Delivering your special collections online: Digitization and CONTENTdm

Jill Fluvog
Director, Sales & Marketing
Digital Collection Services, Seattle

Patrick Jones
Digital Services Consultant
• Services for the digital life cycle
• Excerpts from “Digitization Basics” (Patrick Jones)
• CONTENTdm
• Additional digital collection services
• What is coming?
We provide services for libraries and other cultural heritage organizations to create, manage, preserve and share digital collections.

Our mission is to offer services that:

- Help you create and showcase digital collections on the Web
- Enable the creation of your institution’s digital repository
- Offer your collections maximum Web visibility via WorldCat
- Provide long-term archiving of your digital master files
An integrated set of solutions offering...

Access
Discovery
Creation
Management
Preservation

for your digital collections
A range of options that support the entire digital life cycle

1. Plan
2. Process
3. Manage
4. Preserve
5. Share
The Planning stage

- What are your project’s goals and objectives? Do they support your vision, values and mission?
- Who are your stakeholders?
- Does Web access play a role? ...full-text searching?
- What are your end users’ needs?
- Does preservation play a role?
- How will you fund the project? Do you have the staff, hardware, etc.?

- Determine your organizational priorities first
- Select the technical tools last
Solutions to help you create, manage, preserve and share your digital collections:

1. Digitization services
2. CONTENTdm Digital Collection Management Software
3. Connexion digital import
4. Web Harvester
5. Digital Archive
6. WorldCat
7. User Services
Digitization services

We create high-quality, digital images from...

- Original materials
- Microfilm and microfiche

We provide digital files to meet specifications

- Resolution
- Processing
- File naming
- File format
- Media delivery
Digitization process requirements

1. **Digital conversion** (TIFF creation)
   - Resolution (dpi - dots per inch)
   - Bit depth (bitonal, grayscale, color)

2. **Optical Character Recognition (OCR)**

3. **Metadata**
   - Dublin Core, MARC, METS, ALTO, etc.

4. **File naming**
   - Use ISO date format whenever possible
5. Output/derivative files
   • TIFF archival masters, bound PDF, JPEG2000, thumbnails

6. Image processing
   • De-skew, crop, sharpen, page and article segmentation, etc.

7. Delivery method
   • CD, DVD or hard drive

8. CONTENTdm
   • Digital collection building
What is CONTENTdm Digital Collection Management Software?

• A complete software solution

• Stores, manages and provides Web access to digital collections

• For organizations of all types and sizes
  Academic libraries, public libraries, historical societies, museums, state libraries, cultural heritage organizations and others

• Standards-based for staff efficiencies and cost effectiveness
  Dublin Core, JPEG2000, OAI Harvesting, PHP API and more

• Showcases a wide range of media types
  Photos and documents to audio and video files
• Enables you to share your digital collections on the Web

Maryland Library's Most Wanted - Enoch Pratt Free Library

Collection location: Maryland Department of the Central Library, Enoch Pratt Free Library / State Library Resource Center.

Contributor: Contributed to the MDCH Program by the Central Library, Enoch Pratt Free Library / State Library Resource Center.

Inclusive Dates: undated.

Collection Overview: Calling all good citizens! The Central Library, Enoch Pratt Free Library / State Library Resource Center has a number of unidentified photographs in our digital collection and we are asking your help in identifying them. All the photographs were taken in the state of Maryland in the 20th century, but in many cases that's all we know.

We are asking for your help. If you click the "Wanted Images" button, you'll see all the pictures we have which we need your help in identifying. If you recognize the pictures, please send us an e-mail with your name, the picture number, anything you recognize about the picture, and any sources of information you
Items of all types: images, maps, drawings, text, diaries, audio, video and more....
Organizations of all types

American Society for Cell Biology
- http://cellimages.ascb.org/
- Item link - http://cellimages.ascb.org/u/?p4041coll12,127

Buffalo Bill Historical Center
- http://www.bbhc.org/hmrl/collection.cfm

Arizona Memory Project
- Lists provide range of contributors and shows use of predefined queries
- http://azmemory.lib.az.us./cdm4/topicbrowse.php

George Washington Masonic Memorial Meeting Minutes
- Full-text search - http://cdm2623-01.cdmhost.com/u/?p262301coll1,1394
Building your digital collections

Three ways CONTENTdm fits into your workflow:

1. CONTENTdm Acquisition Station - submit digital files (any format) through this CONTENTdm client

2. CONTENTdm Web-based Add - submit digital files (any format, only simple files) thru a Web browser

3. Connexion digital import - submit digital files through the Connexion client

Optional workflows

- Web Harvester
- Digital Archive
Building your digital collections with CONTENTdm - workflow

(1) Acquisition Station client and (2) Web-based Add

Dublin Core metadata and files from local computer or network

WorldCat

CONTENTdm
Building your CONTENTdm collections

Overview

1. Define your collection
   - Configure your metadata schema

2. Populate your collection
   - Add digital items and metadata

3. Share your digital collections on your CONTENTdm Web site
1. Define your collection

Configure the collection metadata schema

- A group of objects (items) that
  - Share the same metadata scheme
  - Live on the same CONTENTdm server (license)
- 200 collections per license

- Simple or Qualified Dublin Core supported
- Title required
- 100 fields per collection
- 128,000 characters per field
- 16 million items per collection
CONTENTdm and the Dublin Core

Field properties

View, add, edit and delete fields. Enable full text searching and controlled vocabulary. Once you have added, changed, or deleted fields, index the collection to allow changes to take effect.

<table>
<thead>
<tr>
<th>Field name</th>
<th>DC map</th>
<th>Data type</th>
<th>Large</th>
<th>Search</th>
<th>Hide</th>
<th>Vocab</th>
<th>add field</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Title</td>
<td>Title</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>2 Subject</td>
<td>Subject</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>3 Description</td>
<td>Description</td>
<td>Text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>4 Creator</td>
<td>Creator</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>5 Publisher</td>
<td>Publisher</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>6 Contributors</td>
<td>Contributors</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>7 Date</td>
<td>Date</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>8 Type</td>
<td>Type</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>9 Format</td>
<td>Format</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>10 Identifier</td>
<td>Identifier</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>11 Source</td>
<td>Source</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>12 Language</td>
<td>Language</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>13 Relation</td>
<td>Relation</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>14 Coverage</td>
<td>Coverage</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>15 Rights</td>
<td>Rights</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>16 Audience</td>
<td>Audience</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
</tbody>
</table>
2. Populate your collection

- Add digital items and metadata
3. Share your digital collections on your CONTENTdm Web site

- Browse
- Advanced Search
3. Share your digital collections on your CONTENTdm Web site

Digital Archives Initiative
Early Days Exhibit - Health Sciences Construction Photographs

<table>
<thead>
<tr>
<th>Image</th>
<th>Title</th>
<th>Subject</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td>001 Before start of construction. Standing in Leary's Brook looking south east. Arts &amp; Cultural [sic Culture] Centre on left. Physics-Chemistry Building on right.</td>
<td>Health Sciences Centre - construction - Leary's Brook - Arts and Culture Centre</td>
<td>b&amp;w; 7 x 7 cm; negative 187 (digital format only)</td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td>002 Looking south from Leary's Brook</td>
<td>Health Sciences Centre - construction - Leary's Brook</td>
<td>b&amp;w; 7 x 7 cm; negative 188 (digital format only)</td>
</tr>
</tbody>
</table>

Multiple languages
3. Share your digital collections on your CONTENTdm Web site - End user tools

- View in slide show
- Save as Web page
- Export to PowerPoint
- Use for ‘my shopping cart’
Building your digital collections

Three ways CONTENTdm fits into your workflow

1. CONTENTdm Acquisition Station - submit digital files (any format) through this CONTENTdm client

2. CONTENTdm Web-based Add - submit digital files (any format, only simple files) thru a Web browser

3. Connexion digital import - submit digital files through the Connexion client

Optional workflows

- Web Harvester
- Digital Archive
Building your digital collections with Connexion digital import - workflow
Connexion digital import

- Option for getting your digital items into CONTENTdm
- Cataloging workflows
- Integrates cataloging and CONTENTdm
- Via Connexion client, submit digital files stored on your local computer/network to your CONTENTdm collections
- For users of CONTENTdm Hosting Service who also have an OCLC cataloging subscription
Three ways CONTENTdm fits into your workflow

1. CONTENTdm Acquisition Station
2. CONTENTdm Web-based Add
3. Connexion digital import

Optional workflows

- Web Harvester - *Coming soon!* Easily harvest Web content and add it to your CONTENTdm collections
- Digital Archive - Preserve your master files and digital originals
Web Harvester - *Coming soon!*

Easily harvest Web content

- Web pages
- Web sites
- Individual publications

Add Web content to your CONTENTdm collections

Provides a workflow that is integrated into both your normal cataloging and digital content management workflows
Building digital collections with the Web Harvester - workflow

Web Harvester via OCLC Connexion client

Metadata

WorldCat

CONTENTdm

Web sites

DC Metadata and files
Web Harvester: Integrated capture and management of Web content

- Integrated with existing tools and workflows
- Cataloging workflow via OCLC Connexion client
- Digital collection management via CONTENTdm
- Access via 856 to single objects and serials via WorldCat, major search engines, and your local systems
- Customizable end user online access via CONTENTdm
- Digital Archive (optional)
Some challenges of preservation

- Increasing volume of digital materials to manage
- Some solutions for widely published materials
- Few standards and not widely implemented
- Limited resources to apply to preservation

A preservation solution: OCLC’s Digital Archive

- Secure, managed storage for digital originals and master files
- Integrated with your workflow for building digital collections
The Montana Historical Society has chosen the Digital Archive service as the storage facility for our digital collections. The ease of adding materials through Connexion and the secure, managed storage make the Digital Archive service the ideal solution for our needs.”
—Molly Kruckenberg, Research Center Director
Simple, straightforward workflows that fit in with

- CONTENTdm collection building
- WorldCat cataloging
- Web Harvester
- Other digital asset management systems
Sharing your digital collections

Your organization’s digital collections are built and shared via the Web... what’s next?

- Promote your collections

- *Increase discovery* through WorldCat
  - WorldCat.org
  - WorldCat Local
  - Google
  - Yahoo! Search
Increase discovery via Google and WorldCat
Search results for 'Japanese artist Chikanobu'

Sort by: Relevance

Results 1-10 of about 14 (0.13 seconds)

1. Illustration of Kinryuzan Sensoji
   by Museum of Fine Arts, Boston.
   Language: English Type: Internet Resource: 2-D Image: Visual Material
   Publisher: Museum of Fine Arts, Boston, Boston, Massachusetts, USA. 2003 30 Jean S. and Frederic A.
   Sherr Collection 1882 1882 registered February 1882.

2. #30 Yamashiro, Flowers of Kinkaku-ji, Princess Yuki
   by Claremont College Digital Library; Kobayashi Tetsujirō, Ruth Chandler Williamson
   Gallery, Scripps College
   Language: English Type: Internet Resource: 2-D Image: Visual Material
   Publisher: Claremont Colleges Digital Library 1884-12-10

3. A Glimpse of Dignitaries Dancing
   by Claremont College Digital Library:
   Language: English Type: Internet Resource: 2-D Image: Visual Material
   Publisher: Claremont Colleges Digital Library 1888-02-08
### Chikanobu and Yoshitoshi Woodblock Prints

**Yamashiro, Flowers of Kinkaku-ji, Princess Yuki**

<table>
<thead>
<tr>
<th>Accession number</th>
<th>93.8.118</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td>Yamashiro, Flowers of Kinkaku-ji, Princess Yuki</td>
</tr>
<tr>
<td><strong>Title-Alternative</strong></td>
<td>Snow, Moon, Flowers, Setsu gekka, Yamashiro, Kinkaku-ji, Yuki-hime</td>
</tr>
<tr>
<td><strong>Artist</strong></td>
<td>Chikanobu, Yoshu</td>
</tr>
<tr>
<td><strong>Artist Nationality</strong></td>
<td>Japanese</td>
</tr>
</tbody>
</table>
OCLC number: 181589045

Item citation:

"#30 Yamashiro, Flowers of Kinkaku-ji, Princess Yuki"

by Claremont College Digital Library; Kobayashi Tetsujiro, Ruth Chandler Williamson Gallery, Scripps College
Citation Styles for "30 Yamashiro, Flowers of Kinkaku-ji, Princess Yuki"

**APA**

**Chicago (Author-Date)**

**Harvard**

**MLA**

**Turabian**
“It was the solution we needed to not only manage the collection, but to share it with the broadest audience possible.”
Pat Vince, Digital Initiatives Librarian
Claremont Colleges

http://ccdl.libraries.claremont.edu/
Better global discovery with WorldCat

Libraries and other cultural heritage organizations add their CONTENTdm collection metadata to WorldCat

- Through WorldCat.org and WorldCat Local, people on the Web discover—not only traditional library collections—but special collections that have been digitized.
We can help you get to success!

Professional assistance is an available option for digital collection projects:

- Planning and implementation
- Metadata and workflow design
- CONTENTdm collection building
- Training (onsite and online)
The CONTENTdm user community

- More than 400 licensed sites
- More than 1000 users
CONTENTdm users freely share best practices and locally-developed extensions through:

- CONTENTdm listserv
- User Support Center
  - Available online 24x7
  - Tutorials, comprehensive Help files and more
- User groups
- Example: Drake University
  
  http://www.lib.drake.edu/heritage/odm/map/
ALA Midwinter 2008
  • January in Philadelphia

Midwest User Group
  • April in Indianapolis, Indiana

Western User Group
  • June in Claremont, California

ALA Annual 2008
  • June/July in Anaheim, California

Northeast/Mid-Atlantic User Group
  • July at Pennsylvania State University

South/Southeast User Group
  • July/Aug at Mississippi State University
Version 3.4 Jan 2003
OAI support for harvesting of metadata

Version 3.5 Jul / Sep 2003
Multi-Site Server
WorldCat link for metadata harvesting

Customizable Web Templates
Version 3.6 Feb 2004
Batch Add wizard
Advanced Search has a new interface and added functionality
User Support Center introduced
Version 3.7 & 3.8 Jul 2004
Zoom and pan toolbar for viewing images; Compare button
Tab-delimited Text Import
Collection and item-level security
Web browser-based editor

Version 3.8
JPEG2000

Version 4.0 Jun / Oct 2005
EAD support

PHP-based API which supports broader customizations and interoperability
User Interface based on the new PHP API
OCR Extension with ABBYY FineReader

Version 4.1 Mar 2006
Redirects for obsolete URLs

PHP 5.0 supported

Version 4.2 Dec 2006
Multiple Compound Object Wizard
OCR Extension upgraded to ABBYY FineReader
Version 8

Ability to highlight, view and clip individual newspaper articles
Search and browse by date range

Version 4.3 Oct 2007

Easier to manage and access PDF documents, especially multiple-page PDFs

Improved controlled vocabulary

OCLC Connexion digital import
CONTENTdm 5

- In progress!
- Next release will fully support Unicode
  - New search engine (Find)
- Pilot project with Simon Fraser University
  - Chinese language newspapers
Solutions to help you create, manage, preserve and share your digital collections
1. Explore CONTENTdm users’ collections at www.oclc.org/contentdm/

2. Evaluate the CONTENTdm software www.oclc.org/contentdm/evaluate/

3. Call Patrick if you’re starting a digitization project, applying for a grant or have any questions

4. Contact us

   Jill Fluvog       Patrick Jones
   fluvogj@oclc.org  jonesp@oclc.org
   360-866-0579     866-913-8576 (toll-free)