UNIVERSITY of VIRGINIA LIBRARY

University of Virginia Community Digitization Guidelines

This document offers guidance and minimum recommendations that are in line with the UVa Library's current practice for faculty who are planning digitization projects. Inherent or unique characteristics of different source materials necessitate different approaches to scanning and conversion requirements for digital projects should be considered on a case-by-case basis (particularly for grant projects with specific requirements).

These guidelines have been developed in order to:

1. Increase the interoperability and accessibility of digital collections across UVa through the use of accepted standards and formats
2. Ensure a consistent, high level of quality across collections
3. Decrease the likelihood of rescanning in the future by promoting best practices for conversion of materials into digital format and the long-term preservation of these digital resources.

Because technology and industry standards are constantly improving and changing, we view this as a continually evolving document.

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? Does access need to be limited to certain groups? Do different groups needs different types of access?
- What options do you have for delivering the materials?
- 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?

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. Hard 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 Firewire standard allows for faster access to these drives. MiniDV/DVcam is a commonly used medium for storing for video.

Definitions:

Access Quality Collection Master: An uncompressed and uncorrected raw digitized file, digitized to a level that supports the minimum standards for generating quality deliverable files, but not to a higher preservation standard.
Bit-depth: The number of colors/level of grayscale captured in the digitization.
Color Space: A standard describing the values of colors mathematically. Use of a common color space better ensures consistency of images created across multiple environments.
Compression: When a file's size is reduced to save storage space. Much compression is "lossy," which results in permanent loss of some of the data captured at the point of digitization. Master files for all formats but video are not compressed.
Delivery Master: A color-corrected, cropped, edited, or otherwise altered version of the original digital master that will be used to generate all deliverable files.
DPI versus PPI: DPI (dots per inch) refers to the quality of a printed version of an image. Images are printed as a series of dots – the more dots per inch, the colors or grays look better and blends between colors or tones are smoother. PPI (pixels per inch) is used to describe the display quality of an image -- if there are too few pixels per inch, then the pixels will be very large and you will get a very "pixilated" image with jagged edges, especially if you zoom into the image on screen.
File Format: The format in which a file is saved, such as .doc for Word, .jpg for images, or .wav for audio. Different formats can be used by different applications for different purposes.
Metadata: Descriptive information about the content and format of the files.
Pixel: The dots that make up an image, measured in resolution and the height and width of an image.
Preservation Quality Collection Master: An uncompressed and uncorrected raw digitized file, digitized to a very high quality to support preservation.
Preview/Thumbnail: Highly reduced in quality and size or duration, functions as an identifier; has little or no output or editing value.
Resolution: The density of pixels captured in the digitization of an image, or the sample rate at which audio and video are digitized.
Service/Deliverable: Output quality is reduced, to support efficient delivery to users. Multiple files may be created at this level for multiple purposes or user communities. When Service and Deliverable qualities are outlined separately, the distinction resides in the assumption that Service versions are editable by the user, but Deliverable versions are not.
XML: A structured method for encoding content and metadata in a text file. Standards for XML encoding include METS (Metadata Encoding and Transmission Standard), EAD (Encoded Archival Description), TEI (Text Encoding Initiative), and GDMS (Generalized Descriptive Modeling Scheme).

For a more extensive glossary and links, visit:
http://www.lib.virginia.edu/digital/reports/dl_terminology_uva.htm

Images – Bitmap/Raster

Color Space: AdobeRGB <http://www.adobe.com/digitalimag/adobergb.html>. The Adobe RGB Color Space ICC Profile is available for Windows machines and for the Macintosh. You will need to download the profile as part of a free Adobe ICC Profile Pack, and install the profile as described by the accompanying instructions.

Capture: Access Quality				
Type of Original	**Bit depth**	**Resolution**	**Compression**	**File Format**
Books (text pages)	Bitonal (4-bit black and white)	400 ppi	CCITT Group 4 Fax Compression	TIFF
Books (illustrations or figures)	8-bit (grayscale) or 24-bit (color)	400 ppi	Uncompressed	TIFF
Slides (35mm)	24-bit (color)	300ppi @ 900% (2700ppi)	Uncompressed	TIFF
Oversized items (large books, maps, etc.)	24-bit (color)	400 ppi	Uncompressed	TIFF
Project Specific (dictated by desired use)	1-bit, 8-bit, or 24-bit, as appropriate	Resolution should ensure a minimum capture size of 3000 pixels on long side	Uncompressed	TIFF
Capture: Preservation Quality				
Type of Original	**Bit depth**	**Resolution**	**Compression**	**File Format**
Books (text pages)	24-bit (color)	600 ppi	Uncompressed	TIFF
Books (illustrations or figures)	8-bit (grayscale) or 24-bit (color)	600 ppi	Uncompressed	TIFF
Slides (35mm)	24-bit (color)	300ppi @ 900% (2700ppi) ~ 18 Megabytes	Uncompressed	TIFF
Oversized items (large books, maps, etc.)	24-bit (color)	400 ppi	Uncompressed	TIFF

Deliverables				
Purpose		**Resolution**	**Compression**	**File Format**
Thumbnail		120 pixels on the longest side	JPEG is automatically compressed, select High or level 10 compression	JPEG
Screen-sized		1024 x 768 pixels; or 650-850 pixel width with proportional height (as appropriate) for page images	As above	JPEG
Maximum (create only as needed)		3000 pixels on the longest side	As above	JPEG

Illustrations/Graphs/Charts - Vector

Creation					
Purpose	**Format**	**Compression**	**Bit depth**	**Resolution**	**Comments**
Master copy	EPS, SVG, proprietary formats, e.g. Adobe Illustrator	NA	NA	NA	Include color reference whenever appropriate and feasible
Deliverables					
Purpose	**Format**	**Compression**	**Bit depth**	**Resolution**	**Comments**
Deliverable	EPS,SVG, SWF, JPEG	NA	24-bit color, 8-bit grayscale	Appropriate for display of necessary information; 300 ppi if readable printing must be supported	Vector images may be retained in their original format or converted to bitmap/raster formats for delivery; use the chart above as a reference.
Thumbnail	JPEG	JPEG is automatically compressed, select High or level 10 compression	24-bit color; 8-bit grayscale 120 pixels on the longest side	72 ppi	

Electronic Texts

Capture			
Purpose	**Description**	**Format**	**Standard**
Structured Text Transcription	A literal transcription of the text, encoded in XML. Requires additional files and specialized server software to deliver, especially if searching is desired.	XML	TEI P4, with local modifications; follow the DTD available at: http://www.lib.virginia.edu/digital/reports/teiPractices/dlpsPractices_postkb.html
Unstructured Text Transcription	Plain text that may include minimal structural or formatting information.	XHTML, ASCII text, e.g. OCR output	
Page Images	If the text will include references to page images, select the capture specifications from the Image Table above.	As appropriate from above options	
Toolkit	PDF texts for use in Toolkit.	PDF	If PDFs are needed, please contact Instructional Scanning for assistance. http://lib.virginia.edu/leo/iss.html
Deliverables			
Purpose	**Description**	**Format**	**Standard**
Structured Text Transcription	Marked-up to reflect the content and the structure of the original document.	XML	TEI, as documented above.
Unstructured Text Transcription	Plain text that may include minimal structural or formatting information.	XHTML, ASCII text, e.g. OCR output	
Page Image Deliverable(s)	If the electronic text is a transcription with dependent page image deliverables, select the deliverable specifications from the Image Table above.	As appropriate from above options	

Audio

Creation			
Purpose	**Format**	**Resolution & Sample rate**	**Description**
Master	Broadcast WAV	44.1 kHz, 16 bits per sample	Maintain channel pattern of original, e.g. stereo, mono, and multi-channel.
Deliverables			
Purpose	**Format**	**Resolution & Sample rate**	**Description**
Service	MPEG 1/2 Layer 3 (.mp3); MPEG 4/AAC	Appropriate to type and quality or original	Maintain channel pattern where practical.
Deliverable	MPEG 1/2 Layer 3 (.mp3);; MPEG 4/AAC	Appropriate to delivery needs and conditions	
Preview	MPEG 1/2 Layer 3 (.mp3);;		Reduce duration to create a representative sample: a "clip"

Video

QTVR: For QTVR 360º panoramas or object views, capture the images using the image specifications listed above, and stitch the QTVR files together with your preferred tool.

Creation			
Purpose	**Format**	**Compression**	**Description**
Master	NTSC DV, DV-Cam tape, Beta-SP	DV	Media should be stored in an environmentally stable location
Deliverables			
Purpose	**Format**	**Compression**	**Description**
Service	Select as appropriate for use	Appropriate to format; and use	Service, i.e. editable, versions produced as required by "dubbing"; implies change of storage medium and/or format. Very large file sizes; not network distributable.
Deliverable	MPEG1, MPEG2, MPEG4	Appropriate to format and use	Only highly compressed forms, network distributable.
Preview	MPEG4	Appropriate to format and use	Reduce duration to create a representative sample: a "clip."
Thumbnail	120 pixels on the longest side, JPEG	JPEG is automatically compressed, select High or level 10 compression	Representative frame: indication of content.

Statistical/Numeric Data

Purpose	Format	Comments
Master copy	ASCII columnar format SPSS, STATA, SAS program code and/or machine readable text based documentation to define data for analysis	ASCII delimited preferred DDI standard metadata preferred documentation format Following the ICPSR standard for data archiving and preservation.
Service	Data stored in some statistical package format (SAS, SPSS, STATA) or in queryable SQL database system	Storage for access, retrieval, or extraction.
Deliverable	SAS, STATA, SPSS, Excel or delimited ASCII format with data map or variable list.	Excel not advised for very large files. All users get documentation built from DDI records.
Preview	Screen dump of 5% of records, no more than 100	Practice not currently in place.

Spatial Data - Vector

Purpose	Format	Comments
Master copy	ASCII-based exchange format such as SDTS, Arc Exchange (.e00), ArcGenerate (.gen), or delimited text.	Note that two of these are tied to proprietary software formats and are not available for all data models. SDTS is available but rarely used in data distribution.
Service	Industry standard formats such as ESRI shape (.shp) or ArcInfo Coverage model, or CAD format such as Microstation (.dgn) or AutoCAD (.dgw). Possible storage in SQL based system through proprietary middleware (ArcSDE, Oracle Spatial)	Note that ESRI's shapefile model consists of several related files. The ArcInfo Coverage model is directory-based. RDBMS models are still relatively new.
Deliverable	Industry standard formats such as ESRI shape, Arc Exchange, or CAD formats.	
Preview	GIF, JPG or other raster image format.	Preview graphics need to be large enough to convey the general "look" of the data.

Spatial Data – Raster

Capture the images used to generate these files using the image specifications listed above, and generate these files using your preferred tools.

Purpose	Format	Comments
Master copy	Photography or remote sensing imagery: Non compressed TIF+world file or GeoTIFF (preferred), BIL, IMG (Erdas Imagine)	Also applicable for geo-referenced maps. GeoTIFF retains geographic information in TIFF header; world file does same as separate file.
	Non-image raster data: ASCII based storage and exchange format (Arc Exchange .e00; ArcGenerate .gen; Spatial Data Transfer Standard (SDTS))	SDTS is federal standard, but not widely adopted in commercial industry or government; format is cumbersome for further processing.
Service	Photography or remote sensing imagery: GeoTIFF, BIL, IMG, SID + world file	
	Non-image raster data: ArcExchange; GeoTIFF; native data formats (.cdo); native software data models (ArcGRID)	Users will almost always need to process stored data. Tiffs can store pixel value as color value and be converted in GIS software; native data formats are common in federal data. GRID data model is directory, not file-based but could be stored for access purposes.
Deliverable	Photography or remote sensing imagery: GeoTIFF, BIL, IMG, SID+world file, JPG+world file	
	Non-image raster data: Arc Exchange, native formats or models, GeoTIFF	
Preview	JPG, GIF, or SID	Sizes may need to be slightly larger than those outlined for other types of images

Describing Your Digital Resources

The Library suggests a minimum list of categories of information that you should use to describe the content of your resources as well as the nature of the digital files themselves. We recommend that you send email to <lib-metadata-help@virgina.edu> at the start of your project. A Librarian with your area or subject expertise will be happy to work with you in setting up a process and identifying appropriate descriptive terminology.

The guidelines that follow outline the *type* of descriptive information that we recommend you collect and give you some basics for structuring that data. For assistance with creating a database or choosing a metadata format to encode your descriptions, please feel free to contact send email to <lib-metadata-help@virgina.edu>.

There is important descriptive information to be gathered both about the intellectual content of the resource *and* about the digital creation. These elements are outlined below. Some fields are strongly recommended, some are required, and others are optional. In order for the Library to take ownership of your resource and/or commit to digital preservation, we ask that you consider **all** of the fields for describing the intellectual content and the digital resource. The absolutely required fields are marked with asterisks. Please document your practices and standards and be prepared to include that documentation with any data files you deliver to the Library.

The Notes in the third column are included in a Notes Appendix, or are available online at http://www.lib.virginia.edu/digital/metadata/communityguidelines.html .

The entire chart is also available online at http://www.lib.virginia.edu/digital/metadata/communityquickchart.html.

DESCRIBING THE INTELLECTUAL CONTENT

*Title	The actual title of the content of the resource, or a brief descriptive phrase.	See Notes Appendix
*Agent	The name(s) of individuals or organizations that bear some important relationship to the content. At least one agent of some sort is required. Agents have types (creator, publisher, contributor) and one of these types is also required to be specified in the data.	See Notes Appendix
*Date	Date or date range associated with the creation of the content.	See Notes Appendix
Place	A physical location associated with the creation of the content (i.e. the place of publication or the location of a building or of a painting).	See Notes Appendix
Physical Description	The extent of the resource (number of pages of the print book), physical dimensions (for paintings or sculpture), the medium (bronze, oil), etc.	See Notes Appendix
*Content Type	The nature of the content being described.	See Notes Appendix

DESCRIBING THE DIGITAL RESOURCE

*Identifier	A name/code for each resource that is unique within your database.	See Notes Appendix
*Access Rights	The level of access that a member of the UVa community or the general public can have to this resource.	See Notes Appendix
Agent	The name(s) of individuals or organizations that bear some important relationship to the digital resource.	See Notes Appendix
*Resource Type	The type of digital object being described	See Notes Appendix
*Date	The date the digital file was created.	See Notes Appendix

OPTIONAL ELEMENTS

Culture	A culture of origin or context for a given resource.	See Notes Appendix
Style	A style or period associated with the content.	See Notes Appendix
Description	Descriptive text, notes, remarks, or comments about the resource.	See Notes Appendix
Language	The language(s) of the intellectual content of the resource	See Notes Appendix
Subject/Keywords	Topic of the resource. Typically the subject will be expressed as keywords or phrases that describe the subject content of the resource, or terms related to significant associations of people, events, or other contextual information.	See Notes Appendix
Place coverage	A physical location represented by the content (i.e. the geographic subject of a book or the representation of a place within a painting).	See Notes Appendix
Date coverage	Date or date range represented by the content (i.e. the temporal subject of a book).	See Notes Appendix
Relationships	Used to relate two metadata records together, i.e. items in a set, issues of a newspaper, a painting located within a Church.	See Notes Appendix
Mimetype	A standard for the formatting of files so that they can be sent over the Internet.	See Notes Appendix

Where to Get More Help

Digital Media Lab
Clemons Library, 3rd Floor
Judy Thomas, jthomas@virginia.edu
Jama Coartney, jama@virginia.edu
http://lib.virginia.edu/clemons/RMC/dml.html

Digital Scholarship Services

Scholars' Lab
Alderman Library, 4th Floor
Donna Tolson, dtolson@virginia.edu
http://www.lib.virginia.edu/scholarslab/

Rare Materials Digital Services
Small Library, 2nd Floor
Bradley Daigle, bjd2b@virginia.edu
http://www.lib.virginia.edu/rmds/

Fiske Kimball Fine Arts Library
Campbell Hall
Liz Gushee, egushee@virginia.edu
http://www.lib.virginia.edu/fine-arts/collections/visual_res.html

Instructional Scanning Services
Alderman Library, 3rd Floor
Mitch Farish, ISS Coordinator, lib-iss@virginia.edu
http://lib.virginia.edu/leo/iss.html

Brown Science & Engineering Library Research Computing Lab
Clark Hall
Andrew Sallans, sallans@virginia.edu
http://www.lib.virginia.edu/science/rescomp/

UNIVERSITY of VIRGINIA LIBRARY

Appendix – Notes on Describing Your Resources

http://www.lib.virginia.edu/digital/metadata/communityguidelines.html

Intellectual Content Field Notes

Title *(Required)*

The actual title of the content of the resource, or a brief descriptive phrase.

- If you have multiple titles, you may qualify them (i.e. Primary Title, Variant Title, Series Title, etc).
- Please be sure to specify the primary title, as this is the title that will be used for search result displays and prominent headings.

Agent *(At least one agent of any type is required)*

The name(s) of individuals or organizations that bear some important relationship to the content.

- Agents have types (creator, provider, contributor) and one of these three types is also required to be specified in the data.
- Agents may also have roles (author, editor, publisher, etc.). We recommend that you specify an agent role, but it is not required. Develop a standard list of terminology for roles and use the terms consistently.

  Examples:

Title	Agent Name	Agent Type	Agent Role
Rotunda	Jefferson, Thomas	creator	architect
The letters of Thomas Jefferson	Jefferson, Thomas	creator	author
	Betts, Edwin Morris	contributor	editor
Mrs. Putnam Catlin *[painting]*	Smithsonian American Art Museum	provider	repository
The nutcracker : a dimensional storybook	Simon & Schuster	provider	publisher

- Personal names generally should be last name, first name, middle name, or initial.
- Organizations or corporations should be in natural name order.

- Part of the purpose of a name is to identify the resource bibliographically. The other purpose is to *distinguish* one person/entity from another with the same name. We highly recommend, therefore, that you use an authorized list of names:
    - The Library of Congress Name Authority File - http://authorities.loc.gov/
    - The Getty Union List of Artist Names - http://www.getty.edu/research/conducting_research/vocabularies/ulan/
- From those websites, you can search for a personal or corporate name. Use the form of the name found in the authority file.
- If you cannot find an already established form, you can use the item in hand to establish a form and the form you establish should be unique (i.e. not conflict with a form already established for a person/organization/corporation by the same name). Add as much information as needed to establish uniqueness. For personal names, this might include full middle names, Jr./Sr., dates, etc. For organizations/corporations, this might include place information. Spell out abbreviations or include a key with your documentation.
- Document the scheme used (i.e. LC or Getty). If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- You may also have dates for your agents (i.e. author birthdate/deathdate, an organization's date founded, etc.). Put dates in their own field and formulate them according to the guidelines for **Dates**.

Agent Name	**Agent Dates**
Jefferson, Thomas	1743-1826
Cruise, Tom	1962-

- An agent of some type is required. If you do not have an agent, you may use "unknown". Do not leave this field blank.

Date ***(At least one date of any type is required)***

Date or date range for the intellectual content. This is the date when something was created, built, or published, *not* the date represented by the content.

- A novel published in 2004 about 14th century France -- Date is 2004.
- A painting created in 1896 depicting an event in 1545 -- Date is 1896.

See **Date Coverage** for dates represented *by* the content. Also see **Digital Resource Date** for encoding the date on which your digital file was created.

- Specific days (as for a letter), specific years (as for creation or publication dates), date ranges, or textual date strings all fall under Date. Please, however, distinguish numeric dates in different fields from textual date information. In order to migrate data later this distinction is crucial.
- Numeric dates should be in YYYYMMDD format *(MMDD are optional).*
- Dates may also have types (publication, creation, etc.). We recommend that you specify a date type, but it is not required. Develop a standard list of terminology for types and use the terms consistently.

Examples:

Date Numeric	Date Type	Date Textual	
1954	publication		
18600112	creation		*(for January 12, 1860)*
190010	creation		*(for October 1900)*
1990-2004	publication		*(for date ranges)*
	creation	18th century	
	creation	spring 2004	

- If you need to indicate certainty or eras, please also do so in separate fields. Unless otherwise specified, dates will assumed to be AD.

Date Numeric	Date Textual	Date Certainty	Date Era
1954		circa	
	18th century	uncertain	
333		circa	BC

- A date of some type is required. If you do not have a date, do your best to estimate a date range (i.e. 12th-14th century). If absolutely necessary, you may use "unknown". Do not leave this field blank.

Place

A physical location associated with the creation of the content. This is the place where something was published, created, or is located, *not* the place represented by the content.

- A novel published in New York about 14th century France -- Place is New York
- The Parthenon -- Place is Athens.

See **Place Coverage** for places represented *by* the resource.

- Part of the purpose of a place name is to identify the resource bibliographically. The other purpose is to *distinguish* one place from another with the same name. We highly recommend, therefore, that you use an authorized list of place names.
    - The Library of Congress Name Authority File - http://authorities.loc.gov/
    - The Getty Thesaurus for Geographic Names - http://www.getty.edu/research/conducting_research/vocabularies/tgn
    - The Geographic Names Information System (GNIS) -- for U.S. places - http://geonames.usgs.gov/
    - GEOnet -- for international places - http://earth-info.nga.mil/gns/html/

- From those websites, you can search for place names. Use the form of the name found in the authority file.
- If you cannot find an already established form for a place, you can use the item in hand to establish a form and the form you establish should be unique (i.e. not conflict with a form already established for a place by the same name). Add as much information (state, country, etc.) as needed to establish uniqueness.
- Document the scheme used (i.e. LC, Getty, GNIS, GeoNet). If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- Places may also have types (publication, creation, former location, current location etc.). We recommend that you specify a place type, but it is not required. Develop a standard list of terminology for types and use the terms consistently.
- If you wish to use longitude and latitude, please send email to lib-metadata-help@virginia.edu for assistance in structuring your data.

Physical Description

The extent of the resource (number of pages of the print book), physical dimensions (20 ft. x 10 ft.), the medium (bronze), materials (oil, watercolor), etc.

- Physical Description may have types (extent, medium, dimensions, materials, etc..). We recommend that you specify a type, but it is not required. Develop a standard list of terminology for types and use the terms consistently.

Content Type / Resource Type *(Required)*

Content type: The nature of the content being described.
Resource type: The type of digital object being described

- If all of the resources in your database are of the same Content Type/Resource Type, please just make a note of that information in your documentation.
- If you are mixing Content Types/Resource Types in your database, each record will need to specify the resource's Type.
- Choose from the following:

Type	Definition	Examples	May be used as:	
			Content Type	Resource Type
text	A resource whose content is primarily words for reading	Books, letters, dissertations, poems, newspapers, articles, archives of mailing lists, TEI markup.	X	X
image	A static visual representation.	Paintings, drawings, photographs, graphic designs, plans, and maps, as well as digital page images of books or musical scores.	X	X

movingimage	A series of visual representations that, when shown in succession, impart an impression of motion.	Animations, movies, television programs, videos, zoetropes, or visual output from a simulation.	X	X
physicalobject	An inanimate, three-dimensional object or substance.	A building, a computer, the great pyramid, a sculpture.	X	
dataset	Information encoded in a defined structure (for example, lists, tables, and databases), intended to be useful for direct machine processing.	GIS data	X	X
event	A non-persistent, time-based occurrence.	Exhibition, web-cast, conference, workshop, open-day, performance, battle, trial, wedding, tea-party, conflagration.	X	
sound	A resource with content that is primarily intended to be rendered as audio.	A music playback file format, an audio compact disc, and recorded speech or sounds.	X	X
collection	An aggregation of items of varying types	Archival records which are not described item-by-item but rather at a collection level.	X	X
interactiveresource	A resource that requires interaction from the user to be understood, executed, or experienced.	Forms on web pages, applets, multimedia learning objects, chat services, virtual reality		X

- This is a standardized list. If your content is not represented by one of these terms, please send email to lib-metadata-help@virginia.edu for assistance.

  Examples:

	Content Type	**Resource Type**
A letter: the intellectual content described is text; the digital resource is TEI markup.	text	text
A letter: the intellectual content described is text; the digital resource is a page image.	text	image

A building: the intellectual content described is the physical building; the digital resource is a photograph of the building.	physicalobject	image
A painting: the intellectual content described is the original painting; the digital resource is a photograph of the painting.	image	image

- If needed, you may provide additional information to describe the form of your content more specifically, but it is not required.

Type	**Form** *(can be repeatable)*
image	photograph
text	letter
physicalobject	church, cathedral

Digital Resource Field Notes

Identifier *(Required)*

A name/code for each resource that is unique within your database.

- Follow a consistent naming scheme and document your scheme.
- Your filenames should correspond to the identifier for your digital resource.

Examples:

Title	**Identifier**	**Filenames**
Rotunda	UVA7654	UVA7654_thumbnail.jpg
		UVA7654_screensize.jpg
Thomas Jefferson's Letter to James Monroe	X123	X123_markup.xml
		X123_imagePage1.jpg
		X123_imagePage2.jpg

Access Rights *(Required)*

The level of access that a member of the UVa community or the general public can have to this resource.

- If all of the resources in your database have the same access rights, please just make note of that information in your documentation.

- If you are mixing Access Rights in your database, each record will need to specify the resource's rights.
- Choose one of the following. You may use the codes or the text strings as you prefer.

public	Publicly accessible
uva	Accessible to UVa community only
viva	Accessible to VIVA community only (Virtual Library of Virginia consortium -- http://www.vivalib.org/cfapps/lib1.cfm)
restricted	Restricted to Library staff

Date *(Required)*

The date the digital file was created.

- This must be a specific numeric date in YYYYMMDD format.
- If you do not know the date when your file was created, use the earliest known date. If necessary, use the date the metadata record was created.

Optional Field Notes

Culture

A culture of origin or context for a given resource, i.e. Minoan, Haida, Asante, etc.

- We recommend that you use an authorized list of terms, i.e.
    - The Getty Art & Architecture Thesaurus (AAT) - http://www.getty.edu/research/conducting_research/vocabularies/aat/
- Document the scheme used. If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- If you choose to create your own list of terms, develop a standard list of terminology and use the terms consistently.

Style

A style or period associated with the content, i.e. Rococo, Art Deco, Archaic, etc.

- We recommend that you use an authorized list of terms, i.e.
    - The Getty Art & Architecture Thesaurus (AAT) - http://www.getty.edu/research/conducting_research/vocabularies/aat/
- Document the scheme used. If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- If you choose to create your own list of terms, develop a standard list of terminology and use the terms consistently.

Description

Descriptive text, remarks, or comments about the resource. This might include abstracts, summaries, notes, content descriptions, or other descriptions not included elsewhere.

- We recommend that you categorize your descriptions and qualify them in your database. Do not have single fields holding varying types of descriptions. For example, rather than having one Description field, use:

Title	**Description--Summary**	**Description--Contents**	**Description--Notes**
Gone with the wind	Set during the American Civil War, this story focuses on the lives and loves of Southerners during this period and the hardships they endured.		Title page autographed by Margaret Mitchell.
Andy Warhol : the late work		Vol. 1. Paintings & wallpapers -- v. 2. Photographs/films/videos/books/interviews -- v. 3. Texts.	Exhibition catalog.
Lewis and Clark and me: a dog's tale	Seaman, Meriwether Lewis's Newfoundland dog, describes Lewis and Clark's expedition, which he accompanied from St. Louis to the Pacific Ocean.		Includes bibliographical references.

- Also, descriptions may be about the intellectual content (i.e. abstracts or summaries) or about the digital resource (i.e. "p. 10-20 were not scanned in this digital edition."). In your database, separate descriptions of intellectual content from descriptions of digital representations.

Description--ContentNotes	**Description--Digital ResourceNotes**
Catalog of an exhibition held at the Centre d'art contemporain Geneève, Jan. 25-April 28, 1996, and at the Museum für Gegenwartskunst Zürich, May 5-Sept. 15, 1996.	Specialized software required to view this resource.

Language

The language(s) of the intellectual content of the digital resource (languages(s) in which the text is written or the spoken language(s) of an audio or video resource). Visual images do not usually have a language unless there is significant text in a caption or in the image itself.

Subject/Keywords

The topic of the resource. Typically the subject will be expressed as keywords or phrases that describe the subject content of the resource, or terms related to significant associations of people, events, or other contextual information.

- Please use a field labeled "Subject" only if your terms are drawn from an a standardized list. We recommend that you use either:
    - The Library of Congress Subject Authority File (LCSH) - http://authorities.loc.gov/
    - The Getty Art & Architecture Thesaurus (AAT) - http://www.getty.edu/research/conducting_research/vocabularies/aat/
- Document the scheme used (i.e. LC or Getty). If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- If you prefer to develop your own list of terms, use a field labeled "Keyword." We recommend that you develop a standard list of terminology and use the terms consistently.
- We encourage you to send email to lib-metadata-help@virginia.edu for assistance either in choosing/assigning LCSH or AAT terms or in developing your own list.

Place Coverage

A physical location represented by the content (i.e. the geographic subject of a book or the representation of a place within a painting).

- A novel published in New York about 14th century France -- Place Coverage is France.
- A landscape painting of a building in Athens -- Place Coverage is Athens.

See **Place** for the place associated with the creation of the content.

- Part of the purpose of a place name is to identify the resource bibliographically. The other purpose is to *distinguish* one place from another with the same name. We highly recommend, therefore, that you use an authorized list of place names:
    - The Library of Congress Name Authority File - http://authorities.loc.gov/
    - The Getty Thesaurus for Geographic Names - http://www.getty.edu/research/conducting_research/vocabularies/tgn
    - The Geographic Names Information System (GNIS) -- for U.S. places - http://geonames.usgs.gov/
    - GEOnet -- for international places - http://earth-info.nga.mil/gns/html/

- From those websites, you can search for place names. Use the form of the name found in the authority file.
- If you cannot find an already established form for a place, you can use the item in hand to establish a form and the form you establish should be unique (i.e. not conflict with a form already established for a place by the same name). Add as much information (state, country, etc.) as needed to establish uniqueness.
- Document the scheme used (i.e. LC, Getty, GNIS, GeoNet). If you are mixing schemes within your database, each record will need to specify the scheme you have used.
- If you wish to use longitude and latitude, please discuss database structure with a representative from one of the ECenters.

Date Coverage

Date or date range represented by the content (i.e. the temporal subject of a book). It is the date represented by the content, *not* the date something was created, or built, or published.

- A novel published in 2004 about 14th century France -- Date Coverage 14th century.
- A painting created in 1896 depicting an event in 1545 -- Date Coverage is 1545.

See **Date** for dates associated with the creation of the content.

- Specific days, specific years, date ranges, or textual date strings all fall under Date. In your database, however, please distinguish numeric dates in different fields from textual date information. In order to migrate data later this distinction is crucial.
- Numeric dates should be in YYYYMMDD format *(MMDD are optional).*

  Examples:

Date Coverage Numeric	**Date Coverage Textual**	
1954		
18600112		*(for January 12, 1860)*
190010		*(for October 1900)*
1990-2004		*(for date ranges)*
	18th century	
	spring 2004	

- If you need to indicate certainty or eras, please also do so in separate fields. Unless otherwise specified, dates will assumed to be AD.

Date Coverage Numeric	**Date Coverage Certainty**	**Date Coverage Era**
1954	circa	
1920	uncertain	
333	circa	BC

Relationships

Used to relate two metadata records together, i.e. items in a set, issues of a newspaper, a painting located within a Church.

Relationships are very complicated! Generally, there should be a textual relationship note and an explicit link between the identifiers of both records. We strongly encourage you to send email to lib-metadata-help@virginia.edu if you would like to relate your metadata records.

Mimetype

A standard for the formatting of files so that they can be sent over the Internet.

- Please document how your files are formatted, examples:
    - text/xml
    - image/jpg
    - image/tif