

UNC Digital Library Services

Digitization Guidelines [Quick Reference]

Introduction	2
Using these Guidelines	2
Before You Begin Digitizing	3
Images	4
Photographs, Paintings, Slides	4
Images for Archiving	4
Images for Web Delivery	4
Images for Printing.....	5
Logos, Line Drawings, Geometric Shapes	5
Audio.....	6
Music Audio.....	6
Archiving.....	6
Web Delivery	6
Voice Audio	6
Archiving.....	7
Web Delivery	7
Video.....	8
Video for Archiving.....	8
Web and DVD Delivery.....	8
Text.....	9
Web Display	9
Printing	9
Archiving.....	10
Presentations.....	11
Other.....	11
Storage	12
Looking for Something Else?	13
Campus Help.....	13

Introduction

The purpose of this document is to help support common digitization needs at UNC-Chapel Hill. The recommendations included here may not be appropriate for all digitization projects, especially those with special needs. The recommendations are also likely to evolve with advances in digitization techniques and underlying technologies.

These guidelines are part of an effort to encourage campus digitization projects to adopt standards that will facilitate resource-sharing within and across departments. A central infrastructure for storing, managing and sharing media collections is currently being piloted on campus. For more information, see the Storage section of this document or visit the UNC Digital Library Project website at <http://www.unc.edu/projects/diglib/>.

The recommendations in this document were developed by a working group of media users and campus support representatives. The following organizations participated, and are adhering to these guidelines in their support of departments and individual instructors and researchers:

- Academic Affairs Library
- Center for Instructional Technology
- Center for Teaching and Learning
- College of Arts and Sciences (OASIS)
- Health Sciences Library

Using these Guidelines

This is the Quick Reference version of the UNC Digital Library Services' Digitization Guidelines. The Quick Reference version provides the standards for each media type along with brief explanations. The full version provides richer background information and explanations, and explores a number of emerging digitization solutions not covered in the Quick Reference version.

Please keep in mind that the media editors and viewers recommended in this document do not represent all the useful software available. The focus is on software commonly used here at UNC-Chapel Hill.

Before You Begin Digitizing

Before you digitize anything, take some time to consider your needs. The worst possible outcome is to spend time digitizing materials that end up being inappropriate for the goals of your project. To avoid this scenario, consider a number of issues ahead of time.

- For what purposes will the materials be used?
- What level of media quality is necessary to achieve your goals for the project?
- Who needs to have access to your digital media?
- What options do you have for making the materials available to your audience?
- Who owns the copyright to the materials you are digitizing?
- What options, both short-term and long-term, are available to you for storing your digitized media files?
- Would it be cost-effective to outsource the digitization of your media?

If you have any questions or would like to discuss a digitization project with someone on campus, contact one of the organizations listed under Campus Help or send email to mediasupport@unc.edu.

For additional information on planning digitization projects, you might also consult one of many online resources available. Several are listed here:

- Digital Library of Georgia Digitization Guide (<http://dlg.galileo.usg.edu/guide.html>)
- Washington State Library's Digital Best Practices (<http://digitalwa.statelib.wa.gov/newsite/projectmgmt/planning.htm>)
- Digitization Guidelines at Harvard University (<http://preserve.harvard.edu/resources/digital.html>)
- National Digital Library Program Checklist (<http://memory.loc.gov/ammem/techdocs/prjplan.html>)

Images

Photographs, Paintings, Slides

Images for Archiving

When you try to enlarge digital versions of complex and colorful images like photographs, paintings and slides, the quality of the scan degrades. Having a larger archive version will provide you with the most flexibility when you need to go back and create new versions of an image. It is recommended that you save archive versions of images in the TIFF format. PNG is an alternative to TIFF, but it is not yet as widely accepted and supported. While Photoshop's PSD format is widely used for editing, it is proprietary and thus not recommended for archive use. Because TIFF files are large, you will probably want to use your smaller web versions on a day-to-day basis. Storage options for your archive images should be considered in advance.

Images for archiving

Format	TIFF (no compression), PNG
Size	4000x2500 to 6000x4000 pixels
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Adobe Photoshop, Adobe Photoshop Elements, Microsoft Photo Editor
Recommended Viewers	Any web browser

Images for Web Delivery

For images that you plan to make available via the Web, it is recommended that you save these files in the JPEG format. Ideally, your JPEG images would be derivatives of your archive version, or original scan (see Images for Archiving). You should also have two versions of each image, one larger than the other. There are several advantages of having two different versions of your images for use on the web. Having a smaller version makes it easier for users with slower speed Internet connections to download the image. The larger version allows for examining details of the image.

If you are affiliated with the College of Arts and Sciences and are having media scanned through OASIS Media Services, your web-ready versions will be created for you. If not, some scanning software and common image-editing software like Photoshop can be used to create these derivatives from your archive version.

Images for web delivery (smaller size)

Format	JPEG
Size	600x400 pixels
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Adobe ImageReady, Adobe Photoshop, Adobe Photoshop Elements, Macromedia Fireworks Microsoft Photo Editor
Recommended Viewers	any web browser

Images for web delivery (larger size)

Format	JPEG
Size	1800x1200 pixels
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Adobe ImageReady, Adobe Photoshop, Adobe Photoshop Elements, Macromedia Fireworks Microsoft Photo Editor

Images for Printing

If you want to print images for your personal use, or use in a class, we recommend that you follow the Web Delivery guidelines (above).

If you are readying images for publication in a journal, book, magazine, etc., publishers generally have much higher standards for image quality. The standard size for formal publications usually varies anywhere between 150 dpi and 300 dpi. You should contact your publisher and inquire about its preferred specifications.

Images for printing

Format	JPEG; TIFF; PSD
Size	300 dpi
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Adobe Photoshop, Adobe Photoshop Elements, Microsoft Photo Editor

Additional information regarding dpi or ppi and image size:

Image resolution is a tricky and often confusing concept. Typically, image resolution is defined by dots or pixels per inch (dpi or ppi respectively). However, this terminology is only useful for determining the *quality* of an image when it is *printed*. The more dots or pixels per inch, the more complex and detailed the image will appear *in print*. But if you are simply displaying an image on a *computer screen*, the resolution is only referring to the *size* of the image on the computer screen. When you change the resolution of a digital image, it does not become more or less complex - it only gets bigger or smaller. The standard 300 dpi for formal publications ensures a level of quality for the printed image, regardless of its eventual size on the printed page. Below you will find a table of inch-sizes, dpi, and resulting pixel length and width, and a table with recommendations for some common photographic images.

Original	DPI	Result
Slide (1" x 1.5")	1200	1200 x 1800 pixels
Print film (4" x 6")	600	2400 x 3600 pixels
Oversized print (8" x 10")	300	2400 x 3000 pixels

Logos, Line Drawings, Geometric Shapes

Images based on geometric shapes, also known as vector-based images, can be enlarged without loss of quality and thus do not require different sizes for different output formats. If you need to display a logo or simple line drawing on the web, GIF is the recommended format.

Vector-based Imagery

Format	GIF*, Flash, SVG
Size	N/A
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Adobe Illustrator, Macromedia Flash/Fireworks
Requirements	Any web browser (with Adobe SVG or Macromedia Flash plug-in)

Audio

Before digitizing your audio files, determine what level of audio quality is most appropriate for your audience and the materials you are digitizing. The higher the quality, the larger the file size. File size will have some bearing on your options for storage and delivery.

In addition to the length of your original recording, the three variables that will determine your digitized file size are: bit-rate (defined by your choice of format), frequency (KHz, which is format-independent), and bits-per-channel (mono vs. stereo).

Music Audio

Archiving

You can not convert lower-quality digital audio back into a higher-quality version. Keeping a high quality archive version will provide you with the most flexibility when you need to go back and create new versions of a sound file. The base unit for archiving music remains the uncompressed WAV (for PCs) or AIFF (for MACs) file, at the highest frequency and bits-per-channel possible.

Music Audio for archiving

Format	Windows Waveform (WAV); Macintosh AIFF; Shorten (SHN); Monkey's Audio (APE)
Size	44.1-48KHz; 16-24 bit; Stereo
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Goldwave or Soundforge (editing), RealMedia's production applications (encoding), Digidesign ProTools
Recommended Players	Any audio player

Web Delivery

If you are going to deliver music over the web, we recommend encoding as MP3, or a comparable compressed format (like Real Media or Quick Time).

Music Audio for web delivery

Format	MP3, RealAudio, QuickTime
Size	128-224 Kbps / 22-44.1 KHz; Stereo
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Goldwave or Soundforge (editing), CoolEdit, RealMedia's production applications (encoding), Digidesign ProTools
Recommended Players	RealOne Player, WinAmp, Windows Media Player
Requirements	MP3 player

Voice Audio

Voice audio does not require the frequency or bit-rate that music does. You can save both storage space and download time by compressing voice audio at lower frequencies. Single voice recordings should also be recorded in monophonic rather than stereophonic sound, which immediately decreases the file size by half.

Archiving

See Archiving for Music Audio above. We also recommend archiving voice audio in WAV or AIFF formats.

Voice Audio for archiving

Format	Windows Waveform (WAV); Macintosh AIFF; Shorten (SHN); Monkey's Audio (APE)
Size	22 KHz; 8 -16 bit; Mono
Campus Help	CIT, CTL, OASIS, UNC Libraries
Recommended Editors	Goldwave or Soundforge (editing), RealMedia's production applications (encoding), Digidesign ProTools
Recommended Players	RealOne Player, WinAmp, Windows Media Player

Web Delivery

Voice Audio for web delivery

Format	MP3, RealAudio, QuickTime
Size	28 - 128 Kbps / 11 KHz; Mono
Campus Help	??
Recommended Editors	Goldwave or Soundforge (editing), RealMedia's production applications (encoding), Digidesign ProTools
Recommended Players	RealOne Player, WinAmp, Windows Media Player

Video

Digital video creation is one of the most complex multimedia activities you can undertake. Consider some of the issues raised in *Before You Begin Digitizing* as you plan your project.

Video for Archiving

Archiving raw video is very storage-intensive. Raw digital video AVI files require roughly 10 to 15 *gigabytes* of storage per hour! More reasonable is MPEG2-based DVD, which allows for a high quality video stream and a high quality audio stream. DVD-quality video is lossy and therefore not a traditional archival format, but its high quality makes it a reasonable alternative to the raw AVI format.

Video for archiving

Format	MPEG2/DVD; raw Microsoft AVI (DV-compatible); miniDV/DVcam
Size	720x480 pixels or higher
Campus Help	CIT (Video Services), OASIS
Recommended Editors	Adobe Premiere, TMPGEnc, FinalCut Pro (Mac)
Simple Editors	Macintosh iMovie, Windows Movie Maker
Recommended Viewers	RealOne Player, Windows Media Player, QuickTime Player
Requirements	drive space (these files are very large)

Web and DVD Delivery

Because RealMedia files are small, retain excellent video information, and are platform independent, we recommend you use this format for web delivery of video.

For higher speed networks (including intracampus and Internet2) and DVD media, we recommend MPEG2-based digital versatile disc (DVD) video. This video has a higher resolution and better audio quality. Since it is also recommended for archival video, choosing this format may cover your delivery and archiving needs, depending on your project.

Video for web delivery

Format	MPEG1/VCD (DivX codec); RealMedia; QuickTime (Sorenson codec)
Size	160x120 to 320x240 pixels
Campus Help	CIT (Video Services), OASIS
Recommended Editors	Adobe Premiere, TMPGEnc, FinalCut Pro (Mac), RealMedia's production applications (encoding)
Simple Editors	Macintosh iMovie, Windows Movie Maker
Recommended Viewers	RealOne Player, Windows Media Player, QuickTime Player

Video for high-speed networks and DVD

Format	MPEG2/DVD
Size	320x240 (Web); 640x480 (VHS); 720x480 (DVD)
Campus Help	CIT (Video Services), OASIS
Recommended Editors	Adobe Premiere, TMPGEnc, FinalCut Pro (Mac), RealMedia's production applications (encoding)
Simple Editors	Macintosh iMovie, Windows Movie Maker
Recommended Viewers	RealOne Player, Windows Media Player, QuickTime Player
Requirements	greater bandwidth (e.g., Internet2); DVD-ROM

Text

Many options exist for the creation and presentation of text. Text documents in their simplest form (plain ASCII, for example) have no formatting applied, but most of the commonly used document formats mix together the structural and presentation aspects of their documents. This is true of, e.g., Microsoft Word, WordPerfect, RTF, PDF, and HTML. There are also formats, particularly SGML and XML markup languages, which allow for the separation of these aspects. In the latter document types, the structure and semantics of the document are dealt with in the document itself, and the presentation is generally delegated to other software or documents. For the purposes of presentation, SGML and XML files may be transformed into any of the common formats mentioned above.

Anyone who wishes to produce a collection of texts should answer the following questions before deciding on what format(s) to use:

- 1) Who is the audience for this collection?
- 2) What is the level of complexity of the collection?
- 3) What text-creation capabilities do I possess or have access to?
- 4) How static are the documents in the collection and the collection itself?
- 5) How sustainable should the collection be?
- 6) To what kinds of uses might the collection be put in the future?

See the full version of the Digitization Guidelines for more on deciding about text formats.

Web Display

Modern web browsers can handle most text formats. ASCII (plain text) text files require the least amount of space (one byte per character), but they cannot handle any advanced formatting, including italicization and bold marking. Rich Text Format (RTF) allows for this formatting, but not much more. Microsoft Word documents and HTML documents allow for advanced features including internal and external hypertext referencing. And Adobe PDF is useful for publication but not for shared editing of documents. The most innovative solution is the use of XML documents, which can be rendered into any of the other formats mentioned below using XSL stylesheets, or they can be displayed directly in some browsers with an XSL or CSS stylesheet. Knowing who you are publishing for and why will help you determine the most appropriate format.

Text for web display

Format	Adobe PDF; Microsoft Word (or other word processor); RTF; XML+stylesheet; HTML; ASCII/Unicode(TXT)
Campus Help	CIT, CTL, ITRC, OASIS, UNC Libraries
Recommended Editors	Notepad, Wordpad, Word, Adobe Acrobat; XMLSpy, JEdit, Xmetal, OpenOffice Writer, Adobe FrameMaker, Macromedia HomeSite, Macromedia Dreamweaver
Recommended Viewers	Web browser, Acrobat Reader, see above

Printing

For most text documents, you will probably find that the formats mentioned under Web Display (above) are adequate.

There are, however, other formats that allow for more control over layout and printer-specific commands. PostScript (PS) and Encapsulated PostScript (EPS) are both such formats, as well as

LaTeX, which even has its own language used for formatting and rendering text documents. XML can be rendered into a number of print-ready formats. Unfortunately, all of these print-specific methods require some expertise to execute correctly.

If you are readying text for publication in a journal, book, magazine, etc., you should contact your publisher and inquire about its preferred specifications.

Text for printing

Format	(Encapsulated) PostScript; XML+XSLFO; LaTeX
Size	150-300 dpi
Campus Help	CIT, CTL, ITRC, OASIS, UNC Libraries
Recommended Editors	Adobe FrameMaker; TeX/LaTeX; OpenOffice Writer, Microsoft Word (for smaller documents), Microsoft Publisher, Adobe InDesign, Adobe Pagemaker, Adobe Framemaker

Archiving

If you think you will need to access a text document again over a longer period of time, it is recommended you keep an archive version. How you archive text depends directly on what the text is intended to do. Certainly if you're writing a simple document without any necessary formatting, archiving text as a simple TXT document, edited in something as simple as Notepad, is fine. But as documents get more complex, both in contents and formatting, it is important to archive as much information as possible. We recommend using XML along with a document type definition (DTD) or schema, which will allow you to define your document's structure and can be used to validate your final document. Because XML is both standardized and human-readable, encoding your document in this manner will ensure its longevity even as the editing programs and formats themselves change. For archiving text, we strongly recommend avoiding proprietary formats such as Microsoft Word

Text for archiving

Format	XML+DTD (e.g., MathML, TEI, DocBook, etc.); SGML+DTD; RTF; ASCII/Unicode(TXT); LaTeX; (E)PS
Campus Help	OASIS
Recommended Editors	Notepad, Wordpad, XMLSpy, JEdit, XMetal, OpenOffice Writer, any other text editor or SGML/XML editor that does not save to a proprietary format.
Recommended Viewers	Any XML/XSL compatible Web browser

Presentations

There are a number of options for converting analog presentations (e.g. overhead slides) into digital formats. Due to its low cost and wide availability on campus, Microsoft PowerPoint is a popular option. If you want to share presentations with people who may not have PowerPoint, saving PowerPoint presentations as web pages is a standard PowerPoint feature. For those more comfortable with multimedia authoring, programs like Macromedia's Flash can be used to create interactive presentations.

Presentations

Format	Microsoft PowerPoint; Macromedia Flash; Adobe Persuasion; SVG; SMIL
Campus Help	CIT, help.unc.edu
Recommended Editors	Microsoft PowerPoint; Macromedia Flash; Adobe LiveMotion; GRiNS (SMIL)
Requirements	Web browser (and plug-in)

Other

Computer Code. Similar to the reasons for retaining both original and output versions of presentations, it is important to retain the original computer code alongside its binary derivative.

Data Sets. Our only recommendation for data sets is again similar to that above: save the original format of the data if at all possible. Because the data sets used on campus vary widely both in origin and purpose, no one program can be recommended for handling them. There are, however, "expert pockets" on campus which should be consulted when working with particular data sets: e.g., the Odum Institute for Research in Social Science, the Department of Geography or the Ancient World Mapping Center for geographical information systems (GIS) data, the School of Information and Library Science for bioinformatics data, etc. This document defers to these groups as experts in these areas and will not discuss them any further here.

Storage

Storage options for your digitized media should be considered *before* you begin digitizing. Storage space needs vary significantly, depending on file formats and the quality of media desired. Backup policies should *always* be implemented.

There are numerous solutions for storing media. Computer hard drives, ZIP drives, and CD or DVD offer local but limited storage space for many media types. For storage of larger media like raw digital audio and video, you might consider an external hard drive. The new Firewire standard allows for faster access to these drives. DAT is also a commonly used medium for storing audio, and miniDV/Dvcam for video.

If your media are directly related to a particular department or project, you can store the media in networked space that can be accessed from remote locations. Projects are currently allotted up to 10GB per project or department, with more available to purchase. Of course, you may also use your own personal space, currently limited to 250MB.

If you have a fairly large number of media files or expect to build a collection over time, UNC Digital Library Services may be an appropriate option for storing, managing and sharing your instructional or research collections. For more information on the project and who to contact with questions, see the project home page at www.unc.edu/projects/diglib/, or send email to diglib@unc.edu.

ATN also offers university projects a remote tape backup system called "mass storage," which allows for a more permanent method archiving. It is not, however, as quickly and easily accessible as local media and thus should be used more as a strict archive than a working space. Additional information on this service is available at: http://www.unc.edu/atn/mass_storage/

Looking for Something Else?

If you have multimedia needs that are not addressed in this document, send email to mediasupport@unc.edu, or contact a campus service organization directly. See Campus Help below.

Campus Help

General Digitization Support

Academic Affairs Library

- Media Resources Center (<http://www.lib.unc.edu/house/mrc/index.php>)
- Collaboratories (<http://www.lib.unc.edu/house/index.php?display=collaboratories>)
- Photographic Services Section (<http://www.lib.unc.edu/ncc/copies.html>)

Phone: 962-1355

Email: kimv@email.unc.edu

Center for Teaching and Learning (instructors only)

Teaching Resource Lab (<http://ctl.unc.edu/csssml.html>)

Phone: 966-1289

Email: ctl_unc@unc.edu

Center for Instructional Technology (instructors only)

- Audio/Video Services (<http://www.unc.edu/cit/vidserv/index.html>)
- Other services (<http://www.unc.edu/cit>)

Phone: 962-6042

Email: cit@unc.edu

College of Arts and Sciences (College-affiliated faculty/instructors only)

OASIS Media Services (http://oasis.unc.edu/services/media_services.html)

Phone: 843-2205

Email: diglib_media@unc.edu

Health Sciences Library

Media Kitchen (<http://www.hsl.unc.edu/mk/MKslices.htm>)

Phone: 962-0800

IT Response Center

<http://help.unc.edu>

Phone: 962-HELP

Other Campus Resources

The following campus services, projects and repositories may also serve as valuable references for your digitization goals:

- Ancient World Mapping Center
<http://www.unc.edu/depts/awmc/>
- Documenting the American South
<http://docsouth.unc.edu/index.html>
- Ibiblio
<http://ibiblio.org/>
- Medical Illustrations and Photography
<http://www.med.unc.edu/wrkunits/4serv/medill/>

- Odum Institute for Research in Social Science
<http://www2.irss.unc.edu/irss/home.asp>
- OpenVideo Archive Project
<http://www.open-video.org>
- University Design Services
962-7123