MINITEX/LDS JOINT STANDARDS REVIEW TASK FORCE
Standards and Guidelines for Automated Library Systems

GUIDE TO DIGITAL PROJECTS

This guide has been prepared by the Digitization Working Group of the MINITEX/LDS Joint Standards Review Task Force Working Group in collaboration with the Minnesota Digital Library (MDL) with the goal of providing guidance to MDL participants and to all Minnesota cultural heritage institutions that undertake digitization projects.

The Standards Task Force thanks the members of its Digitization Working Group for their participation in the preparation of this document. The Working Group includes: Michael Kathman, College of St. Benedict/St. John's University, Chair; M. Keith Ewing, St. Cloud State University; Kathleen Flynn, St. Paul Public Library; Jason Roy, Minnesota State Historical Society; Daardi Sizemore, Minnesota State University, Mankato; and Carla Dewey Urban, MINITEX.

I. INTRODUCTION

This document is meant to be a companion and a guide to:

- *Western States Dublin Core Metadata Best Practices*
- *Western States Imaging Best Practices,*

which have been adopted by the Minnesota Digital Library, the MINITEX Advisory Committee, and the Advisory Council to the Minnesota state library agency, Libraries and School Technology (a unit of the MN Dept. of Education). The documents are available through the website of the MINITEX/LDS Joint Standards Review Task Force at:

http://www.minitex.umn.edu/committees/standards/documents/digitization.asp

Because of the complexity of digitization projects, we are providing:

- a digitization FAQ
- information about descriptive, structural, and administrative metadata
- a glossary of digital terms
- a listing of relevant standards

Included are a large number of URLs to locations that we think will be of assistance to you. In addition to the sites in the body of the document, the Colorado Digitization Program (http://www.cdpheritage.org/resource/introduction/questions.html) has an excellent site that will be of value to anyone considering a large digital project.

Many Minnesota institutions have ongoing digitization projects; the University of Minnesota, Twin Cities; the Minnesota Historical Society; the Hill Monastic Manuscript Library at St. John's University; and other institutions large and small are engaged in digitization of resources of many kinds. On a statewide level, the Minnesota Digital Library (MDL) Coalition, which is made up of representatives from content providers including Minnesota libraries, archives, and museums, is working to create the Minnesota Digital Library (http://www.mndigital.org). Part of the goal of the MDL Coalition's first major project is to provide a server and database environment to serve as a foundation for future activity. The Coalition proposes to work with local, regional, and statewide collections to identify and digitize unique and special resources to support education, scholarship, and entertainment through Internet-access by the citizens of Minnesota — and for a global audience. "Minnesota Reflections," funded by a Minnesota LSTA grant, is the initial statewide digitization effort sponsored by the MDL Coalition.
As organizations begin to consider digital projects, it is important to think through the whole project — especially as to how the images will be used in the future. There have been many digital projects started in the past that either have not been completed or are comprised of a large number of images that are difficult or impossible to find. In addition, very little thought has been given to how to archive

both the images and the metadata. A careful review of the two Western States Best Practices documents should help those responsible for new digital projects to avoid these problems in the future.

Because there is not a long history of capturing and maintaining digital images, this is an area that is in flux. This document is intended to point project directors to the most current information available. By using URLs to organizations that are major players in this area, the information they provide will constantly be updated as more is know about digitization.

II. DIGITIZATION: QUESTIONS & ANSWERS

Question: What are the standards I need to know?

Although the MINITEX/LDS Joint Standards Review Task Force has endorsed both the Western States Dublin Core Metadata and Imaging Best Practices, it is important to know the other types of standards that are being implemented at other institutions. Standards for metadata and imaging, be they local, regional or national, allow institutions to share this information about their collections in a more centralized way. The emergence of the Internet as a vehicle to distribute this information quickly and efficiently has made the need for standards all the more important. It is quite costly for many smaller institutions to go it alone. By harvesting this data and placing it in a shared environment, we can distribute the cost of scanning and managing the images, increase their visibility and elevate their research value. A few of the more prominent standards are identified below, there are many more (see "V. Digitization: Basic Standards" on p. 17). And, of course, many more will emerge in the next few years.

Dublin Core
http://www.dublincore.org/

Visual Resources Association (VRA) Core
http://www.vraweb.org/vracore3.htm

Open Archive Initiative (OAI)
http://www.openarchives.org/

Metadata Encoded Transmission Standard (METS)
http://www.loc.gov/standards/mets/

Encoded Archival Description (EAD)
http://lcweb.loc.gov/ead/

Question: What types of equipment do I need to consider?

Beyond choosing what items to scan, individuals and organizations must also determine how best to build the infrastructure necessary to successfully complete the project. Hardware/Software components as well as other aspects involved in building a basic and quite serviceable digital imaging lab are listed below.

Scanner
The most obvious piece of equipment in any digital imaging lab, the scanner, is also the most important. That is why it is imperative that individuals and organizations choose the correct scanner for the job. Flatbed, drum and film scanners are all optimized to work best for certain types of materials. In general, however, a good flatbed scanner would provide the most flexibility for scanning collection materials. Since most institutions involved in digitization projects are scanning hundreds, even thousands of images, it is vital that an investment be made on a production level scanner. Contact local vendors and other institutions to find out about models that may be right for your project. Product reviews found on the Internet can also be a good starting point.

Computer
The hub of any scanning station is the computer. Because we are dealing with files that are very large in size and can be quite taxing on computers, purchasing one with adequate RAM and disk space is strongly encouraged. Remember that image-processing speed will have a direct effect on your workflow and your ability to produce scans in a timely fashion. For those not in a networked environment, purchase enough computer disk space to store your files prior to copying to offsite media. As with scanners, contact vendors and other institutions for help in choosing the correct system.

Monitor
A good monitor is one of the best ways to ensure that the scanned image is a good facsimile of the original. A large-screen monitor that supports high-resolution display will aid in the image editing and quality control stages of your imaging project. The ability to calibrate and control monitor functions (e.g. contrast, brightness, color temperature) is also highly desirable.

Software
Be it freeware, shareware or commercial imaging software, choosing the right image editing tools is important. There are lots of products available to those undertaking a digitization project. Do a little research to determine what is right for your project and your budget. Common products include Adobe PhotoShop, Paintshop Pro, and ImageMagik.

Storage
Digital images take up a large amount of space on any computer or network. The most common way to store them is on offline media. This includes CD, DVD or other optical media. You will need to determine which option is best for your needs and purchase the necessary equipment (e.g. CD or DVD recorder). Remember to make multiple copies or "backups" of all you images.

Work Environment
If at all possible, a separate workspace should be allocated for the scanning station. Since scanning is such a specialized task, often involving the use of collection materials, it is a good idea to reserve enough space to adequately meet the needs of scanning staff.

Training
Buying the right tools is not enough. Training staff on how to use these tools is often forgotten. Often staff is forced to learn "on-the-fly." This approach can cause a serious lag in project workflow. A better method would be to plan in advance which staff will be assigned to which task and to ensure that

they are trained on the equipment and software prior to beginning the project. A number of companies and organizations offer computer classes and training throughout the state. The Minnesota Digital Library Coalition has begun to offer a number of workshops and more will be scheduled in the future. These will include workshops on planning digitization projects, best practices for metadata, and best practices for imaging; workshops are announced through the MINITEX and other statewide listservs. (Also, a useful, currently available online resource is the Cornell Digital Imaging Tutorial: Moving Theory into Practice, http://www.library.cornell.edu/preservation/tutorial/contents.html)

Question: How do you sustain your images over time?

Sustaining your images over the long term should be a goal of any digital imaging project. Just how you go about it should be discussed as early as the initial planning meeting. Here are some things to think about when planning to manage your images over time.

Metadata
Long-term sustainability begins with quality metadata. If you don't know what you have, there is no way to know what you will need to do with it down the road. Luckily, by following an accepted standard for metadata cataloging institutions should have the necessary information to manage the data long-term.

Backup
Never keep only a single instance of any digital image. It is always important to maintain at least two copies of every digital image created. Should anything happen to the original, you will still have another copy available. Backing up your scans is considerably cheaper than having to rescan the objects.

Refreshment
From time to time it will be necessary to check the integrity of your offsite storage media. The integrity of CDs, DVDs and other optical media over time is still somewhat of a question mark. For this reason as well as others (e.g. disk error resulting from poor climate controls or storage environment), it is necessary to copy your information from older disks to newer versions. This regeneration acts as a further quality control and ensures that the images you think are there actually are.

Migration
At some point in time, all of your files will need to be transferred to a new file format. Technological changes, standards and protocols that you follow today may become obsolete in the future. By following the steps laid out above, however, you should be in a good position to migrate your information. Be warned, though, that migration should be discussed during the project-planning phase. The worst-case scenario would be to find out that at the point of migration that you lack the funds to transfer the data. You would be left obsolete media unreadable by current equipment (e.g. laserdiscs and 8-tracks).

III. WHAT METADATA SHOULD I INCLUDE?

(Keeping in mind that *Western States Dublin Core Metadata Best Practices* have been identified as guidelines documents by the MINITEX/LDS Joint Standards Review Task Force, the references that follow refer to sections of that document. Since *Western States Dublin Core Metadata Best Practices* is an evolving document and pagination may change, we have not included specific page references.)

Metadata is the "cataloging piece" of a digitization project. The purpose of metadata and the benefits of the following guidelines and best practices in applying metadata are to create standardized records that:

- Enhance online search and retrieval accuracy in local and shared databases
- Improve resource discovery capabilities
- Improve quality control of metadata records
- Facilitate inter-institutional interoperability

For a definition of metadata and explanation of its purpose, see **"What is Metadata?"**

The Western States Digital Standards Group has chosen Dublin Core as the basis for its recommended metadata scheme with a few added elements and refinements. For a description of the Dublin Core standard, recommended additions, and general input guidelines (initial articles, punctuation, capitalization, etc.), see **"What is Dublin Core and why use it?"**

There are three categories of metadata: Descriptive, Structural, and Administrative. Descriptions and guidelines for each metadata element can be found in **"Dublin Core Metadata Element Set."**

Descriptive metadata
Information used for the indexing, discovery, and identification of a digital resource

Mandatory elements:
- Title
- Creator (if available)
- Subject (using controlled vocabulary from established thesauri or discipline-related word lists)
- Description
- Date.Original (if applicable)
- Date.Digital
- Identifier

Optional elements:
- Publisher
- Contributor
- Source
- Language
- Relation
- Coverage

Structural metadata
Information used to display and navigate digital resources; also includes information on internal organization of the digital resource. Structural metadata might include information such as the structural divisions of a resource (i.e. chapters in a book) or sub-object relationships (such as individual diary entries in a diary section)

Mandatory elements:
Format.Use
Optional elements:
Type
Relation

Administrative metadata
Represents the management information for the digital object, which may include information needed to access and display the resource, as well as rights management information. Administrative metadata might include the resolution at which the images were scanned, the hardware/software used to produce the image, compression information, pixel dimensions, etc.

Mandatory elements:
Rights (if available)
Format.Creation
Date.Digital
Holding.Institution (for collaborative projects)
Optional elements:
Publisher

An institution might have some proprietary or confidential information, such as provenance, location, or donor information that shouldn't be publicly accessible. Although such information should be tracked and retained locally, best practice is to eliminate proprietary or confidential data in a shared catalog.

IV. DIGITIZATION: DIGITAL IMAGING GLOSSARY OF TERMS

Access Image (see also Derivative Image)
A term used to denote low-resolution images that are often intended for distribution via the Internet.

Algorithm
The specific process in a computer program used to solve a particular problem.

Archival Image
An image meant to have lasting utility. Archival images are usually kept in a secure environment. Archival images are of a higher resolution and quality than the digital image delivered to the user online. The file format most often associated with archival images is TIFF, or Tagged Image File Format.

Archival Scans
Digital images serve as surrogates of the original. At this point in time, there is no such thing as an Archival or Preservation scan that acts as an exact replica or replacement of the original, as it is not yet possible to record every piece of information found in the original with today's scanner technology.

Artifacts
Visual digital effects introduced into an image during scanning that do not correspond to the original image being scanned. Artifacts might include pixellation, dotted or straight lines, regularly repeated patterns, moiré, etc.

Bandwidth
In digital systems bandwidth is a measure of data speed in bits per second. A higher bandwidth network is required for fast transfer of image files as they typically contain large amounts of data.

Batch Process
A method of processing digital images in large numbers such that minimal staff time is required. This becomes important when performing the same task on multiple image files.

Bit Depth
The bit depth of an image refers to the number of bits used to describe the color of each pixel. Greater bit depth allows more colors to be used in the color palette for the image.

> **1-bit color** is the lowest number of colors per pixel in which a graphics file can be stored. In 1-bit color, each pixel is either black or white.
>
> **8-bit color/grayscale** has eight bits assigned to each pixel, providing 256 colors or shades of gray.
>
> **24-bit color** has 24 bits assigned to each pixel, representing 16.7 million Colors. 8-bits are assigned to each of the red, green, and blue components of a pixel.

Calibration
The act of adjusting the color of one device relative to another, such as a monitor to a printer or scanner. Or, it may be the process of adjusting the color of one device to some established standard.

Capture Device
These include flatbed scanners, drum scanners, film scanners, and digital cameras. They use electronic devices to capture images rather than photographic film.

CCD Array
Charge-Coupled Device array. Light sensitive diodes used in scanners and digital cameras that sweep across an image during capture and, when exposed to light, generate a series of digital signals that are converted into pixel values.

Checksum
A value that is computed based on the contents of a set of data. A unique checksum value is generated for every electronic file. Any change in the file would result in an alteration of the checksum value. Therefore, the checksum can be used to detect if the data has been altered during transmission or when being stored and retrieved. Examples of checksums include MD5.

CMYK (Cyan, Magenta, Yellow, Black)
One of several color encoding systems used by printers for combining primary colors to produce a full-color image. In CMYK, colors are expressed by the "subtractive primaries" (cyan, magenta, yellow) and black.

Color Correction
The process of altering colors as they appear in a digital image or in print to ensure they accurately represent the work depicted.

Compression
The reduction of image file size for processing, storage, and transmission. The quality of the image may be affected by the compression techniques used and the level of compression applied. There are two types of compression:

> **Lossless** compression is a process that reduces the storage space needed for an image file without loss of data. If an image has undergone lossless compression, it will be identical to the image before it was compressed.
>
> **Lossy** compression is another process that reduces the storage space needed for an image file, but it discards information (often information that is "redundant" and not perceptible to the human eye).

There are both standard and non-standard compression techniques available. In general, it is better to employ a compression technique that is supported by standards, non-proprietary, and maintained over time. In selecting a compression technique, it is necessary to consider the attributes of the original. Some compression techniques are designed to compress text; others are designed to compress pictures.

Continuous Tone
Smoothly varying gradation of tones. Includes photographs, postcards and some original artwork, as well as graphic art that is produced with continuous tone-like attributes.

Controlled Vocabulary
A collection of terms compiled with control over form, format, inclusion and exclusion of terms. Catalogers would be required to only select terms from this predefined list. Use of a controlled vocabulary helps to group similar items in a way that allows for better sharing and retrieval of information.

Derivative Image
An image that has been created from another image through some kind of automated process, usually involving a loss of information. Techniques used to create derivative images include sampling to a

lower resolution, using lossy compression techniques, or altering an image-using image processing techniques.

Digital Camera
A camera that does not contain any film but records the image as a digital object. The digital camera allows for the scanning of a range of differently sized materials. Resolution of digital cameras is usually fixed and is expressed as a pixel ratio.

Digital Image
An electronic photograph scanned from an original document, made up of a set of picture elements ("pixels"). Each pixel is assigned a tonal value (black, white, a shade of gray, or color) and is represented digitally in binary code (zeros and ones). The term "image" does not imply solely visual materials as source material; rather, a digital image is simply a representation of whatever is being scanned, whether it be manuscripts, text, photographs, maps, drawings, blueprints, halftones, musical scores, 3-D objects, etc.

Digitization
The process of converting analog information into digital format.

Dots per inch (dpi) (See also pixels per inch)
A measurement of the scanning resolution of an image or the quality of an output device. Can also be expressed as pixels per inch.

Dublin Core
A metadata element set intended to facilitate discovery of electronic resources. Two features of the Dublin Core metadata element set are simplicity and extensibility. Dublin Core is intended to be usable by both non-catalogers and specialists alike, to provide an economic alternative to more elaborate descriptive models such as full MARC cataloging, and be flexible enough to encode the structure and semantics inherent in rich descriptive standards. The Dublin Core seeks to promote a commonly understood set of descriptors to help facilitate interoperability across disciplines. The Dublin Core can be mapped to the MARC record and a variety of output structures can be generated. More information is available at the Dublin Core Initiative web site (http://dublincore.org/).

Encoded Archival Description (EAD)
An SGML based encoding standard developed by the Society of American Archivists for you in archival finding aids. More information is available at the Encoded Archival Description web site (http://www.loc.gov/ead/)
.

Export
The process of transporting data form one computer, program, type of file format, or device to another.

File Format
A type of data file. Some common image file formats include TIFF, JPEG, and TIFF.

File Size
The file size of an image is proportional to its resolution. The higher the resolution, the larger the file size. File size is different from image size.

File Transfer Protocol (FTP)
The universal format for transferring files over the Internet.

Flatbed Scanner
An image capture device much like a photocopier. The object to be scanned is placed face-down on a glass plate. The CCD array passes beneath the glass.

Graphic Image File Format (GIF)
A widely supported image storage format promoted by CompuServe for use on the web.

Grayscale
A range of shades of gray in an image. Grayscales of scanners are determined by the number of grays, or values between black and white, that they can recognize and reproduce.

Hypertext Markup Language (HTML)
An encoding format for linking and identifying electronic documents and used to deliver information on the World Wide Web.

Image Capture
Using a scanner or other device to create a digital representation or electronic photograph of an image.

Image Processing
The alteration or manipulation of images that have been scanned or captured by a digital recording device. Can be used to modify or improve the image by changing its size, color, contrast, and brightness. This is done through the use of image editing software.

Image size
Describes the actual physical dimensions of an image, not the size it appears on a given display device.

Interpolated/Uninterpolated - Most scanners have a maximum pixel-per-inch resolution before they start guessing or interpolating the data. Interpolated files require the computer to simulate data in an image file, while uninterpolated files hold only data that is accurate to the original. Uninterpolated resolution is, therefore, preferred for accurate scanning.

Joint Photographic Experts Group (JPEG)
A compression algorithm for compressing the size of image files. JPEGs are helpful in allowing access to full screen image files online because they require less storage and are therefore quicker to download into a web page.

Kilobyte (KB)
An amount of computer memory consisting of about one thousand bytes (1024 bytes).

Lempel-Ziv-Welch (LZW)
A common lossless data-compression algorithm often associated with the TIFF file format.

Machine Readable Cataloging (MARC)
"The MARC formats are standards for the representation and communication of bibliographic and related information in machine readable form." The MARC formats contain an explicit set of rules for the structure of fields and the content values within those fields. More information about MARC is available at the MARC Standards web site (http://lcweb.loc.gov/marc/).

Metadata
Data about data, or information known about the image in order to provide access to the image.

Metadata Encoding and Transmission Standard (METS)
The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. More information is available at the Metadata Encoding and Transmission Standard web site (http://www.loc.gov/standards/mets/).

Metadata Harvester
Operated by a service provider, a metadata harvester is a client application that is used to collect metadata from repositories.

Megabyte (MB)
An amount of computer memory consisting of about one million bytes (1,048,576 bytes).

Migration
Preserving the integrity of digital images by transferring them across hardware and software configurations and across subsequent generations of computer technology.

Noise
Data or unidentifiable marks picked up in the course of scanning or data transfer that do not correspond to the original.

Open Archives Initiative (OAI)
The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. More information is available at the Open Archives Initiative web site (http://www.openarchives.org/).

Permanence - In the digital world, permanence is a complicated term that requires not only the accurate refreshing of files, but the continued updating of metadata.

Pixel
Pixel is short for picture elements, which make up an image, similar to grains in a photograph.

Pixel per inch (ppi) (See also dots per inch)

A measurement of the scanning resolution of an image or the quality of an output device. PPI expresses the number of used to capture the digital image. Can also be expressed as pixel per inch.

Portable Network Graphics (PNG)
A new image file format that has been approved by the World Wide Web consortium. PNG is non-proprietary.

Preservation
(As it relates to scanning) Digitizing an original photograph, document, or three-dimensional object is only a method of preservation if the digital file becomes the access tool and the original is no longer available for use. Although high resolution scanning (i.e., scan at the highest resolution possible appropriate to the type of media you are scanning) is recommended for all materials in order to achieve the highest quality possible and to ensure that information held in the original is not lost in the scan. However, the digital file, as of yet, should not serve as a replacement of the original for preservation purposes.

Production-level Scanner
Any scanner capable of handling the high throughput of images typically associated with digital imaging projects. In general, these scanners are more expensive than "home-use scanners" and often perform better with regard to speed, quality of scan and durability.

Project Management
Digitization projects require substantial planning. A project manager is responsible for coordinating components of the project including goals, budget, and the project team. The project manager should be familiar with all aspects of a digitization project and will be able to either make decisions about technical and procedural questions that arise or will be able to identify the person responsible for making those decisions. Without a project manager, there is no mechanism for making decisions that fit into the overall goals of the project and the budget.

Quality Control
Techniques ensuring that high quality is maintained through various stages of a process. For example, quality control during image capture might include comparing the scanned image to the original and, then, adjusting colors or tonal values, or checking for artifacts. Quality control is vital for a project's success. Both expertise about technical digitization issues and knowledge about collection materials being digitized are important, and project should seek to build adequate quality checks into project workflow.

Random Access Memory (RAM)
The most common type of computer memory. This is where the CPU stores software, programs and data currently being used. A large amount of RAM usually offers faster manipulation and processing of image files.

Refreshment
The transfer of digital files to a new media on a regular basis. This is the most important part of an institution's long-term commitment to digitization. In order to ensure long-term access to the data, it must be transferred to the most recent and stable type of media storage.

Resolution
The number of pixels (in both height and width) making up an image. The more pixels in an image, the higher its resolution. The higher the resolution of an image, the greater its clarity and definition (and the larger the file size). Resolution can also refer to the output device, such as a computer monitor or printer, used to display the image. Image file resolution is often expressed as a ratio (such as 640x480 pixels), as is monitor resolution; however, resolution is also expressed in terms of dots per inch (dpi) or pixels per inch (ppi).

RGB
Short for Red, Green, and Blue. This is an "additive" method used to simulate natural color on computer monitors and television sets.

Rights Management
Information relating to the ownership and use of an object or resource. Rights may include, copyright, intellectual, and property. Often this takes the form of a brief statement or URL to an institution's broader rights statement.

Scanner
A device for capturing a digital image. There are many types of scanners, such as flatbed scanners, drum scanners, slide scanners, and microfilm scanners.

Spatial Resolution - The number of pixels held in a file, for example 640 pixels across by 480 pixels down (640x480).

Standard Generalized Markup Language (SGML)
An international standard for the definition of device-independent, system-independent methods of representing texts in electronic form. SGML emphasizes descriptive rather than procedural markup. SGML is a system for describing structural divisions in a text (title page, chapter, scene, stanza), typographical elements (changing typefaces), and other textual features (grammatical structure, etc.). More information is available at the SGML Standards Generalized Markup Language web site (http://sunsite.berkeley.edu/SGML).

Tagged Image/Interchange File Format (TIFF)
A file storage format implemented on a wide variety of computer systems, usually used for archival scans.

Thumbnail Image (see also Access Image)
Small, low-resolution images often hyper linked to a larger version of the same image. Users can often view multiple thumbnail images on a single screen.

Uniform Resource Locator (URL)
A standard addressing scheme used to locate or reference files on the Internet. Used in World Wide Web documents to locate files. A URL gives the type of resource being used and the path to the file. The syntax used is: scheme://host.domain/path filename

VRA Core Categories
A metadata element set established by the Visual Resources Association. The *VRA Core Categories, Version 3.0,* consists of a single element set that can be applied as many times as necessary to create records to describe *works* of visual culture as well as the *images* that document them. More information is available at the VRA Core Categories, Version 3.0 website (http://www.vraweb.org/vracore3.htm).

Watermarks
Bits altered within an image to create a pattern, which indicates proof of ownership. Unauthorized use of a watermarked image can then be traced.

World Wide Web (WWW)
An interconnected network of electronic hypermedia documents available on the Internet. WWW documents are marked up in HTML. Cross references or hyperlinks between documents are recorded in the form of URLs.

eXtensible Markup Language (XML)
XML is a meta-language (a way to define tag sets) that allows you to design your own customized markup language for many classes of documents. XML is designed for easy and straightforward use on the web, ease of use in authoring and managing documents, and ease of transmission and sharing of electronic documents across the web. XML is intended to deliver information, not just pages.

Z39.50
A protocol that is an international standard for information retrieval. More information is available at the official Z39.50 maintenance agency web site (http://www.loc.gov/z3950/agency/).

Glossary terms taken from the following sources:

Besser, Howard and Jennifer Trant, *Introduction to Imaging: Issues in Constructing an Image Database.* The Getty Art History Information Program Imaging Initiative, 1995.
http://www.getty.edu/research/conducting_research/standards/introimages/

Colorado Digitization Program. "Digitization Glossary."
http://www.cdpheritage.org/resource/introduction/rsrc_glossary.html

Cornell University Library. Research Department. "Moving Theory Into Practice: Digital Imaging Tutorial," 2000-2003.
http://www.library.cornell.edu/preservation/tutorial/contents.html

Kodak. "Digital Learning Center Glossary of Terms."
http://www.kodak.com/US/en/digital/dlc/book4/chapter2/index.shtml

Library of Congress. Metadata Encoding and Transmission Standard: Official Web Site.
http://www.loc.gov/standards/mets/

Nebraska State Historical Society. Digital Imaging Laboratory. "Glossary of Digital Imaging Terms." http://www.nebraskahistory.org/lib-arch/research/photos/digital/glossary.htm

Open Archives Initiative. The Open Archives Initiative Protocol for Metadata Harvesting, Protocol Version 2.0 of 2002-06-14. http://www.openarchives.org/OAI/openarchivesprotocol.html#DefinitionsConcepts

Technical Advisory Service for Images (TASI). "Glossary of Technical Terms." http://www.tasi.ac.uk/glossary/glossary_technical.html

V. DIGITIZATION: BASIC RESOURCES

In addition to the *Western States Dublin Core Metadata Best Practices* and the *Western States Digital Imaging Best Practices*, the following standards and documents about project management and best practices can serve as important guides:

Best Practices - Imaging

Besser, Howard and Jennifer Trant, *Introduction to Imaging: Issues in Constructing an Image Database.* The Getty Art History Information Program Imaging Initiative, 1995. http://www.getty.edu/research/conducting_research/standards/introimages/

Cornell University Library. Research Department. "Moving Theory Into Practice: Digital Imaging Tutorial," 2000-2003. Available online at: http://www.library.cornell.edu/preservation/tutorial/contents.html

Digital Library Federation. *Guides to Quality in Visual Resource Imaging*. July 2000. Available online at http://www.rlg.org/visguides/. Includes the following articles:

1. Planning an Imaging Project, by Linda Serenson Colet, Museum of Modern Art
2. Selecting a Scanner, by Don Williams, Eastman Kodak Company
3. Imaging Systems: the Range of Factors Affecting Image Quality, by Donald D'Amato, Mitretek Systems
4. Measuring Quality of Digital Masters, by Franziska Frey, Image Permanence Institute Rochester Institute of Technology
5. File Formats for Digital Masters, by Franziska Frey

Kodak. *Digital Learning Center*. Available online at: http://www.kodak.com/US/en/digital/dlc/index.jhtml. A collection of information, tips, and techniques about digital imaging.

Technical Advisory Service for Images (TASI). 2002-2004. *An Introduction to Making Digital Image Archives*. Available online at: http://www.tasi.ac.uk/advice/overview.html. Includes the following units:

1. Introduction

2. Managing Digitisation Projects
3. Creating Digital Images
4. Delivering Digital Images
5. Using Digital Images

Best Practices – Metadata

Baca, Murtha and Patricia Harpring, eds. Art Information Task Force, J. Paul Getty Trust. *Categories for the Description of Works of Art*. Available online at http://www.getty.edu/research/conducting_research/standards/cdwa/index.html.

CIMI Consortium for the Computer Interchange of Museum Information. *Guide to Best Practice: Dublin Core*. Version 1.1, April 2000. Available online at http://www.cimi.org/old_site/documents/meta_bestprac_v1_1_210400.pdf

Digital Library Federation and the Library of Congress. *Metadata Encoding & Transmission Standard (METS)*. Available online at http://www.loc.gov/standards/mets/. The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language.

Digital Library Federation. *TEI Text Encoding in Libraries: Guidelines for Best Encoding Practices*. Version 1.0. Available online at http://www.diglib.org/standards/tei.htm.

Library of Congress MARC Standards Office. *Metadata Object Description Schema (MODS)*. Available online at http://www.loc.gov/standards/mods/.

Research Library Group. *RLG REACH Element Set for Shared Description of Museum Objects*. Available online at http://www.rlg.org/reach.elements.html

Society of American Archivists and the Library of Congress. *Encoded Archival Description (EAD)*. Available online at http://lcweb.loc.gov/ead/.

Visual Resources Association Data Standards Committee. *VRA Core Categories*. Version 3.0. Available online at http://www.vraweb.org/vracore3.htm

Project Management

Chen, Ching-Chih and Alberto Del Bimbo, eds. *Report of the DELOS-NSF Working Group Digital Imagery for Significant Cultural and Historical Materials.* Available online at http://delos-noe.iei.pi.cnr.it/activities/internationalforum/Joint-WGs/digitalimaging/DigitalImaging.pdf

Cornell University. *Digital Preservation Management: Implementing Short-term Strategies for Long-term Problems*. Available online at http://www.library.cornell.edu/iris/tutorial/dpm/. A useful tutorial and links to additional resources.

Fernie, Kate, and Julian D. Richards, eds. *Creating and Using Virtual Reality: a Guide for the Arts and Humanitie*s. Available online at http://vads.ahds.ac.uk/guides/vr_guide/index.html

Gillings, Mark, and Alicia Wise, eds. *GIS Guide to Good Practice*. Available online at http://ads.ahds.ac.uk/project/goodguides/gis/.

Harvard Law School Library. *Guidelines for Digital Reformatting Projects*. Available online at http://www.law.harvard.edu/library/digital/guidelines.htm. Good outline of questions and decisions. Includes a good glossary.

Institute of Museum and Library Services. *A Framework of Guidance for Building Good Digital Collections*. Available online at http://www.imls.gov/pubs/forumframework.htm.

National Initiative for a Networked Cultural Heritage. *The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials*. Available online at http://www.nyu.edu/its/humanities/ninchguide/.

Sitts, Martha, ed. Northeast Document Conservation Center. *Handbook for Digital Projects: A Management Tool for Preservation and Access*. Available online at http://www.nedcc.org/digital/dighome.htm.

Visual Arts Data Service. *Creating Digital Resources for the Visual Arts: Standards and Good Practice*. Available online at http://vads.ahds.ac.uk/guides/creating_guide/contents.html.

Washington State Library. *Digital Best Practices: Planning a Digital Project*. Available online at: http://digitalwa.statelib.wa.gov/newsite/best.htm. Topics covered include:

1. Project Planning (cost, training, marketing, vendors)
2. Collections (preparation, rights, preservation)
3. Technology (hardware, software, scanning, delivery)
4. Funding (pursuing funding, resource management)