DIGITIZATION PROJECT

Easy as 1-2-3

1…Select the item

2….Scan the Item

3…Create the Metadata

Table of Contents

Section I - Overview

Section II - Worksheets

Appendix

Throughout this document we have tried to offer specific instruction, concrete best practices, and guidelines whenever possible. However, certain sections do present either a range of choices or a spectrum of best practices from ideal to acceptable. These often represent an acknowledgement of the diversity and range of capabilities within the Connecticut Libraries Community, and, in some cases, areas where a best practice has not clearly evolved.

SECTION I
OVERVIEW

Step 1 - Selection

CORNERSTONE OF THE COLLECTION

Successful materials selection is the fundamental element in digital library projects, the one on which all other phases depend. Selecting what you want to digitize is one of the most time-intensive tasks. Materials selection includes judging each item based on several factors, and deciding which item would be an excellent candidate for digitization can be overwhelming.

The first step in the selection process is to determine whether the item is in the public domain (all public domain items may be added to the collection) or is currently under copyright protection. You will need to obtain permission (through license or otherwise) for almost all copyrighted materials that you want to include in your collection. Deciding if an item is copyright projected can be a tedious process, although to save time, a few general rules can be applied based on one simple question: Is the item published or unpublished?

For **published works,** the safe rule of thumb is**:** "Sail the ocean blue through 1922"**.** This phrase means that in general, items published in the United States in1922 or earlier are in the public domain. If there is a little staff (or volunteer) time, check to see if there is a proper copyright notice on works **published** before 1978. If not, it's in the public domain. If you have some additional time and money, investigate the items published before 1964 that have a copyright notice on them to see if the copyright was renewed. If it wasn't renewed properly, the item is in the public domain. For **unpublished** items that were created in 1977 or earlier and the author died prior to 1936, it is safe to assume that the item is not copyright-protected (and therefore is in the public domain) and no permission is necessary. To assist you in this process a Public Domain 1-2-3 worksheet (section II of this document), and a more extensive table (appendix) has been provided as part of this document.

For further instructions, the On-Line Books Page (http://www.cs.cmu.edu/~spok/renewals.html) contains an excellent (and more thorough) explanation of the search process used to determine if the copyright has been renewed, and their site is worth a visit. Another way that doesn't involve an exhaustive copyright record search is to write to the author, their agent, estate, or to the last publisher of the book, and see if they can tell you whether the book's copyright was renewed. Of course, you might not always be able to reach them, and they might not always cooperate, but people who research copyright issues have reported on websites that they have had success with this technique. Sometimes, even if the copyright *has* been renewed, by contacting them you might be able to convince them to let an edition go online anyway.

Once you obtain permission or determine the item is in public domain, the next step consider the value of the item to the end user, the availability of sufficient accurate information for cataloging, and the condition of the item. Attached is a Selection 1-2-3 worksheet which has been developed to give you a systematic approach to this part of the selection process.

Once you have made the selections…

Rank the selected materials based upon their ease of cataloging, so that you start with the easiest items to catalog and save the hardest materials until the end.

Attach a barcode (place it on the protected sheet of the image). This barcode number will be used as the file name. You may add a location identifier to further help you keep track of the item.

Step 2 - Digitizing your Item *(This step can be out sourced to Library Connection)*

SCAN ONCE, SCAN WELL!

Each item starts with your best scan which is referred to as your archival image or digital master (TIFF). From this master, library Connection will then create your access or display images (JPEG).

A small JPEG cannot be turned into a high-quality print, but you can take a high-quality TIFF and turn it into a small JPEG. The key to image quality is to match the conversion process to the informational content of the original, and to scan at that level. During the scanning process, you will make decisions on size of the image color resolution and file size.

This step can either be done by your library or Library Connection. If you prefer to do this in house please use the attached is a Scanning 1-2-3 worksheet which has been developed to give you a systematic approach to the scanning process. We recommend that you use this worksheet along with the Image Quality Calculator located at http://images.library.uiuc.edu/calculator/image_calc.asp,

Summary of Guidelines in the Worksheet:

➢ The scanned image should not be manipulated to increase or decrease its output size. Scan Size (ratio) of Original Item: 1:1 ratio.

➢ Originals should be depicted entirely, including blank margins, against a suitable background paper (usually white or gray) so that the digital image documents the physical artifact, as well as reproducing the imagery that the artifact portrays

➢ The number of colors (or shades of gray) available in the image is determined by the bit depth you choose during the scanning process. For true color photos 16.7 million colors or 24-bits, for black and white photos they require 256 shades or 8-bits.

➢ The resolution (the number of dots per inch) you choose determines the clarity of the image and to what degree it can be magnified and still remain clear. The higher the number, the smoother the picture - and the more it can be magnified.

➢ The quality of the images and the size of the originals both affect the size of the file. The higher the resolution, the greater the file size; the higher the bit depth, the greater the file size. Larger originals are usually captured at a lower resolution to keep the file size in the same range, and smaller originals may be captured at higher resolution.

Step 3- Create the Metadata and the Dublin Core record

THINK ACCESS!

The process the Digitization Project uses to create a Dublin Core record is very similar to MARC cataloging.

To create a MARC record, the cataloger can gather all the information needed to catalog by the examining the item. However, the creator of a Dublin core records needs to know more about the objects (photographic, letter etc.) in order to properly document them, this information is called metadata.

Once you have the metadata, creating the Dublin core record is very similar to creating a MARC record. The creator fills in specific information in certain elements, much the same as the fields in MARC records.

Creating these records does not have to be time-consuming; the solution in most cases is to redesign the work flow rather than to develop new tools or training. You may want to break it out into two steps: data producers should provide informal, unstructured documentation, but they should not necessarily have to go through the rigors of fully-structured formal metadata (leave that to Library Connection). For the historian/library staff, it simply isn't worth their time to learn the Dublin Core standard. Instead, they should be asked to fill out a less- complicated form or template that will be rendered in the proper format by a cataloger who is familiar (but not necessarily an expert) with the subject, and is well-versed in the metadata standard.

As you create the metadata, continually ask the question: "Will this help someone find this resource?" Keep in mind the following characteristics of quality information (metadata) about the object:

> **Completeness & Accuracy** - Include as many fields (elements) as your time and factual information will allow. Good descriptive metadata records can be lean, but take care to include the right amount of information. If too little information is included, records are not discoverable. If too much information is included, users of the shared metadata must wade through extra information in order to target items that interest them.

> **Accessibility** - Include enough information in the searchable fields (title, subject, and description) to allow a user to find the record.

> **Conformance to Best Practices** - Follow the rules and formatting guidelines. This way each item in the collection will appear in search results alongside related resources from other Connecticut libraries

Attached is a Cataloging 1-2-3 worksheet which is designed to help the creator of the record decide what information is required, which element is the best choice for the information, and the format in which the information should be entered.

SECTION II
WORKSHEETS

PUBLIC DOMAIN 1-2-3 Worksheet

Step 1 - Identify if the Item is
☐ Published in United States
☐ Unpublished in United States
☐ Not created or Published in United States

If you answer "**Published"** please proceed to **Section A of this worksheet**

If you answer "**Unpublished"** please proceed to **Section B of this worksheet**

If you answer **"Not created or Published in United States"** stop not in Public domain, get permission from current copyright owner or creator

Section A - Published Works

STEP 1a – Identify if the Item is copyrighted and date of copyright

☐ Published in United States in 1922 or earlier with a copyright - In the public domain, no permission necessary.

☐ Published in United States between the dates1923-1963 with a copyright and it can be determined that the copyright has been renewed - Not in public domain, permission necessary.

☐ Published in United States between the dates1923-1935 with copyright and it can be determined that copyright has not been renewed - In the public domain, no permission necessary.

☐ Published in United States after 1963 with copyright - Not in the public domain, permission necessary

☐ Published in United States without a copyright proceed to **Step 2**

STEP 2a – Identify the date of Publication for items without copyright

☐ Published in United States in 1922 or earlier with out a copyright - In the public domain, no permission necessary.

☐ Published in United States before 1989 with out a copyright - Probably in public domain, may need permission

☐ Published in United States before 1978 without copyright - In the public domain, no permission necessary.

☐ Published in United States after 1989 with out a copyright - Is probably is under copyright, permission is necessary.

Section B - Unpublished Works

□ Unpublished in United States written by employee for the employer, and more than 120 years have past since date of creation - probably in public domain

□ Unpublished in United States, anonymous authors: and more than 120 years have past since date of creation - Probably in public domain.

□ Unpublished in United States and the personal authors died more than 70 years ago - Probably in public domain

MATERIALS SELECTION 1-2-3 Worksheet

STEP 1 – Identify the value of the item

- Is the item rare or unique?
- Is the item (or type of item) frequently requested by the public?
- Is there limited access to this type of item on the web?
- From the point of view of the public, is this material of great importance?

If you answer "**Yes" to any** of these questions, please proceed to **Step 2.**

If you answer **"No" to all** of these questions, please **STOP**; this item is not a good candidate for digitization.

STEP 2 – Examine copyright and legal issues

- Does the Copyright Status fall into one of the following criteria (*see attached*)?
    - ☐ Public domain
    - ☐ Owned and controlled by [your library]
    - ☐ Owned by someone else, but permission secured

If you answer **"Yes" to all** of these questions, please proceed to **Step 3.**

If you answer **"No" to any** of these questions, please **STOP**; this item is not a good candidate for digitization.

STEP 3 – Determine viability of digitization

- Is there enough accurate information to add context to the content (information on subjects or key people, places, events, objects, periods, activities, or projects)? Ex: Official names of all buildings and the groups which use the buildings at the time the photo was taken (example Masonic Hall, Freemasons)
- Is the condition of the original item stable or good enough to be scanned without risk of damage to the item?
- Can this item be delivered from the library to Library Connection without damage to the item?
- Is the item no bigger than 11 X 17? Items which are oversized, unusual, or varying in format may only be included with permission from Library Connection due to the limitations of the processing.
- Would the item retain its informational value without precise duplication of color or detail?

If you answer **"Yes" to all** of these questions, this item is a **good candidate for digitization**.

If you answer **"No" to any** of these questions, please **STOP**; this item is not a good candidate for digitization.

SCANNING 1-2-3 Worksheet

STEP 1 – Identify the Format & Find Scanning Guidelines

Is the item a photograph?

    If you answer "**Yes",** please proceed to **Item A.** If you answer **"No"** please proceed to next question.

Is the item a manuscript?

    If you answer "**Yes",** please proceed to **Item B.** If you answer **"No"** please proceed to next question.

Is the item text?

    If you answer "**Yes"**, please proceed to **Item C.** If you answer **"No"** please proceed to next question.

Is the item a map?

If you answer **"Yes"** to**,** please proceed to **Item D**. If you answer **"No"** please call Library Connection (860-298-5322 ext 1004).

Item Type A - Photography	
File Format	TIFF
Bit Depth	8 bit grayscale; 24 bit color
Spatial Resolution	Approximately 3,000 to 4,000 pixels across the long dimension
Spatial Dimension	Approximately 4X6
File Size	One mega byte or smaller

Item Type B - Manuscripts	
File Format	TIFF
Bit Depth	8 bit grayscale; 24 bit color
Spatial Resolution	Approximately 2,000 to 3,000 pixels across the long dimension
Spatial Dimension	Approximately 4X6
File Size	One mega byte or smaller

Note: for particularly hard-to-read materials such as manuscripts, please consider providing transcriptions of the materials in addition to the digital image.

Item Type C-Text	
File Format	TIFF
Bit Depth	1 bit bitonality; 8 bit grayscale; 24 bit color
Spatial Resolution	600 ppi
Spatial Dimension	100% of original
File Size	One mega byte or smaller

Alternative format: PDF (Portable Document Format) from Adobe is an alternative file format for creating and displaying files on the web. PDF format provides the ability to construct complex multi-page objects such as books, journals, or diaries. See http://www.adobe.com for further information. Requires Adobe Acrobat software to create and manipulate files. The Adobe Acrobat viewer is free to download so users can view documents on their computers

Note: When scanning text documents, spatial resolutions should be based on the size of text included in the document and resolutions should be adjusted accordingly. Documents with smaller printed text may require higher resolutions and bit depths than documents that use large typefaces.

Item Type D-Map	
File Format	TIFF
Bit Depth	8 bit grayscale; 24 bit color
Spatial Resolution	3000 pixels across the long dimension
Spatial Dimension	100% of original

Other considerations: When scanning maps, spatial resolutions should be based on the size of smallest detail included in the map and resolutions should be adjusted accordingly. Documents with smaller details may require higher resolutions and bit depths than maps that use larger details and typefaces. Quality Calculators http://images.library.uiuc.edu/calculator/image_calc.asp) can provide help in selecting the best resolution for scanning maps

STEP 2 – Review Image
If you are not satisfied, consult Image Calculator at: http://images.library.uiuc.edu/calculator/image_calc.asp,

STEP 3 –Name & Save TIFF and JPEG File
All files must have unique names or they will be overwritten.

- Attach a barcode to the acid-free envelope containing the photograph. The file name is the barcode number of the label you have attached to the item. You may add (a suffix) a three-letter location code (to identify your library) if desired.

Library Connection is not responsible for archiving your files. Below is a list of best practices for archiving locally:

- ➢ Writing data to two different storage media and archiving (CD or DVD)

- ➢ Hold archive copies in various locations

- ➢ Store under correct environmental conditions and handle properly

- ➢ Subject media to management routines

CREATING METADATA RECORDS 1-2-3 Worksheet

STEP 1 – Fill in the Required Fields

Using the Excel Worksheet fill in the following required fields:

- Title
    - If there is already a title provided on the item, use it.
    - If not, you will need to compose a title. Start with a generic description of the item, such as letter, postcard, or generic description of the item in the picture, such as "Gravestone". Follow this generic description with a space-dash-space and add a more specific identification of the subject, such as: letter - To Jane Doe from John Doe or Gravestone - Hart, Sara. Finally, add the date (if known). Try to keep the title as brief as necessary to include this information.
    - Leave off all initial articles such as: the, a, an, etc. in the title field, if possible.
    - Dates should be formed using the following order: year, month abbreviation, and day [don't use commas to separate portions of the date]. Use the following abbreviations: Jan., Feb., Mar., Apr., May, June, July, Aug., Sept., Oct., Nov., and Dec.
    - Names should be Last Name, First.
- Description
    - Anything significant about the original object digital form or original resource: its history, physical appearance, contents, abstract, etc. Include information on the image if it provides documentation about the image. This is a searchable field so continually ask the question, "Will this help someone find this resource?" and "Am I using the correct vocabulary for the patron?"
- Publisher
    - Name of the entity that created the digital version of the resource (to clarify that this is the publisher of the digital object, not the original)
- Rights /Copyright information
    - Describe any use or reproduction restrictions, for example: Digital image © Farmington Library. All rights reserved. Images may be used for personal research or non-profit educational

STEP 2 –Review non required files and add information if available

- Creator
    - The creator of the item (such as the author, photographer)
    - Use of controlled vocabularies : Library of Congress Name Authority File
    - The order should be Last name, First name
- Contributor
    - Always put Library Connection Inc. as the first name in this field.
    - May also include any person or entity that contributed to the information about the resource, in addition to the person listed as creator
    - The order should be Last name, First name
    - Give corporate names in full
    - Separate multiple contributors with semicolon and space

- Date
    - Date of the original resource was created. For published objects use publication date, for unpublished objects, use the date it was written, drawn, or photographed. If the date is uncertain use the term circa.

- Language
    - The language of any text or sound in the object. Use uppercase, do not abbreviate.

- Source
    - Name of the institution, name of the collection the item belongs to, location of the item, barcode number and the physical characteristics of the object. such as dimensions, for example: Farmington Library, Brady Collection of Civil War Photographs, Box 1, Barcode no. 2501000123, Map; b&w; 5 X 5 in
- Relation
    - Describe the relationship this object has to another, for example: Research Guide to Aerial Photographs at the Connecticut State Library
- Towns Covered
    - The Connecticut Town depicted in a photograph, etc.
    - Use of controlled vocabularies: http://www.cslib.org/cttowns.htm

STEP 3 – Proofread the record & email the Excel worksheet to Library Connection

- ➢ Spell-check the record.

- ➢ Check each field to verify that the field (element) is appropriate for the information you are cataloging, and all required fields are included (title, description, publisher and rights).

- ➢ Check the format of names and dates to ensure they are in the format shown in the worksheets.

- ➢ Save the file using library name and date sent.

- ➢ E-mail to estidum@libraryconnection.info

APPENDIX

Copyright Term and the Public Domain in the United States
1 January 2009[1]

Never Published, Never Registered Works[2]		
Type of Work	**Copyright Term**	**What was in the public domain in the U.S. as of 1 January 2009[3]**
Unpublished works	Life of the author + 70 years	Works from authors who died before 1939
Unpublished anonymous and pseudonymous works, and works made for hire (corporate authorship)	120 years from date of creation	Works created before 1889
Unpublished works when the death date of the author is not known[4]	120 years from date of creation[5]	Works created before 1889[5]

Copyright Term and the Public Domain in the United States
1 January 2009[1]

Works First Published in the U.S.		
Date of Publication[6]	**Conditions**[7]	**Copyright Term**[3]
Before 1923	None	None. In the public domain due to copyright expiration
1923 through 1977	Published without a copyright notice	None. In the public domain due to failure to comply with required formalities
1978 to 1 March 1989	Published without notice, and without subsequent registration within 5 years	None. In the public domain due to failure to comply with required formalities
1978 to 1 March 1989	Published without notice, but with subsequent registration within 5 years	70 years after the death of author. If a work of corporate authorship, 95 years from publication or 120 years from creation, whichever expires first
1923 through 1963	Published with notice but copyright was not renewed[8]	None. In the public domain due to copyright expiration
1923 through 1963	Published with notice and the copyright was renewed[8]	95 years after publication date
1964 through 1977	Published with notice	95 years after publication date
1978 to 1 March 1989	Created after 1977 and published with notice	70 years after the death of author. If a work of corporate authorship, 95 years from publication or 120 years from creation, whichever expires first
1978 to 1 March 1989	Created before 1978 and first published with notice in the specified period	The greater of the term specified in the previous entry or 31 December 2047
From 1 March 1989 through 2002	Created after 1977	70 years after the death of author. If a work of corporate authorship, 95 years from publication or 120 years from creation, whichever expires first _
From 1 March 1989 through 2002	Created before 1978 and first published in this period	The greater of the term specified in the previous entry or 31 December 2047
After 2002	None	70 years after the death of author. If a work of corporate authorship, 95 years from publication or 120 years from creation, whichever expires first

Metadata Fields: General Instructions for All Formats & Fields

Element	Required	Notes	Controlled Vocabulary
Title	Yes	**Information that goes in this Field:** Take title from the digitized item when possible, or create one if none exists. **Format:** The order should be: type, subject, date (year, month abbreviation, and day). **Example (s):** Essay on funding for the University of Georgia, [1853-1854?] Letter: Sparta, to Callie King, 1852 Sept. 28	No
Creator	No	**Information that goes in this Field:** The creator of the item **Format:** Last, First **Example:** Brady, Mattew, photographer, or his associates Connecticut Dept. of Environmental Protection	Yes: Library of Congress Name Authority File (LCNAF)
Subject	No	**Information that goes in this Field:** Terms that describe what the original object is about. Include general topical; chronological; geographic (i.e. place names, states, towns, named geographic features); personal; or corporate name headings. **Format:** Use Library of Congress Subject Headings [LCSH]. If the LCSH are not effective, use the Library of Congress Thesaurus for Graphic materials Separate subheadings in the LCSH structure with a space dash space. Separate multiple headings from the same controlled vocabulary with a semicolon (;) and a space. **Example:** Historical museums--California—Arcadia	Yes Library of Congress Subject Headings (LCSH) Library of Congress Thesaurus for Graphic Materials (TGM) GNIS DEP Geographic Names

Metadata Fields: General Instructions for All Formats & Fields

Name of Field	Required	Notes	Controlled Vocabulary
Description	Yes	**Information that goes in this Field:** Anything significant about the digital resource not covered elsewhere: its history, physical appearance, contents, abstract, etc **Format:** Use standard punctuation and grammar. If the date is uncertain use the term circa Use hyphen to separate a span of years. **Example(s):** Circa 1867, circa 1861-1865, 1934 Apr. 29	No
Publisher	Yes	**Information that goes in this Field:** Name of the entity that created the digital version of the resource (to clarify that this is the publisher of the digital object, not the original) **Format:** Full name of the library	
Contributor	No	**Information that goes in this Field:** Any person or entity that helped create or contributed to the information about the resource, in addition to the person listed as creator **Format:** Last name, First name. Give corporate names in full. Separate multiple contributors with semicolon and space.	Library of Congress Name Authority File (LCNAF)
Date	No	**Information that goes in this Field:** Date the original resource was created. For published objects use publication date. For unpublished objects, use the date it was written, drawn, photographed. **Format :** Years (four digits) Months (abbreviate) Day. Use a space between the parts of the date. Use the following: Jan. Feb. Mar. Apr. May June July Aug. Sept. Oct. Nov. Dec. If the date is uncertain, use the term circa. Use a hyphen to separate a span of years. When the date is unknown, use "undated" (not capitalized). **Example:** 1934- when only year is known ,1934 Apr. 29 circa 1867, circa 1861-1865,	

Metadata Fields: General Instructions for All Formats & Fields

Name of Field	Required	Notes	Controlled Vocabulary
Format	No	**Information that goes in this Field:** Describes the extension of the access file. This is provided by the scanner **Format:** Use: image/jpeg; image/jp2; text/html; video/mpeg; audio/mp3. Otherwise construct an appropriate designation, type/file format.	No
Identifier	Yes	**Information that goes in this Field:** File name for digital object, this is the barcode number of the item and is supplied by the scanner. **Example/Format:** 25010092.jpg	No
Source	No	**Information that goes in this Field:** Description of the physical characteristics of the original object such as dimensions, publisher, or manufacturer information, and where the item is located. **Format:** Name of the institution, name of the collection the item belongs to, location of the item, barcode number. Use item type, color, size when describing the item. Separate portions of the description information with a semicolon (;) and a space. For size description, use inches (in.) or feet (ft.)- # X # in. **Example:** Farmington Library, Brady Collection of Civil War Photographs, Box 1, Barcode no. 2501000123, Map; b&w; 5 X 5 in	No

Example of Dublin Core Record:

Title   Anthony A. Borgatti (1916-1996) standing in front of SPAG'S, Shrewsbury, Mass.

Subject   SPAG'S, Shrewsbury (Mass.) circa 1940 ; Hardware stores—Shrewsbury (Mass.) ;
                   Stores, Retail—Shrewsbury (Mass.) ; Borgatti, Anthony A., (1916-1996) ; SPAG'S
                   Supply, Inc. ; Discount housed (Retail trade—Shrewsbury (Mass.).

Description   One of Shrewsbury's most famous residents in the twentieth century was Anthony A.
                   Borgatti, Jr. (Spag). Born the oldest of four children on February 29, 1916, in Worcester
                   (Mass.), he opened a tire and battery store on Route 9 in Shrewsbury (Mass.) in 1934
                   that by 1939 had expanded to hardware and other items, and since then, was known by
                   Anthony's nickname, Spag's. He married Olive Lutz on November 24, 1940 in St. Anne's
                   Church. Together they raised three daughters, Carol, Jean and Sandra. Olive died on
                   December 28, 1990, and Spag died on February 23, 1996. Oil portraits of both Borgattis
                   by Worcester artist Leon Hovesepian now hang in the Shrewsbury Public Library.

Publisher   C/WMARS http://www.cwmars.org/

Rights Management Digital image © Farmington Library. All rights reserved.   Images may be used for
                   personal research or non-profit educational uses without prior permission. For requests
                   to publish or exhibit, see Farmington Library

Identifier   25010092.jpg

Example of record in Treasures Collection:

Gravestone - Hart, Ruth

Farmington Library

Title	Gravestone - Hart, Ruth
Subject – LCSH	Sepulchral monuments--Connecticut--Farmington;Hart Ruth
Description	Ruth's epitaph is carved within a heart-shaped outline. The head, angel or cherub, shows a benign fa and large wings. It has a crown of five round balls. The stone hasa typical arched tympanum with sh life is not carved within panels but surrounds the heart, rising to the tops of the shoulder curves, enc of five leaves.
Creator	Whitaker, Lindsay
Contributors	Library Connection, Inc.
Date – Created	1987
Date – Digital	2008
Type	Image
Format	Image/jpg
Language	English
Source – Original	Farmington Library; Old Cemetery Gravestone Collection; b&w; 4x6
Publisher	Farmington Library
Rights	Digital image © Farmington Library. All rights reserved. Images may be used for personal research c educational uses without prior permission. For requests to publish or exhibit, see Farmington Library