

"Digital imaging projects are complex, time-consuming, and costly. The success of a project is generally in proportion to the time spent in planning the project."

Planning for Successful Digital Imaging Projects

Alyce Scott Digital Imaging Program, Illinois State Library Coordinator

Digital imaging technology has become a more practical method of providing a greater level of citizen access to the intellectual content of primary source materials held in libraries of all types around the world. A library digitizing primary source material provides access to its community's cultural heritage. Digitization most certainly increases an appreciation of the material and how it relates to the community and the nation.

Digital imaging projects are complex, time-consuming, and costly. The success of a project is generally in proportion to the time spent in planning the project. To help you avoid some of the pitfalls here are some recommendations and resources for planning a digital imaging project.

Some key components of a digital imaging project:

- Selection
- Standards
- Access

Selection - Issues to consider when selecting material for digitization:

- Collection development plans your library may already have in place
- \cdot $\;$ Intellectual value of the collection to researchers
- \cdot $\;$ Demand from current (or potential) users
- Historical or geographic area covered by the collection
- Has another institution digitized the same, or similar, materials?
- Physical condition of the collection, is the material suitable for digitization? (Issues to consider: will preservation work need to be done prior to digitization; 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)
- Copyright permission (if the materials are not in the public domain you MUST have permission from the

copyright owner to digitize the material)

Standards for digitization-

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

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

Archival images:

- File saved in uncompressed TIFF format
- Printed black & white text or maps: bitonal, 600 ppi
- Black & white photographs: 8-12 bit grayscale or 24-36 bit color, 300-600 ppi
- · Color photographs, manuscripts: 24-36 bit color, 300-600 ppi

Access (or display) images:

- File saved in JPEG format, with medium quality compression
- · 150 ppi
- 1024 pixels in length

Thumbnail images:

- File saved in Compuserve GIF format
- 72 ppi
- \cdot 150-200 pixels in length

Access - issues in organization, management, and delivery of your digital image collection:

- Metadata: cataloging and technical data associated with digital images either embedded or as associated text, crucial for searching and access
- Storage: where will the images reside, will you need to purchase a server? Backup/disaster recovery: two copies of all digital image files are recommended, one stored off-site
- Rights management: copyright notices, licensing agreements, digital watermarking
- Viewing software: will your users need to download a plug-in to view the images?
- Finding aids/indexing: will creation of indexes/finding aids be necessary to assist users in accessing information?
- Reformatting/media refreshing to avoid data degradation: plan ahead, it's inevitable

In closing, it should be noted that digitization is not preservation. While there are elements of preservation involved in the digitizing process, the purpose of creating digital images is to provide access to primary source materials that might otherwise be unavailable to researchers, scholars, and the public.

Resources

Books

Besser, Howard and Jennifer Trant. *Introduction to Imaging: Issues in Constructing an Image Database*. Santa Monica, CA: The Getty Art History Information Program, 1995.

Ester, Michael. *Digital Image Collections: Issues and Practice*. Washington, D.C.: Commission on Preservation and Access, 1996.

Kenney, Anne R. and Oya Y. Rieger. *Theory into Practice: Digital Imaging for Libraries and Archives. Mountain View*, CA: Research Libraries Group, 2000.

Northeast Documents Conservation Center. *Handbook for Digital Projects*. Andover, MA: NEDCC, 2000.

Smith, Abby. *Why Digitize?* Washington, D.C.: Council on Library and Information Resources, 1999.

Web sites

Benchmark for Digital Reproduction of Monographs and Serials (revised January 2002) available at the Digital Library Federation Web site.

Digital Conversion of Research Library Materials by Stephen Chapman and Anne E. Kenney (Cornell University)

(D-Lib Magazine, October 1996 – old but still useful) available at the Digital Library Magazine Web site.

Digital Formats for Content Reproductions (Library of Congress) July 1998 available at The Library of Congress American Memory Web site.

A Framework of Guidance for Building Good Digital Collections (National Information Standards Organization) available at the National Information Standards Organization Web site.

Image Quality Calculator (use to determine the best scanning resolution for your digital images) available at the University of Illinois Library Web site.

Moving Theory into Practice: Digital Imaging Tutorial (Cornell University) available at the Cornell University Library PReservation and Collection Maintenance Web site.

Metadata

Introduction to Metadata (Getty Institute) available at the Getty Research Institute Web site.

Understanding Metadata (National Information Standards Organization) available at the National Information Standards Organization Web site.

Copyright

Fishman, Stephen. *The Public Domain: How to Find & Use Copyright-Free Writings, Music, Art & More.* Berkeley, CA: Nolo Press, 2001.

Copyright (OCLC Digitization & Preservation Online Resource Center) available at the OCLC Web site.

Crash Course in Copyright (University of Texas) available at the University of Texas Web site.

When Works Pass Into the Public Domain, 2003 available at the University of North Carolina Law Library Web site.

Copyright Term and the Public Domain in the United States available at the University of Cornell Copyright Information Center.

Glossary of Scanning and Digital Imaging Terms

Bit

The smallest unit of computing information.

Bit depth (1-bit, 8-bit, 24-bit)

The amount of information (black and white or color) a computer can discern for each bit of an image. 1-bit is black and white (off or on), 8-bit is 256 "shades", "values" or "levels" of gray or 256 colors, 24-bit is millions of colors.

CD-ROM (Compact Disc-Read Only Memory)

A storage disk for computer files; a CD-ROM can hold about 650 megabytes of data; you cannot replace the information on a CD-ROM as you can on a floppy disk or hard disk.

Crop

To select out an area of an image. Once an image is cropped, save the cropped version with a different name, retaining the original image.

Digital image

A computer file which, when used in conjunction with the proper software, will display a picture on the computer screen or print out to a digital device such as a laser printer.

Dither

A way of arranging the dots in a digitized image that creates an optical illusion of more continuous colors or gray tones than the computer or device can actually display or print.

Download

To "get" a file; to move a file electronically from one place (such as a Web page or server) to your machine (such as onto your hard drive or floppy disk).

Downsize

To reduce the file size of an image, by lowering the resolution and/or reducing the square measurement of the file.

Dpi (dots per inch) Measure of resolution for a laser printer. See also: Ppi (pixels per inch)

File format

The specific way digital information is made and stored by the computer. Not all software applications can read and/or manipulate all file formats. (See: GIF, JPEG, TIFF.)

GIF (Graphics Interchange Format)

A common graphic file format on the World Wide Web; used by online services and Web browsing software, GIFs contain information compressed into a relatively small file size and may display faster than other formats.

Grayscale

A system of displaying images in gray tones (or "levels of gray"), simulating the continuous gray tones of a photograph. To achieve grayscale, a monitor must be able to display 2 to 16 bits of information per pixel. This allows the monitor to display a black or white pixel as well as several values between black and white.

Image file size

The amount of computer storage space a file requires; usually measured in kilobytes (K) or megabytes (M, MB, mgs or "megs"). An image file that is 5 x 7 inches, 8-bit gray (as in a black and white photo), resolution 300dpi, is 3M in size. (A floppy disk holds 1.3M.)

Image size

The physical dimensions of the image as measured in the small squares (pixels) of a computer screen; an image filling a "typical" computer screen (13 inch diagonal) would be 640 x 480 pixels; compare to image file size above.

JPEG (Joint Photographic Experts group)

Pronounced "JAY-peg", a graphic file format that compresses information about many colors (up to 16 million) in the image into a smaller file.

Line art

Black and white art, usually some type of line drawing (such as that produced by pen and ink).

"Manipulate the image"

Change the image electronically in some way—resize, change the resolution, remove color, sharpen, clean up, edit, convert the file format, etc.

Output resolution

The detail and clarity (achieved by closeness of dots) with which the image will be displayed or printed (dependent on the capability of the display or printing device).

Photoshop

Software developed by Adobe, which allows manipulation and editing (enhancing, resizing, cropping, etc.) of digital images.

Platen The glass surface of a flatbed scanner.

Ppi (pixels per inch) Measure of resolution for a monitor. See also: Dpi (dots per inch)

Resize

To change the size of an image by reducing or increasing the resolution and/or the square measurement of the file. [Note: it is not possible to add more data to an image after it is scanned. It is always preferable to scan an image at the size needed rather than to try to increase the size or resolution later.]

Resolution

An expression of image size; the sharpness and clarity of an image, achieved by the closeness of the dots that make up the image. Resolution is expressed for the scanner as samples per inch (spi), for the screen as pixels per inch (ppi), for the printer as dots per inch (dpi). Most people say "dots per inch" when speaking of scanning resolution, (although technically this is not accurate). The more data per inch (samples, pixels, dots) the higher the resolution of the image and the better looking the image will be. Most screens display at a resolution of 72 pixels per inch. Most laser printers print at 300 or 600 dpi. Higher resolution image files are much larger than low resolution image files, so only save a high resolution image if you need to (such as for archiving). You will need a high resolution image if you are going to print the image in a paper publication and/or enlarge all or any part of the image on screen or on paper.

Scanner

A device that takes a picture of an image, breaks it down into dots and records it as a digital file for use with a computer. Some types of scanners are:

- flatbed scanner: A device for converting paper images (photographs, drawings, printed images) into computer graphic files.
- slide scanner: A device for converting 35mm slides into computer graphic files.

Server

A computer on a network that can be accessed by other computers on the same network; a server can hold software for several people to use and/or space for people to save and access files.

Size (see "Image file size")

TIFF (Tagged Image File Format)

A type of graphic file format developed for scanning. TIFFs are bitmapped graphics that can contain lots of information about each bit or pixel. TIFFs can be read by both Macintosh and PC/Windows applications, such as PageMaker and QuarkXpress. If you think you will ever print your image in a book or publication of any kind, you will want to save a copy of your image as a TIFF. Because TIFFs save a lot of information about each pixel, they can be very large files.

Value

The lightness or darkness of a gray or a color. The darkest level or value of gray is black and the lightest level of gray is white.