Metadata Uses for Networked Information Used with Digital Images at the Austin History Center

October 2003

Developed by:
Grace McEvoy
Margaret Schlankey
Sue Soy
Rachel Schoonover

Table of Contents

Metadata Uses for Networked Information Used with Digital Images at the Austin History Center

Introduction

This report offers some suggestions for use of metadata at the Austin History Center. Today's digital environments with networked information resources offer new ways to extend metadata that enables users to find materials and archivists to store information about digital objects in resource files.

Lagoze and Payette, (2000) offer four key questions applicable to the History Center to consider in applying metadata to the description of our materials; particularly to AHC images that will be indexed in online tools and may be used by many on the Internet:

> How will our customers locate our digital images?
> How will they use the images or collections of images?
> What policies do we need to protect both privacy and intellectual property rights and provide access controls for the images?
> How will our program and policies assure permanence?

Definition for metadata

Metadata has been in use for a long time, but within the past 15 years, it has become connected with networked information resources and resource discovery. Gilliland-Swetland (1998) provides us with a useful "big picture" way of defining and thinking about metadata describing it as "The sum total of what one can say about any information object at any level of aggregation." Information objects are anything that can be addressed and manipulated by a human or a system as a discrete entity. The object may be a single item or it may be an aggregate of many items. All objects can be described using their content, context, and structure. More precisely, it is described by Glogoff and Forger (2001)as "the indexing that is applied to electronic information."

Types of metadata

Common types of metadata include administrative, preservation, descriptive, technical, and use (Gilliland-Swetland, 2000). Within these broad categories many subdivisions can be created that can fully and accurately describe the object. For example within the technical metadata category, we can describe the grayscale properties of the image as well as the size, how the scanning was accomplished, what standards were used for quality, software used, scanners used, etc. Within descriptive metadata, we could describe the address, location, people, date, building name, and other keywords or subject headings that would help locate the image.

Metadata already in use at the History Center

MARC

The Austin History Center uses a number of metadata schemas including the traditional library-related MARC (Machine Readable Cataloging) record, card catalog files, and paper indices. This machine readable format is used in the Austin Public Library online catalog, OCLC, and in nearly all library settings.

Finding Aids

Finding aid development at the History Center is largely influenced by the instruction provided at the University of Texas at Austin and the Society of American Archivists' Technical Manual Series. Generally the finding aid includes a biographical or creator sketch, content and scope note, and a container list along with the MARC record. Beginning in 2001, finding aids and preliminary inventory work at the History Center has been formatted to facilitate EAD implementation.

Encoded Archival Description

Beginning in the summer of 2002, the Austin History Center embarked on a project to place finding aids for many of its collections on the World Wide Web. This project, the Texas Archival Resources Online (TARO), is administered by the Texas Digital Library Alliance (TDLA) and will facilitate descriptions of the rich archival, manuscript, and museum collections in repositories across the state available to the public, along with digitized images of selected holdings. Working with the University of Texas at Austin, a TDLA member, the Austin History Center is selecting finding aids for conversion to encoded archival description using XML mark-up language. When this project is complete, researchers will be able to access information contained in these archival collection finding aids from their home or office.

Dublin Core Metadata Use at the Collection Level

With the digital image project funded by the Texas State Library and Archives Commission (Austin Treasures, August, 2001) Dublin Core and MARC records were produced so that the metadata are available for each of the ten exhibits in the project, but not each of the four hundred individual images used within the exhibits. The exhibits are listed in the online public access catalog and in the Web-based Austin Public Library catalog, a hypertext link exists to link through to the MARC record. An example is shown at "Green Growth" in the APL Web-based catalog. Click through to the web presentation http://www.cityofaustin.org/library/ahc/begin/.

Proposed Metadata Schemes and Visual Resource Database Design Considerations

Beginning with 2003, planning began to standardize procedures and processes for creating item level metadata for each digital image scanned. Each staff member involved in the scanning process is to provide metadata according to the standards and procedures produced in the planning process. This metadata is to be placed on the server space set aside for the AHC digital image collection. Three categories of metadata are proposed for the Austin History Center.

Resource Discovery Metadata Category

The Resource Discovery metadata is used by search engines and makes information about the digital image searchable so users will "discover" the image itself. The History Center uses Dublin Core metadata elements as a basis for resource discovery data element construction.

Dublin Core originated with the Dublin Core Metadata Initiative (DCMI), an open forum constituted to develop useful interoperable online metadata standards. It is a simple element set that allows the curator, author, archivist, etc. to use 1 to 15 elements and to repeat the elements if necessary for the individual object being described. Each element or field is optional and repeatable. The data elements appear the same to every user, can be used to complement more complex (local) systems for maintenance purposes, represent a 1:1 relation between a metadata set and a resource, and allow management of versions, copies, transformations, etc. through the DC Relation and DC Source elements.

Dublin Core can be researched at the Web Site for the Dublin Core Metadata Initiative at http://dublincore.org. The fifteen data elements are described in detail there as well as at the OCLC Dublin Core Web Site (http://purl.oclc.org/dc/documents/rec-dces-19990702.htm).

Dublin Core has the advantage of enabling information to be placed in code designed to facilitate resource discovery on the Web. It is also compatible with XML, or Extensible Markup Language. When embedded in HTML header documentation Dublin Core looks like:

```
<HEAD>
...
<META NAME="DC.Title" CONTENT="Syntax/encoding of Dublin Core">
<META NAME="DC.Creator" CONTENT="Simon Cox">
<META NAME="DC.Date" CONTENT="1997-10-29">
<META NAME="DC.Rights" CONTENT="http://www.agcrc.csiro.au/general/copyright.html">
<META NAME="DC.Subject" CONTENT="metadata, indexing, world wide web;>
<META NAME="DC.Type" CONTENT="web page">
<META NAME="DC.Identifier"
CONTENT="http://www.agcrc.csiro.au/projects/3018CO/talks/dc5/dcsyntax.html">
<META NAME="DC.Format" CONTENT="html">
<META NAME="DC.Language" CONTENT="eng">
...
</HEAD>
```

The Austin History Center, as part of the City of Austin, is not able at this time to embed resource discovery metadata in its Web site. The City of Austin's information systems office does not permit any additional information other than the City's header. Because of this limitation, the History Center will enter its Resource Discovery Metadata into a Microsoft Access database referred to as the Visual Resource Database, until we are able to make use of the information on our Web site.

Technical [Administrative] Metadata Category

Another role for metadata is to help manage the storage of digital images using file storage system directories. All of the structural information about the images can be stored in a relational database (like Microsoft Access ™). This type of metadata might contain information about characteristics, technical attributes, and features of the images such as file format, spatial resolution, and compression. It might contain information about image production or reformatting, the make and model of scanning equipment used, and image identification data such as linkage to other metadata (MARC records).

These data help support preservation strategies even before the strategies have been put into practice, but they do not cover the important issues of authenticity, provenance, or intellectual property rights. It is intended that this metadata will exist separate from the image itself, but will reside in the Visual Resource Database collecting metadata for all images in the Austin History Center collection. It is also assumed that as the collection ages and grows at the History Center, changes will occur and these changes will be documented for each of the images using additional metadata.

The precedent for this metadata was set with the TexTreasures Grant funded Austin Treasures Project work (August, 2001). Standards based on the NISO, Draft Data Dictionary: Technical Metadata for Images, Version 1.0, July 5, 2000 edition, were implemented by Diva Imaging, the vendor selected to scan the images used in the project. Diva Imaging, a local company, created a separate database of technical metadata for each of the digital image files they produced. The technical metadata chosen for the Visual Resource Database is based on the work that Diva Imaging accomplished.

Use Metadata Category

Since many of the images digitized for internal and external customers are used in publications, exhibits and displays, it is helpful to track instances of use for later reference by staff. Some examples of the uses to be recorded include: Austin History Center Association Newsletter, Austin American-Statesman newspaper, in-house exhibit, and Parks and Recreation Department Exhibit.

Balancing Data Entry with Staffing Requirements

At some point in the future, more automated systems will be available for metadata creation, but at this point recording metadata requires human intervention and customarily, a limited amount of time is available for completion of the work. The AHC's initial data dictionary considers the future needs of the History Center customer, but is also balanced with realism concerning the amount of staff and volunteer time available for data entry. Form design and database design are intended to be self explanatory and user-friendly.

Before entering data into the database, three critical components identified by Moen (2000) are defined and applied specifically to fit the Austin History Center.

♦ Elements & Semantics—definitions for the content of the elements selected for use. These elements are divided into three data dictionary sections at the AHC which correspond to the categories of metadata used at the AHC: Resource Discovery, Technical/Administrative, and Use. For example, the Dublin Core metadata element TITLE, part of the Resource Discovery metadata, is used at the AHC. Its label is Title. The semantics of this title is "The name

given to the resource by the Creator or Publisher." From the Technical Metadata category, the AHC uses the metadata element COMPRESSION. Its label is Compression. The definition of this element is "Designates the compression scheme used to store the image data." In the Use Metadata category, the AHC uses the element SCANNING REQUESTED BY. The instructions for choosing a value for this element are to "select the name of the AHC unit or customer that initiated the request [for the digitized image]."

♦ Content Rules—guidelines and directions for inputting the content. These are called procedures for data entry at the AHC. For example, content entered into the "Creator" field must indicate the name of the person or organization or service primarily responsible for creating the intellectual content of the resource. This information differs from a later field, "Person Scanning," that requires the name of the person responsible for creating the digital image. The "Creator" field is not a required field, but the "Person Scanning" field is required.

♦ Syntax—rules for structuring and expressing the elements, abbreviations allowed in the DRA catalog using MARC or the Web using HTML or XML in the case of EAD finding aids. These rules are included in the procedures.

In addition, as the database is constructed a series of basic questions is posed to guide the user to create appropriate database entries.

> Can there be only one entry in this field?
> (single=Yes, not single=No)
> Does there have to be some data in this field?
> (required=Yes, not required=No)
> What type of data is entered in this field?
> (text, memo, date, logical, object or link, number, or auto number)
> Can this data appear in the same field for more than one record?
> (not unique=Yes, unique=No

How to Create and Input In-House Metadata

In order to facilitate consistent and complete input for each image scanned from the collection, a form is filled out as completely as possible by the individual doing the scanning. As the form is completed, data recorded on it will become the metadata used for (a) resource discovery, (b) collecting technical information, and (c) recording intended use of the images. These metadata categories will help customers locate History Center digital images, identify the technical specifications attached to each image, and help record use data that will identify past use of images. At the same time, data collected during the scanning process will help formulate future preservation strategies.

The most important feature of any metadata use is consistency. Spellings, use of terms, title conventions, naming standards all must be set and followed consistently.

Step 1   Search the Visual Resources Database to determine if the image has been previously scanned and now exists in derivative form. The derivative form of the image is the digitized version of a hard copy or is another copy of an already digitized image.

Step 2   If not, create the new derivative image using the Scanning Procedures titled Scanning Photos located in the appendix of this manual [waiting for new scanner to be set up to write scanning directions]

Step 3 Fill in the fields of the Visual Resource Database, including the three categories of metadata - Resource, Technical, and Use Metadata - and the file path to the small jpeg image for use in the database. The database is located in S:\SHARED\Operational Records\Reproduction\Metadatadocument\Visual Resource Database.

Step 4 Click on save disk icon and close database.

Metadata Collection Form
Data Dictionary:
The items shown in all caps appear as field names in the Visual Resource Database. This Data Dictionary explains what type of information will appear in each field and what format the information should take when being entered by staff and volunteers. A number of the fields have pull down menus from which to choose data to be entered.

Resource Discovery Metadata Category:
Most of the metadata in this section of the database is taken from the Dublin Core Initiative suggested fields. Additional fields were added after consulting the work of the University of Texas Humanities Research Center's image database guide, as well as considering compiling the AHC's in-house metadata needs.

1. IMAGE ID
        Unambiguous digital resource identifier. This identification number is peculiar to the digital version of the original image. We use the filename of the digital image; example: PICA 01300.jpg. If there is more than one digital form of the original, then add a lower case letter (a, b, c, etc.) to the identification number of the original item to distinguish it from other digital versions of the original, such as a tif or thumbnail file.

2. RECORD DATE

Automatic entry of date on which record was originally input. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes among others the form yyyy-mm-dd. This field displays the date in the following format: YYYY-MM-DD.

3. RECORD COMPLETE

Check box format - when required information is entered into the entire database, click inside box to make check mark indicating record is complete.

4. COLLECTION

Name of Austin History Center collecting unit from which the original resource comes; choose from pull down list and highlight appropriate selection.

5. TITLE

A name given to the digital resource. While this field may not apply to photographic images, it could apply to a manuscript collection, literary work or report that has been digitized or is born digital. Capitalize only the first letter of first word of the title or of any proper names contained within the title. Do not underline or write the title in italics. When possible, exclude initial articles from title. Use two spaces after a period, one space after a comma, colon or semi-colon. Examples of titles include:

Four great tragedies

Balcones Canyonlands Preserve management handbook

Report of the Texas Equal Suffrage Association

Austin, Texas. Police Department Records of the Charles Whitman mass murder case

6. CREATOR

Name of the person or organization primarily responsible for creating the intellectual content of the resource. In this field, we are concerned with the original resource that is being digitized. Examples of creators include authors of written documents; artists; illustrators; photographers; or organizations that generate archival collections. Determine the correct form of the name when possible, using the History Center's AF-Biography File, DRA catalog or the Library of Congress Authority File (http://authorities.loc.gov) as a guide whenever possible. Enter personal names in inverted form: last name comma space first name comma space middle name or middle initial. Follow middle initial with a period. Do not include junior, senior or Roman numerals after name. Example: Oliphant, William James. Enter group or organization names in full, direct form. Example: Balcones Canyonlands Preserve Coordinating Committee.

7. SUBJECTS:

What the content of resource is about or what it is, expressed by headings, keywords, phrases, or names. The best place to start in entering a subject is to use the AF--Subject heading, the AF—House/Building address or the AF--Biography name of the image or document you are scanning. Include names of persons pictured or mentioned in the resource. Other subjects for inclusion might be elements of the image or document. For instance, if the photo shows an automobile (and automobile is not the heading under which the resource is filed), then you might want to include automobile as a subject heading in this field. Recommended best practice when using non-AHC subject headings is to select a heading from a controlled vocabulary or formal classification scheme such as the Library of Congress Subject Heading list

(http://authorities.loc.gov/) or the Art and Architectural Thesaurus (http://www.getty.edu/research/conducting_research/vocabularies/aat/). Enter multiple subjects in the order of their importance (often based upon how much of the entire content is devoted to a particular subject). Separate each subject entry by a semi-colon and space. Subjects may be personal or organization names as well as topics, places, genres, forms and events. Enter subjects as they appear in the AF-Subject Headings list or on the label of the photo, including the same use of parentheses, spaces, punctuation, etc. The examples below do not cover all eventualities; if unsure, check the subject headings list.

Personal Name: Enter personal names in inverted order, placing a comma space between the last name and first name middle name or initial. Example: McCallum, Jane Y.

Organization name: Enter organization names in direct order. Example: National Organization for Women.

Topics: If a topic includes dates given in parentheses, follow the topic with two spaces then the dates in parentheses. If the topic has a several parts, separate with each part with a space hyphen space. Example: Capitol (1888-) – Construction

Places: Enter place names as subjects in direct order, following the first word with a comma space second word. Example: Manor, Texas

Address: Enter name or number of street, using Arabic numerals for street numbers followed by "th" or "st" then E period or W period for direction then three spaces then number of address. Example: Congress Avenue   1000; 6th Street E.   100

Genre: Enter name of genre in direct order. If there are several parts to a genre listing, separate with a space hyphen space next word. Example: Popular music – 1981-1990

8. DESCRIPTION (Abstract or content description)

A textual description of the content of the resource such as an abstract, table of contents, full-text, or free text account of the resource.   A description is helpful when the resource that is digitized is a book or manuscript with multiple pages; or when the resource is a photograph or drawing that needs more description than simply the subject headings in the previous field. Enter descriptive text, remarks, and comments about the object. This information can be taken from the object or be provided by the record creator. You can also enter specialized information not included in other elements, e.g., measurements of a depicted object, description, provenance, technique, distinguishing features, inscription, condition, and history of the work. For photographs in the Austin History Center collection, information for the description field can be gathered from the back of the photograph.

Examples of descriptions include:

Black and white photograph of horse and buggy, in front of the J.C. Penney store, Longmont, Colorado, ca. 1901

Typewritten letter from Mary Heard Ellis to Miss Blanton, January 2, 1919

9. PUBLISHER

The publisher is the entity that created the digital resource. For all digital files created by Austin History Center staff, the default entry in this field is Austin History Center. In the case of an object that existed in another form before being digitized, the publisher of this earlier form may be given in the *Source* element.

10. DATE OF ORIGINAL

Creation dates for the original resource from which the digital object was derived or created. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes among others the form yyyy-mm-dd. For many items, only the year will be known, so it is permissible to enter only the year (yyyy). Use a single hyphen (no spaces) to separate the year, month, and

date components. To show a date is approximate, follow it with a question mark as in ***1890?***. If date is unknown, leave blank.

11. IMAGE PATH

When a photo or document is scanned, staff members will save two files on the server. One of these files will be a higher resolution image to be used for customer orders or other staff purposes. The second low resolution jpeg file will be used for display in the database. The filename for this low-resolution image will have a letter following the image identification number to distinguish it from the larger resolution image. Example: PICA 01300b.jpg. The file path to this image should be entered in this field so the database knows what image to show in the frame on the right side of the page. If the image described does not appear in the frame on the right, then you should check the file path to make sure it is correct. The image path will look something like this: I:/archives/AF Biography/Colquitt, Oscar Branch/C00311a.jpg

12. FILE PATH

File path to the higher resolution jpeg saved on the server. This image is considered the digital resource about which metadata is being entered in the Visual Resource Database.

13. TYPE

Type includes terms describing general categories, functions, genres, or aggregation levels for content of the original document. Recommended best practice is to select a value from a controlled vocabulary. Examples: Image (Map, stereograph, photograph, painting, engraving); Text (Scrapbook, diary, poem, home page, manuscripts, music score. Note that page images are text); Object (Museum piece, architectural structure, monument); or Moving Image with Sound (video). This field has a pull-down menu from which to choose the type of resource.

14. SOURCE

Reference to the original resource from which the present resource has been digitized. Recommended best practice is to reference the resource by means of a string or number conforming to a formal identification system. For images or documents digitized from the AHC collections, use the image identification number (PICA, Chalberg, Statesman or Jones number), the archives collection number (FP or AR number), the call number for a book, or folder title for clipping files (AF--Subject, AF--Biography or AF—House/Building).

15. FORMAT

The digital manifestation of the resource. The Format field in the AHC's Visual Resource Database has a pull-down menu from which to choose the electronic format of the digital file. These choices include Bitmap, Gif, Jpeg, and Tif. Commonly, Jpeg will be the first choice.

16. LANGUAGE

Language of the intellectual content of the resource. The default for this field is "en" or English. This abbreviation conforms to the ISO 639 standard codes for languages. Since most of the materials digitized from the AHC's collection are in English, use the "en" default. However, as more materials in other languages (especially Spanish) are added, the need will arise to use codes for other languages. For other standard letter language codes, consult http://www.loc.gov/standards/iso639-2/englangn.html#ab.

17. RELATION

Reference to a related resource. Use the refinement "Is part of" to indicate if resource is part of a larger resource. For instance, images used in the "Austin Treasures" Web exhibits can be referenced in the Relation field as "Part of the Austin Treasures Web Exhibit 'Green Growth', http://www.ci.austin.tx.us/library/ahc/green/default.htm. " Another example is items that are part of an archival collection, in which the entry would read "Part of FP.E.4 Jane McCallum Papers".

18. COVERAGE

Coverage will typically include spatial locations, period date or date range, or jurisdiction label. The category spatial refers to the location(s) covered by the intellectual content of the resource not the place of publication. Temporal coverage refers to the time period covered by the intellectual content of the resource not the publication date. The best practice is to use a controlled vocabulary list, for example, the Thesaurus of Geographic Names, http://www.getty.edu/research/conducting_research/vocabularies/tgn/. Since most materials digitized from the AHC's collection relate to Austin, the default spatial coverage is "Austin, Texas". Separate spatial and temporal with semi-colon. For temporal coverage, use the YYYY-MM-DD format if the year, month and day are known. If only the year is known, then enter that information. For a range of dates, enter the dates on the same line, separating them with a space, hyphen, and space as in 1900 – 1950. To show a date is approximate, follow it with a question mark as in 1997?

19. RIGHTS MANAGEMENT

Information about rights held in and over the resource, and any restrictions placed on access or use of the resource. The Rights Management field has a pull-down menu that lists the options available for entry. These options include: Austin History Center owns copyright; Restricted due to confidentiality restrictions not yet expired; AHC and creator must be credited; AHC and collection must be credited; No photoduplicating without written permission of creator; No photoduplicating without written permission of donor; Not AHC photo - must get permission from owner of image; and No commercial use; noncommercial use only if AHC and creator credited.

20. DONOR NUMBER

Gives the AHC Donor Number for materials that have been donated. The donor number is typically written on the item, for example the donor number is in the lower left hand corner of materials in the photo collection. In the future, this field will link to the Donor Registration Database.

21. CD NUMBER

Because of server space considerations, when high resolution scans are created, they will be stored on CD-ROM rather than the server. Each CD-ROM will be given a number so that tracking the high resolution resources will be easier. The number of the CD-ROM will be entered into this field. A lower resolution scan will be derived from this CD-ROM and saved to the server.

Technical Metadata Category:
The metadata captured in these fields record how the image was captured, information that will be necessary in future migrations of the images to other systems and platforms. Many of these fields will have default values or pull-down menus from which to choose values.

1. COMPRESSION

Designates the compression scheme used to store the image data. In other words, it identifies whether resource is compressed or not. Default is "jpeg", but can substitute other compression schemes (png, lzw, etc.)

2. RESOLUTION

Indicates the resolution of the image in dots per inch measurement. Default value is "300 dpi". If this standard is not followed, insert other resolution value.

3. FILE TYPE

An additional letter is added to the record number in order to differentiate the file names of thumbnails, watermarked, and other variant scans. Add a "t" to filename if image is a thumbnail; add a "w" to filename if image is watermarked.

4. FILE SIZE

Extent of image in bytes; see image size in PhotoShop.

5. COLOR SCAN

Most images will not be scanned in color, so the default value is "No". If this standard is not followed, change the entry by selecting "Yes".

6. BIT DEPTH

The default value is "8". If this standard is not followed, select "Other" and indicate the bit depth used.

7. PERSON SCANNING

This field has a pull-down menu of the names of individuals who perform scanning. Choose name from the list.

8. DEVICE SOURCE

Classification of device used to create the image data. Since most of our images will be scanned on the flatbed scanner in the lab or the exhibits area, the default value for this field is "reflection print scanner". If other (i.e., digital still camera), insert appropriate source.

9. SCANNING SYSTEM

Choose the scanner manufacturer and model name or number and name and version of the capture software from the pull-down menu.

10. DATE CREATED

Date the digital image was created. Use date format yyyy-mm-dd. This information should be entered when capture process started.

11. DATE PROCESSED

Date image was processed. Use date format yyyy-mm-dd. Date will refer to the final processing action. The value for this field will be null for images that receive no processing following image conversion. Processing includes rotating or resizing images, or changing the image color or exposure.

12. PROCESSING SOFTWARE

The name and version number of the image processing software used to edit or transform the image data, i.e., Adobe PhotoShop V.5.

13. PROCESSING ACTIONS

An ordinal listing or script of the image processing steps performed by way of processing software. Processing includes rotating or resizing images, or changing the image color or exposure. Examples of values are: Rotate 90 degrees clockwise, PhotoShop actions script, etc.

14. CAPTURED FROM

Default value "Original". If not original, insert copy print, slide, or transparency from pull-down menu.

Use Metadata Category:
These fields describe how the image will be used and who will be using the image. Both fields have pull-down menus from which to choose values.

1. USE

How will image be used? Make a choice from the pull-down menu of values: Austin History Center Association newsletter, Austin American Statesman, in-house exhibit, etc.

2. SCANNING REQUESTED BY

Select name of AHC unit or customer that initiated request; pull down list of AHC units and customer options.

References

Gelaw, D. (2000). Metadata for description of images. Ebook. Arlington, TX: University of North Texas. http://digitallab.lis.unt.edu/revelations/daniel/E-book/Metadatachapter.htm

Gilliland-Swetland, A. J. (1998). *Introduction to metadata: Setting the stage.* In M. Baca, Ed*., Introduction to metadata: Pathways to digital information. Los Angeles, CA: Getty Information Institute.*

Glogoff, S. J. & Forger, G. J. (2001) Metadata protocols and standards: Bringing order to our digital objects. *Internet Reference Services Quarterly,* 5 (4): 5-14.

Harry Ransom Humanities Research Center Digitized Item Database Scanning and Input Manual. Draft November 26, 2002.

Lagoze, C. and Payette, S. (2000) Metadata: Principles, practices, and challenges. In A. R. Kinney and Rieger, O. *Moving theory into practice: Digital imaging for libraries and archives* Mountain View, CA: Research Libraries Group.

Moen, W. (2000). *Metadata: A tool for managing digital objects.* Denton, TX: University of North Texas. http://digitallab.lis.unt.edu/revelations/daniel/E-book//ManagingDigObjWithMetadataJun2000.ppt

NISO, Draft Data Dictionary: Technical Metadata for Images, Version 1.0, July 5, 2000 edition
White, H. J. (2001). Metadata elements in a document imaging environment: An AIIM recommended practice prepared by the Association for Information and Image Management International. Available online: Metadata stnds. http://stnds.aiim.wegov2.com/file_depot/0-10000000/0-10000/1462/folder/15044/Metadata+Standard+AIIM+International.doc

Western States Dublin Core Metadata Best Practices. Version 1.2. January 2003. Western States Digital Standards Group Metadata Working Group. http://www.cdpheritage.org/westerntrails/wt_bpmetadata.html

Other definitions:
In the document imaging context, metadata is " Informational data which provide the administrative, descriptive, preservation, technical, and use descriptions regarding the digital or analog object, within a document imaging environment." Page 6 Herbert J. White

Metadata is a summary of information about the form and content of objects, documents or services. In its broadest sense, metadata is defined as "data about data". Metadata may be a part of the resource itself, or be kept separately from it. In the Internet and Web environment, use of metadata is easy and convenient in the sense that it allows authors or creators of hypertext markup language (HTML) documents to describe the documents at the time of creation by including metadata elements within the documents. (Gelaw, 2000) http://digitallab.lis.unt.edu/revelations/daniel/E-book/Metadatachapter.htm

Gelaw (2000) illustrates how metadata touches all aspects of our lives.
Since the underlying principle for metadata is to improve the management and retrieval of information, it can support many needs in the Internet environment, such as: digital signatures, content rating (to disclose the nature of particular page's contents), intellectual property rights (including the contractual terms related to the document's use and distribution), standardizing for semantic interoperability (by describing resources in same metadata scheme and mapping their resources to a common scheme for search and retrieval), electronic commerce (to encode price, terms of payment, etc.).