These policies concern digitization projects undertaken by the library using library resources for the purpose of providing the public with free and immediate access to digitized materials. Projects initiated by outside parties, in which the library is a partner, are also included in the scope of these policies.

Currently, these policies only include the digitization of print materials (books, photographs, manuscript collections, maps, etc); audio and moving images may be addressed at a later time.

Policies

Digitization Project Prerequisites:

- Digital collections created by the library should conform to the mission of the University as defined in the WIU Strategic Plan and/or in the mission of the WIU Libraries.
- Digital collections created by the library should follow the library's collection development policies.
- Digital collections created from the Archives and Special Collections should follow the Archives and Special Collections' collection development policies.
- Digitization projects should meet with the approval of the Dean of the University Libraries.
- Digitization projects should be initiated by or partnered with an employee of Western Illinois University Libraries.
- Digital collections should consist of a cohesive set of items that are related to each other in some way.
- Those undertaking a digitization project should submit a digitization project request form and checklist to the Digitization Unit. This form will include a description of the project.
- Additionally, project sponsors should coordinate their projects with the Digitization Unit:
    - Project sponsors should consult with the Digitization Unit and with other units in the library as appropriate (for example, Acquisitions/Cataloging for cataloging issues, Computer Services for hardware issues and, Information Systems for CONTENTdm issues).

    - Project specifications should be coordinated with the Digitization Unit in advance.
    - Desired completion date of project is suggested, but may be subject to workflow of impacted units.
    - Additional Digitization Unit policies may apply.
- In the case of projects to be placed in CARLI's CONTENTdm interface, CARLI's "Prerequisites for Publishing a Collection in CARLI Digital Collection" should be followed, including submission of a project description for use as an "About this Collection" page.
- Projects to be placed in CARLI's CONTENTdm interface should follow CARLI's Collection Development Policies.

Selection of Items for Digitization:

- Items selected for digital collections should fit into the scope, purpose, and audience of the project itself as stated on the project description form.
- All digitized items should comply with copyright law.
- Originals of digitized textual items owned by the library should be represented in the library's online catalog.
    - Once the project is complete, each digitized item's catalog record should be modified to include a link to the digital surrogate of the item.
- The following criteria for item selection should be considered:
    - **Uniqueness**: Items that are unique or rare should be considered high-priority candidates for digitization, as the likelihood of duplicating previous digitization efforts will be low, and digitizing these materials will greatly increase the public's access to them as they are not widely available.
    - **Historical Significance:** Items of significant historical value should be given high priority.
    - **Regional Significance:** Items documenting the West-Central Illinois region should be given high priority, given our mandate to serve this region.
    - **Pedagogical value:** Items of significant educational value to students at WIU and beyond should be given high priority, given WIU's educational mission.
    - **Past use or potential use by patrons:** Items known to have high use should be given high priority, as should items known to have a high *potential* for use once they are digitized. For example, items with a specifically identified and interested audience.
    - **Context**: Only identified photographs should be digitized.
    - **Size/Format:** The size and/or format of the item to be digitized may limit the library's ability to digitize it given the capabilities of our scanners.

        - In some cases, it may be advisable to outsource the scanning of such materials.
    - **Fragility:** Fragility may either increase or decrease an item's potential for digitization. Fragile items that are able to be digitized with minimal damage should be given high priority, as digitization will reduce handling of the originals. However, some items may be so fragile that they cannot be digitized without significant damage; these should not be scanned without first stabilizing the item.

<u>Scanning/imaging</u>

Handling of Materials

- Rare and fragile materials, including items from the Archives/Special Collections, should be handled carefully with clean hands. Photographs should be handled with gloves. All items should be kept away from ink pens, food, and drink. Additionally, archival materials should be maintained in the order in which they arrived.
- Materials from the Archives and Special Collections should be kept in a secure and locked area when not in use.

Scanning

- See Appendix A for a chart describing the technical requirements for most common forms of scanned material.
- Some materials, such as photographs, may be outsourced for scanning. All policies in this document still apply to those materials.

File Naming

- Consistent, well-considered, and well-documented file naming conventions should be followed within each project. Each collection will receive a unique project accession identifier. This identifier will appear as a prefix to each file name within a project. Each file name should be unique and easily tracked.

Master Files

- Preservation master files should be created at the time of scanning; please see the Preservation Policies section of this document for policies regarding master file creation and maintenance.
- Service master files may be created as a working copy of a master file for image manipulation and conversion.

Conversion

- Service master files should be converted to JPG or JPG2000 access files for presentation on the Web.

Quality Control

- Scanned images, including master files and derivatives, should be checked for image quality and accuracy in file naming.
    - (Further policies regarding quality control will be forthcoming, after Linda Wade and Heather Richmond attend an April 15 webinar on the topic).
- Computer monitors used for checking images should be set to the following calibrations:
    - Set to 24 millions of colors.
    - Set monitor Gamma at 2.2.
    - When using a CRT monitor, calibrate color temperature to 6500° K.
    - If using an LCD monitor, calibrate color temperature to 5500° K.

(Information taken from BCR-CDP website)

Metadata:

Definition: BCR-CDP defines metadata as "descriptive information about digital resources." As defined by various people and practiced at different libraries, there are several types of metadata, often broken down into elements such as descriptive, structural, rights, and administrative. Descriptive metadata describes the content of an object; structural metadata describes the organization and display of a resource; rights metadata describes a resource's relationship with copyright law; and administrative metadata describes requirements for management and use of a resource, including technical information. For the purpose of this document, we are concerned with descriptive, rights, and administrative metadata.

Descriptive and Rights Metadata

- All Digital Collections should use Dublin Core as a minimum metadata standard for the creation of descriptive and rights metadata .
    - Metadata created in CONTENTdm or another platform should be mapped to unqualified Dublin Core fields in order to ensure that collections can be harvested by the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).
    - Projects should follow CARLI's Best Practices for Descriptive Metadata guidelines, including CARLI's Required Metadata Fields.
        - Currently, WIU Libraries will only require the five metadata fields required by CARLI; this number may increase as the policies are revised.
- When subject headings are used, they should utilize controlled vocabularies, whether from an established thesaurus such as LCSH, Art and Architecture

Thesaurus, or TGM-I; or from a vocabulary set devised for the purpose of the project.
- Other means of subject access, such as keywords, user-added content, or narrative description, need not be standardized.
- The choice of controlled vocabularies, classification schemes or depth of keyword assignments may depend on intended use and audience.
- All typed text collections should undergo Optical Character Resolution (OCR) conversion.
- While not required, transcription is encouraged for handwritten manuscript collections.

Administrative Metadata

Administrative metadata should be created for the preservation digital master files; see Preservation Policies for details.

Accessibility

- All digital collections should conform to WIU accessibility standards as applicable.
- Digital projects receiving grant money should conform to accessibility standards required by the granting agency.

Marketing and Assessment

- Digital collections, once complete, should be marketed to the collections' target audience.
    - Though the help of the library's marketing librarian can be enlisted, the project manager(s) should be primarily responsible for marketing the collection.
- If a defined audience can be easily identified (such as teachers), digital project managers should make every effort to seek their feedback.
- Digital collections will be periodically reevaluated for relevance, quality, enhancement/update needs, and possible weeding; user feedback may be a considered as part of this evaluation.

Preservation

Documentation

- In order to keep a record of a project and to establish institutional memory in the case of staff turn-over, projects should document the following:
    - Project description (see form)
    - File naming conventions

    - Subject standards used
    - Scanning settings
    - File formats of scanned images, both master files and access images
    - Metadata encoding standards used for both for master files and for access images
    - Location of master files
    - Location of metadata
    - Date of project completion
    - Migration information
    - Name of project contact
- All project information should be maintained in the library for the duration of the project.
    - Once a project is complete, project information should be transferred to the University Archives.

Master Files

- Digitization projects should create digital master files of scanned objects for long-term preservation. These masters should be stored in secure, dedicated digital storage space. In most cases, these master files will be preserved indefinitely, although there may be exceptions.
- Format:
    - For the time being, master files of images should be created as TIFFs.
        - While TIFFs are currently considered the standard for preservation master files, JPG2000 is being investigated as an alternative by organizations such as the Library of Congress; once the results of these investigations are available, the WIU library may begin producing all master files in JPG2000.
    - Preservation master files should be preserved as they were originally scanned, before any image manipulation occurs. Image manipulation can be performed on service master files (see Scanning Policies).
        - In cases where service master files have been substantially altered, they may need to be preserved as secondary master files.
- Metadata
    - Master files should include preservation metadata documenting the following:
        - Digital Collection Accession (project) Number
        - Object Identifier
        - Location, call number, size, and format of original (analog) source
        - Copyright information

        - Scanner used in image capture
        - Software used in image capture
        - Image capture date
        - Bit depth (1-bit bitonal, 4-bit grayscale, 24-bit RGB, etc)
        - Resolution (dpi)
        - Digital file format
        - Compression information
        - Image enhancement, if any
        - Migration information as files are converted to new formats and migrated to new hardware for preservation purposes
        - Descriptive metadata as specified in Metadata policies
        - Quality control information
    - Metadata encoding for master files should be standardized. We are investigating the use of the Metadata Encoding and Transmission Standard (METS) for use as a metadata standard for master files.

Storage

- Currently, preservation master files and their accompanying metadata should be stored on secure, dedicated server space, available through University Technology.
    - This server space cannot be used for anything but the long-term storage and preservation of digital objects.
    - In the future, this server space may be used to store a more complex, OAIS-compliant digital repository and thus should have capabilities that extend beyond simple file storage.
    - Master files stored on servers will be periodically backed up on tape as part of University Technology's backup program.

Review and Migration

- On an annual basis, all projects should be reviewed in order to refresh and migrate master files and their accompanying metadata as necessary.
    - Master files and metadata may be migrated to new file formats and standards, and all files may be migrated to new storage hardware, as technology becomes obsolete.
    - Any future format used for master files and their metadata should be open source, self-supporting, robust, and sustainable; any future storage hardware or preservation systems should be secure.

<u>Policies Review and Update</u>

- Portions of these policies are currently in flux and will be modified as needed.
- These policies should be reviewed on an annual basis and updated according to evolving best practices and standards in the digitization field.

-- Submitted by the WIU Libraries Digitization Policies Committee, March 13, 2009
Revised, August 28, 2009 -- HR

Appendix A
Image Quality Standards by Document Type

Document Type	Expected Outcome	Image Parameters Standards		Digital Master (Preserva-tion)	Display Format	Printing & Reproduc-tion	Thumbnail Image	Display Software
		Resolution	**Bit Depth**					
Local Copying, including ILL and reserves	Accurate representation	200 to 300 ppi			TIFF or PDF			TIFF or PDF image viewer
Printed text: books w/illustrations, pamphlets, typed pages, newspapers,	Image of text*	300 ppi minimum	8-bit grayscale	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
	OCR'ed text	300-400 ppi	8-bit grayscale	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
	Black & white documents that require color to produce the most accurate representation	400 ppi	24-bit color (RGB)	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Music: sheet music, annotated scores, music	Access to content	300 ppi minimum	8-bit grayscale**	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer

manuscripts	Recognition of artifactual features	400 ppi	8-bit grayscale	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Manuscripts: handwritten, typewritten copies	Access to content	300 ppi minimum	8-bit grayscale**	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
	Recognition of artifactual features	400 ppi	8-bit grayscale	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Maps: printed tones printed color up to D-size 22” x 34” *oversized	Content Research	250 ppi	24-bit color	TIFF	GIF, JPG, or JPG 2000 (500 ppi)	PDF or GeoPDF (500 ppi)	72 ppi	GeoPDF viewer, GIS software or Image viewing software
	Map reproduction	400† ppi	24-bit color	TIFF	GIF, JPG, or JPG 2000 (500 ppi)	PDF or GeoPDF (500 ppi)	72 ppi	GeoPDF viewer, GIS software or Image viewing software
Photographs (film) (35mm slides)	Accurate representation	2800 ppi	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer

Photographs (film) (4" x 5" negatives)	Accurate representation	800 ppi	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Photographs (film) (8" X 10" or larger)	Accurate representation	800 ppi	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Photographs (prints) (8”x10” or smaller)	Accurate representation	600 ppi (8" X 10") to 800 ppi (3" X 5" prints)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Photographs (prints) (8”x 10” to 11" x 14")	Accurate representation	600 ppi (8”x10”) to 430 ppi (11”x14”)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Photographs (prints) (larger than 11" x 14")	Accurate representation	570 ppi	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF (300 ppi)	GIF or JPG (72 ppi)	Adobe reader or Image viewer
Aerial Photographs (film) (70 mm roll)	Accurate representation	2700 ppi	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image viewing software
Aerial Photographs (film) (127mm roll, 4”x5” and	Accurate representation	2100 (127 mm) 1600 ppi (4”x5”) to 1100 ppi (5”x7”)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image

5”x7”)								viewing software
Aerial Photographs (film) (Larger than 127mm roll and 5”x7”)	Accurate representation	2000 ppi for 5”x5”	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image viewing software
Aerial Photographs (print) (Smaller than 8”x 10”)	Accurate representation	400 ppi (8”x10”), 570 ppi (5”x7”) and 800 (4”x5”)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image viewing software
Aerial Photographs (print) (8”x10” to 11”x14”)	Accurate representation	600 ppi (8”x10”) to 570 ppi (11”x14”)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image viewing software
Aerial Photographs (print) (Larger than 11”x14”)	Accurate representation	570 ppi (11”x14”)	8-bit grayscale or 24-bit color	TIFF	GIF or JPG (200 ppi)	PDF or GeoPDF (500 ppi)	GIF or JPG (72 ppi)	GeoPDF viewer, GIS software or Image viewing software

* Resolution depends on character size (1 mm)
** 24-bit color should be used where color is an important attribute of the document. In most cases, manuscripts should be scanned in color.
†Minimum resolution standards for maps will increase once the library has the capability to scan images over 400 ppi.

Information based on:

Library of Congress, 2006. Library of Congress Technical Standards for Digital Conversion of Text and Graphic Materials. Washington, D.C.: The Library. URL: http://lcweb2.loc.gov/ammem/about/techStandards.pdf

Puglia, Steven, Reed, Jeffrey and Rhodes, Erin, 2002. Technical Guidelines for Digitizing Archival Materials for Electronic Access. Washington, D.C.: National Archives & Records Administration. 87 pages. URL: http://www.archives.gov/preservation/technical/guidelines.html