

ILLINOIS STATE LIBRARY DIGITAL IMAGING PROGRAM- BEST PRACTICES (Rev. 2018)

Illinois Digital Archives (IDA)

Since 2000, the Illinois Digital Archives (IDA) has provided free access to the digital collections of libraries and cultural institutions throughout Illinois. Some examples of the images that comprise these collections include manuscripts, photographs, slides, maps, letters, posters, videos, oral histories and historical publications. The Illinois State Library (ISL) recognizes an ongoing need to preserve this content and to continue to add regional digital library collections to IDA. All digital content that increases awareness and access to local history adds enormous value to IDA and enhances the connection to Illinois' cultural heritage.

The ISL, through hosting IDA and partnering with the multitype Illinois library community plays a proactive role in furthering online access to historically significant Illinois materials held by libraries throughout the state. These collections will benefit both Illinois citizens and the broader internet community for future generations.

Planning for Digital Imaging Projects

A successful digital imaging project requires substantial planning, as these projects can be very complex, time-consuming, and costly. The following principles and assumptions provide an excellent framework to begin planning a digitizing project:

- 1) Define clear goals and milestones, especially the end point of the project.
- 2) Creation of digital files makes economic sense for reasons of creating access, not for reasons of preservation of collections.
- 3) Digitization can create more rather than less demand to use the original documents.
- 4) Digital projects are as much about cataloging as they are about creating electronic images of pictures and documents.
- 5) The work of preparing collections and finding aids to digital collections involves much more time and effort than the scanning of objects.
- 6) Organizing materials, creating indexing and making the collection searchable by users will require 2/3rds of the project time and funding.
- 7) Digital collections should be created with a context in mind. Design a set of outcomes for the presentation, creating a context for the collection that is deliverable to a wider audience.

Selection Criteria

One of the first and most important components of a digitization project is selection of material to digitize. The choice to digitize a collection can be made to 1) bring little used portions of the collection to a wider audience; 2) enhance access to fragile or restricted collections; or 3) enhance and unite collections on a common subject or theme for materials that are held at multiple organizations. Regardless of which of these functions a digitizing project serves, the following considerations need to be addressed when selecting materials to digitize: 1) Collection development plans your library may already have in place

- 1) Intellectual or cultural value of the collection to researchers
- 2) Demand from current (or potential) users
- 3) Historical or geographic area covered by the collection
- 4) Has another institution digitized the same, or similar, materials?
- 5) What is the physical condition of the collection?
- 6) Is the material suitable for digitization?
- 7) Will preservation work need to be done prior to digitization?
- 8) Bound volumes should be able to be opened to at least a 90 degree angle to be scanned; maps may need to be significantly reduced to display online resulting in a loss of fine detail and spatial context.
- 9) Copyright permission, since materials not in the public domain MUST have permission to digitize from the copyright owner.

Digital Imaging Guidelines

There are many best practices recommendations for digitizing materials. These guidelines may require adaptation to particular projects, dependent upon source document characteristics such as font size, photographic detail, and physical size.

The following general principles provide a starting point for creating digitized collections:

- 1) Capture once, use many times. The creation of high quality master images is key and should contain all the important information from the original material.
- 2) Create a faithful reproduction of the original. So as not to diminish the historical, cultural accuracy and value of the material, don't attempt to correct or improve the original content.
- 3) Scan from the earliest generation practical. Because each generation of an image loses some detail, always use the earliest generation of the original material that is practical to use. Typically, scanning from negative rather than prints or the original material rather than microfilm or photocopies is preferable.
- 4) Scan only at the resolution needed to capture the detail contained in the original; and therefore, when setting technical specifications, a higher resolution is not always the right choice.
- 5) Ensure that scanners, monitors and printers represent image colors accurately.
- 6) Ensure that digitizing equipment is appropriate to the materials that are being scanned.

The Illinois State Library Digital Imaging Program uses the following best practices for scanning:

ILLINOIS DIGITAL ARCHIVES GENERAL GUIDELINES FOR IMAGES			
MEDIA	SCAN FORMAT	MASTER IMAGE	ACCESS IMAGE
Printed text	8-bit grayscale	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi (grayscale) Resize images to 1024 pixels across long dimension
Handwritten text	8-bit grayscale or 24-bit color	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi Resize images to 1024 pixels across long dimension
B/W photo	8-bit grayscale or 24-bit color	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi Resize images to 1024 pixels across long dimension
Color photo Color slide B/W slide Map Illustration etc.	24-bit color	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi Resize images to 1024 pixels across long dimension
3-D artifact	24-bit color	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi Resize images to 1024 pixels across long dimension
Newspapers	8-bit grayscale or 24-bit color	TIFF Uncompressed Resolution: 300 ppi	JPEG Medium quality compression Resolution: 150 ppi As a general rule, do not resize the image

Metadata Standards

Metadata is a structured set of data that describes a resource for the purpose of discovery, use, management and preservation. The Dublin Core (DC) is an internationally agreed upon basic metadata scheme that defines 15 descriptive elements (e.g., Title, Creator, Subject). Detailed information about DC is available at <http://dublincore.org/documents/dces>

The complete metadata element set in IDA includes 24 fields which are based upon but are not limited to the DC metadata scheme. The metadata records in IDA are created using OCLC's CONTENTdm software. IDA metadata fields are, for the most part, mapped to DC metadata fields, which enables cross database searches in multiple collections.

Some metadata may be for use only within your institution or may be included in the metadata records for administrative purposes only. Any metadata fields *not* intended for public view can be hidden using the Administration in the CONTENTdm software.

The following table provides a brief overview of the fields contained in IDA metadata records. For detailed instruction, refer to the *Illinois Digital Archives: Metadata Guidelines (revised: 2018)* at: <http://www.cyberdriveillinois.com/departments/library/grants/pdfs/ida-metadata-guidelines.pdf>

IDA FIELD NAME	REQUIRED OR RECOMMENDED	MAP TO IN CONTENTdm	DEFINITION
1. Title	Required	Title	Name of the object, such as the title of a book or a painting
2. Creator	Required, if known	Creator	Name of the primary person or organization that produced the object, such as writer, photographer, artist or manufacturer
3. Subject	Required	Subject	What the content of the resource is about or what it is, expressed by headings, keywords, phrases, names, or other terms for significant people, places and events; includes keywords and phrases that describe a topic of the resource.
4. Description	Required	Description	A narrative of the content of the resource. Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content, a tape log or a free-text account of the content
5. Date Original	Required	Date Created	Date of creation of the original object.
6. Format	Required	None	The physical manifestation of the resource.
7. Identifier	Required	Identifier	Unique numeric or alpha-numeric character string used to label or classify a resource.
8. Rights	Required	Rights Management	Information about rights for access and reproduction held in and over a resource and may include copyright, citation, use or reprint information.
9. Type	Required	Type	The nature or genre of a digital resource.

10. Geographic Coverage	Required, if known or applicable	Coverage – Spatial	Geographic location relevant to the resource.
11. Contributing Institution	Required	None	Full name of the organization that owns the original object or digital resource.
12. Collection Name	Required	Relation-Is Part Of	Locally defined sets or collections to which the original source belongs.
13. Full Text/Transcript	Required for documents, audio and video files	Description	An exact typewritten account of the contents of a document, audio or video file.
14. File Size	Required for audio and video files	None	File size in KB and duration of the digital audio file.
15. Language	Required (excludes image only)	Language	Indicates the language(s) of the intellectual content of the resource. This implies the language(s) in which a text is written or the spoken language(s) of an audio or video resource.
16. Time Period	Recommended, as appropriate	Coverage – Temporal	A period represented by, relevant or related to the intellectual content of the resource
17. Browse Topic	Recommended, as appropriate	None	Selected topics pertaining to the general content of the original object or digital resource.
18. Contributor	Required for oral histories, otherwise recommended when available	Contributor	The person(s) or organization(s) that made significant intellectual contributions to the resource but whose contributions is secondary to any person(s) or organization(s) already specified in a Creator element.
19. Publisher	Recommended, when available	Publisher	Person or corporate/organizational entity responsible for producing a resource or a digital copy of a resource.
20. Source	Recommended, when available	Source	Information about the original object from which the digital resource has been derived. Whenever possible, use a standard identifier. If no standard exists, use a control number, accession number, barcode, locally derived call number or other local convention.
21. Date Digital	Recommended, when available	None	The date the object was digitized. The contributing institution may approximate the date a resource was digitized.
22. Digitization Specifications	Recommended, as appropriate	None	Describes the process used to capture and create the digital derivative of the original object. Record technical digitization information including the hardware, software and processes used. Include any information that will aid in improving access to and migration of the digital resource.
23. Digital Format	Recommended, as appropriate	Format-Medium	The digital manifestation or form of the resource.
24. Serial Information	Recommended, as appropriate	Relation-Is Part Of	Information to enable users to identify, cite and locate continuing publications issued in installments, which are typically numbered and dated.

Digital Preservation

The Digital Preservation Coalition describes digital preservation as, “all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change.”

The preservation of digital content is a vital component to any digitization project and needs to be factored into the workflow of the digitization process, rather than implemented on an ad hoc basis. It is the responsibility of the contributing institution to provide long-term storage for their digital collections that are included in IDA; and by developing a sustainable digital preservation strategy, you can ensure that the digital items found in your collection will be accessible now and in the future.

A successful strategy needs to address the creation, integrity, and maintenance of your digital items.

An example strategy might look something like this:

Creation

- Use common archival formats and recommended metadata for preservation.
 - Archival formats: TIFF for image files, MPEG-4/H.264 for video, and MP4 for audio files
 - Metadata for preservation: Digital Format, Identifier, File Size, Date Digital, Digitization Specifications

Integrity

- Assign someone with the role of data steward and carefully define their responsibilities.
- Use persistent identifiers for off-site storage of files.
- Run fixity checks your files on a set schedule to ensure none have become corrupt and replace corrupted files from a backup source when necessary.

Maintenance

- Maintain at least 3 copies of the same data. This includes copies for preservation (long-term storage) and Web optimized copies for access (often on different mediums).
 - Long-term storage format for digital images = TIFF file
 - Web optimized format for digital images = JPEG file
 - Burn files to DVDs, save files to a portable hard disk drive, and FTP files to a remote storage server
- Evaluate storage hardware at least every five years.

- Format obsolescence can render content useless and you will need to migrate data when necessary.

See links in Selected Resources for more detailed preservation strategies.

Selected Resources

Project Planning

Digital Public Libraries of America (DPLA), *Planning for Digitization* (2015):
<https://dp.la/info/about/projects/public-library-partnerships/planning-for-digitization/>

Federal Agencies Digitization Guidelines Initiative (FADGI), *DIGITIZATION ACTIVITIES, Project Planning and Management Outline* (2009): <http://www.digitizationguidelines.gov/guidelines/DigActivities-FADGI-v1-20091104.pdf>

Cornell University, *Moving Theory Into Practice: Digital Imaging Tutorial* (2003):
<http://preservationtutorial.library.cornell.edu/contents.html>

Oregon Parks and Recreation Department (OPRD): Heritage Program, *Planning a Digitization Project* (2013):
http://www.oregon.gov/oprd/HCD/docs/Heritage_Bulletins/HB21_Digitization_Projects_Planning.pdf

Standards/best practices/recommendations for digitization

Federal Agencies Digitization Guidelines Initiative (FADGI), *Guidelines: Technical Guidelines for Digitizing Cultural Heritage Materials* (2016): <http://www.digitizationguidelines.gov/guidelines/digitize-technical.html>

Bibliographical Center for Research (BCR): Collaborative Digitization Program (CDP), *BCR's CDP Digital Imaging Best Practices Version 2.0* (2008):
http://www.mwdl.org/docs/digital-imaging-bp_2.0.pdf

Consortium of Academic and Research Libraries in Illinois (CARLI), *GUIDELINES FOR THE CREATION OF DIGITAL COLLECTIONS* (2013):
http://www.carli.illinois.edu/sites/files/digital_collections/documentation/guidelines_for_text.pdf

National Information Standards Organization (NISO), *A Framework of Guidance for Building Good Digital Collections, 3rd edition* (2007):
<http://framework.niso.org/5.html>

Metadata

OCLC, *Best Practices for CONTENTdm and other OAI-PMH compliant repositories: creating sharable metadata, Version 3.1* (2013):
<http://www.oclc.org/content/dam/support/wcdigitalcollectiongateway/MetadataBestPractices.pdf>

Illinois Digital Archives (IDA), *Metadata Guidelines* (2018)

<http://www.cyberdriveillinois.com/departments/library/grants/pdfs/ida-metadata-guidelines.pdf>

Dublin Core Metadata Initiative (DCMI), *User Guide*:

https://github.com/dcmi/repository/blob/master/mediawiki/wiki/User_Guide.md

Preservation

Library of Congress, *Sustainability of Digital Formats: Planning for Library of Congress Collections* (2017):

<https://www.loc.gov/preservation/digital/formats/index.html>

State Archives of North Carolina, *File Format Guidelines for Management and Long-term retention of Electronic Records* (2012):

http://digitalpreservation.ncdcr.gov/file_formats_in-house_preservation.pdf

ICT Policy Support Programme, *Best practices for a digital storage infrastructure for the long-term preservation of digital files* (2013):

http://www.dca-project.eu/images/uploads/banners/DCA_D62_Best_practices_for_a_digital_storage_infrastructure_2013_0506_Version1.pdf

UMass Amherst Libraries, *Digital Preservation Policy* (2011):

<http://www.library.umass.edu/assets/aboutus/attachments/University-of-Massachusetts-Amherst-Libraries-Digital-Preservation-Policy3-18-2011-templated.pdf>