



Digitization Training Manual

Table of Contents

In this Manual

1. Project Overview
2. Recommended Equipment and Software
3. Software Toolkit Overview
4. Selection and Copyright
5. Metadata Standards and Creation
6. Image Theory and Specifications
7. Scanning
8. Image Manipulation
9. Performing Quality Control

Credits

Funding for this *Manual* was made available in part by the United States Department of Education's *Technological Innovation and Cooperation for Foreign Information Access* (TICFIA) Program.

For more information about the TICFIA Program, see the Program's web site at <http://www.ed.gov/programs/iegpsticfia/index.html>

Funding for content contribution and other services supporting the *Digital Library of the Caribbean* (dLOC) was made available by institutions through out the Caribbean Basin.

For more information about dLOC, see its web site at <http://www.dloc.com>



SECTION 1

Project Overview

Bienvenidos! Bienvenue! Welcome!

As a member of the *Digital Library of the Caribbean* (dLOC), you are part of a consortia of agencies working to increase access to valuable resources for the study of the Caribbean. Together, institutions from across the Caribbean providing content will be building a truly international library to be used by young school students, college-age researchers, and university professors. dLOC resources will serve the continuing learning interest of average citizens, government officials, and tourists. The content we provide will tell the stories of our peoples, their history, the lands they populate, and the cultures they've constructed and maintain.

For more information about the *Digital Library of the Caribbean*, visit the dLOC web site at <http://www.dloc.com>

About the Digital Library of the Caribbean

The *Digital Library of the Caribbean* embodies research interest in cultural identity. The collections it comprises and the research it is intended to engender illuminate the confluence of cultures, languages and governmental systems in the Caribbean basin.

Caribbean culture, with localized variation, reflects the world and portends the world's future. Three major cultures (races) converge here: Indigenous; European; and African. And, within locales, Semitic, and Asian cultural influences have been formative.

The region's languages represent the principal cultures of these areas. While the region linguistically is more patchwork than shared ground, its creoles and patois - and, perhaps, none more so than Papiamentu - reflect cultural convergence, and it is not at all uncommon for an individual to speak more than one tongue. Language, and specifically translations, can often be used to track the introduction of new concepts and cultural cross-currents.

Regional customs and practices are indicative of the means and methods by which cultures come to live together. The language of cooking for example is intimately tied to the land as much as to culture, and can reveal "genetic" markers based on race and ethnicity that may indicate - first - where cultures are influenced by environment and - second - where experience and adaptation are shared among cultures. Agrarian life-styles converge and diverge throughout the region, their junctions and conjunctions mapping cultural highways. While concrete highways may map cultures in regions beyond the Caribbean, such tangibility masks the making of culture that lays bare in the Caribbean.

The cultural map is nowhere more diverse than in our urban centers. Mexico City, for example, draws together indigenous and European cultures, mixing a working class raised in the fields with a white collar elite. No city in North America is home to more immigrants than Miami, Florida where more than two-thirds of the population has been born elsewhere. Both cities, with all their ills and innovation, represent the future of North and South America. Within their fragile ecosystems, each can provide vital clues to the stresses that population growth places on the ability of natural systems



to sustain life. In this sense the Caribbean is truly a biosphere. Climatologists and archaeologists have begun to publish findings that help us understand the fate of the Maya world's great cities. Mindful that only a handful of ancient codices survive, these scientists in particular have suggested that we mine our cultural heritage for evidence of the past that would serve us into the future.

These examples merely touch the surface of the content available in the Digital Library of the Caribbean. Increasing contributions from the rich landscapes and diverse populations of the Caribbean bring forth an every increasingly detailed understanding of the region.

Project Team & Contact Information

A complete list of the project team is available online at <http://www.dloc.com>.

Project co-Directors:

Judith Rogers, Manager,
Library & Faculty Technology Services
University of the Virgin Islands
Phone: (340) 692-4132
Email: jrogers@uvi.edu

Gayle Williams, Head,
Latin American and Caribbean Information Ctr
Florida International University
Phone: (305) 348-2228
Email: gayle.williams@fiu.edu

Technical Director:

Laurie Taylor, Digital Library Center
University of Florida Libraries
Phone: (352) 273-2900
Email: kesse@ufl.edu

Project Coordinator:

Brooke Wooldridge
Florida International University
Phone: (305) 348-3008
Email: bwooldri@fiu.edu

dLOC Programmers and Trainers:

Mark Sullivan, Systems Department
University of Florida Libraries
Phone: (352) 273-2900
Email: marsull@uflib.ufl.edu

Lourdes Santamaría, Digital Library Center
University of Florida Libraries
Phone: (352) 273-2900
Email: lousant@uflib.ufl.edu

Advisory Board

The Advisory Board is comprised of scholars of Caribbean Studies working in the Humanities, Social Sciences, and Sciences. Members are appointed to direct content development, suggest granting opportunities, identify other sources of funding, and assist in the development of both promotional materials and educational resources.

Members serve a two-year term with possibility for reappointment and attend the dLOC Advisory Board meeting at the annual Caribbean Studies Association conference.

For a list of current members see the dLOC web site (<http://www.dloc.com>).



SECTION 2

Recommended Equipment and Software

Recommended Minimum Equipment

- Computer workstation
 - PC with Windows 2000, NT, or XP ; 1 GB RAM ; CR-RW; 17 inch monitor
- Scanner Recommendations
 - Large format, graphic arts quality scanner that will support digitization of photographs, book pages, maps, etc. Scanner will need to be calibrated periodically.
 - Recommended Scanners
 - Epson 10000XL flatbed scanner or similar
 - Panasonic KV-S2046Cor KS-S2026C sheet-feed scanner

Recommended Minimum Software

- Image Capture Software
 - Digital image capture software is supplied by scanner manufacturers. Not all image capture software is the same however. Epson image capture software supports a full range of image settings for flatbed scanning, where as most consumer market scanner software does not.
- Image Processing Software
 - Used to perform image editing (crop and deskew digital masters images) and to perform batch conversion of derivative images.
 - Recommended Image Processing software:
 - Adobe Photoshop (v.7, CS, CS2, CS3, or CS4) or Adobe Elements
- dLOC Applications
 - A toolkit was written specifically for the Digital Library of the Caribbean, and is freely available to all partners. See the next section for detailed information about this software.
 - The latest version of this software can be downloaded from the following URL:

<http://www.dloc.com/downloads/toolkit/>
- Archive Applications
 - Archiving ensures that digital files may be protected. dLOC institutions are encouraged to archive both locally on CD or DVD and remotely in the dLOC Archive. Recommended software both copies files from PC to CD or DVD *and* verifies full and accurate copy.
 - Recommended Local Archiving Software.
 - RecordNow
 - Nero



SECTION 3

Software Toolkit Overview

Introduction

- The freely distributed dLOC Toolkit allows each partner to track the digitization of items and collect data about the digital resources. It is available at:

<http://www.dloc.com/downloads/toolkit/>

- For instructions on installation and configuration, see Appendix 3 or the included installation CD.
- The toolkit includes the following:
 - A local database
 - A central tracking and processing application
- The application tracks the digitization of a resource through each of the required milestones.
- Due to the local database, only one machine in each institution is likely to have the dLOC Toolkit installed. This should be the machine attached to the scanners.
- The application will check the web for updates on a regular schedule.

Directory Structure

- During initialization of the toolkit, the following directory structure is created:
 - C:\DLOC\ - Root folder. You should scan into folders at this level.
 - C:\DLOC\FTP - Once an item is completed, it will be moved here
 - C:\DLOC\Archive - Once an item is FTP'd from the Complete folder, it will be moved here.
- Using this directory structure allows the tracking database to locate each item, and suggest the next processing step.
- Each of these folders will be discussed during the workflow sections.



SECTION 4

Selection and Copyright

Selection

The Digital Library of the Caribbean seeks to build a cooperative digital library among partners within the Caribbean and circum-Caribbean that will provide users with access to Caribbean cultural, historical and research materials currently held in archives, libraries, and private collections. In addition to facilitating access, the dLOC will also provide an additional layer preservation for the information submitted to the library since the high resolution files associated with the contributions will be archived in the Florida Digital Archive.

dLOC welcomes collections that address the histories, cultures, languages and governmental systems of Caribbean countries. The types of collections appropriate for dLOC include but are not limited to: newspapers, archives of Caribbean leaders and governments, official documents, documentation and numeric data for ecosystems, scientific scholarship, historic and contemporary maps, oral and popular histories, travel accounts, literature and poetry, musical expressions, and artifacts. If you have another collection that you think may be appropriate, please contact us to discuss the collection.

We encourage partners to develop a digital collection plan to ensure that the content is converted into digital format based on its anticipated user demand and its need for digital preservation.

Copyright

Copyright law varies from country to county. Though many of the Caribbean Basin countries are Berne International Copyright Convention signatory nations, it is often difficult to determine which set of laws to apply or what the term of protection may be from one country to another or, for that matter, from one year to another.

The copyright laws of the European Union and the United States represent the Caribbean's most challenging and longest lasting copyright protections. When in doubt it is often a safe bet to apply the laws of those countries. For assistance understanding United States copyright legislation and duration of protections see: Cornell University's Copyright Term and the Public Domain in the United States at

http://www.copyright.cornell.edu/training/Hirtle_Public_Domain.htm

When an item selected for digitization is scanned in its country of origin, the laws of the country of origin should be understood to apply. However, when an item selected for digitization is scanned away from its country of origin, consider both the laws of the country of origin and the laws of the location from which the digital resource will be made accessible via the Internet and *apply those laws that afford the item the longest protection*. For content providers using the centralized services of the *Digital Library of the Caribbean*, the laws of that location are the laws of the United States. Applying the longest protections may



not be to the advantage of archives, libraries and museums, but this policy affords the *Digital Library of the Caribbean* the greatest protection under the international law.

Resources Protected by Copyright

Before copyrighted works can be digitized, Internet Distribution Rights must be procured from the copyright holder assuming the work has not passed into the Public Domain. Distribution rights transfer a privileged use of a copyrighted work rather than the copyright, which remains with the copyright holder.

Note: Publications often include separately copyrighted illustrations. Check illustration credits, verify their copyright status, and seek additional permissions as necessary.

The digitization agency most frequently seeks "*non-exclusive*" "*Internet Distribution Rights*" for an unlimited term. A grant of non-exclusive rights leaves the copyright holder in control of his or her rights as guaranteed under national and international copyright laws and with the ability to grant distribution rights to others. Internet Distribution Rights limit dissemination to the Internet.

If rights to other distribution formats are sought or anticipated, those rights should also be requested with specific mention of the distribution format, e.g., "CD-ROM/DVD Distribution Rights", "Print Distribution Rights", or more broadly, "Electronic Distribution Rights".

Sample *permissions request letter* and *grant of permissions* documents are available online at the project website.



SECTION 5

Preparing for Digitization

Now that you have selected your items for conversion to digital media, we will prepare for the digitization. This section will cover:

- Starting a new item
- Describing your item (creating the bibliographic metadata)

Before continuing you will want to collect any information you have about the items chosen: catalog records, spreadsheets, cards, finding aids.

Starting a new item

We will need to enter some basic information about your item.

1. Open the dLOC toolkit.
2. Create a new item by selecting 'Start a new item' from the main from
3. A blank item form will appear

The screenshot displays the 'Digital Library of the Caribbean - Digitization Toolkit' window. The title bar shows the application name and standard window controls. The main header features the dLOC logo and the text 'Digital Library of the Caribbean'. Below the header, there are five numbered steps: 1. Describe Item (highlighted in yellow), 2. Scan Item, 3. Prepare for QC, 4. Quality Control, and 5. Submit Item. The 'Describe Item' step is active, showing a form with two tabs: 'General Information' and 'Subjects and Keywords'. The 'General Information' tab is selected, revealing several input fields: 'Title Information' with a 'Title' field and buttons for 'Add Alternate Title', 'Add Series Title', and 'Add Uniform Title'; 'Physical Description' with 'Resource Type' (set to 'Book'), 'Physical Description', 'Language', 'Identifier' (with a 'Type' dropdown), and 'Holding Location' (set to 'University of Florida'); and 'Name Authorities' with 'Creator', 'Publisher', and 'Publication Date' fields. At the bottom of the form, there are 'Help', 'Cancel', 'Back', 'Save', and 'Forward' buttons.



4. On this first tab, fill in this basic bibliographic information

Title	Holding Location
Resource type	Author or Creator
Physical Description	Publisher
Language of the material	Date of Publication
Identifiers	

Digital Library of the Caribbean - Digitization Toolkit

Digital Library of the Caribbean

1. Describe Item | 2. Scan Item | 3. Prepare for QC | 4. Quality Control | 5. Submit Item

General Information | Subjects and Keywords

Title Information

Title:

Add Alternate Title | Add Series Title | Add Uniform Title

Physical Description

Resource Type:

Physical Description:

Language: +

Identifier: Type: +

Holding Location:

Name Authorities

Creator: +

Publisher:

Publication Date:

[Help](#) | | | |

5. On the next page, you can enter additional subject information about your material

Subject Keywords	Abstract(s)
Geographic Coverage	Descriptions
Coordinate (latitude and longitude)	



Digital Library of the Caribbean - Digitization Toolkit

Digital Library of the Caribbean

1. Describe Item | 2. Scan Item | 3. Prepare for QC | 4. Quality Control | 5. Submit Item

General Information | **Subjects and Keywords**

Subjects

Subject Keyword: Caribbean Scheme: [v]
Dance Scheme: [v] +
Geographic Coverage: Haiti +
Coordinates: Latitude: [] Longitude: [] +

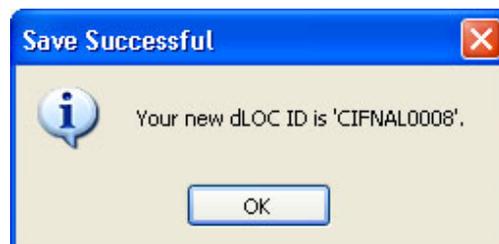
Notes

Abstract: Language: [v] +
Detailed description of the basic steps to the Haitian Meringue, with illustrations and footwork.

Note: (Funding) Digitized with funding from the Digital Library of the Caribbean grant awarded by TICFIA. +

Help | Exit | Back | Save | Forward

- You can hit the 'Save' button at any time to save your work. You will be given the dLOC identifier for your item.





7. Press 'Forward' to move to the next step.



If the item has already been digitized, you can use the links to either copy or move the existing images into this new folder.

If the item has not been digitized, continue to Section 7 for instructions on scanning your item.

Tracking Report

1. If desirable, you can print a tracking report which can be kept with the item to be digitized.
2. A sample tracking report is available online at <http://www.dloc.com>.



SECTION 6

Basic Theory and Specifications

dLOC Requirement for Digital Master Files.

8-bit Grayscale or 24-bit RGB Color (depending on whether original has significant color)
300 dpi for standard text or 600 dpi for stand-alone images (photographs, maps)
Save archival files as uncompressed TIFFs

Bits Depth

In digitization three levels of *Bit Depth* are widely used: 1 Bit, 8 Bit, and 24 Bit images.

A **1 Bit** image is referred to as “bi-tonal” or, less precisely, as “black-and-white”. The picture elements of a 1 Bit image are expressed in strings of one bit. That bit may be either one color or an alternate and, frequently either black or white.

An **8 Bit** image is referred to as “grey-scale”, though an 8 Bit image may represent a very limited color spectrum as well. Most scanning equipment defaults 8 Bit imaging to grey-scale. The picture elements of an 8 Bit image are expressed in strings of eight bits, for example: 00001111. 8 Bit images allow for as many as 255 shades or colors. (N.B. Technically, 8 Bit images allow for 256 shade/color values, but one of these is reserved as a check-digit and is not used to express a shade/color value.)

An **24 Bit** image is referred to as “true color” or, less precisely, as a “color” image. The picture elements of a 24 Bit image are expressed in strings of twenty-four bits. 24 Bit images allow for as many as 16,777,216 shades or colors. You may hear digitization specialists using the short-hand “sixteen million colors”. The 24 bits are divided into three 8 Bit channels, one for each of three composite colors (Red, Green, and Blue.)

Color Space

Color fidelity is fundamental to accurate reproduction of source.

Digitization, faithful to original colors, requires a basic understanding of color and how color reproduction differs from printing technology to digital technology. Fundamental to these differences is the media on which a color image is *printed*.

The color space most commonly used by digitization projects and required by dLOC, is a standardized Red/Green/Blue (*sRGB*) color space.



Choosing the Appropriate Bit Depth and Color Space

1 Bit Image



8 Bit Image



24 Bit Image



dLOC recommends that 1 Bit imaging should **not** be used. 1 Bit images, even at very high resolution (see, *Resolution* below), tend to *pixelate* text. Imperfections on the page or artifacts of age may read as *black*, obscuring text in 1 Bit images. In the 1 Bit page image above, bleed through from the text printed on the inverse page as well as artifacts of age obscure the text. Obscured text will introduce imperfections that reduce the accuracy of text conversion by optical character recognition (OCR) software.

The 8 Bit grayscale image above captures the textual information. And, the reader of the page can make sense of the text.

Readers of Latin religious texts, such as that seen above, will recognize the red text as instructions to the faithful, commentary on the spoken text of a religious service, or the narrative of the priest as opposed to that of the congregation's response.

dLOC advocates preserving *meaningful* color. Meaningful color is color required to interpret the text. In the case of a newspaper with colored images, a color image accompanying an article demonstrates meaningful color, while a color advertisement may not.

It is true that "*The greater the Bit Depth the greater the size of the digital image file*". But, digitization technicians are encouraged to produce images that meet the reader's needs rather than the needs of the digitization technician to conserve space.

Resolution

The resolution of digital images is expressed in terms of pixels. A pixel is a picture element or, simply, a block of solid shade or color that, together with other picture elements comprises a digital image.



Trinidad and Tobago's Coat of Arms



(Zoom area in black box.)

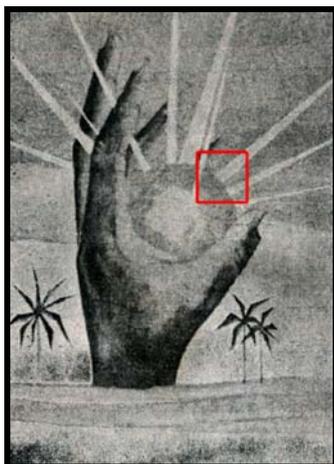
	RESOLUTION	USE FOR
OR	300 pixels per inch (ppi) 118 pixels per centimeter (ppc)	Printed text with <i>normal</i> sized fonts Oversized documents and maps Manuscripts with legible script
OR	600 pixels per inch (ppi) 236 pixels per centimeter (ppc)	Photographs and select graphic arts Printed text with <i>very small</i> fonts Manuscripts with difficult scripts

The dLOC's minimum digital resolution standard for printed text with normal sized fonts is 300 pixels per inch (ppi) or 118 pixels per centimeter (ppc). This threshold is based on both the characteristics of printed graphics and optical character recognition (OCR) tests.

300 ppi / 118 ppc

The Rationale for Printed Graphics

In general, the resolution of printed graphics does not exceed 300 dots per inch (dpi) or 118 dots per centimeter (dpc). Dots per inch/centimeter are rough equivalents of pixels per inch/centimeter; so comparison is appropriate.



Carifesta '72 logo printed in Guyana's Sunday Post and Weekend Argosy (Zoom area in red box.)



Graphics printed in newspapers, for example, often have 80 to 100 dpi (32 to 40 dpc). Most graphics in magazines are printed with 120 dpi (48 dpc) print resolution while graphics in high-end magazines and on post-cards are printed with 300 dpi (118 dpc) print resolution.

Digitization of printed graphics at resolution greater than 300 ppi (118 ppc) would be excessive.

The Rationale for Optical Character Recognition (*Text Generation*)

When a document page is digitized an image of the page is created. All text page images sent to the dLOC's central servers are subject to Optical Character Recognition (OCR).

OCR is a process by which page images are converted to searchable text. Several OCR programs are in common use. Most are optimized for the conversion of images digitized with 200, 300, 400 or 600 ppi (80, 118, 158 or 236 ppc). Images created with other resolution can be converted to searchable text but, generally, with less accurate results.

Resolution and OCR Accuracy in high contrast images			
75 ppi Image	150 ppi Image	300 ppi Image	600 ppi Image
OCR results	OCR results	OCR results	OCR results

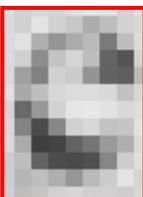


**The Importance of Bit-Depth on Text Recognition:
the Latin word *Feltis* = Goodness**



1 Bit Image

This letter may be any of the following: *c - e - o - 0*



8 Bit Image

This letter may be any of the following: *c - e - o - 0*



24 Bit Image

The letter *e* appears now to be more probable.

dLOC central servers use the Prime Recognition OCR system, configured with six OCR engines to ensure a high level of accuracy in text generation. For printed texts with normal size fonts, whether plain (*sans serif*) or embellished (*serif*), tests demonstrate that the average modern printed document is accurately recognized at 200 ppi (80 ppc).

dLOC sets a slightly higher standard, 300 ppi (118 ppc), for printed texts with normal size fonts to compensate for occasional uses of small fonts or colored, aged (discolored), or blemished paper.

Digitization of normal printed texts at higher resolution (e.g., 600 ppi/236 ppc), in tests, generally showed no increase in text conversion accuracy. 600 ppi/236 ppc images result in higher conversion accuracy only when the source document is printed with very small fonts.

600 ppi / 236 ppc

dLOC recommends digitizing at 600 ppi (236 ppc) only when working with printed texts with very small fonts; photographs and other continuous-tone graphics, and manuscripts with difficult scripts.

Photographs

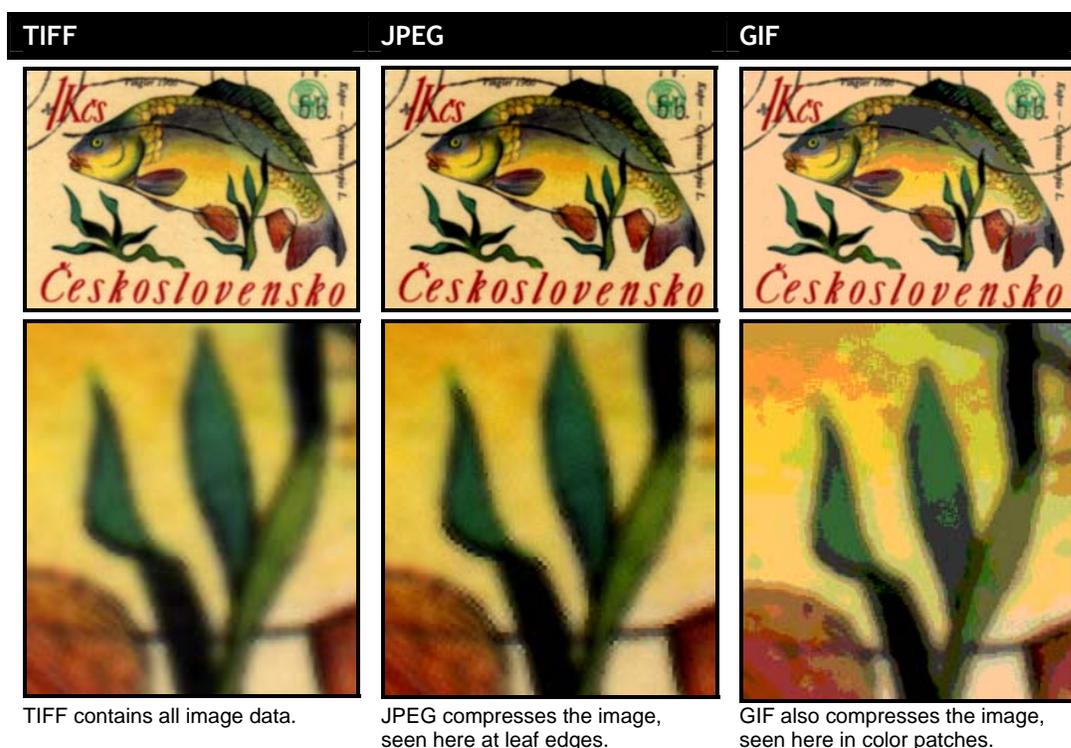
Photographs, unlike printed graphics, have continuous-tone. In the source document, one shade or color blends into adjacent shades and colors. Continuous-tone images may be digitized at any resolution. dLOC recommends 600 ppi (236 ppc) resolution to facilitate special uses of images.



Users of digital photographs frequently consult images for their various subjects as for the whole image. A user may want to zoom on the jewelry or hair braids in the photograph of a woman or on shop sizes in the photograph of a street scene. dLOC central servers use JPEG 2000 technology to facilitate zoom. Images digitized at 600 ppi (236 ppc) produce clearer, sharper, and more *readable* images than do 300 ppi (118 ppc) images.

Saving Files and Image Compression

Once the digital image is created, there remains the issue of saving or archiving the file. The digitization technician prefers not to lose a quality image to the imperfections of file saving and image compression routines.



Saving Files

When saving an image file, the technician has a choice of file types, commonly including GIF, JPEG and TIFF. GIF and JPEG (sometimes: JPG) are Internet deliverable file formats. dLOC creates these *derivative* or secondary file formats for participating institutions from a digital master. Institutions either not participating in dLOC or not using dLOC's central servers, should observe similar practice.

Only the TIFF (sometimes: TIF; *Tagged Image File Format*) is considered archival within the international digital library community. It alone serves as a digital master. There are several reasons for this, primarily: image compression. The illustration above demonstrates image quality issues as a factor in file choice.



Image Compression

When saving an image file, often regardless file type, the technician will be given the opportunity to compress the image. Compression saves file space but has produces other and unwelcome artifacts.

There are two classes of compression: *lossy* and *lossless*.

Lossless compression is an oxymoron. Technically, a lossless image has no compression. A lossless image contains every bit of information created during the scanning process. Here is another simplification: when the scanner captures the bit-stream $1111 \rightarrow$ the lossless file saves 1111 . Though this makes for large files, it also makes for an ideal archival format and, therefore, optimal for file recovery should the digital master ever be damaged in use or degrade in storage.

Lossy data compression technologies attempt to eliminate redundant or unnecessary information, storing a mathematical representation of the eliminated data. Here is yet another simplification: when the scanner captures the bit-stream $1111 \rightarrow$ the lossy file saves a representation of 4. Because lossy images generate smaller files, they can be delivered to readers via the Internet quickly. The human eye compensates for image loss by filling in the gaps. But, because there is image loss, recovery from damage or degradation is more difficult and, in many cases, may be impossible without great expense.

Effects of Compression on Image Quality



0% Compression
9 KB file size

No image artifacts

Shown here:
Both as scanned and
color enhanced
versions.



50% Compression
5 KB file size

Image artifacts appear
as dark discoloration at
the bridge of the nose,
and lightening together
with slight blockiness at
the temple.



Shown here:
Both as scanned and
color enhanced
versions.



85% Compression
3 KB file size

Image artifacts appear
as blocky discoloration.
Compression brings
similar colors together,
resulting in the block
effect.



Shown here:
Both as scanned and
color enhanced
versions.



SECTION 7

Scanning

Creating Directories

- Before scanning, the toolkit should create the folder(s) in which you save the scanned images. For each item a separate folder is created at C:/DLOC/ with the appropriate dLOC ID, for example
 - C:/DLOC/CA00000001/00001
 - C:/DLOC/CA00000002/00001
 - C:/DLOC/CA00000003/00001

dLOC Requirement for Digital Master Files.

8-bit Grayscale or 24-bit RGB Color (depending on whether original has significant color)
300 dpi for standard text or 600 dpi for stand-alone images (photographs, maps)
Save archival files as uncompressed TIFFs

Flatbed Scanning: Epson Expression 10000 XL

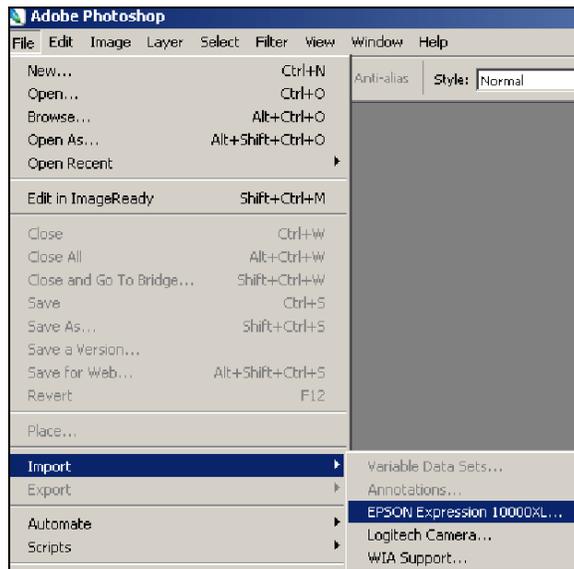
The following screenprints are specific to the Epson Expression scanner, but the same settings apply to any flatbed scanner.

Scan Settings

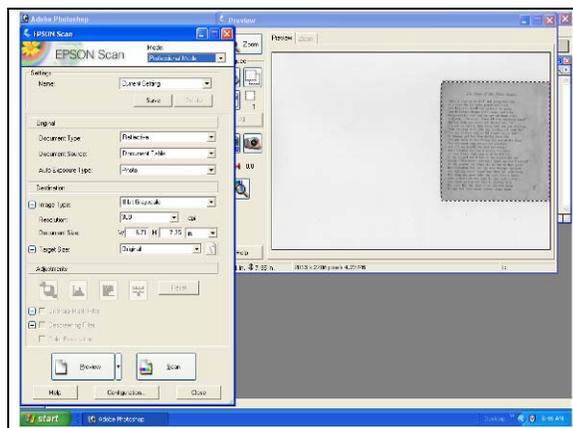
Scan documents using Adobe Photoshop rather than using the scanner's stand-alone image capture software and check the scanner settings with each new document.

Before opening Adobe Photoshop, turn on the scanner and make sure that the bed is clean and free of any dust, debris, etc. If necessary clean the glass with a lint free cloth and a very small amount of glass cleaning fluid.

1. Launch Adobe Photoshop
2. Select: *File* → *Import* → *Epson Expression 10000 XL*



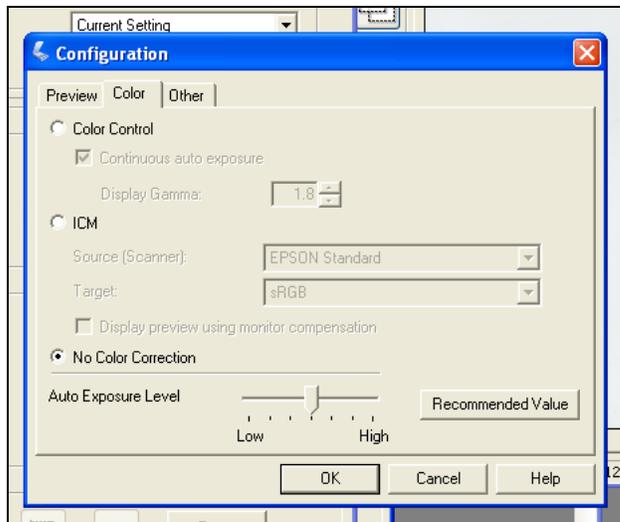
The scanning interface will then open two windows: a scan settings window and a preview window (*as seen here*)



3. Select the appropriate settings for your document
 - a. At the top of the scan settings windows select PROFESSIONAL MODE
 - b. Always select the following settings
 1. Documents Type: REFLECTIVE
 2. Document Source: DOCUMENT TABLE
 3. Auto Exposure Type: PHOTO
 4. Document Size: DO NOT ADJUST
 5. Target Size: ORIGINAL
 - c. Select the appropriate color space and bit depth
 1. 8-bit grayscale for items without significant color
 2. 24-bit RGB color for all other items
 - d. Select the appropriate resolution
 1. 300 dpi for mostly textual items
 2. 600 dpi for stand-alone image items (photographs, maps, etc..)
 - e. Click the **CONFIGURATION** button (*below the Preview and Scan buttons*)



1. Click the COLOR tab
2. Select NO COLOR CORRECTION
3. Click OK



Scanning

1. Place item, image down, on scanner glass. Be careful to place item as straight as possible in order to save time later. Close the scanner lid as much as item permits.



2. Click the **PREVIEW** button in the Scan Settings window. A small preview of your image will appear in the preview window. Make sure the entire document is visible, if not reposition on glass and re-preview.
3. Draw a bounding box around your entire image. If your original has 2 pages facing each other, draw a second box by selecting the dual marquee button . Arrange each box to completely include each side of the item. Once you are satisfied with the boxes positioning click the **ALL** button. DO NOT move the boxes or change settings after pressing this button!
4. Click the **SCAN** button



Saving Files

1. Save your image by selecting: *File* → *Save*
2. Select the dLOC ID folder that corresponds to which the image being saved belongs. E.g., In separates folder at C:/DLOC/ with the appropriate dLOC ID, for example
C:/DLOC/CA00000001/00001
C:/DLOC/CA00000002/00001
C:/DLOC/CA00000003/00001
3. Type in a sequential four digit file name, such as *0001*, *0002*, *0003*, etc.
4. Select *TIFF* from the file format drop down menu
5. Always uncheck the ICC profile box



6. Click Save
For TIFF Options select same as below





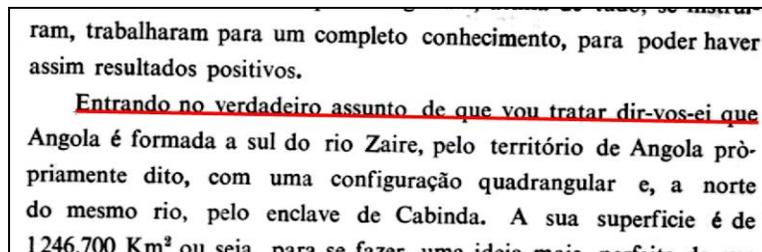
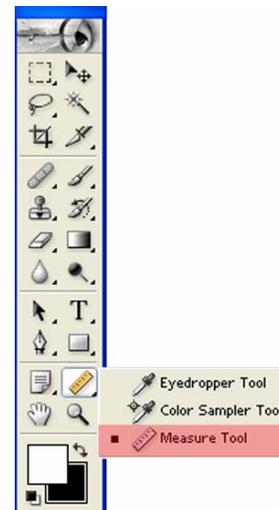
SECTION 8

Image Correction

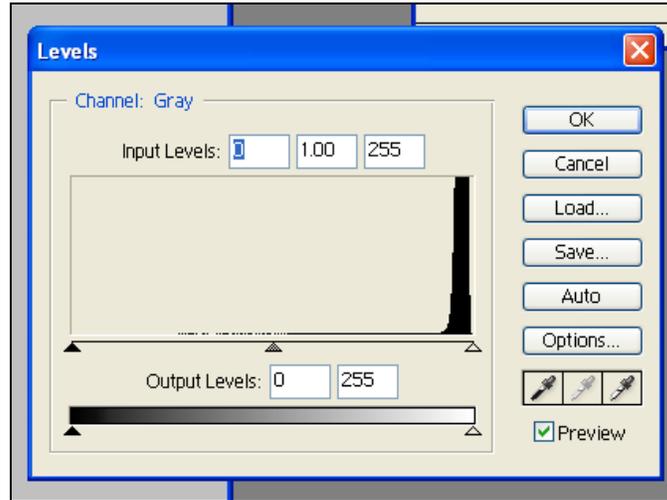
The intent of any digitization should be a faithful reproduction of the original document. Toward this goal, images will need to be deskewed and cropped to fit the in-hand original. In addition, it may be desirable to perform color correction either to reproduce the in-hand original, or the original state of the document. Applying these techniques in Adobe Photoshop is the topic of this section.

Image Correction in Adobe Photoshop

1. To straighten drastically skewed images:
 - a. Click and hold the Eyedropper Tool in the Photoshop tool box -
Select the Measure Tool
 - b. Click and draw a line to follow the bottom of any printed text, line or image
(*line is red, here, for purposes of illustration*)



- c. Select: *Image* → *Rotate Canvas* → *Arbitrary (DO NOT change the angle)* → click *OK*
2. Crop the image to remove any excess borders added during straightening using the crop tool 
3. If necessary (e.g., if the image is muddy), adjust the levels/histogram by selecting *Image* → *Adjustments* → *Levels*

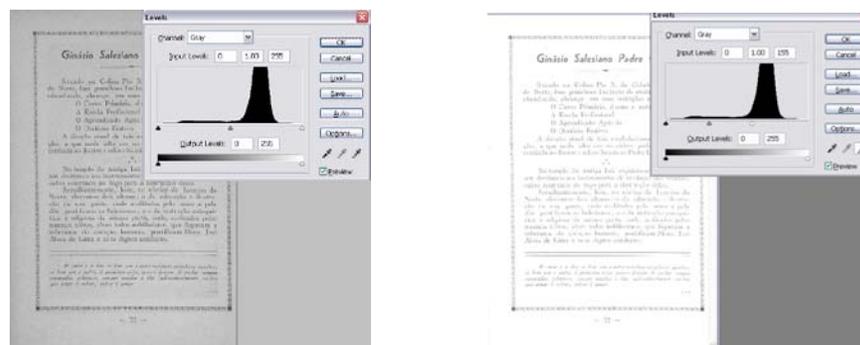


Adjust the black, white and midpoints to improve your image quality and contrast.

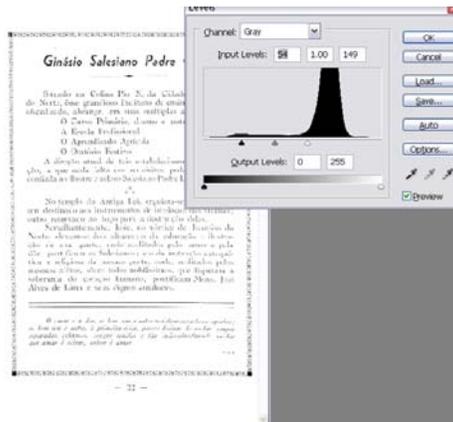
If the image is COLOR, you may make histogram adjustments for each RGB channel: Red, Green and Blue.

But, do not over correct and eliminate detail. A histogram shows the distribution of tones over a range. The image characterized by the histogram above is predominantly white. While the image contains shades of gray, deeper tones of black are almost entirely absent.

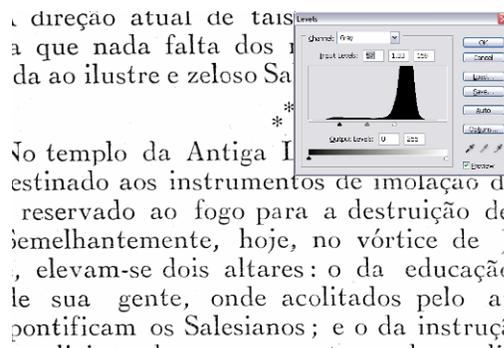
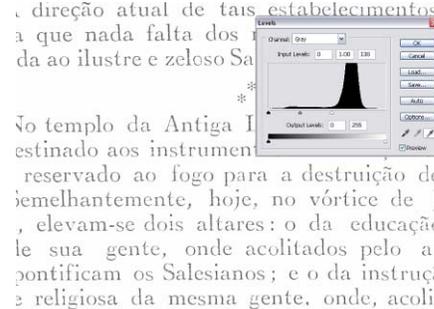
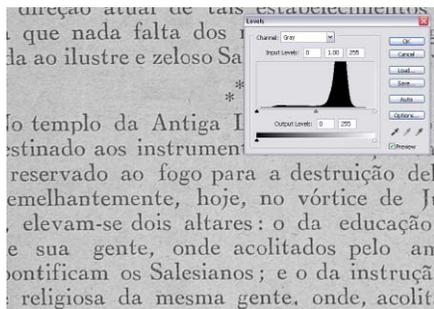
4. Images with good, thick printed text can also be quickly corrected by selecting the document's white point. This is done by opening the levels/histograms by selecting *Image* → *Adjustments* → *Levels*. In the levels window select the eyedropper furthest to the right and then select the point in your image that should be the brightest white. The images below show this effect before and after the white point selection.



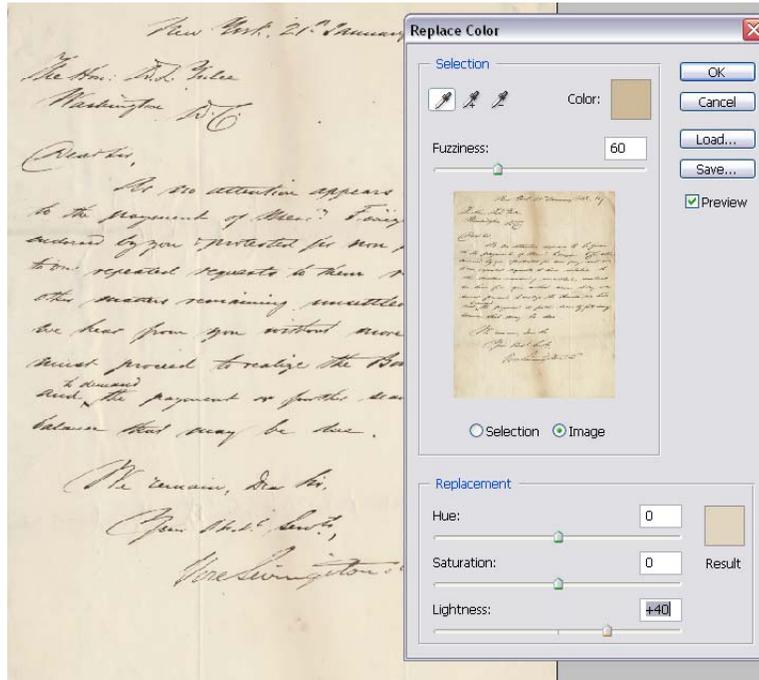
You will notice that the background becomes almost uniformly white, but the text is also lightened. Before selecting OK in the levels/histograms you will need to bring in the black point in order to improve the text. This is done by moving the arrow furthest to the left, in towards the right. You will notice that the numbers in the Input Levels boxes increase.



It is helpful to perform this correction while zoomed in to 100% on your image, as shown below.



5. If the image is extremely stained the document should be scanned in RGB and if possible, the stains should be lightened using *Image* → *Adjustments* → *Replace Color*



Select "Image" and not "Selection" in the Replace Color Window. Then using the eyedropper tool select the darker color of the stain. Adjust the Lightness, Saturation and Hue slider bars as needed to minimize the stains. The fuzziness meter indicates how closely a color must match the selected color to be replaced. Be aware that stains may be similar in color to text and therefore too much manipulation is undesirable in order to not lose information.

Often it is useful to zoom into one section of text while performing the color replacement. One must be careful not to make the text harder to read for the OCR engine.

- Remember that any adjustments done to images can be undone as long as the file remains open. Maintain your history window open by selecting *Window* → *History* in Photoshop, then simply select the previous step done. You can always go back several steps and re-correct your image.

Other Adobe Photoshop Resources

The original Adobe Photoshop installation package should include a tutorial of the software you purchased.

In addition, Adobe has an on-line resource at the following URL:

<http://www.adobe.com/products/tips/photoshop.html>

Adobe, the Adobe Logo, and Photoshop are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.



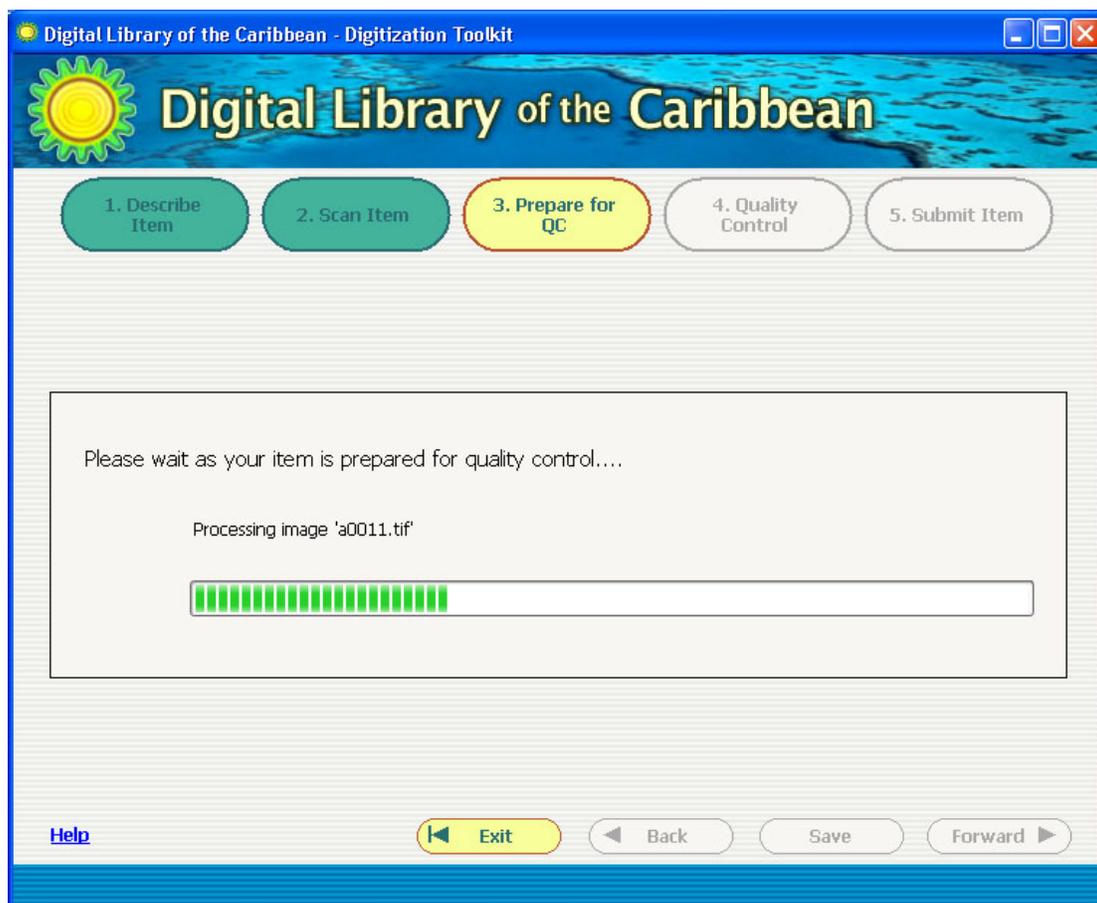
SECTION 9

Performing Quality Control

Prepare for QC

Once the scanning is complete, the next step is to prepare your item for QC processing.

This assembles the metadata file for the package, and also creates the JPEG derivatives. The JPEG derivatives are used for the internet, as well as the QC step.



Once complete, your item will be moved to the next step.

Quality Control



DIGITAL LIBRARY OF THE CARIBBEAN (dLOC)

The Quality Control step allows you to view all the images and add structural metadata.

Once the resource is opened for this step, you will see an image of each page from the original resource. The first time a package is QC'd, all of the pages will be added to a single, main chapter.



- The toolbox in the upper right hand corner gives the user access to a number of features.

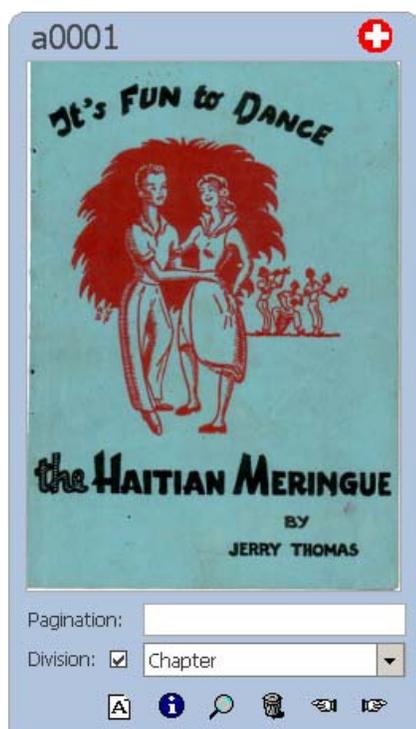




- The table on the next page shows the function of each of the icons.

	This saves all of the entered data into the metadata file for this digital resource.
	This selects the size for the thumbnail images.
	This is the default cursor.
	This cursor allows you to zoom into any page image.
	This cursor allows you to select two pages to switch, or to insert one page in front of another.
	This cursor allows you to set the image that will be used as the main thumbnail for your item.
	These arrows take you to the previous and next page image that already had a QC error associated with it.
	This views the metadata for the current digital resource

Next, you will assign division information and page numbering information to each page of this book. When you mouse over each image, a toolbox will appear below that image.

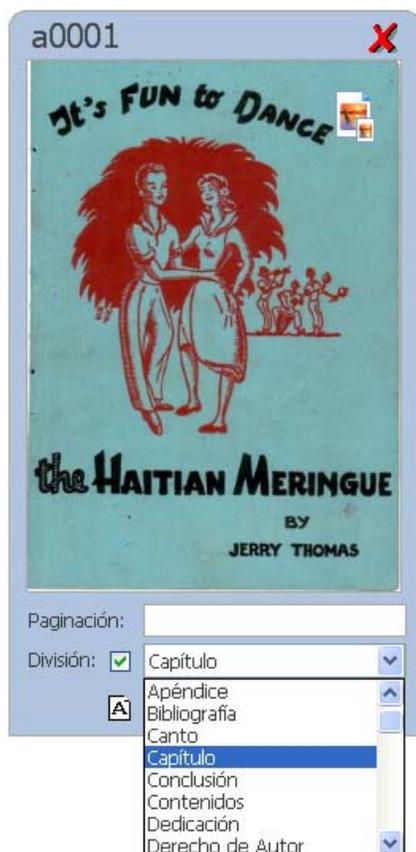


This toolkit can be used to get additional information about this image, zoom into the image, or delete the image. The table below shows the function of each element of this toolbox:

	Indicates that there is text on this page. Clicking this icon toggles the text-indicating 'A'
	This provides technical details about the original page image.
	This zooms into this page image
	This deletes this page image and remove it from the list of images
	This indicates that there was a page in the original which was not scanned either before or after this page.

The pagination text box allows you to name, or number, each page.

The division check mark indicates this page is the beginning of a new division. Once that is selected, you can choose the division type from the combo box.



The following are the main division types available:

Abstract	Errata	List of Tables
Acknowledgement	Foreward	Methodology
Advertising	Front Cover	Poem*
Appendix*	Front Matter	Preface
Back Cover	Frontispiece	Prelude
Back Matter	Half Title	Reference
Bibliography	Index*	Spine
Chapter*	Interview	Table of Contents
Conclusion	Introduction	Title Page
Copyright	List of Figures	
Dedication	List of Illustrations	



In addition to the divisions listed above, there are seven more available divisions. These are:

Subdivision Level 1*
Subdivision Level 2*
Subdivision Level 3*
Subdivision Level 4*
Subdivision Level 5*
Subdivision Level 6*
Subdivision Level 7*

These elements are available to suggest a deeper hierarchy. For example, if the table of contents of a book looks like:

- Elephants in Nature
 - Species
 - African
 - Asian
 - Diets
 - Vegetarian Diets
 - Carnivorous Diets

These divisions would be entered as:

- Chapter 'Elephants in Nature'
 - Subdivision Level 1 'Species'
 - Subdivision Level 2 'African'
 - Subdivision Level 2 'Asian'
 - Subdivision Level 1 'Diets'
 - Subdivision Level 2 'Vegetarian Diets'
 - Subdivision Level 2 'Carnivorous Diets'

Then, the on-line table of contents of this resource will appear as a hierarchical tree.

TABLE OF CONTENTS

- Front Cover
- Title Page
- Elephants in Captivity
- [-] Elephants in Nature
 - [-] Species
 - African
 - Asian
 - [-] Diets
 - Vegetarian Diets
 - Carnivorous Diets
- References
- Back Cover



Some of the divisions (those with asterisks on the previous tables) can be given a resource-specific name. When you select any of these divisions, the following form is displayed to allow the name to be entered.



Below is an example of a page with a named division.

a0010

**THE STEPS
OF THE HAITIAN MERINGUE**

**BASIC STEPS OF THE QUICK
MERINGUE**

- Sideward step
- Natural and reverse turn
- Country hip movement
- Fancy head movement
- Right turn
- Ho variation
- Caribbean variation
- Country dance variation
- Whirling

**BASIC STEPS OF THE SLOW
MERINGUE**

- Forward and backward movement
- Rectangular steps
- Quarter turn

VARIATIONS

- Simple Balancé Yoyo
- Double Balancé Yoyo

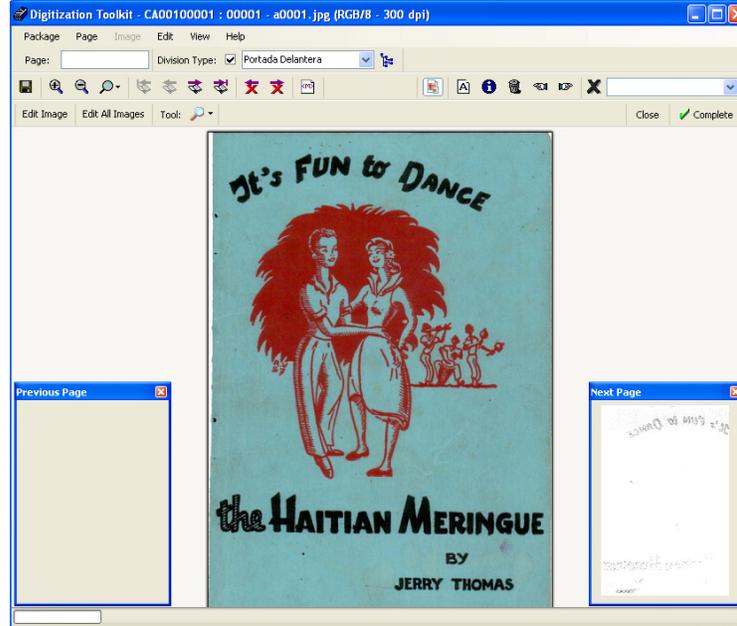


— 8 —

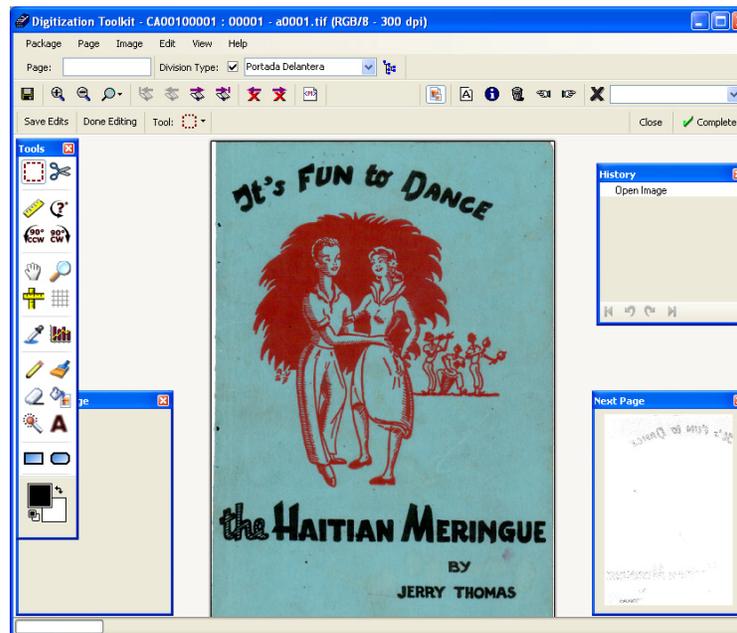
Pagination: Page 8

Division: The steps of the Haitian...

You can view an image zoomed in by double clicking on the image, or selecting the zoom from the image toolbar.



From this form, selecting 'Edit Image' will allow you to edit the original TIFF image.



You can click 'Save' at any time to save your work. Moving to the next, and final, step will also save all of your changes.

Creation of your digital resource is now complete and you will be asked for your preference for submittal. You can either submit immediately or wait to submit later.