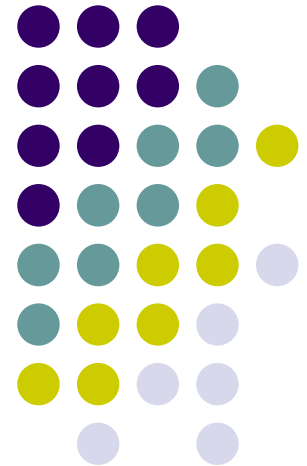


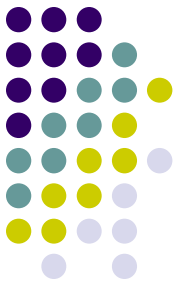
# dLOC Toolkit

**Mark V. Sullivan**

Solutions Developer, University of Florida Libraries  
Programmer and Trainer, Digital Library of the Caribbean

[MarSull@uflib.ufl.edu](mailto:MarSull@uflib.ufl.edu)





- Final Web Product
- UF Digital Library Center Processes
- Digital Library of the Caribbean [dLOC]
- dLOC Toolkit

# UF Digital Library Center



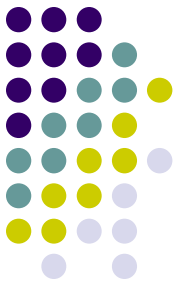
- Established about 1998 from Preservation Dept.
- Slowly increased in capacity and content through grants, endowments, and library funds.
- As of 2007...
  - ~10 TB's of internally created data
  - ~20 TB's of vendor data

# The more things stay the same..



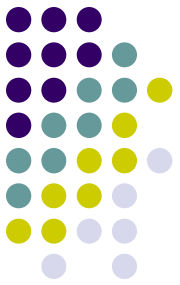
## Format migration

- Image format for service
  - Metadata format
- 
- Other changes
    - Software
    - Hardware
    - Processes



# How to cope...

- Embraced Standards
  - Image Standards
    - GIF → JPEG
    - DjVu, SID → JPEG2000
    - TIFF
  - Metadata Standards
    - Dublin Core
    - METS
    - MODS



# How to cope...

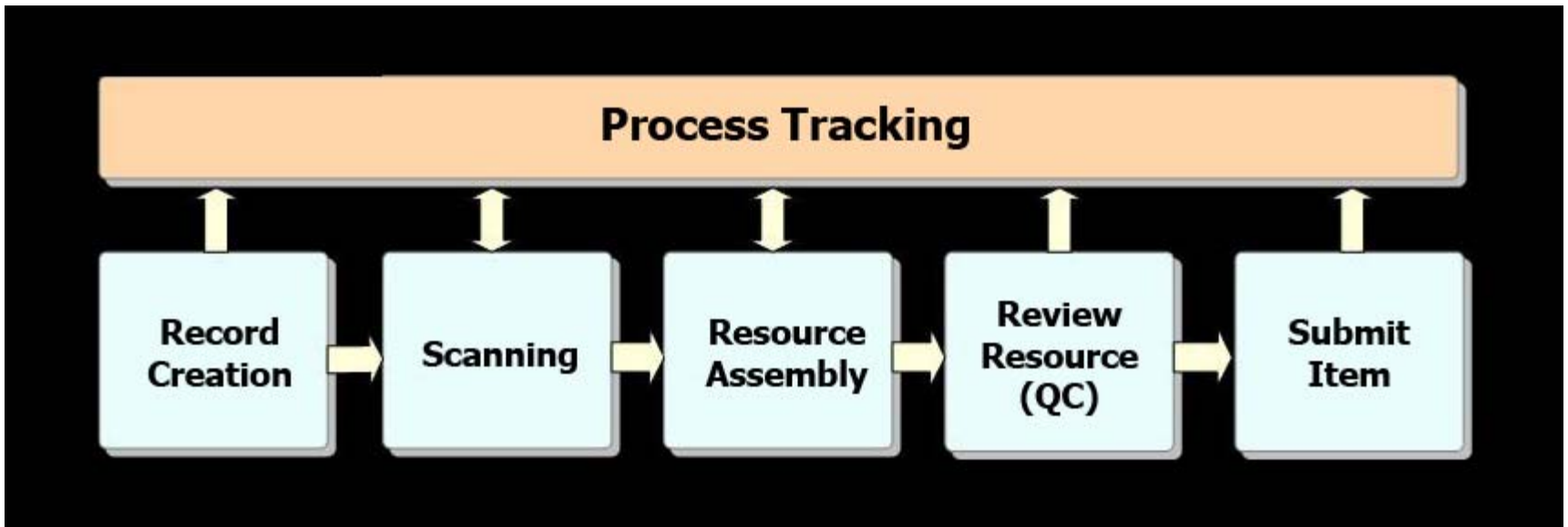
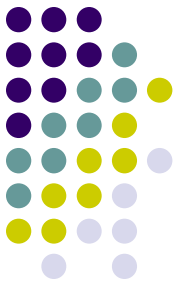
- Embraced Open-Source Software
  - Look for 'market share'
  - Look for active community
  - May have steep learning / customization curve
  - May have to build upon or contribute to software
- Tried to remain agile

# UF Digital Library Center



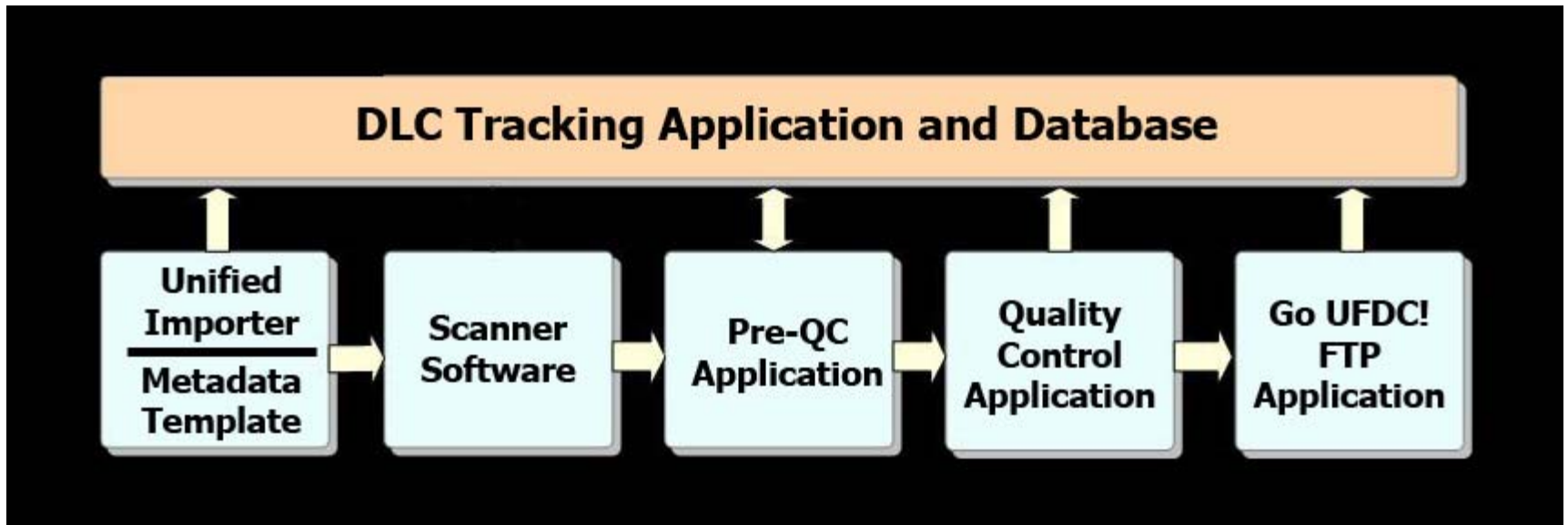
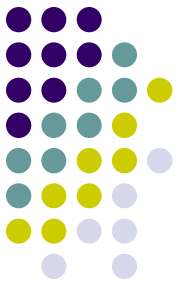
- Standards
  - TIFF archival image
  - JPEG and JPEG2000 service copies
  - METS / MODS metadata formats
- Software
  - Greenstone Digital Library Software
  - No appropriate process software

# Process Development

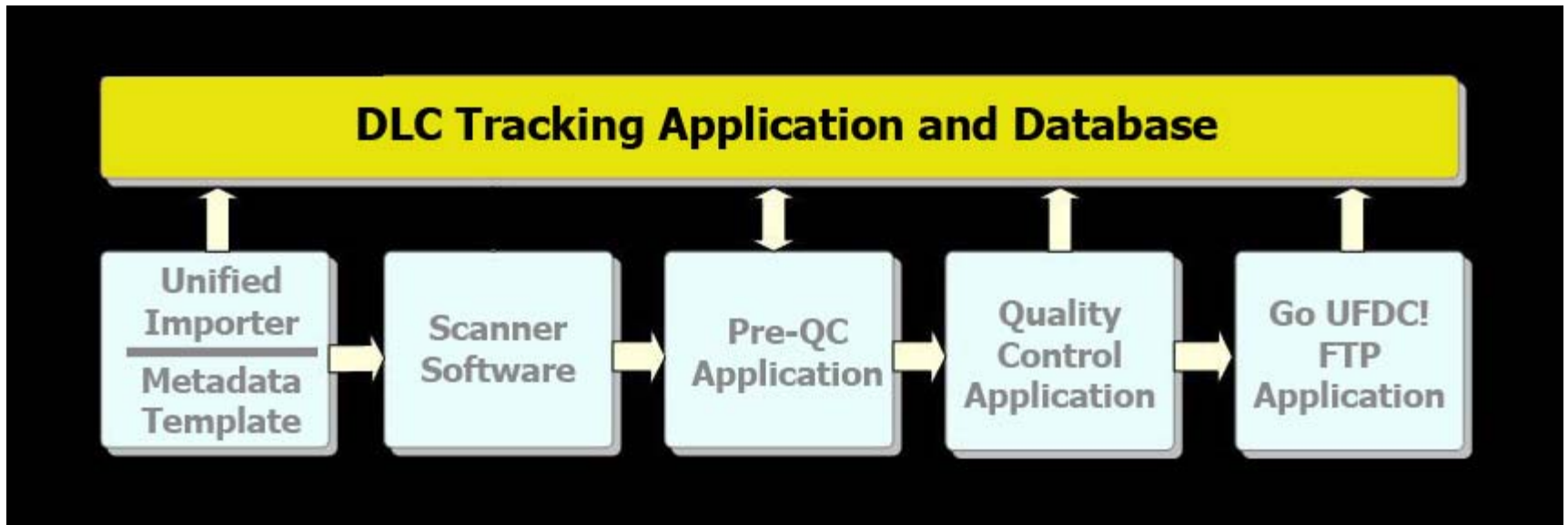




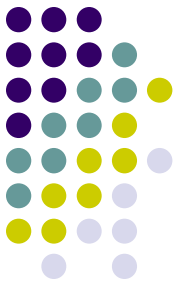
# Application Development



# Tracking Database Application



# Tracking Database Application



- Basic item information
  - Title
  - Author
  - Publisher
  - Copyright Permissions
- Digitization Status and History
- Archive Information (CD → DVD Archives)

## Digital Library Tracking Database Main Form

BibID	LTQF	LTUF	Aleph	Title	Author
UF00003666	AAA5147	NONE		Gainesville, Florida : showing the University of Florida main campus	
UF00004690	AAA5973	APC4753		Map of Gainesville, Florida /	
UF00015236	AAA7853			National Wetlands Reconnaissance Survey. Gainesville, Florida	U.S. Fish and Wil
UF00016763	AAA9430			Greater Gainesville area precinct map, 1986	
UF00016764	AAA9431			Gainesville bikeway system 1989	
UF00017143	AAA9758			Gainesville, Florida, urban area : drawn 1979.	
UF00017145	AAA9760			Urban area : Gainesville, Alachua County 1975	
UF00017146	AAA9761			Urban area : Gainesville, Alachua County 1975	
UF00017147	AAA9762			Urban area : Gainesville, Alachua County 1975	
UF00017252	AAA9881			Alachua County proposed comprehensive plan 1995 unincorporated	
UF00018404	AAB0246			FIRM, flood insurance rate map, Gainesville, Florida, Alachua County	
UF00018405	AAB0582			FIRM, flood insurance rate map, Gainesville, Florida, Alachua County	
UF00018406	AAB0583			FIRM, flood insurance rate map, Gainesville, Florida, Alachua County	

**Limit results by:**

- Institution 
  
 Project 
  
 Show All

**Search by:**

and

Total hits count = 34

# UF00023903

Title:

Author:

Institution

Summary

Source

Monograph

Electronic Resource

Archive

Status

## Bibliographic Records

BibID: **UF00023903**  
Aleph Bib Num: **001745103**  
LTUF: **AJF7877**  
LTQF: **AAB7355**

## Source Information

Material Type: **Monograph**

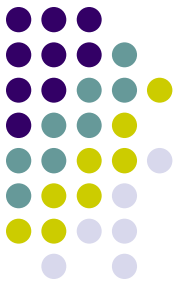
## Electronic Resource

Project Code: **JUV**

## Status

Mark-up: **Completed**

# Older Process



**DLC Tracking Application and Database**

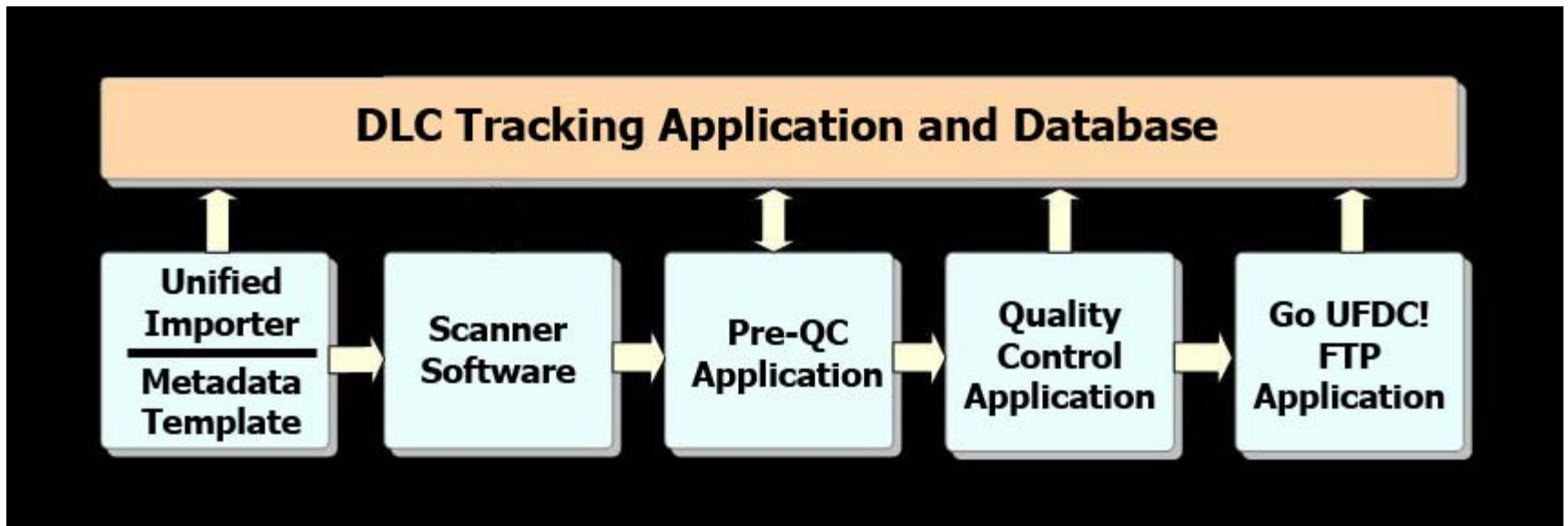
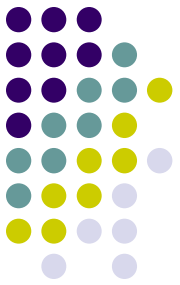
**Scanner  
Software**

**MXF Metadata  
Client**

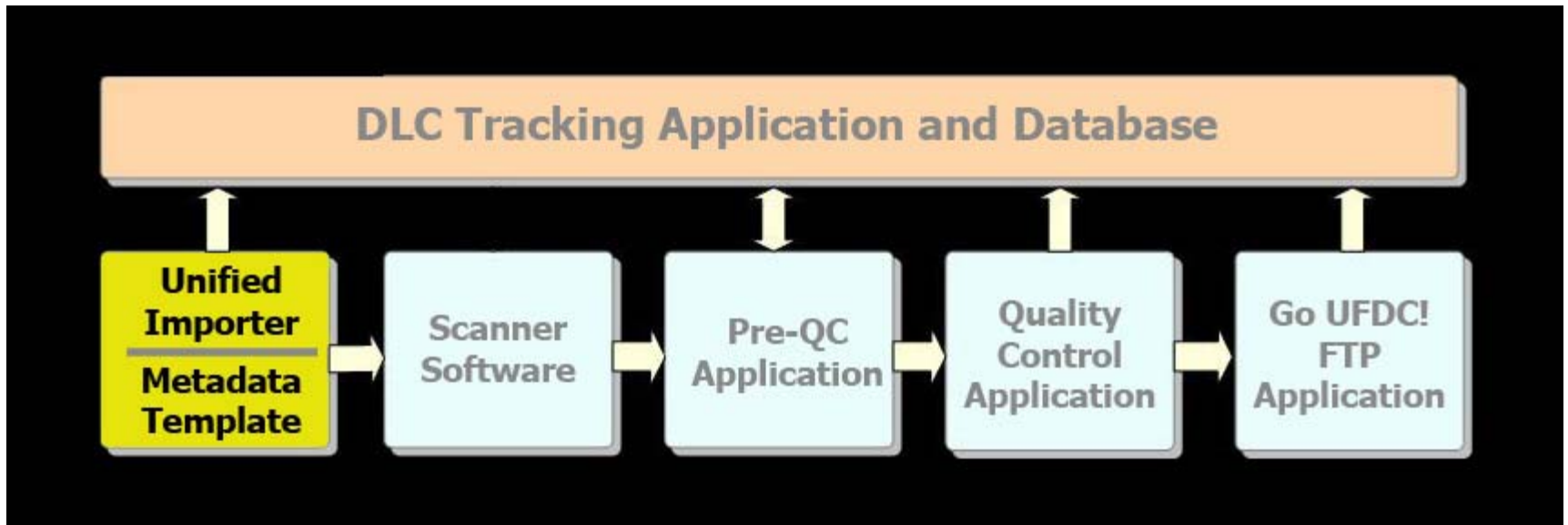
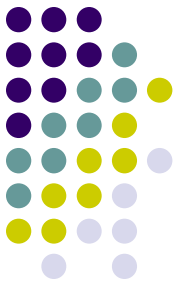
**Any FTP  
Client**



# Newer Process



# Record Creation

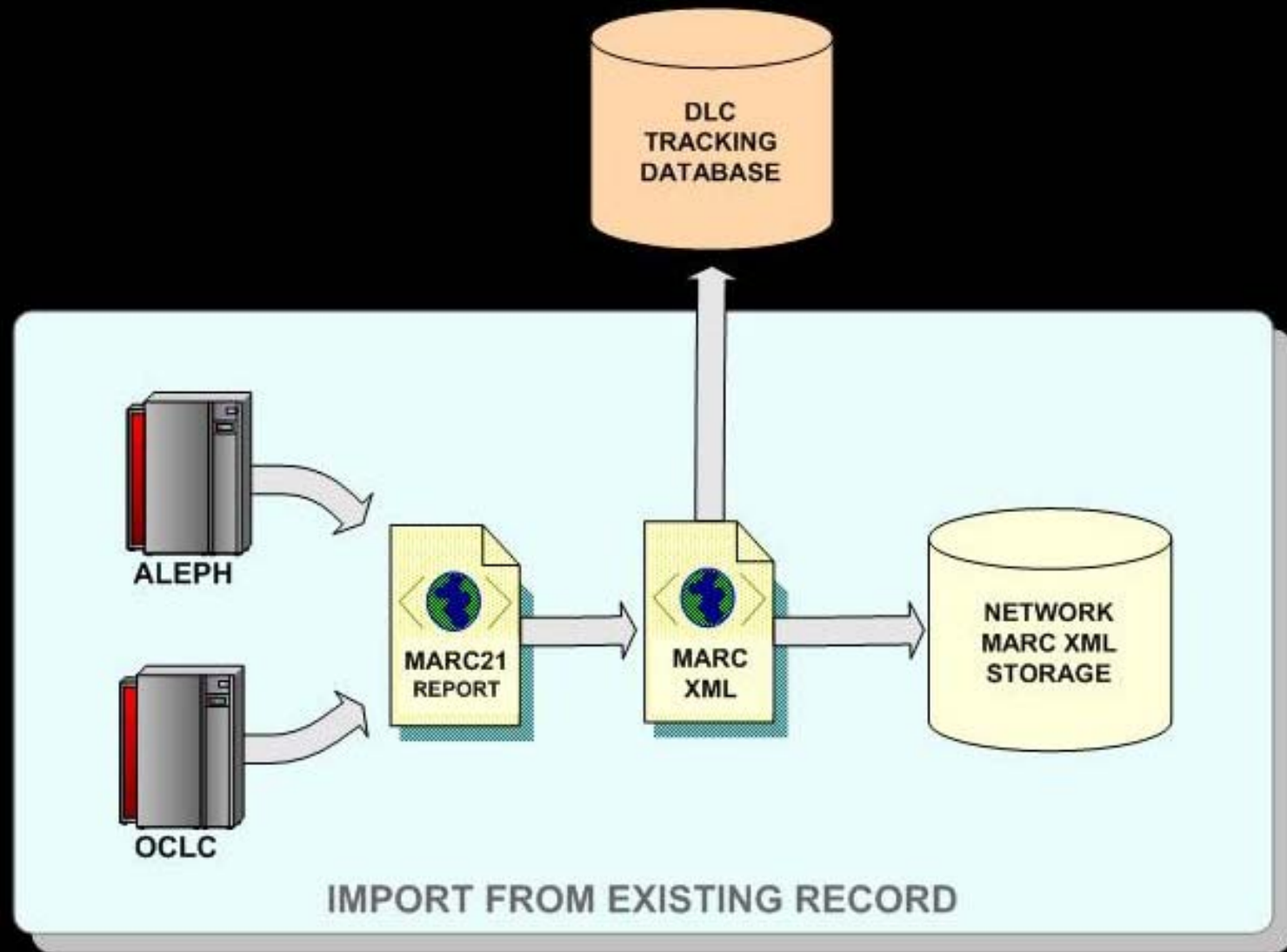
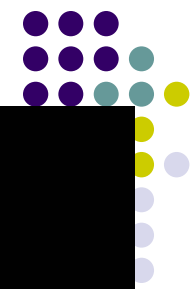




# Unified Importer



- Imports from existing records or data
  - MARC records (NOTIS, ALEPH, OCLC)
  - Excel Spreadsheets
  - Existing metadata (METS) files
- Saves data on network and in Tracking database



# Metadata Template



- Uses
  - Enrich existing records
  - Create new records
  - Edit data for online items
- Saves data as METS file
  - Network location
  - FTP back to online presence

UFDC Metadata Template - CA01000015 : 00001

Action View Options Help

General Information

Source Document

Subjects and Keywords

Behaviors

Serial

Files

Codes

Bibliographic Identifier: CA01000015

Volume Identifier: 00001

Record Status: NEW

Collection: DLOC

SubCollection:

Main Panel

Language: English

Resource Type: TEXT

Physical Description:

Identifier: 733.33 T458t

Type: None

Source Institution: NLJ

National Library of Jamaica (NLJ)

Physical Location:

UF

University of Florida Libraries

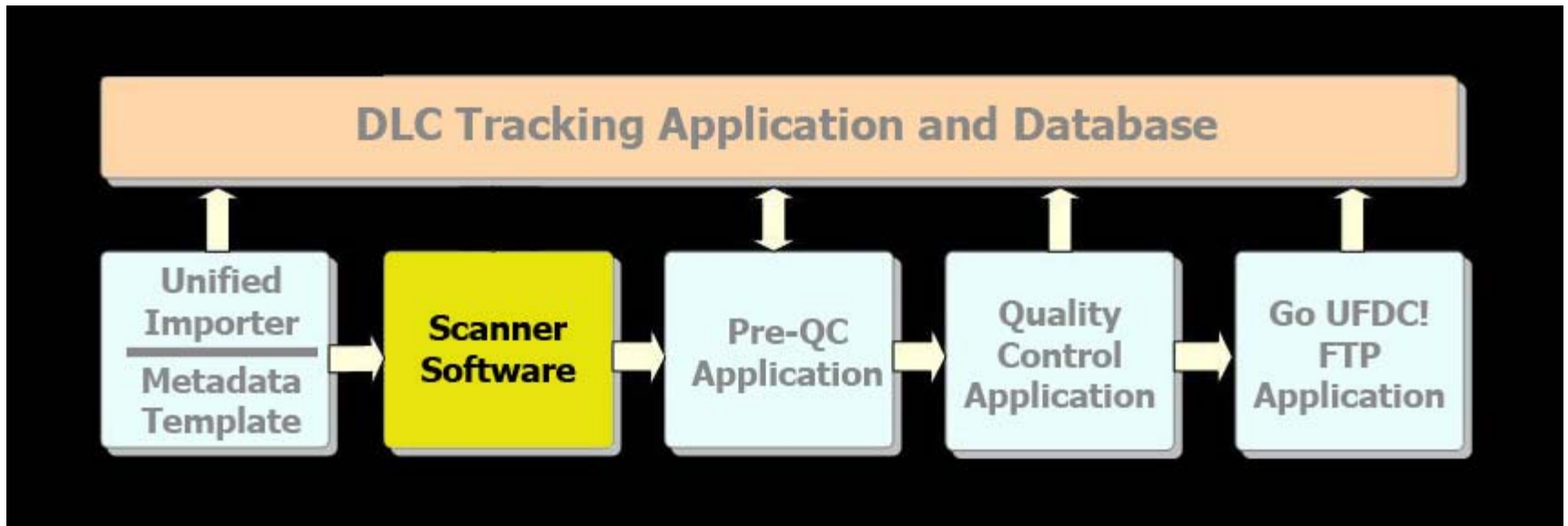
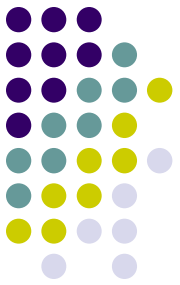
Edit Mode

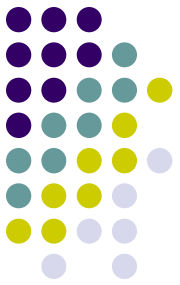
OK

Cancel

Apply

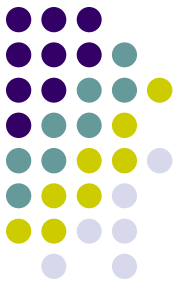
# Digitization





# Digitization Software

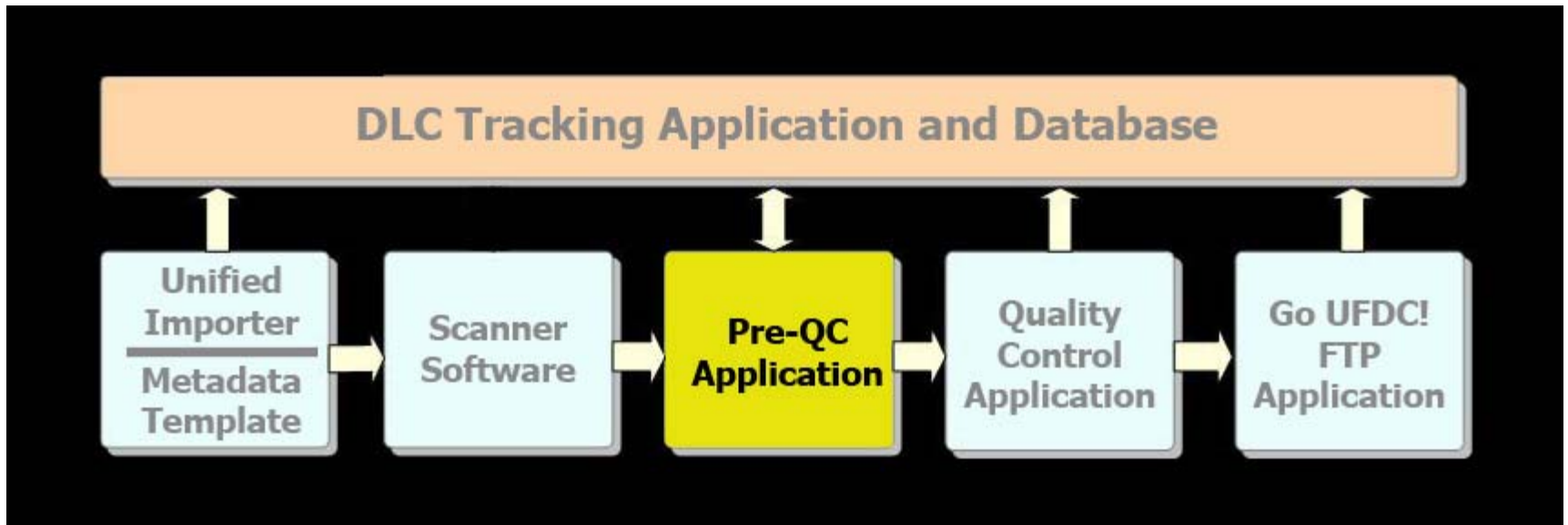
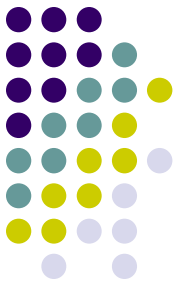
- Scanning
  - Adobe Photoshop
  - Scanner-specific software
- Image Capture
  - Copibook Scanning Software
  - Digital Cameras
- Other types of digitization
  - Audio / Video
  - Vendors (microfilm)



# Scanning Specifications

- Save as uncompressed master TIFF
- Color, Grayscale, Bitmap (black and white)
- DPI
  - Photographs, Postcards, Images : 600 dpi
  - Textual items : 300 dpi
  - Use your judgement

# Resource Assembly





# Pre-QC Application



- Assemble bibliographic metadata
  - MARC XML (Unified Importer)
  - Existing METS (Metadata Template, online)
  - Tracking Database
  - Project specific data



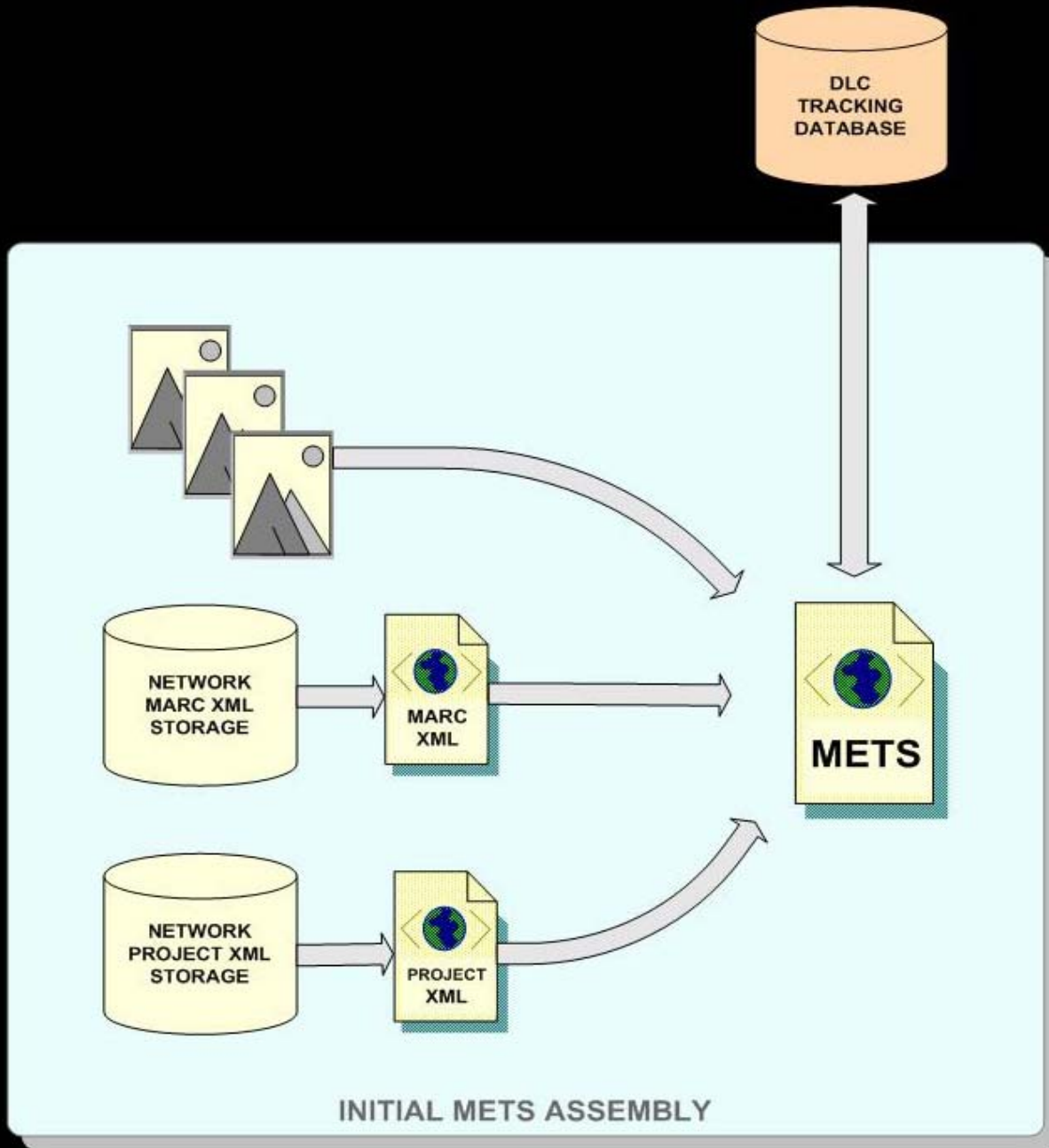
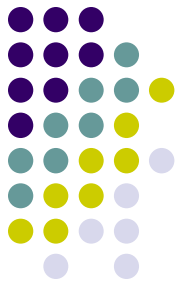
# Pre-QC Application

- Assemble file information
  - Scanned master TIFFs
  - Checksums performed
- Structural Metadata ( Table of Contents )
  - All pages linked to one chapter
  - TOC will be built in next step
- Update Tracking Database

# Pre-QC Application

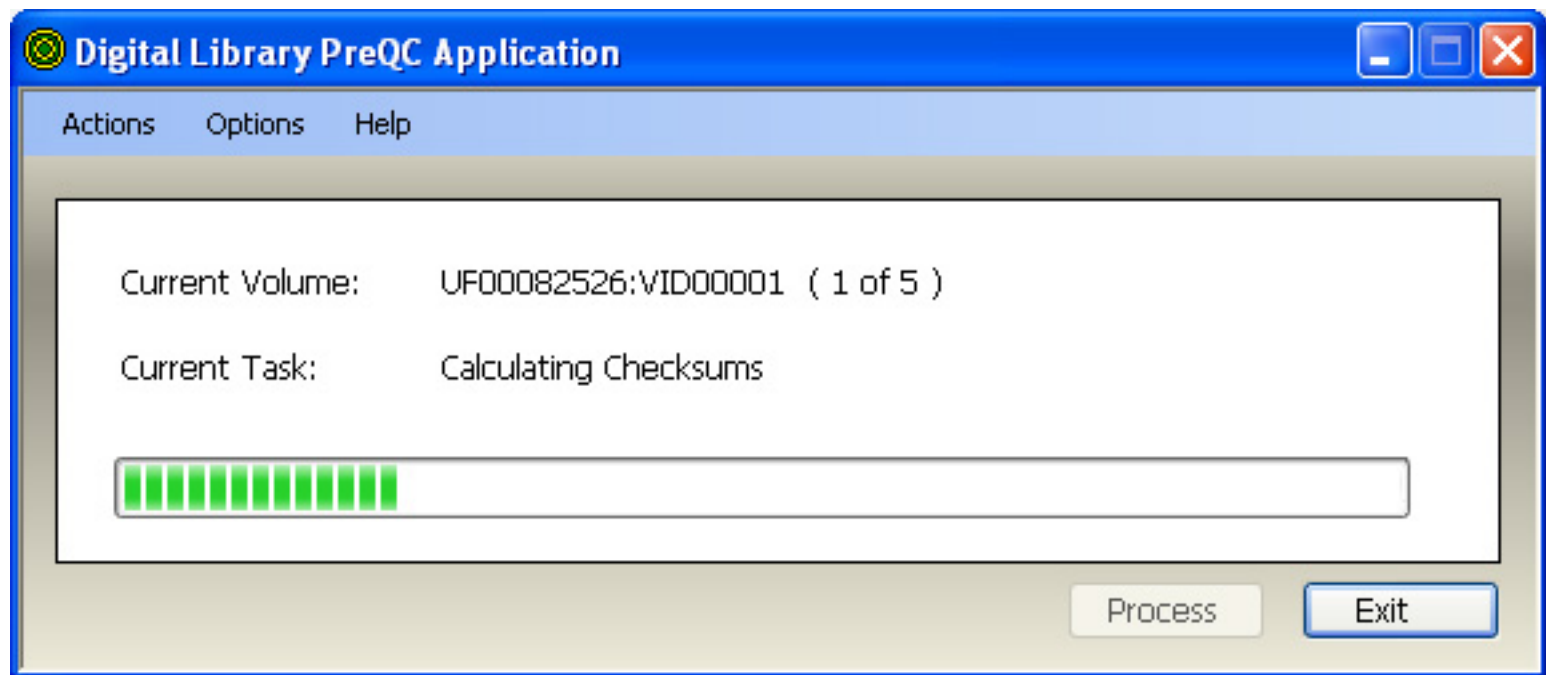
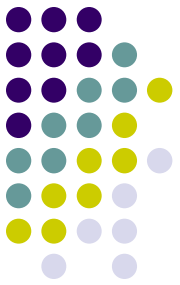


- Create image derivatives
  - .QC.jpg ~315 pixels wide
  - .jpg ~630 pixels wide
  - thm.jpg ~150 pixels wide
  - .jp2 zoomable image

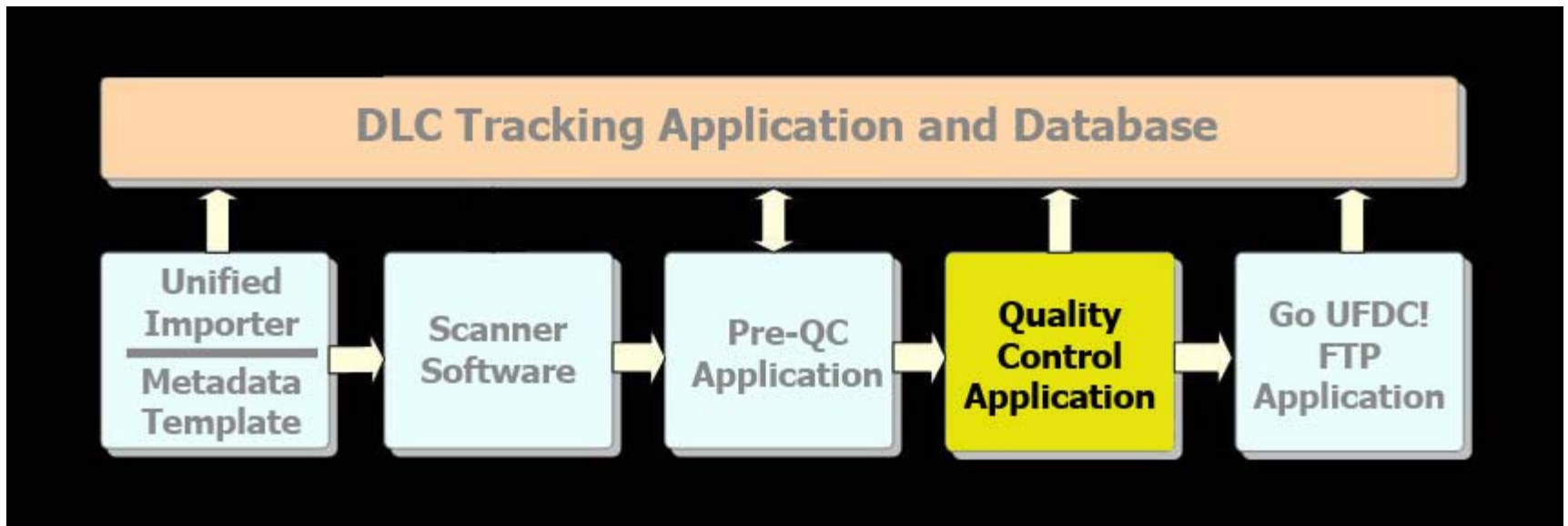
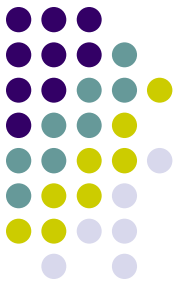


INITIAL METS ASSEMBLY

# Pre-QC Application



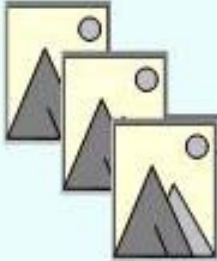
# Review Resource (QC)



# Quality Control Application



- Visually...
  - Inspect (and correct) any image issues
  - Ensure metadata and images match
  - Ensure pages in correct order
- Structural metadata
  - Define page names
  - Define divisions or chapters
- Update Tracking Database



ADD STRUCTURAL METADATA







a0001

It's FUN to DANCE

the HAITIAN MERINGUE

BY JERRY THOMAS

Pagination:

Division:  Front Cover

a0002

Pagination:

Division:  Front Cover

a0003

THEORY OF THE HAITIAN MERINGUE

BY JERRY C. THOMAS

MEMBER OF THE ASSOCIATION OF HAITIAN DANCE ARTISTS OF PARIS AND AFFILIATED ARTISTS, INC.

Pagination: Page 1

Division:  Title Page

a0004

703.93 T4587

Pagination: Page 2

Division:  Title Page

a0005

Foreword

a0006

Notes

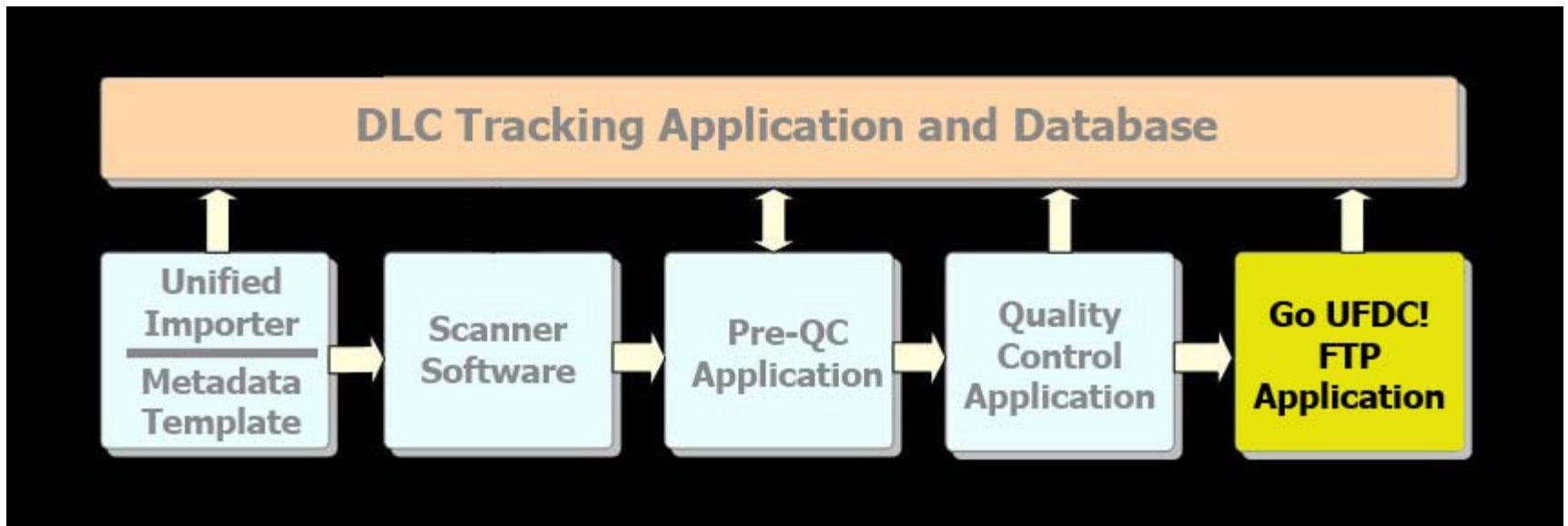
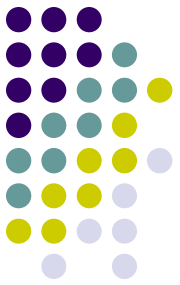
a0007

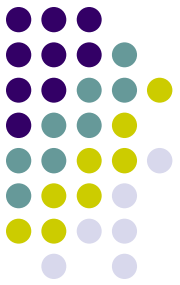
HAITIAN MUSIC stems directly from African Rhythms. Divested of mysticism, its traditions and beliefs from the folklore basis from which the Haitian Meringue, the

a0008

Congo Dance Ibo Dance Mayayo Dance  
Nago Dance Malo Dance Petro Dance  
Yambou Dance Quita Dance Martinique Dance  
Quadrille Bahomey Dance Bards Dance  
Cote Dance

# Submit Item (and more)

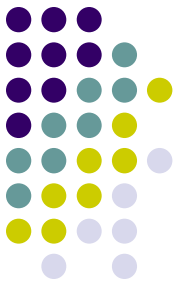




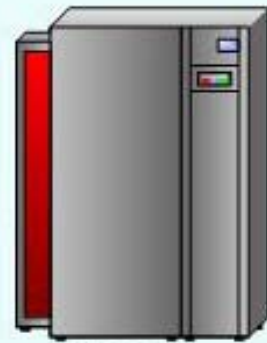
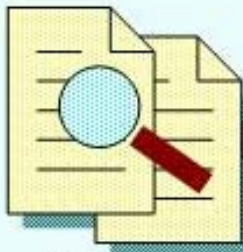
# Final Preparation

- Optical Character Recognition [OCR] performed on the images
- Metadata manually reviewed
- Any additional project-level enhancement completed

# Go UFDC! FTP Client



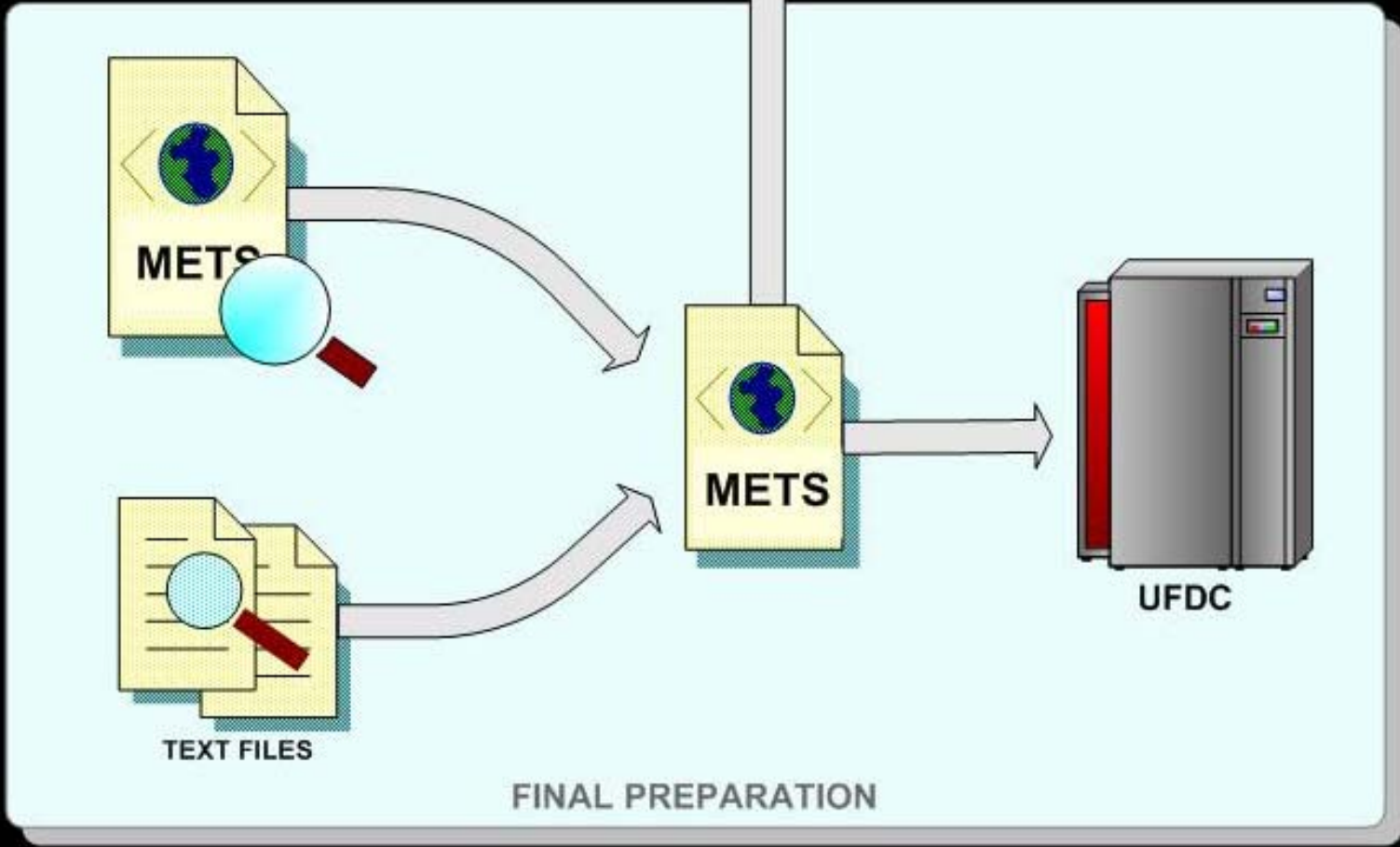
- Automatic final preparation
  - Text images added to the resource package
  - Page naming reviewed and revised
  - Final metadata file validated against schemas
- Item FTP'd to online presence (UFDC)
- Update Tracking Database



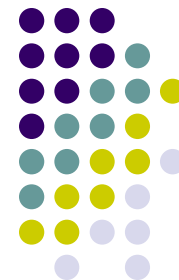
TEXT FILES

UFDC

FINAL PREPARATION



# Go UFDC! FTP Client



Go UFDC!

Actions Tools FDA Preferences Help

Folder: C:\Documents and Settings\Mark\Desktop\TEST\QC\Final

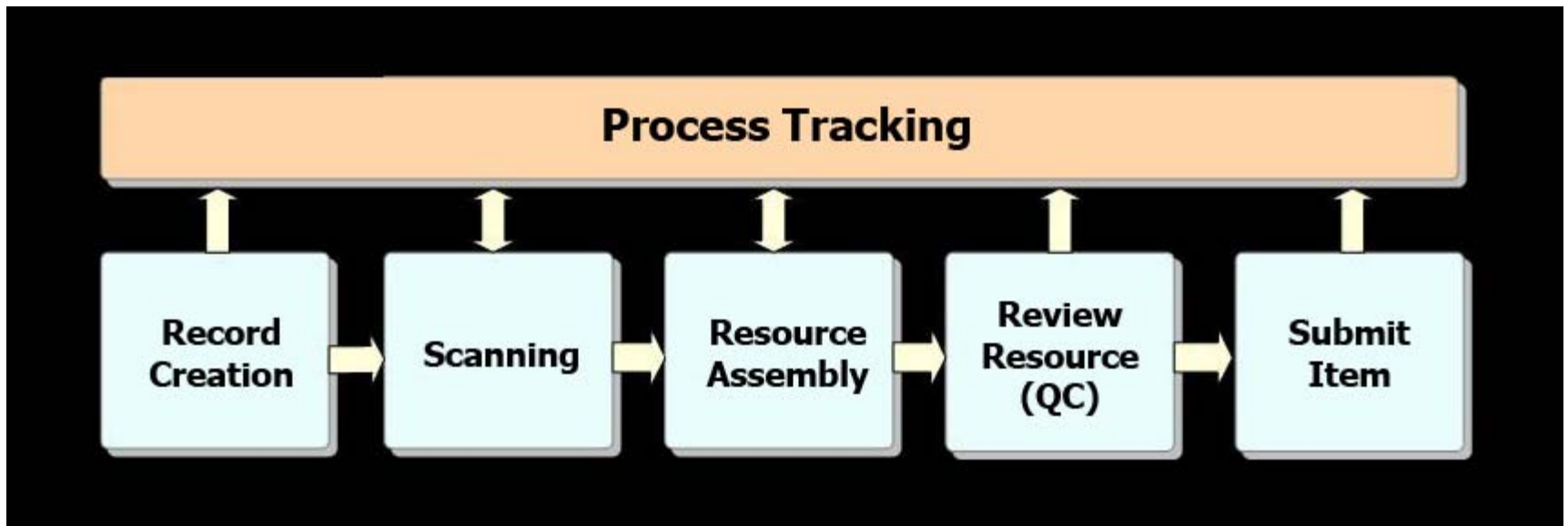
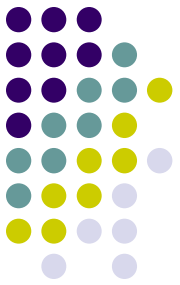
Current Volume: CA01000015\_00001 [ 1 of 1 ]

Current Task: Reading METS File

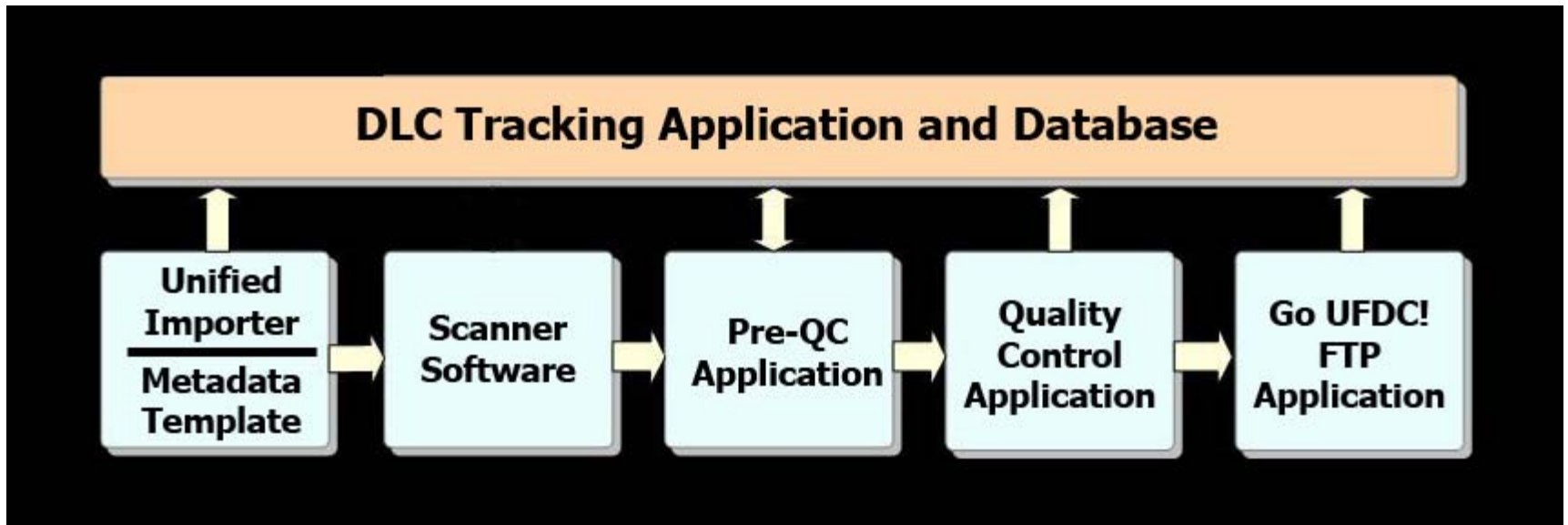
Progress:

Cancel  METS Only!

# Digitization Process



# Application Pathway





# Digital Library of the Caribbean



- Timeline
  - 1998 : Conceived
  - 2000 : Building Partnerships
  - 2004 : Grant received from U.S. Department of Education

# Digital Library of the Caribbean



- Concept
  - Provide a digital voice to those institutions in the Caribbean who wished to contribute
  - Build infrastructure through equipment and training throughout the Caribbean
  - Assist in digital preservation of the resources threatened materials

# Digital Library of the Caribbean



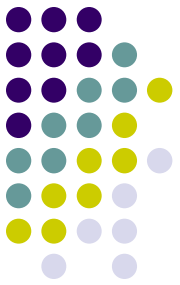
- Original Partners
  - Archives Nationales d'Haiti
  - CARICOM
  - La Fundación Global Democracia y Desarrollo
  - National Library of Jamaica
  - Universidad de Oriente, Venezuela
  - University of Virgin Islands
  - Florida International University
  - University of Central Florida
  - University of Florida

# Digital Library of the Caribbean



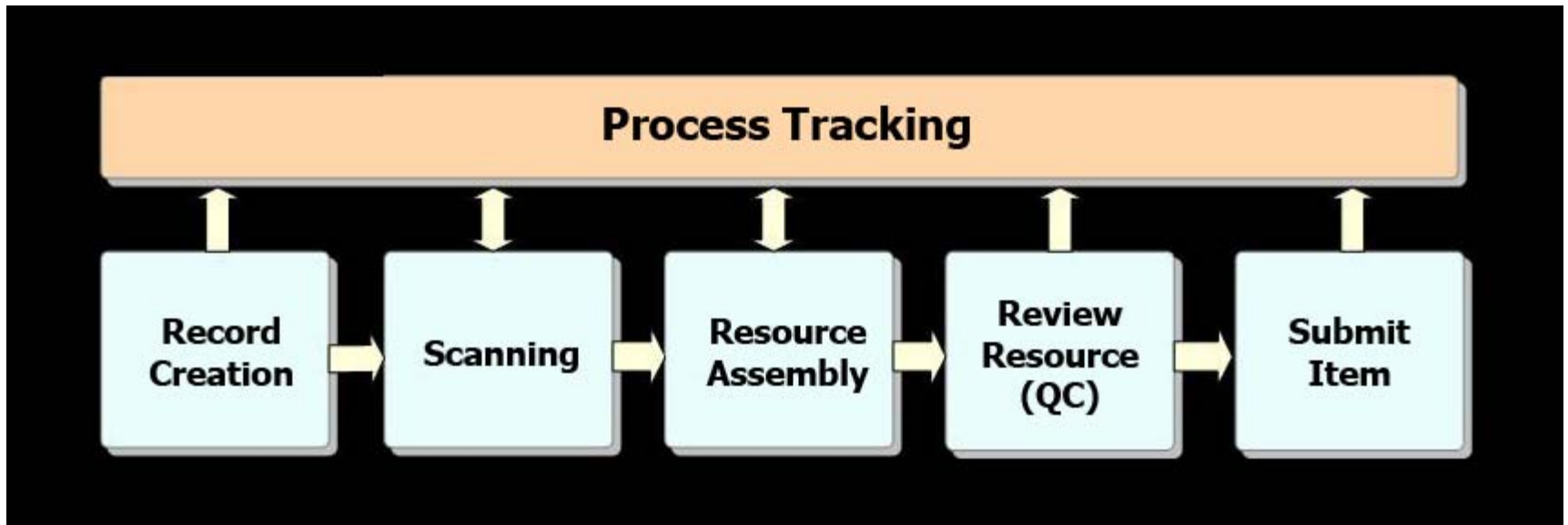
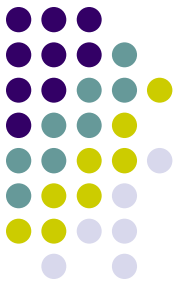
- New Partners
  - Archives Nationales d'Haiti
  - Bibliotheque Hatienne des Peres Saint-Esprit
  - CARICOM
  - The College of the Bahamas
  - La Fundación Global Democracia y Desarrollo
  - National Library of Aruba \*
  - National Library of Jamaica
  - Pontificia Universidad Catolica Madre y Maetsra (D.R.) \*
  - Universidad de Oriente, Venezuela
  - University of Virgin Islands
  - Florida International University
  - University of Central Florida
  - University of South Florida
  - University of Florida

# Digital Library of the Caribbean

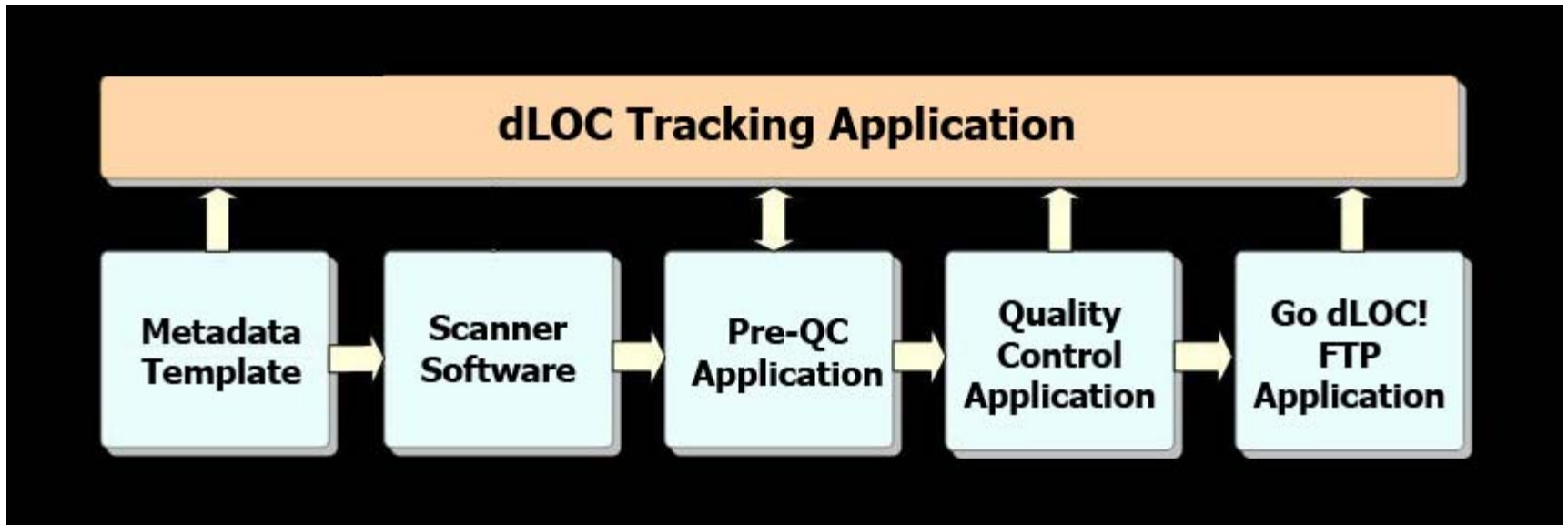
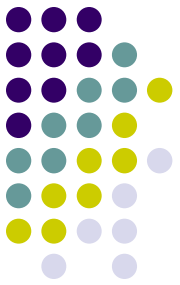


- University of Florida – Technological Role
  - Create distributed tool for resource creation and submission
  - Assist with general digitization training
  - Create and maintain web presence
  - Submit resources to dark archive for partners

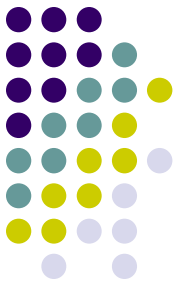
# dLOC Process



# dLOC Applications



# dLOC Toolkit – Version 1



dLOC Tracking Database

File Actions Archives Options Help

## dLOC Tracking Main Form

dLOC ID	Local ID	Title	Author	Material Type
CA01000002		It's fun to dance the Haitian Meringue		IMAGE
CA01000004		Haitian Meringue, part deux	Peter Falk	TEXT

Pre-QC Processor    QC Application    dLOC FTP    C:\DLOC

Search by:

Title

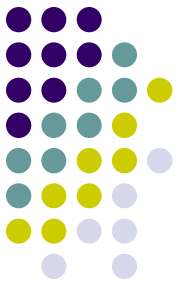
and

Total hits count = 2



# dLOC Toolkit – Version 1



dLOC Tracking Database

**CA01000015 : 00001**

Title:

Details

General Information

Local Identifier:  Material Type:  Year:

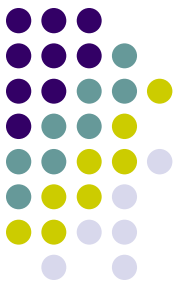
Author:

Publisher:

Comments

Save Cancel

# dLOC Toolkit – Version 1



**dLOC Template**

Action View Options Help

### dLOC Template

General Information | Source Document | Subjects and Keywords

Main Panel

Bibliographic Identifier:  Volume Identifier:

Language:  +

Resource Type:

Physical Description:

Identifier:  Type:  +

Holding Location:

**Edit Mode**

# dLOC Toolkit – Version 1



dLOC Tracking Database

File Actions Archives Options Help

## dLOC Tracking Main Form

dLOC ID	Local ID	Title	Author	Material
CA01000002		It's fun to dance the Haitian Meringue		IMAGE
CA01000004		Haitian Meringue, part deux	Peter Falk	TEXT
→ CA01000015	733.33 T458t	It's fun to dance the Haitian meringue	Thomas, Jerry	TEXT

Pre-QC Processor    QC Application    Go dLOC!    C:\DLOC

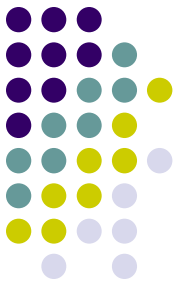
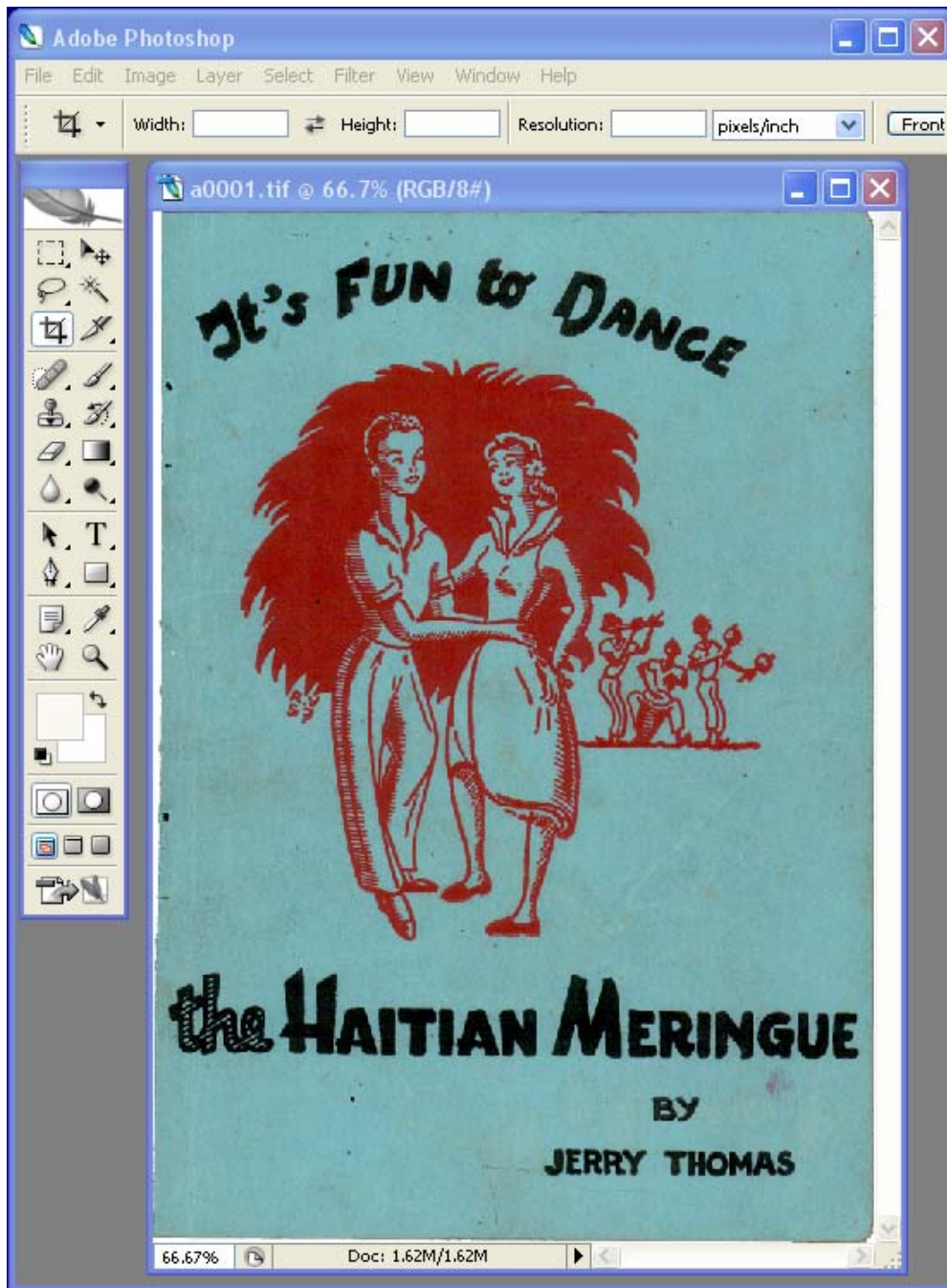
Search by:

Title

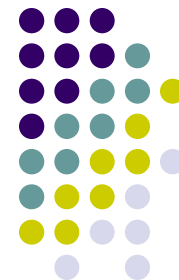
and

Total hits count = 3



# dLOC Toolkit – Version 1



**dLOC Tracking Database** [Window Title Bar]

File Actions Archives Options Help

## dLOC Tracking Main Form

dLOC ID	Local ID	Title	Author	Material
CA01000002		It's fun to dance the Haitian Meringue		IMAGE
CA01000004		Haitian Meringue, part deux	Peter Falk	TEXT
➔ CA01000015	733.33 T458t	It's fun to dance the Haitian meringue	Thomas, Jerry	TEXT

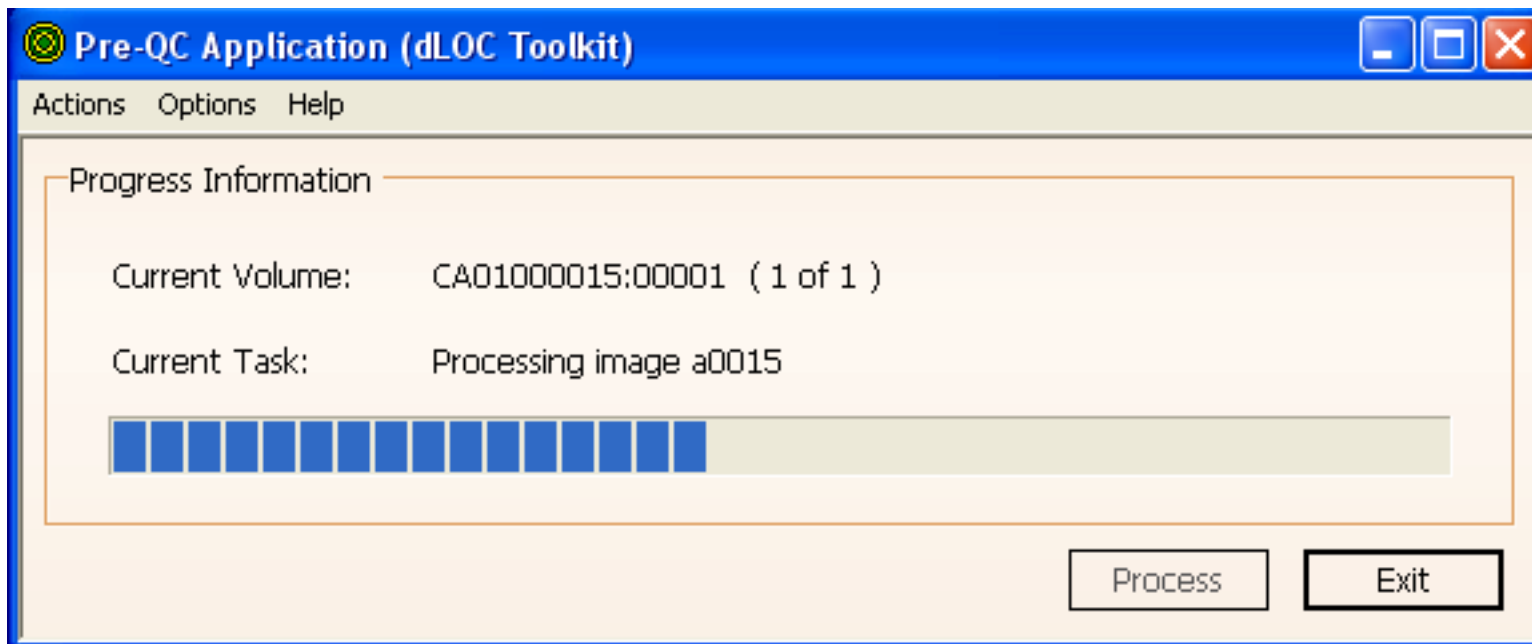
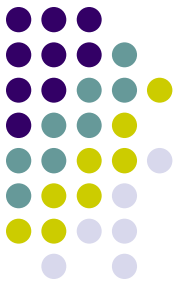
1 Pending

**Total hits count = 3**

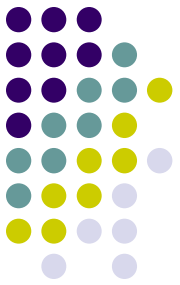
Search by:  
Title [dropdown] [input: meringue]  
and [input] [dropdown] [input]  
[Search] [Clear]

[Exit]

# dLOC Toolkit – Version 1



# dLOC Toolkit – Version 1



**dLOC Tracking Database** [Window Title Bar]

File Actions Archives Options Help

## dLOC Tracking Main Form

dLOC ID	Local ID	Title	Author	Material
CA01000002		It's fun to dance the Haitian Meringue		IMAGE
CA01000004		Haitian Meringue, part deux	Peter Falk	TEXT
➔ CA01000015	733.33 T458t	It's fun to dance the Haitian meringue	Thomas, Jerry	TEXT

Pre-QC Processor

QC Application

1 Pending

Go dLOC!

C:\DLOC

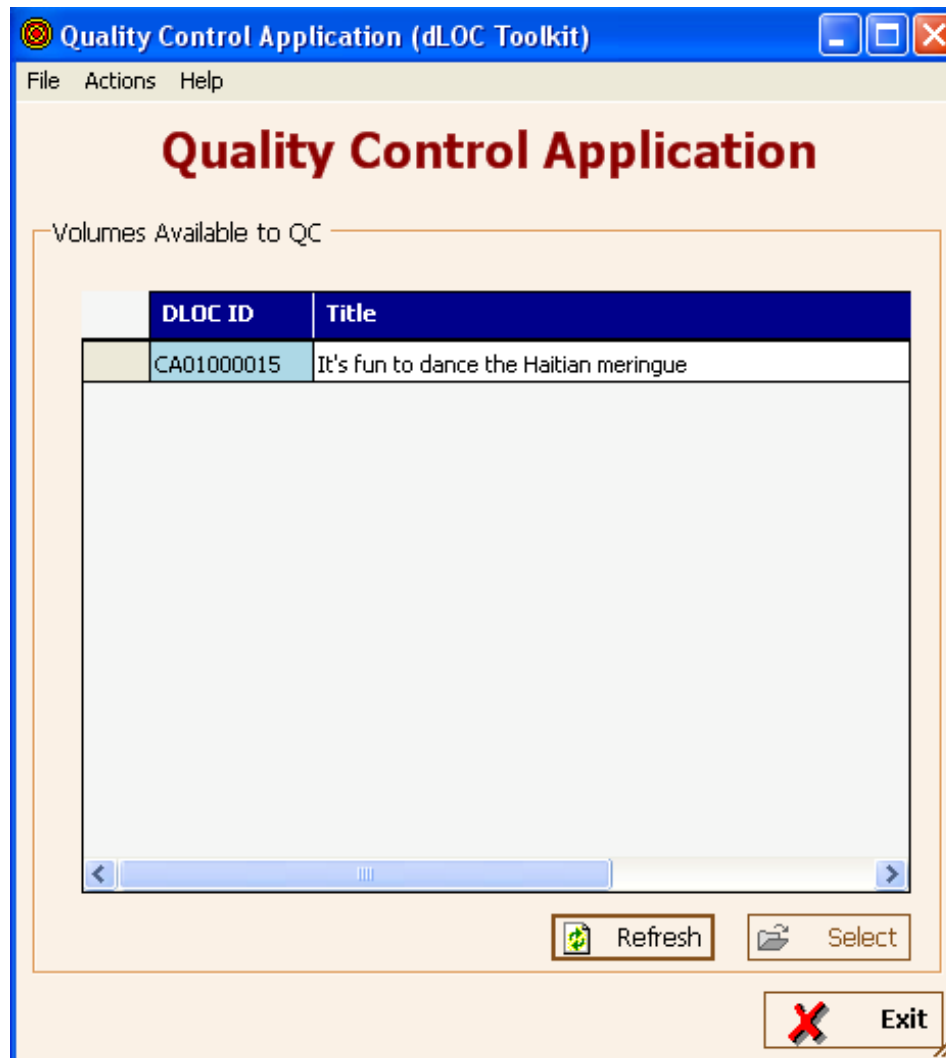
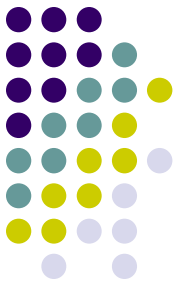
**Search by:**

Title

and

**Total hits count = 3**

# dLOC Toolkit – Version 1





# dLOC Toolkit – Version 1



Quality Control Application (dLOC Toolkit)

File Actions Help

CA01000015 : 00001

Title: It's fun to dance the Haitian meringue

a0001  
  
Pagination:   
Division:  Main

a0002

a0003  
  
Pagination:   
Division:  Main

a0004

a0005

a0006

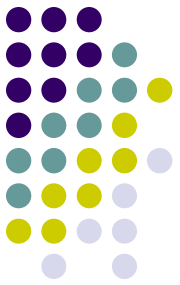
a0007

a0008

Critical Volume Error: No Volume Level Error

Last QC'd:


# dLOC Toolkit – Version 1





**dLOC Tracking Database** File Actions Archives Options Help


## dLOC Tracking Main Form

dLOC ID	Local ID	Title	Author	Material
CA01000002		It's fun to dance the Haitian Meringue		IMAGE
CA01000004		Haitian Meringue, part deux	Peter Falk	TEXT
→ CA01000015	733.33 T458t	It's fun to dance the Haitian meringue	Thomas, Jerry	TEXT

  
Pre-QC Processor

  
QC Application

  
Go dLOC!  
1 Pending

  
C:\DLOC

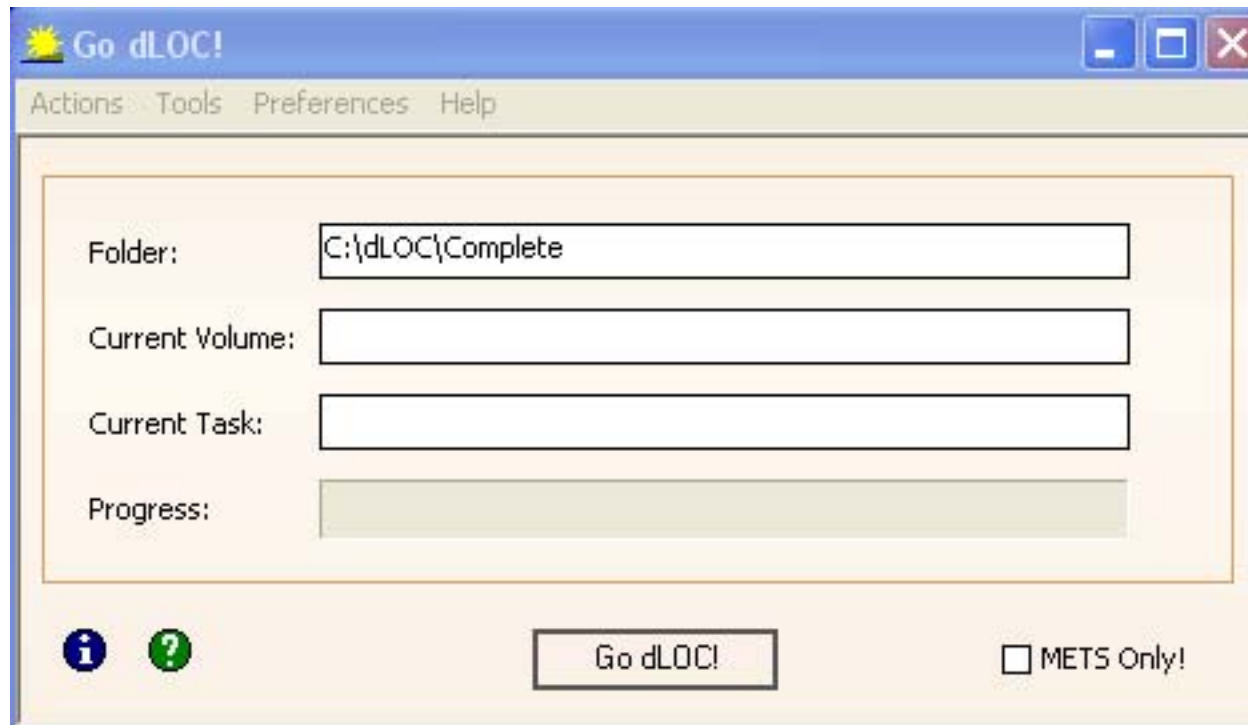
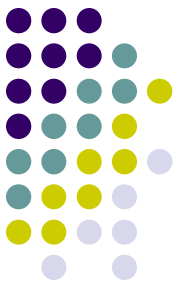
**Search by:**

Title

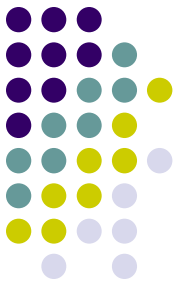
and

**Total hits count = 3**

# dLOC Toolkit – Version 1

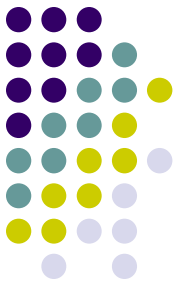


# dLOC Toolkit – Version 1



Item Submission Complete!

# dLOC Toolkit – Version 1









dLOC Tracking Database

**CA01000015 : 00001**

Title:

Details | Status | Archives

Current Status

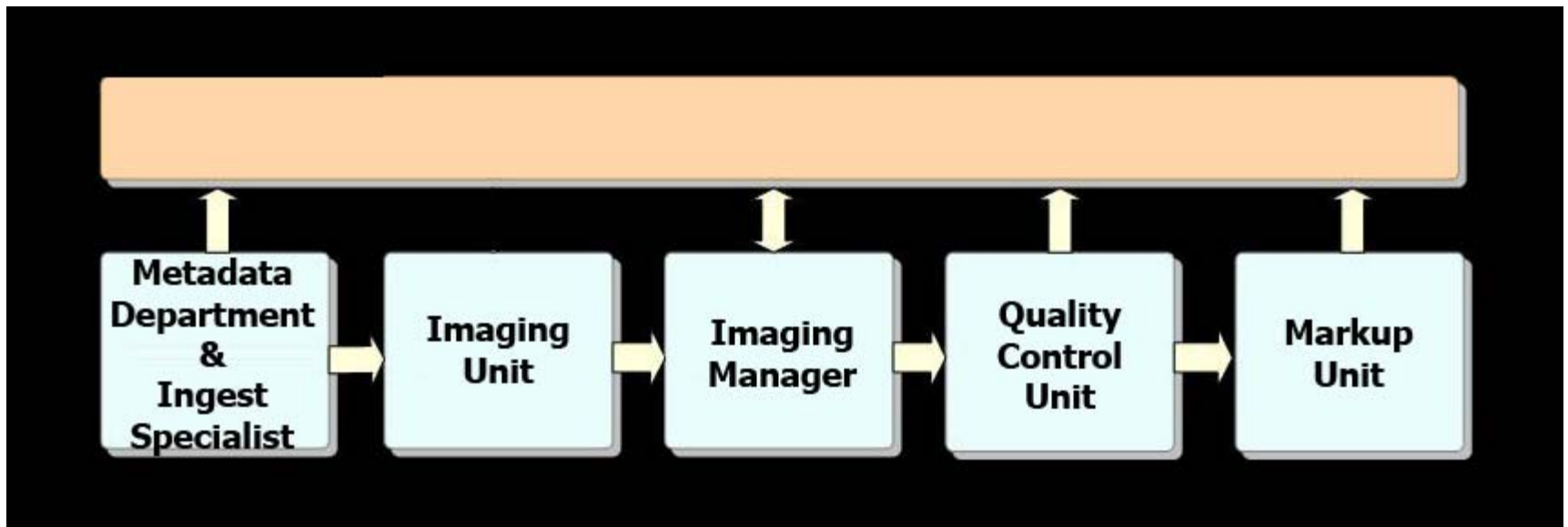
Record Created:	<input type="text" value="Friday, May 30, 2008"/>	<a href="#">Open Directory</a>	
Metadata Created:	<input type="text" value="Friday, May 30, 2008"/>	<a href="#">Edit Metadata</a>	
Scanned:	<input type="text" value="Friday, May 30, 2008"/>		
Quality Control:	<input type="text" value="Friday, May 30, 2008"/>	<a href="#">QC again</a>	
FTP'd:	<input type="text"/>	<a href="#">FTP to dLOC</a>	
Archived:	<input type="text"/>		

Save Cancel

# dLOC Toolkit – Version 1



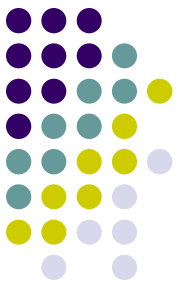
- Problems
  - Level of expertise required missing, even at UF.



# dLOC Toolkit – Version 1

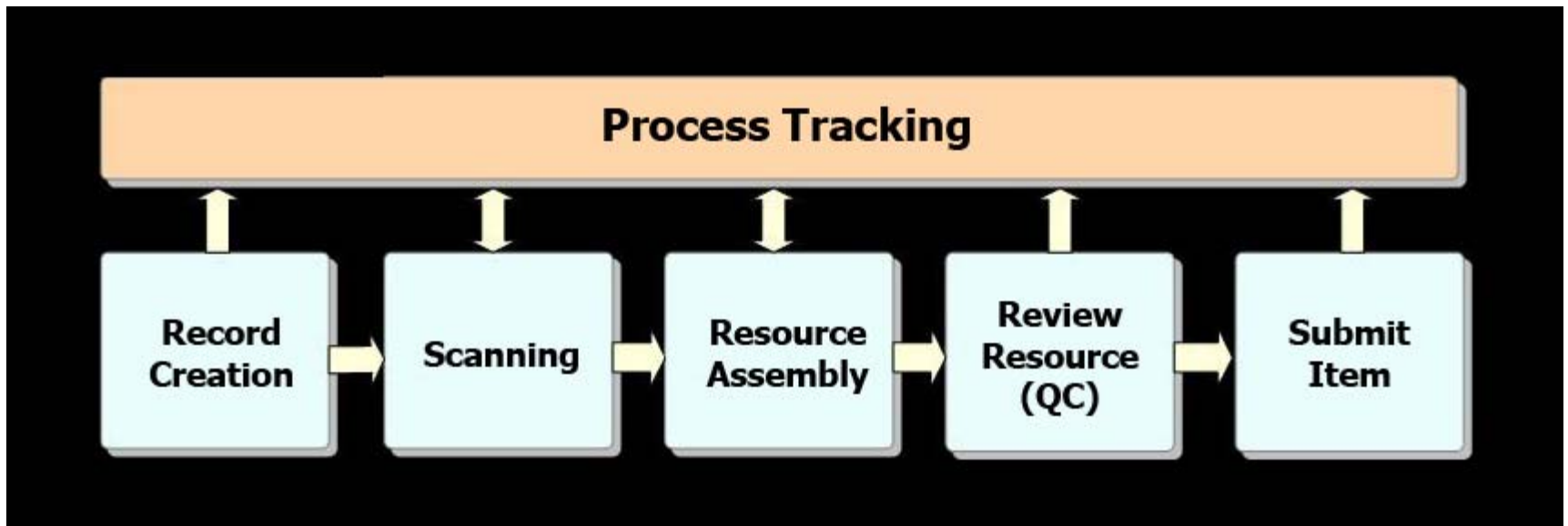


- Problems
  - Frozen “snapshot” of UF DLC applications.
  - Inability to update metadata for items submitted to dLOC
  - Inability to customize sufficiently

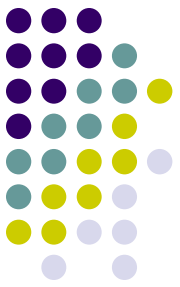


# dLOC Toolkit – Version 2.0

- Same basic process steps



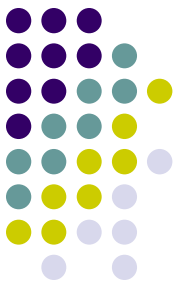




# dLOC Toolkit – Version 2.0

- Integrated applications in one window

The screenshot shows the 'Digital Library of the Caribbean - Digitization Toolkit' window. The title bar includes the text 'Digital Library of the Caribbean - Digitization Toolkit' and standard window control buttons. The main header features a sun icon and the text 'Digital Library of the Caribbean'. Below the header is a navigation bar with five buttons: '1. Describe Item' (highlighted in yellow), '2. Scan Item', '3. Prepare for QC', '4. Quality Control', and '5. Submit Item'. The main content area has two tabs: 'General Information' (selected) and 'Subjects and Keywords'. Under 'General Information', there are two sections: 'Title Information' and 'Physical Description'. The 'Title Information' section contains a 'Title:' label followed by a text input field, and three buttons below it: 'Add Alternate Title', 'Add Series Title', and 'Add Uniform Title'. The 'Physical Description' section contains a 'Resource Type:' label followed by a dropdown menu showing 'TEXT'. At the bottom of the window, there is a 'Help' link and four buttons: 'Cancel', 'Back', 'Save', and 'Forward'.



# dLOC Toolkit – Version 2.0

- More user-friendly



# dLOC Toolkit – Version 2.0



Live Demo

# dLOC Toolkit – Version 2.0



- Advantages
  - Ability to update your items on dLOC
  - Archives Module
  - Customization
    - Language
    - Themes
    - Metadata Templates
    - Output Formats
  - Security
  - Frequent Updates

# dLOC Toolkit – Version 2.0



- Future Enhancements
  - Glossary of Metadata Terms
  - Import existing catalog records
  - Update digitization manual
  - Create videos to demonstrate common tasks
  - Release under GPL open-source license
  - Additional output formats

# dLOC Toolkit – Version 2.0



- Uses
  - Greenstone Metadata Output
  - New Metadata Output Formats
  - dLOC ( [www.dloc.com](http://www.dloc.com) )