OCLC Online Computer Library Center

Digital Collection Services

Digital Collections, Repositories, and Archives

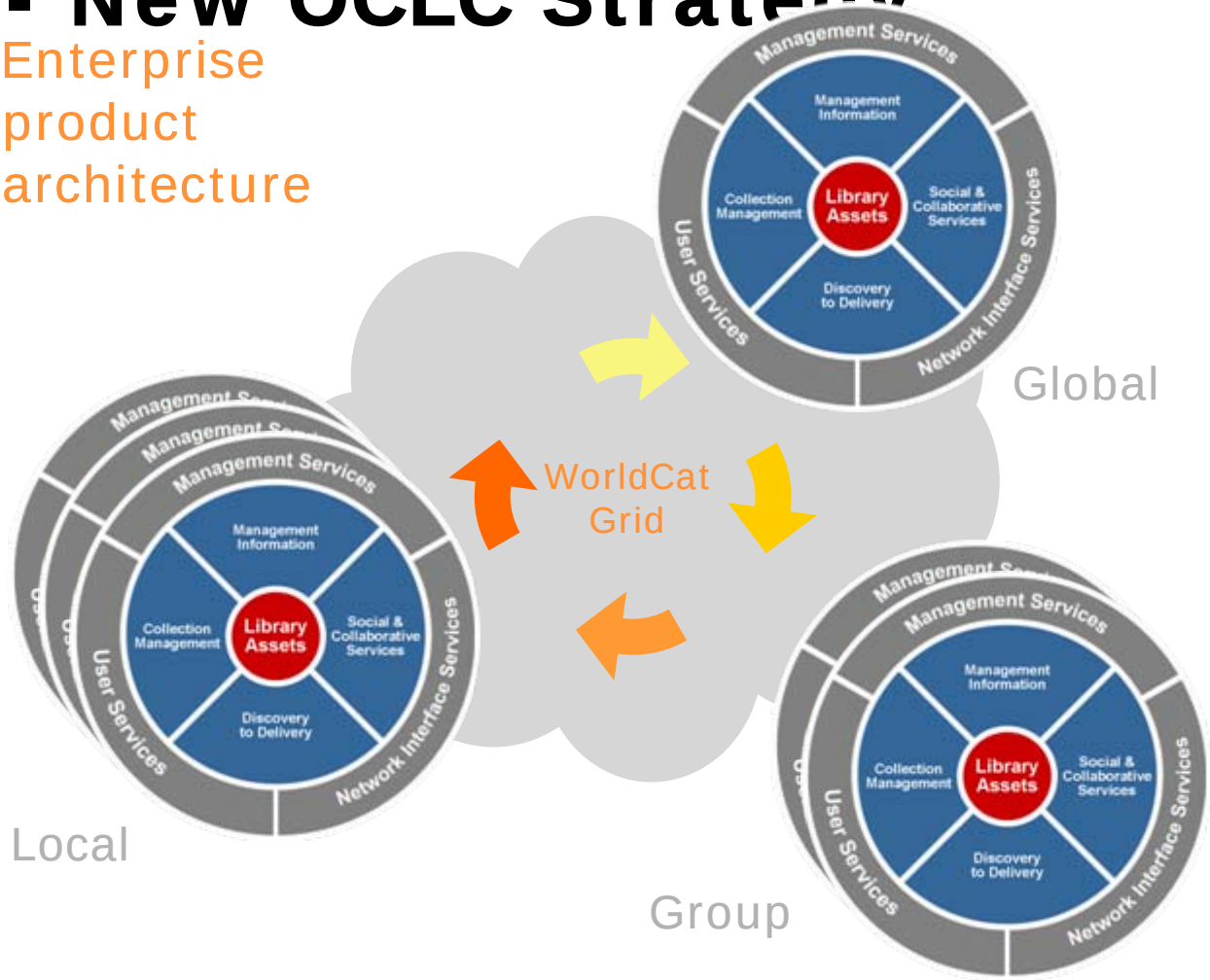
Greg Zick, Vice President
Digital Collection Services

ARL Research Library Leadership Fellows

November 1st, 2007

Future Direction - New OCLC Strategy

Enterprise product architecture



Presentation Outline

- Digital Collection Creation
- Institutional Repositories
- Digital Repositories
- OCLC Future Directions

One way to digitize all the world's collections

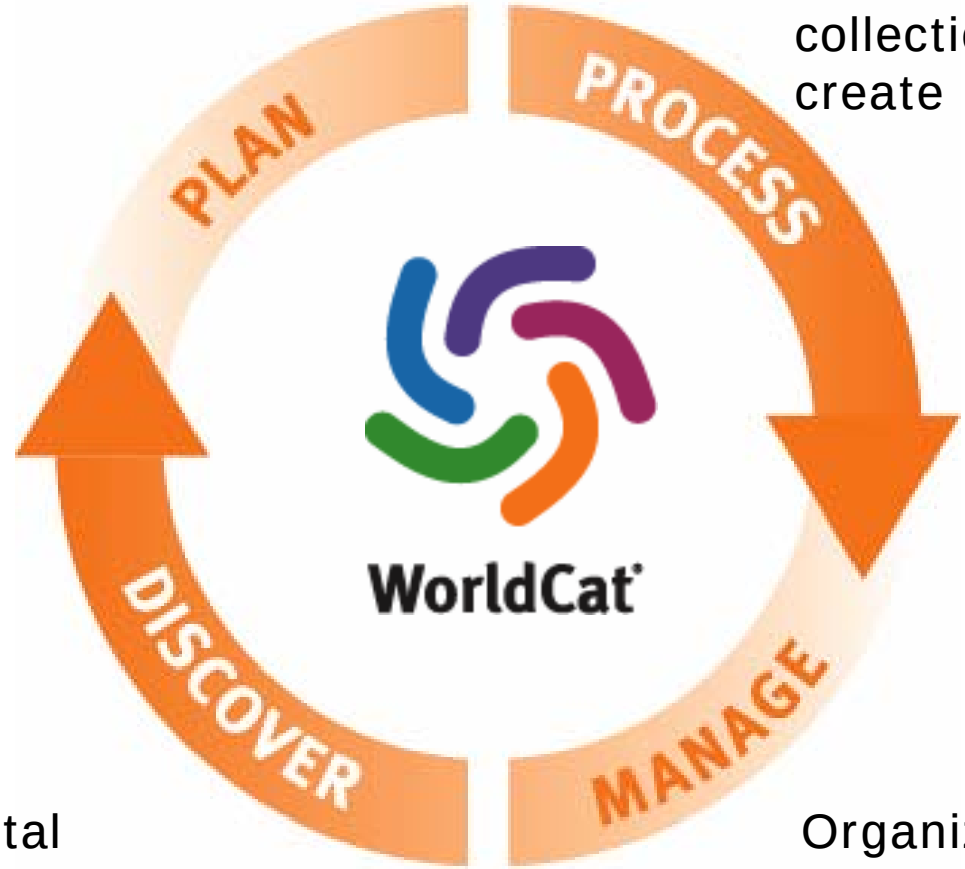
- List all the institutions that have collections?
  - Institutional Registry
- List all the collections at each institution
  - Archive Grid
- Digitize the items in the collections
  - Digital Registry
- Create the Digital Collection
  - CONTENTdm
- Archive original TIFFS
  - Digital Archive
- Add to Global Digital Repository
  - Harvest metadata into WorldCat

Digital Collection Creation

- *CONTENTdm*

Assess users' needs and collections' conditions

Convert your materials to digital collections and create metadata

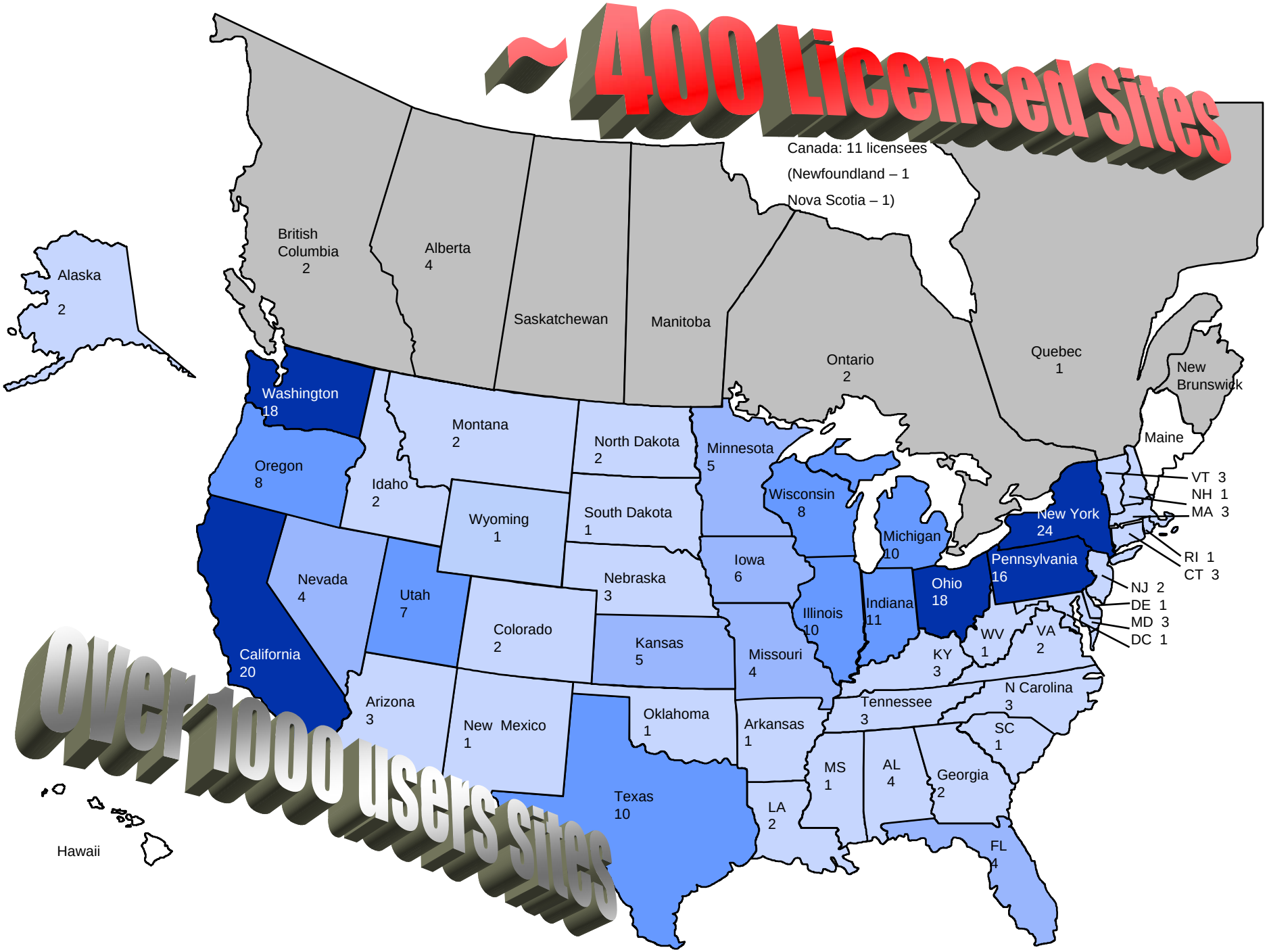


Expose your digital collections while providing Web access

Organize and store your digital collections

What is CONTENTdm?

- A complete software solution developed by DiMeMa
- Store, manage and access your digital collections
- For organizations of all types and sizes from academic libraries, historical societies, public libraries and other collaborators
- Showcase a wide range of media types from photos and documents to audio and video files


~ 400 Licensed Sites
Canada: 11 licensees
(Newfoundland – 1
Nova Scotia – 1)
British Columbia 2
Alberta 4
Saskatchewan
Manitoba
Ontario 2
Quebec 1
New Brunswick
Maine
Alaska 2
Washington 18
Oregon 8
California 20
Nevada 4
Idaho 2
Montana 2
Wyoming 1
Utah 7
Colorado 2
Arizona 3
New Mexico 1
North Dakota 2
South Dakota 1
Nebraska 3
Kansas 5
Oklahoma 1
Texas 10
Minnesota 5
Iowa 6
Missouri 4
Arkansas 1
LA 2
Wisconsin 8
Illinois 10
Michigan 10
Indiana 11
Ohio 18
KY 3
WV 1
VA 2
Tennessee 3
N Carolina 3
SC 1
Georgia 2
FL 4
AL 4
MS 1
New York 24
Pennsylvania 16
VT 3
NH 1
MA 3
RI 1
CT 3
NJ 2
DE 1
MD 3
DC 1
Hawaii
Over 1000 users Sites

Building partnerships and CONTENTdm digital collections

- Multi-regional or statewide consortia: AK, CA, FL, IA, IL, IN, NE, MN, MO, NY,OH, PA, WI
- Mountain West Digital Library – 20 members
- Western Waters Digital Library - 30 members
- Columbia River Basin Ethnic History Archive
- Digital Library of Appalachia -12 members
- Historically Black Colleges and Universities Library Alliance
- Upper Mississippi Valley Digital Image Archive
- REALIA Project –Assoc. Colleges of the Midwest, Assoc. Colleges of the South and the Great Lakes College Assoc.
- Consortium of Liberal Arts Colleges (CLAC) – 62 members

These and many more…

OCLC Online Computer Library Center

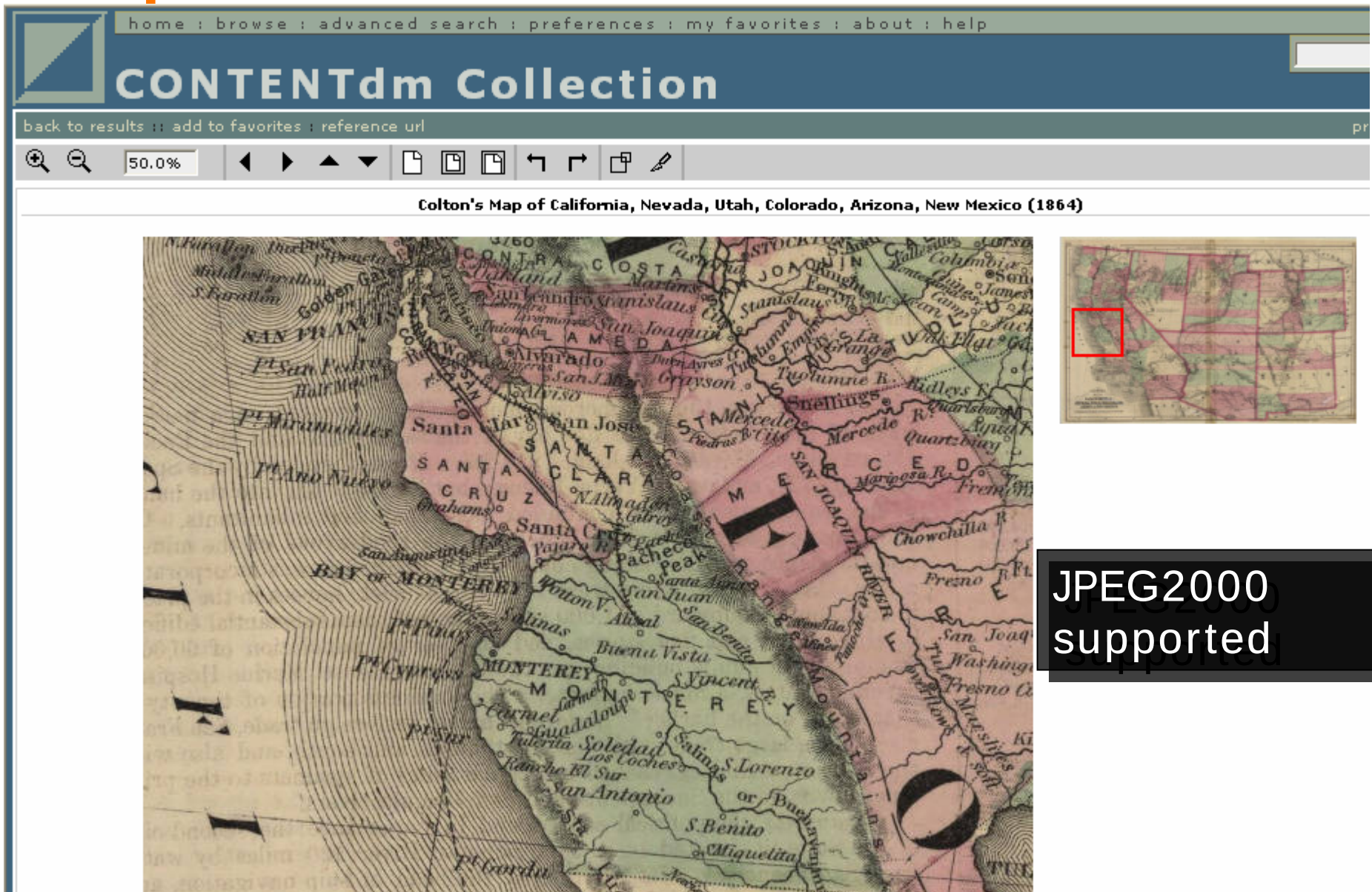
Digital Collection Services

Images: Photographs, posters, postcards

OCLC Online Computer Library Center

Digital Collection Services

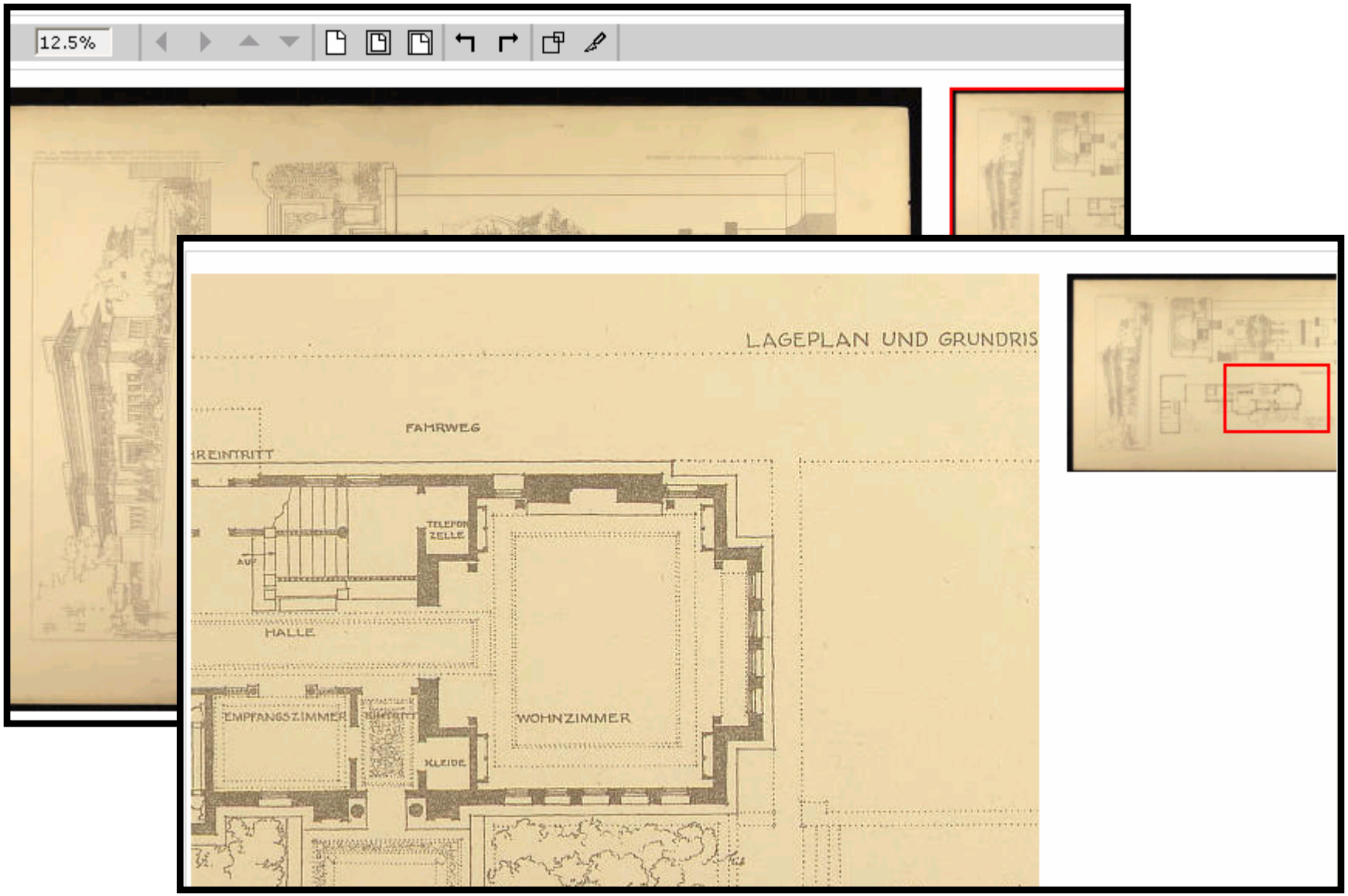
Maps



OCLC Online Computer Library Center

Digital Collection Services

Zoom and Pan

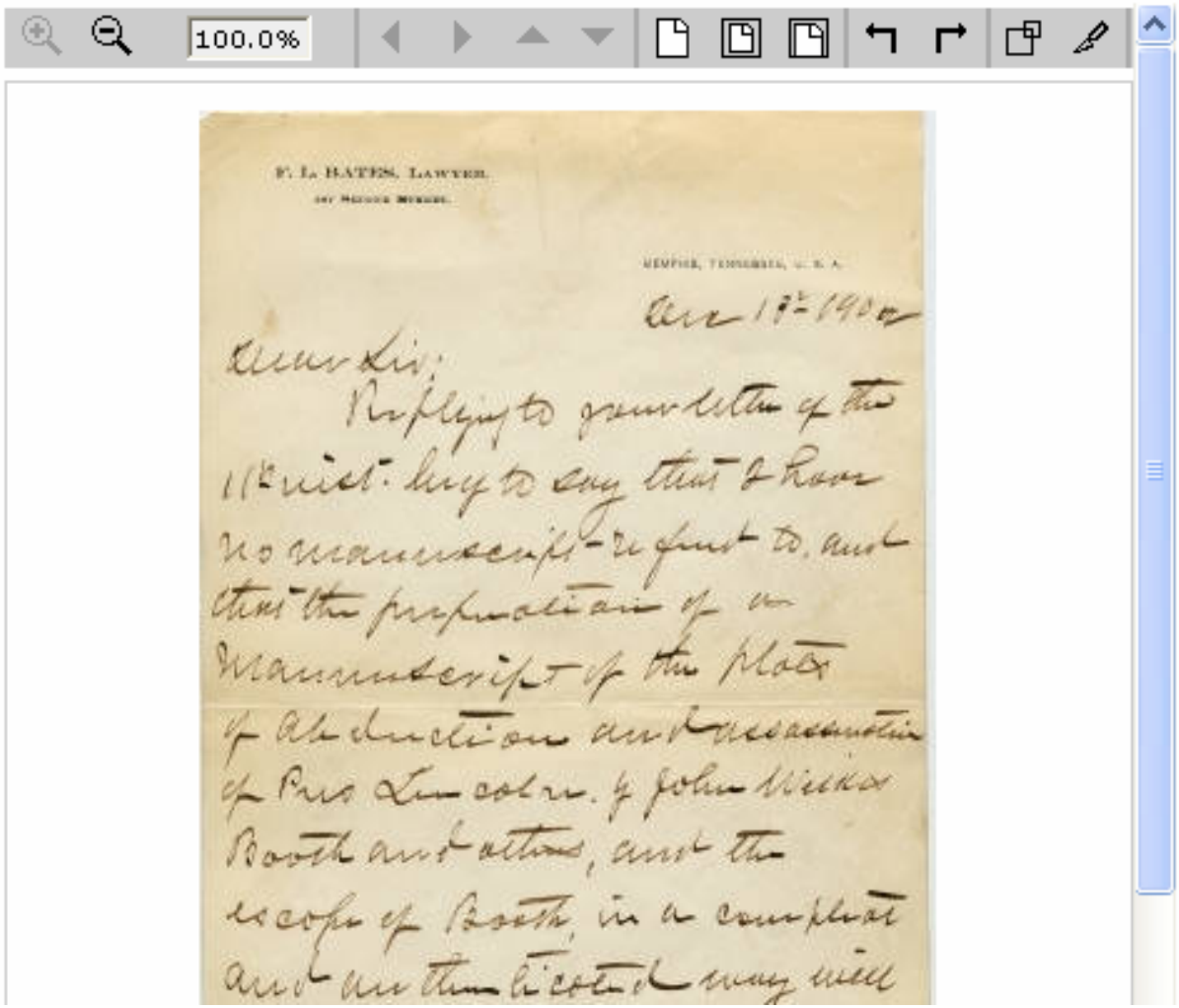


OCLC Online Computer Library Center

Digital Collection Services

Documents:

Text-based letters, journals, diaries and more



Dec. 13, 1902
Dear Sir:
Replying to your letter of the
11th inst. beg to say that I have
no manuscript referred to, and
that the preparation of a
manuscript of the plots
of abduction and assassination
of Pres. Lincoln by John Wilkes
Booth and others, and the
escape of Booth, in a complete
and authenticated way will
require the expenditure of time,
labor and money, and
before beginning the undertaking
should require that we

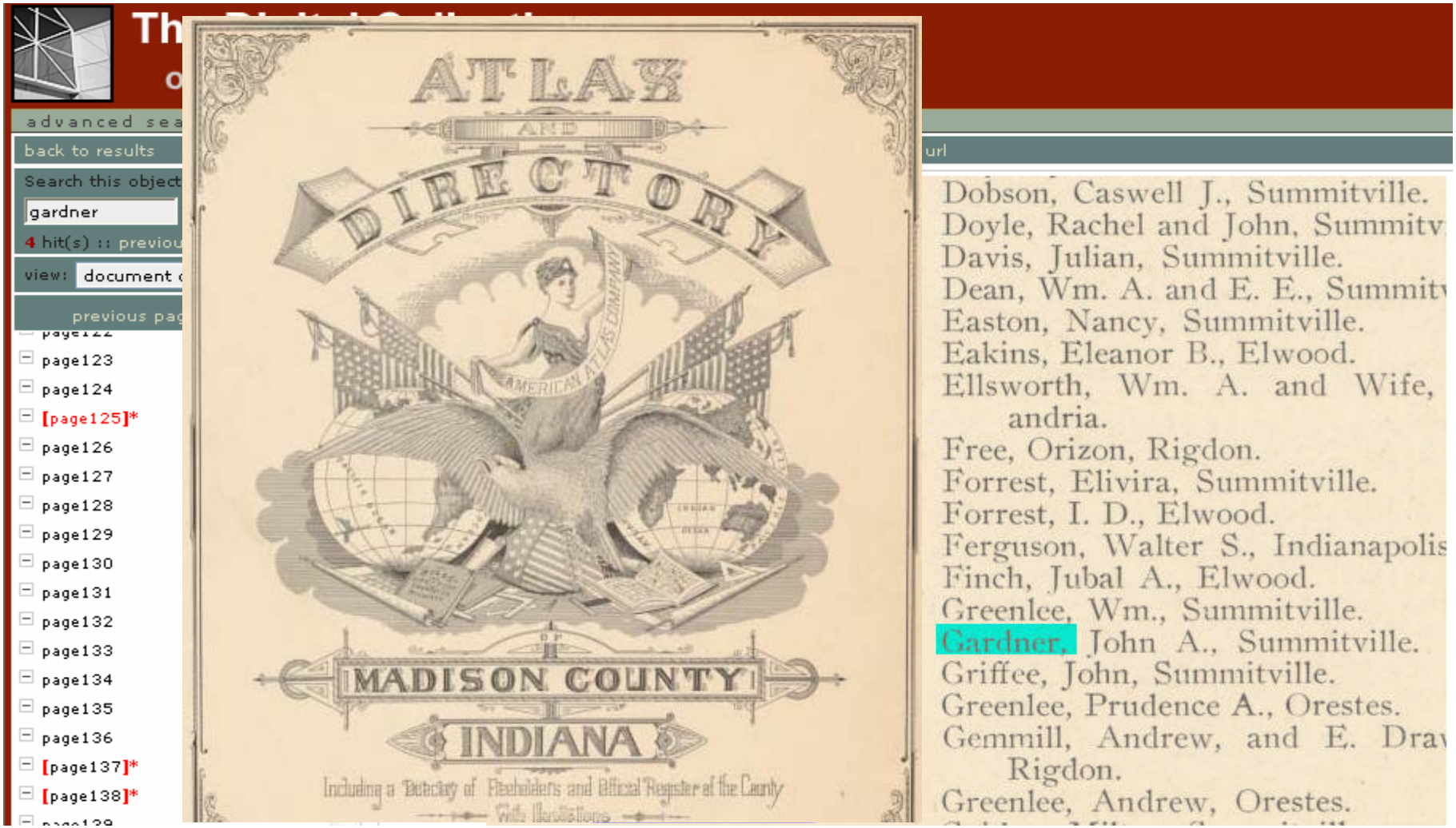
Transcribed text becomes searchable!

OCLC Online Computer Library Center

Digital Collection Services

Documents:

Text-based letters, journals, diaries and more



CONTENTdm's Integrated OCR Capability

OCLC Online Computer Library Center

Digital Collection Services

[Bureau of Economic Analysis]

OCLC Preservation Service Center completed scanning and provided digital images to BEA.

Use of CONTENTdm to search and access.

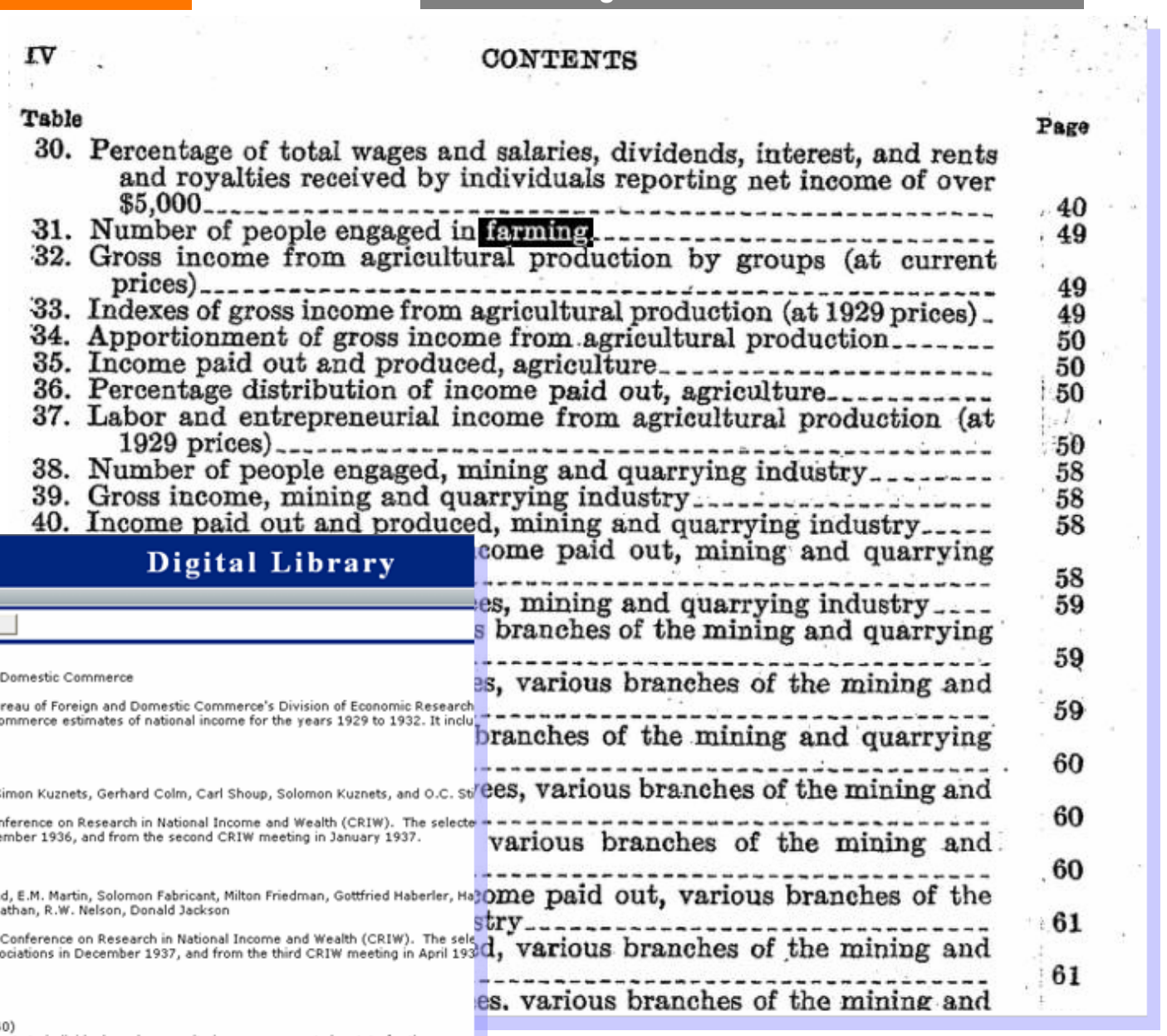

IV CONTENTS

Table Page

30. Percentage of total wages and salaries, dividends, interest, and rents and royalties received by individuals reporting net income of over $5,000 ---- 40
31. Number of people engaged in farming ---- 49
32. Gross income from agricultural production by groups (at current prices) ---- 49
33. Indexes of gross income from agricultural production (at 1929 prices) ---- 49
34. Apportionment of gross income from agricultural production ---- 50
35. Income paid out and produced, agriculture ---- 50
36. Percentage distribution of income paid out, agriculture ---- 50
37. Labor and entrepreneurial income from agricultural production (at 1929 prices) ---- 50
38. Number of people engaged, mining and quarrying industry ---- 58
39. Gross income, mining and quarrying industry ---- 58
40. Income paid out and produced, mining and quarrying industry ---- 58
come paid out, mining and quarrying ---- 58
es, mining and quarrying industry ---- 59
s branches of the mining and quarrying ---- 59
s, various branches of the mining and ---- 59
branches of the mining and quarrying ---- 60
ees, various branches of the mining and ---- 60
various branches of the mining and ---- 60
ome paid out, various branches of the stry ---- 61
d, various branches of the mining and ---- 61
es, various branches of the mining and

OCLC Online Computer Library Center

Digital Collection Services

Documents:

Newspapers

HERRINGTON

KIM QUEY TELLS ON HERRINGTON

Confesses to Dr. Gardner That She Was Coached to Commit Perjury.

Sensation in the Case of the Mayfield Chinese Woman Who Was Delivered Into the Hands of the Highbinders.

ATTORNEY J. N. E. WILSON, who appeared yesterday in United States Commissioner Heacock's court representing Bert Herrington in the case of the United States vs. Kim Quey, was hoisted clean out of the case by his own petard. His client turned the tables on the gang of highbinders that were fighting for the possession of herself as a chattel, and the ring was filled with consternation. If her story is true one lawyer may be disbarred.

Kim Quey was arrested at Mayfield by Chinese Inspector Gardner for being illegally in the country, he having discovered that she was one of the show women who had come here to take part in the exposition at Omaha. Bert Herrington took her

mony on that occasion was false, that th truth was that she had come here wit the Omaha Exposition people and that sh had never been here before.

Dr. Gardner then proceeded to tell th remainder of her confession. The stor had been taught her, he said, by Won Fong, the Chinaman charged with abduct ing her; by Wong Sing, the man wh afterward went through a form of mar riage ceremony with her, and by Lawye Herrington.

According to her statement they said "You must testify thus in order to clea us, for if you don't our lives wouldn't b worth anything. We would be hanged You must say that you were born in Sa Francisco; that you never were out of Sa Francisco; that you were not willing t stay in the Chinese Mission Home; tha

Done

Internet

OCLC Online Computer Library Center

Digital Collection Services

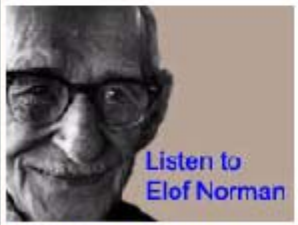
Audio / video

EVERETT PUBLIC LIBRARY Northwest Collection

Northwest Collection | Browse Collections | Contact us | City of Everett - Washington

Search Results: Your search retrieved 3 matching item(s).

Click on the thumbnail or title to view the item and corresponding description.



1. Elof Norman interview

2. George Wardell interview

3. Jack Leonard Miller interview

EVERETT PUBLIC LI

Northwest Collection | Browse Collections |

Title: Jack Leonard Miller interview

Location: United States--Washington (Sta

Date:

Photographer:

Subject:

Notes: Jack Leonard Miller (1889 - 1986) was one of the 74 Wobblies arrested following the Everett Massacre and charged with killing deputy Jefferson Beard. An articulate spokesman for the I. W. W., Jack was interviewed many times for film, video, radio and written histories. Hear an edited segment from an interview done by writer Jim Casey, at that time a columnist for the Everett Herald.

Original Format: Cassette tape interview

Repository: Everett Public Library

Order Number:

Rights & Permissions: Copyright Everett Public Library 2002

Add to My Favorites View My Favorites

OCLC Online Computer Library Center

Digital Collection Services

Browse

home : browse : advanced search : preferences : my favorites : Library Catalog : help

search

Indiana Historical Society - Digital Image Collections

select all : clear all : add to favorites

page 1 of 40 : (<< 1 2 3 4 5 6 7 8 9 10 >>) :: previous : next

Browsing 787 item(s) in: Postcards of Indiana the Jay Small Collection go

	Image:	▲ Title:	Subject:	Description:
1.		"Home Sweet Home", Peru, Indiana	Elephants; Camels; Horses; Baby animals; Circus animals; Drugstores; Banks; Streets	A line of circus animals parade down a street in Peru, Indiana. Adult elephants lead the way followed by a baby elephant and a group of camels. There is also a horse off to the side of the road. The animals pass various businesses including Ager Drug Company and the First National Bank. Several peple have stopped on the sidewalk to watch the animals walk by.
2.		"The Squad 1907" Indiana University, Bloomington, Indiana	Football; Sports; Sport clothes	Indiana University's 1907 football squad is shown. Most are wearing knee-length pants and pads.
3.		"Vandalia" Grounds, North Manchester, Indiana	Railroad locomotives; Railroad crossings; Railroad tracks; Automobiles	View of a railroad crossing in North Manchester, Indiana. A locomotive is shown approaching an intersection. Two lines of tracks cross the street. A car waits along the side of the road for the train to pass. Buildings are visible in the background.
4.		#1 Anderson, Indiana	Bridges; Floods; Roads; Carriages & coaches	Shows people standing by a bridge that crosses high water in Anderson, Indiana. The water appears to be nearing the level of the bridge, and is high up on trees that rise above the bridge. A horse and buggy are standing in the foreground, the road appears to be brick.

OCLC Online Computer Library Center

Digital Collection Services

home : browse : advanced search : preferences : my favorites : Library Catalog : help

search

Indiana Historical Society - Digital Image Collections

Search: **Across all fields** | Selected fields | By proximity

Find results with:

All of the words	ft wayne
The exact phrase	
Any of the words	
None of the words	

search clear all

Advanced search

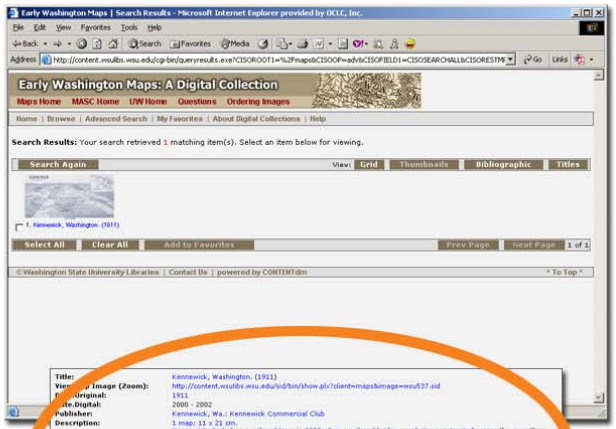
Select available collections

Search

Select specific collections:

- Alexander Gardner Lincoln Glass-Plate Negative
- Art of Mary Lyon Taylor
- Christmas Greetings
- Civil War Materials
- Daniel R. Weinberg Lincoln Conspirators Collection
- Indiana Through the Seasons
- Indianapolis Recorder
- Jack Smith Lincoln Graphics Collection
- James Whitcomb Riley -- Hoosier Poet
- Karl Bodmer Prints
- Lew Wallace in New Mexico
- Madam Walker
- Miscellaneous Images from IHS Photo Collections
- Native American Portraits from the Aboriginal Port folio
- Nineteen's 19
- Notable Hoosiers
- Other Lincoln-Related Manuscripts in the Collection
- Panoramic Photograph Images
- Photography of O. James Fox: 1945-1960
- Postcards of Indiana the Jay Small Collection
- Treasures from the Indiana Historical Society
- W.H. Bass Photo Company
- W.H. Bass Photo Company--Indiana Rail Transportation Images
- W.H. Bass Photo Company--Pamela Tranfield Memorial Collection

Discovery: WorldCat harvesting & Open WorldCat

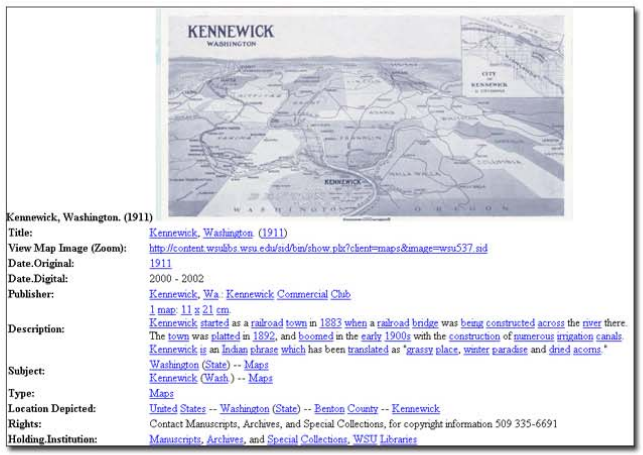


WorldCat

CONTENTdm Server

Metadata from CONTENTdm

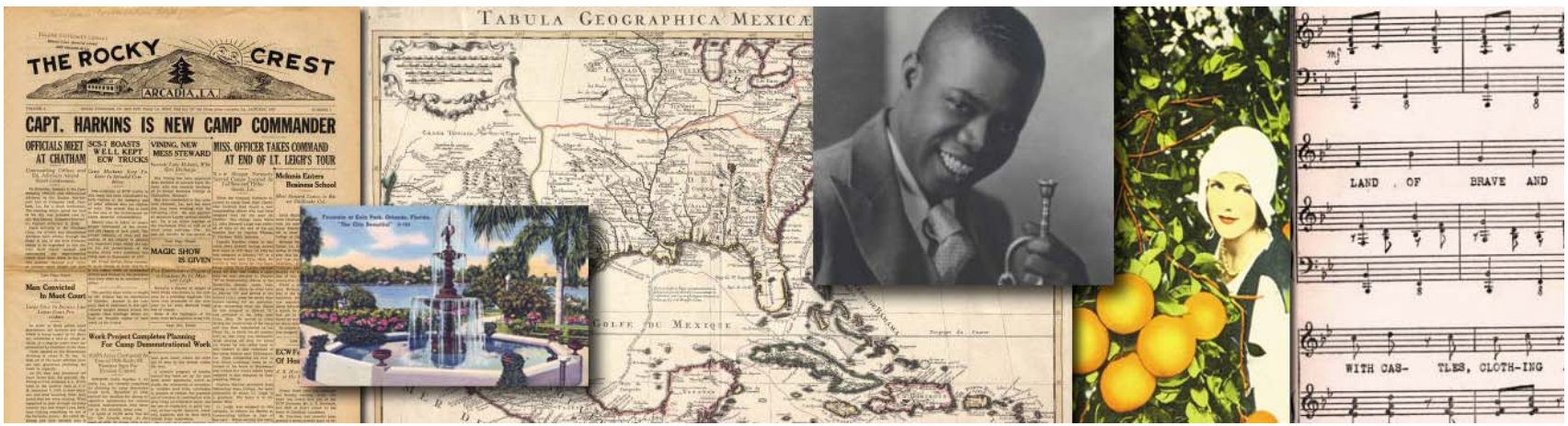
Discovery Options



One system, many solutions

CONTENTdm provides comprehensive solutions for your historical digital archives

Photographs, diaries, slides, newspapers, books, letters, audio and video oral histories, maps –all your digital assets!



http://www.contentdm.com/customers/

CONTENTdm 4.3

- Connexion digital import
  - New option for OCLC catalogers
  - Add digital resources to CONTENTdm via Connexion cataloging process
- Benefit
  - More options for building digital collections
  - Catalog using most convenient process for the organization

CONTENTdm 4.3

Connexion digital import

- Add items to CONTENTdm via the Connexion Client
- Digital collection growth built into cataloging workflow
- WorldCat MARC record crosswalked to Qualified Dublin Core and added to CONTENTdm
  - OCLC number stored in CONTENTdm – a global persistent identifier
- Digital items accessible by FirstSearch, WorldCat.org and WorldCat Local
- Requires OCLC Cataloging subscription, CONTENTdm license and CONTENTdm Hosting Services

CONTENTdm 4.3

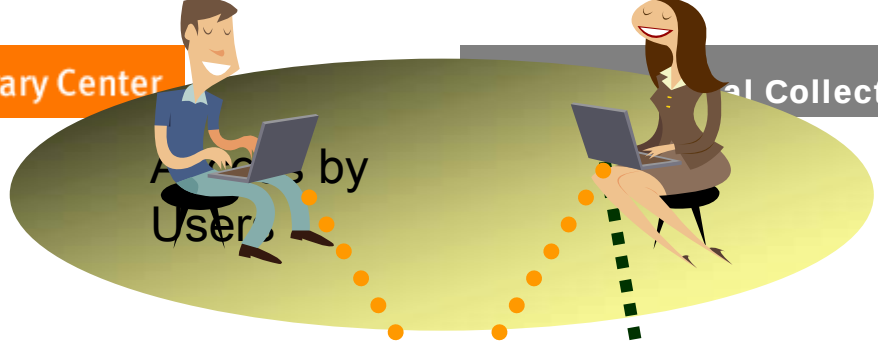
Connexion digital import

- Metadata choices for cataloging
  - Connexion client (MARC)
  - CONTENTdm (DC, QDC, VRA)
    - Acquisition Station
    - Web-based Add option
- Serials support
  - Use “Attach Digital Object” in Connexion client for each issue in a serial item
  - 856 link will automatically retrieve a search results page with links to each issue stored in CONTENTdm

CONTENTdm 4.3

Connexion digital import

- In Connexion Client:
  - Attach Digital Content to existing record
  - Select CONTENTdm collection
  - Select file(s) from local computer/network
  - Replace command
- System processes metadata and file for import into CONTENTdm
  - Digital item sent to CONTENTdm collection
  - MARC metadata mapped to Qualified Dublin Core
  - Compound object creation, JPEG2000 conversion, and OCR or PDF processing, if applicable
  - Thumbnails generated
  - Link added to 856 field in WorldCat record



WorldCat.org

OCLC# hyperlink to digital content

WorldCat

Connexion

CONTENTdm Import

CONTENTdm

Cataloger w/ Connexion Client

Attach digital content to WorldCat record

MARC → QDC
TIFF → JP2
OCR, PDF

Configure CONTENTdm collection with Qualified Dublin Core

CONTENTdm Collection Administrator

CONTENTdm 4.3

- Improved PDF handling
  - Convert multiple-page PDF files to CONTENTdm compound objects
    - Subset print options
  - Search term highlighting within PDF files
  - Creation of thumbnail images from PDF files
  - Improved text extraction from PDF files
- Benefits
  - More efficient processing of PDF documents
  - Single database of **ALL** digital resources
  - Better end user experience

CONTENTdm 4.3

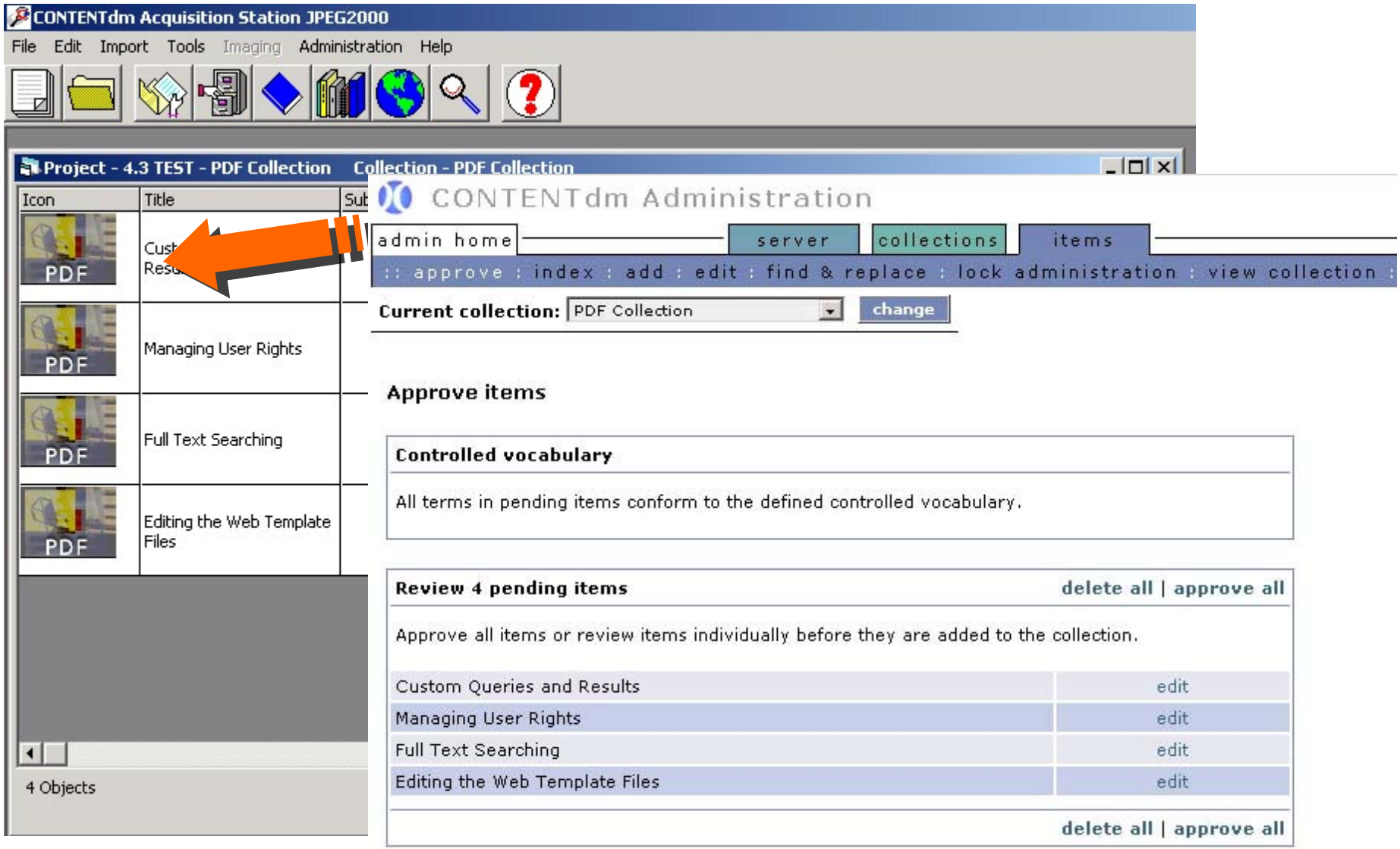
- PDF files can be imported using standard options
  - Single or batch import via Acquisition Station
  - Web-based Add option
  - Connexion digital import
- Thumbnail images are automatically generated from the PDF when the item is added to the collection
- Text is extracted from the PDF and inserted into the full text search field when the item is added to a collection
  - Collection must have full text search field and that field must be empty when PDF is added to the collection

CONTENTdm 4.3

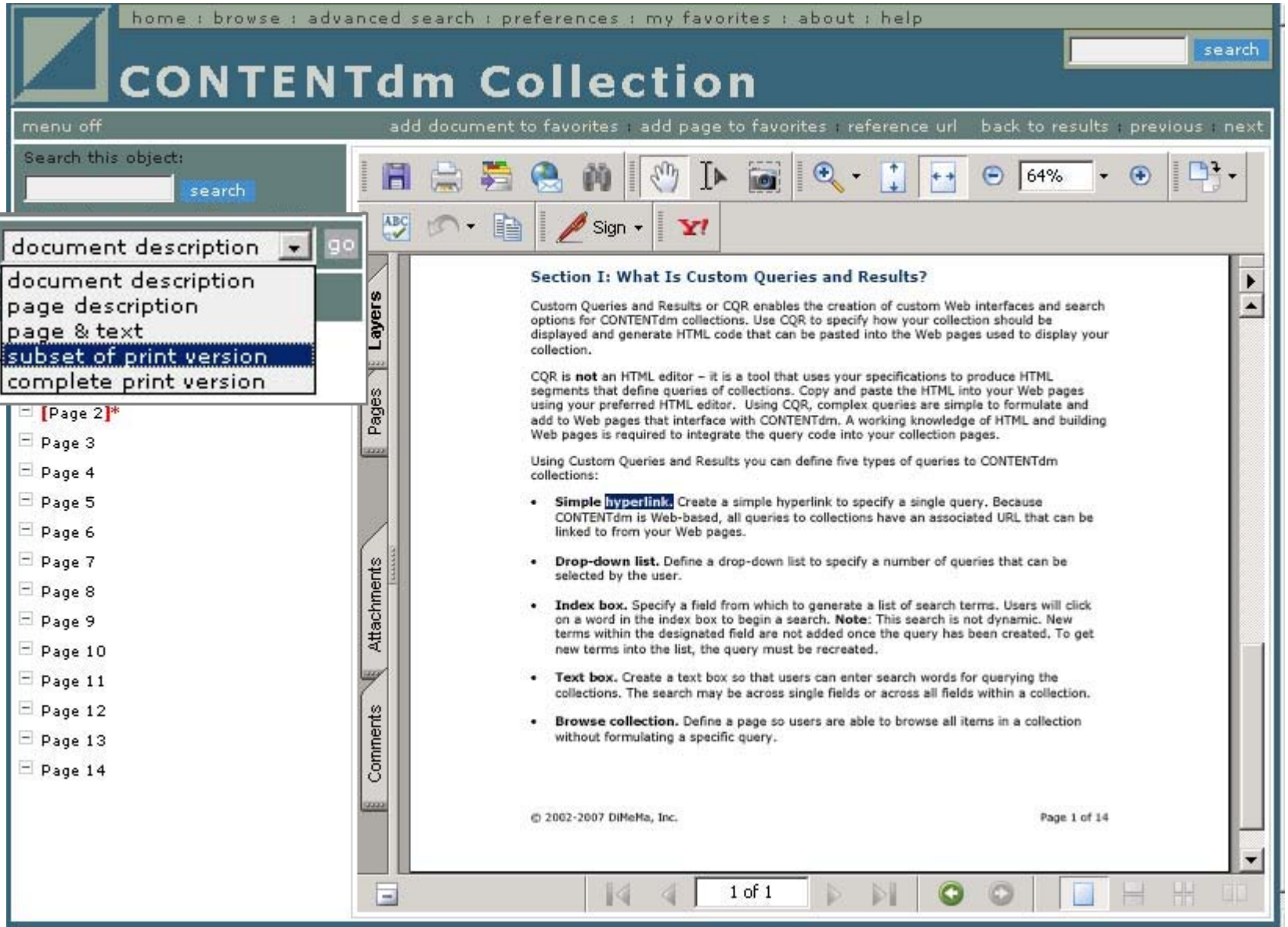
Compound object conversion

- When compound object conversion is enabled, CONTENTdm:
  - Creates a compound object based on the page order of the PDF.
  - Generates a page-level metadata record for each page.
  - Extracts text from the PDF, converts it to UTF-8, and inserts it into the full text field of the associated page level record.
  - Generates thumbnail images of each page of the PDF. The thumbnail image of the first page will also be used for the compound object.
  - Retains the original PDF file for export and printing.
  - Displays the PDF compound object in a compound object viewer with each page of the PDF accessible from the left navigation menu.
  - Highlights search terms in the PDF.
  - Provides an option to select a subset of the PDF to print or save.

Compound object conversion



Compound object conversion

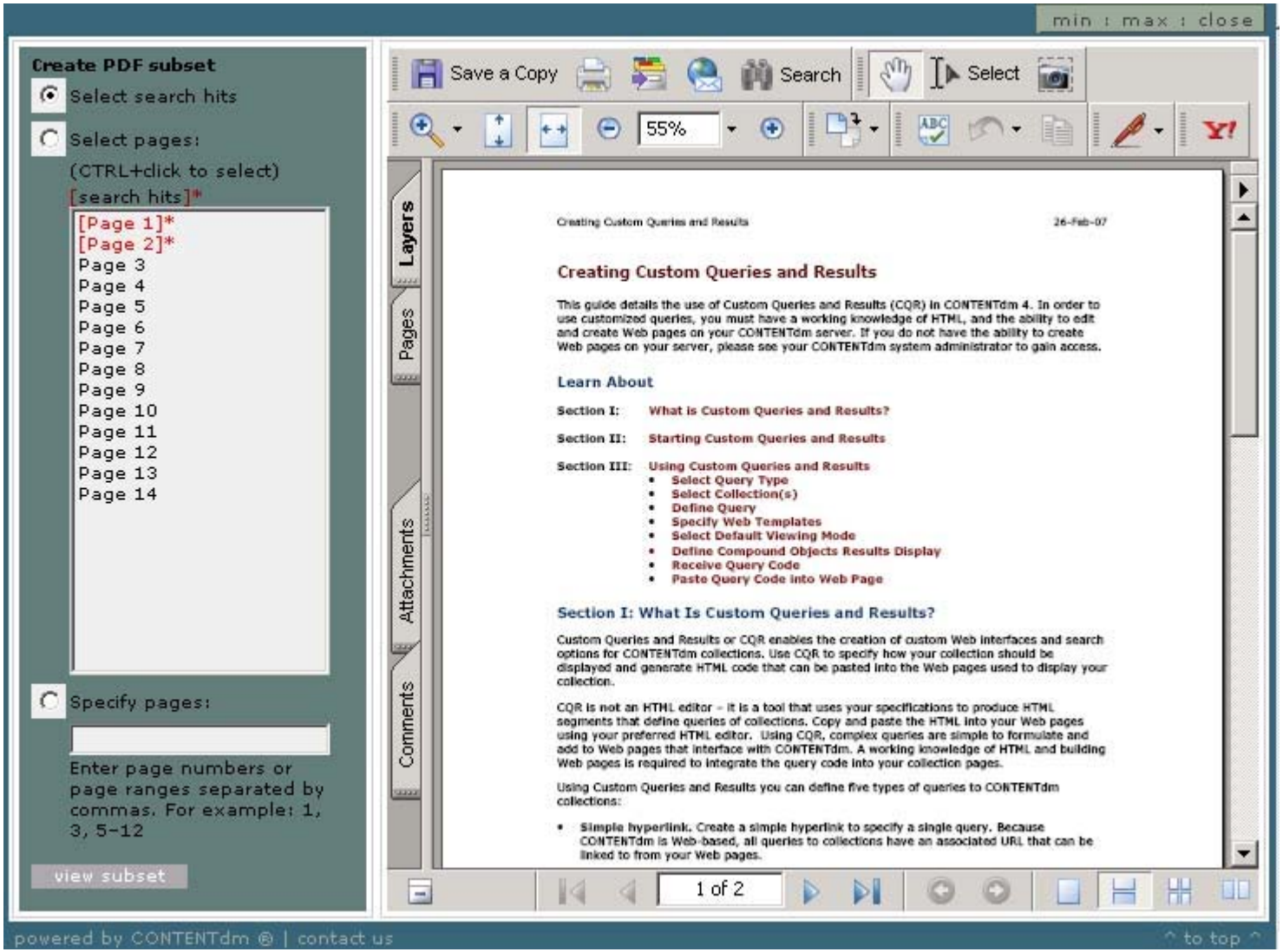


PDF Enhancements

Printing and downloading

- Complete print version
  - Original PDF file retained for printing and saving
- Subset of print version
  - Select a subset of pages from the PDF to view, save, or print
  - Select all pages with search hits or pick individual pages or page ranges
  - Do not have to wait for large download if only need a few pages
  - Also available for non-PDF compound objects when they have been processed using the OCR Extension

Printing and downloading



CONTENTdm 4.3

Compound object conversion

- Reduce the size of file that is downloaded for viewing
  - An entire PDF may be several MB but individual pages are much smaller
  - View a page within large PDF without downloading the full document
  - Increase speed of access to view
- Provide full text indexing by page not document
  - No secondary search required to find specific content in PDF
- Print only the information you need
- Better end-user experience!

CONTENTdm 4.3

Compound object conversion

- Quick and efficient for collection builders!
- PDF pages of compound object do not count against total number of items on the server!
- Ideal for born digital documents!
  - Theses, dissertations, government documents, e-publications, and more…

Digital Repository
- future directions

- Automatic WorldCat registration and links
  - Register and replicate
  - Web-based user controlled mapping DC, MARC, VRA
  - Auto add WorldCat record #
- Seamless/transparent WorldCat ingest
- Update of all digital items with unique WorldCat reference number
- Integration with all features and functions of WorldCat.org
- Local branding and customization

OCLC Enterprise Strategy

"Discovery to Delivery" Collection Curation

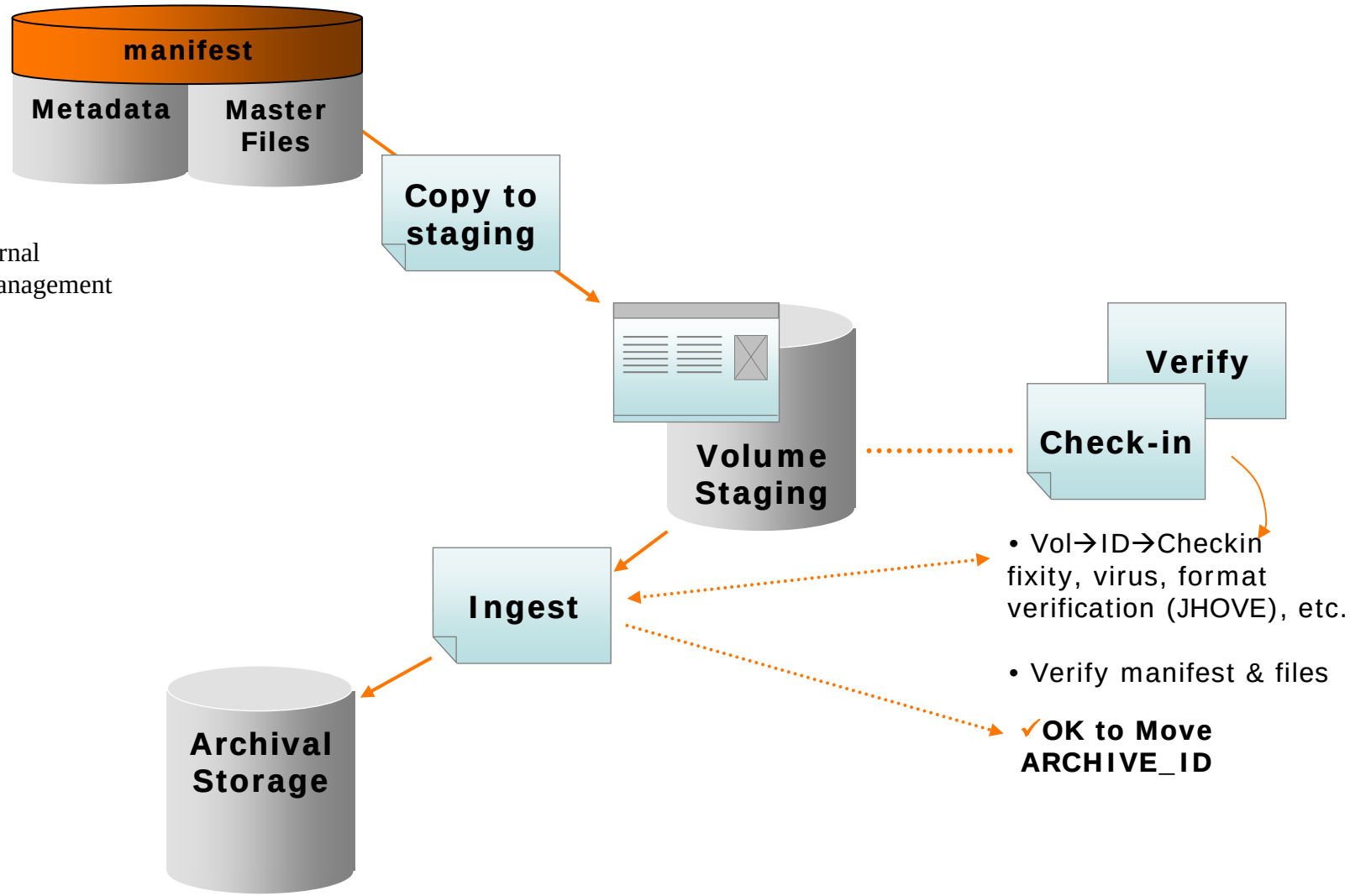
- Consumer Platform
  - WorldCat.org "Indiana Cotton Mills"
- Management Systems
  - WorldCat Local "Hegg Alaska Photographs", "meed"
- Content Platform
  - Digital Repository
  - Digital Archive
- Network Services
  - Use statistics
  - Collection Analysis

Digital Collection Services

Digital Archive – re-engineered

- **Development of storage infrastructure in Dublin**
- **Redefinition and packaging of Digital Archive for market**
  - long term archiving of digital masters
  - opportunity to reprocess as software improves
    - example OCR
- **Significantly reduce cost of service**
- **Two modes**
  - With repository
  - Mirroring

Digital Archive workflow for other systems
Type A service: batch mirror



Digital Repositories
cooperative advantage

- Broaden use of the repository
  - Cultural heritage organizations
  - Institutional Repositories
  - Government documents
- Integrate with library tools and systems
  - Connexion
  - WorldCat.org
  - WorldCat Local

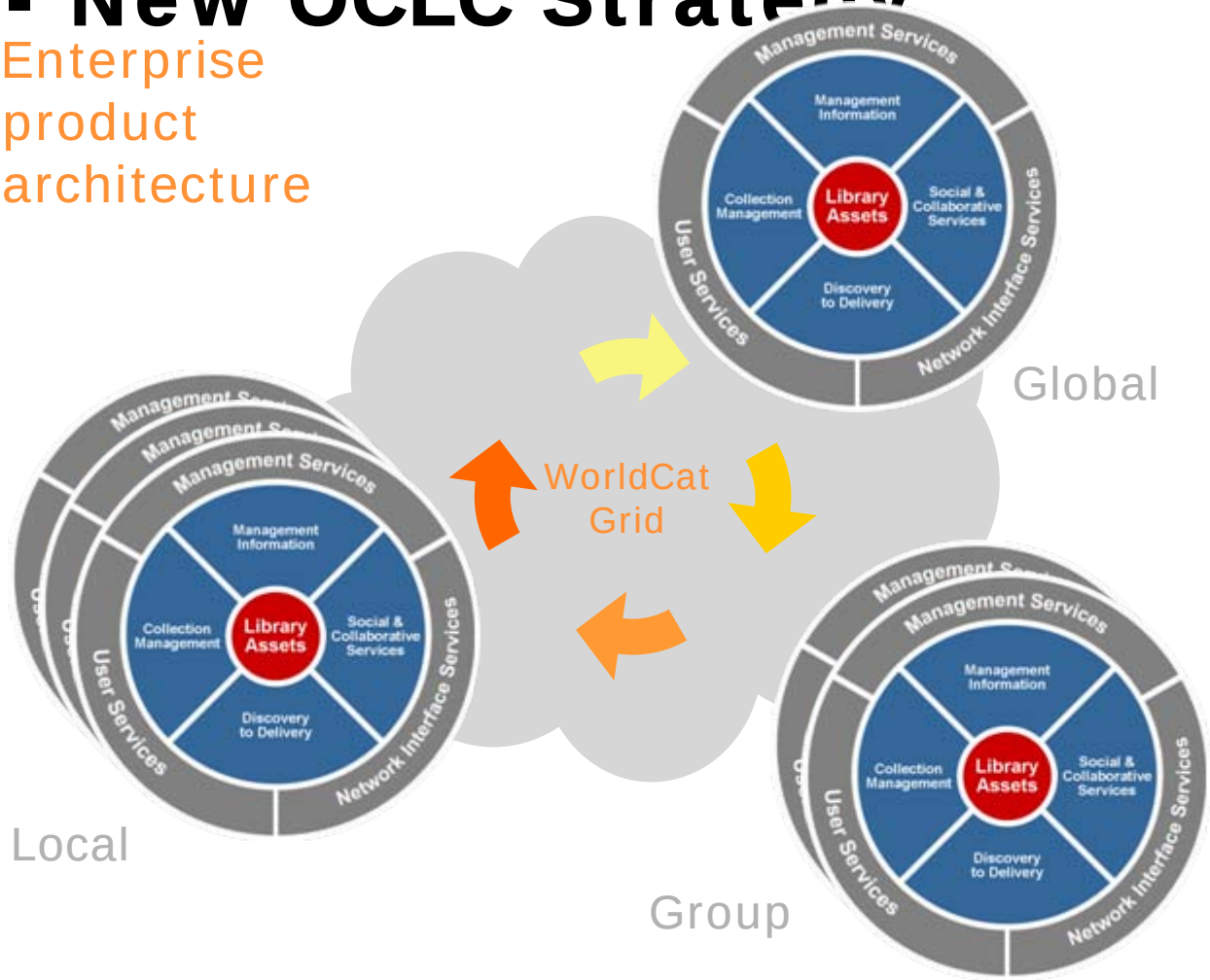
Digital Repositories
cooperative advantage

- Integrate with other content
  - Journals
  - Books, manuscripts
  - Audio and video

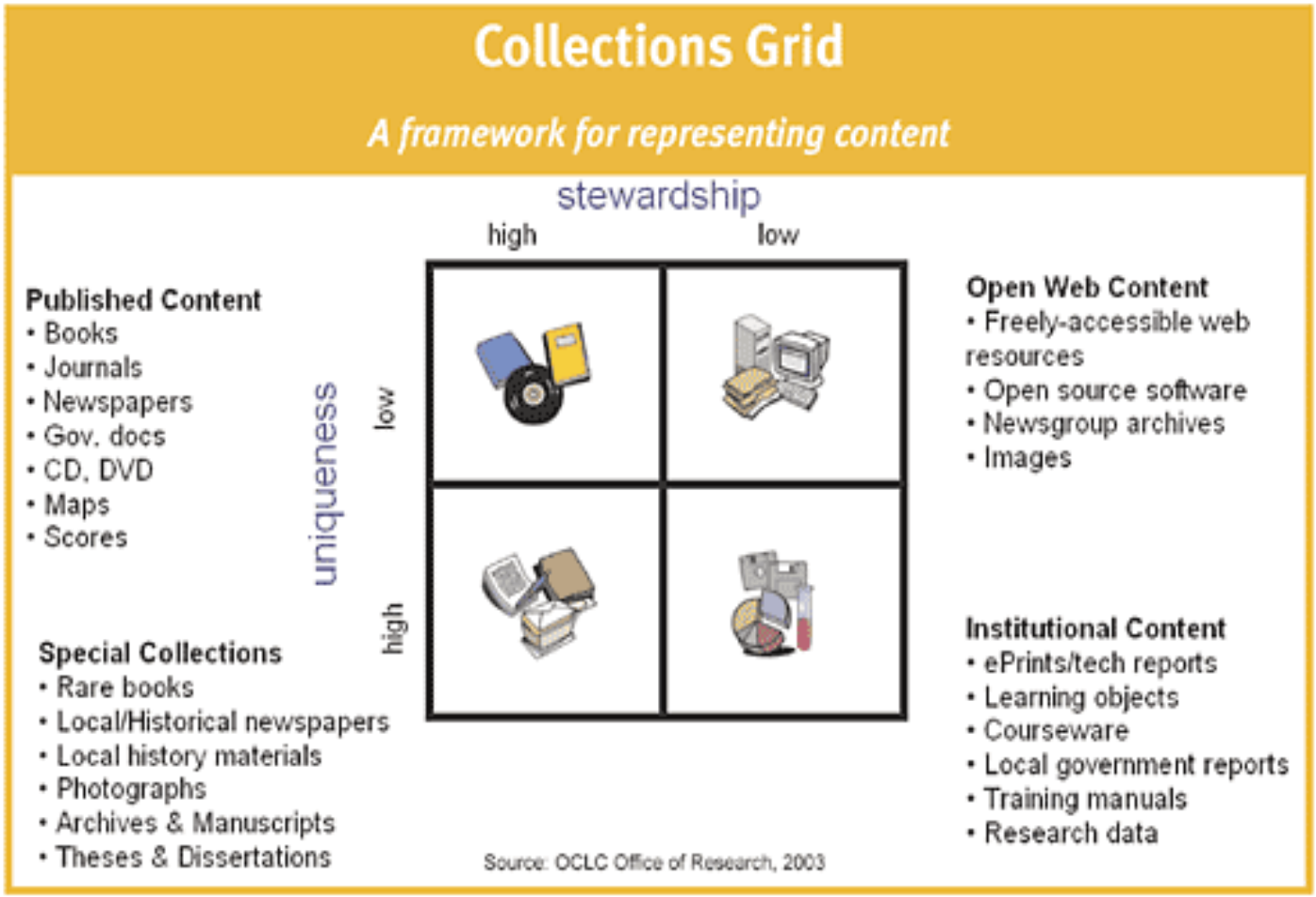
- Increase the value, reduce the costs
  - Search across institutional collections
  - Search within the larger scope of content
    - WorldCat.org
  - Reduce cost through shared services

Future Direction - New OCLC Strategy

Enterprise product architecture



Future direction - Products and services to support the collections grid



Future direction - support for the full range of library content

- dContent
- eContent

- “Digital Stacks”

Thank you

Questions?

email:Greg_Zick@oclc.org