



Colorado's Historic Newspaper Collection: A Digitization Success Story

Brenda Bailey-Hainer
CHNC Program Director
Colorado State Library
June 20, 2006

www.ColoradoHistoricNewspapers.org



Overview

- Description of CHNC program
- Workflow
- Hardware and software
- Demonstration
- Access points
- Collection development
- User response
- Funding
- Future plans



Colorado's Historic Newspaper Collection

- Single source of historic newspapers
 - Statewide coverage
 - Available online
 - Starting with earliest – from 1859
 - Long term goal:
 - Comprehensive as possible
 - 275+ newspapers published before 1923
 - Additional years as copyright is cleared



Colorado's Historic Newspaper Collection Partners

- On-going endeavor of:
 - Colorado State Library
 - Colorado Historical Society
- Collaborative Digitization Program (CDP)
 - Lead agency on original LSTA and IMLS grants



Colorado's Historic Newspaper Collection

- As of June 2006
 - 93 newspapers
 - 50 cities represented
 - 36 counties represented
 - Around 327,000 pages digitized
 - Coverage from 1859 to 1930
 - Includes papers in English, German, Spanish, and Swedish



Processing Workflow Overview

- Digitization based on newspaper microfilm held centrally at Colorado Historical Society
- State Library staff manages software, hardware, telecommunications, user support
- Program Director at State Library provides general oversight, fund raising, financial management
- Advisory Group makes recommendations
- Olive Software used for processing, software interface



Processing Workflow Detail

- CSL - secures funding
- CSL - requests master negative from State Archives
- CSA – pulls master negative
- CHS – picks up negatives at Archives
- CHS – delivers negative to microfilm duplication
- Duplicate negative made
- CHS – picks up original negatives and duplicates
- CHS – returns original to Archives
- CHS – ships duplicates to Olive in Israel



Processing Workflow Detail

- Olive Software in Israel
 - Creates profile for each unique newspaper title
 - Scans the microfilm at 300 dpi
 - Creates metadata
 - Performs “distillation” on images
 - Creates article segmentation
 - Creates repository and index data files
 - Ships data and raw TIFFs to CSL on DVDs
 - Ships microfilm to CHS



Processing Workflow Detail

- CSL loads data and index files
- CSL sends raw TIFFs to offsite storage
- CSL updates "What's New" and "Titles Available" list
- CSL creates mastheads for new titles
- CSL conducts quality assurance and verification of loaded data
- CSL notifies donors of availability of newspapers




Hardware

- Original server configuration designed with expansion in mind
- Windows 2000 Dell PowerEdge server
 - 2 CPUs
- DAS (direct attached storage)
 - 2 terabytes of storage
 - Approx 200,000 pages per terabyte
- T1 internet access
- 8-slot LTO1 back-up solution



Software

- Olive Software
 - ActivePaper Platform
 - ActivePaper Archive
 - ActivePaper Administrator
 - ActivePaper Librarian
 - ActivePaper Repository Indexing Agent
 - Pipex processing software



Olive Software Searching Features

- Select newspapers from list, then by date
- Leaf through an issue page by page
- Search full text of newspapers
 - Topics, names, companies, organizations, etc.
 - Search one newspaper or all papers
 - Limit searching by date
 - Limit searching to articles, pictures or ads
- Look at article by itself
- Look at article in context
- View featured articles on specific themes
- Supports broadband or dial-up connection



Olive Software Features Customized for CHNC

- Large number of newspapers made navigation challenging
- Map and regional approach developed
- State divided into 7 regions
- Separate skin developed for each region
- Users may search or browse newspapers within a region
- Ability to select any combination of papers



User Challenges

- Frequent title changes
- County boundaries change over time
- Some towns extinct
- OCR "distillation" challenging
 - Broken type face
 - Blurry type
- Article segmentation challenging in oldest publications
 - Articles often not clearly delineated
 - Text ads mixed in with news stories
- Archaic/historic language
- Size of the database



Colorado's Historic Newspaper Collection

- Demonstration



Access to CHNC

- Cataloged and included in OPAC
- Link on library website
- Featured on Colorado Virtual Library homepage

SKYLINE

Result page: 1 2 3 4 5 Next

Save Maked Records Save All On Page

Name	Mark	TITLE (1 - 12 of 57)	Year
Colorado's Historic Newspaper Collection			
1	<input type="checkbox"/>	Manitou Journal (Manitou, Colo. : 1984) Internet; F771 .A43	
2	<input type="checkbox"/>	The Bayfield Blade (Electronic Resource) Internet; F771 .B29	
3	<input type="checkbox"/>	The Boulder County Courier (Electronic Resource) Internet; F771 .B38	
4	<input type="checkbox"/>	The Boulder County News (Electronic Resource) Internet; F771 .B385	
5	<input type="checkbox"/>	The Boulder County Bonnet (Electronic Resource) Internet; F771 .B507	
6	<input type="checkbox"/>	Boulder News And Courier (Online) Internet; F771 .B598	
7	<input type="checkbox"/>	Burlington Free Press (Burlington, Colo. : 1871) Internet; F771 .B97	
8	<input type="checkbox"/>	The Colorado Antelope (Electronic Resource) Internet; F771 .C652	
9	<input type="checkbox"/>	Colorado Banner (Electronic Resource) Internet; F771 .C65	
10	<input type="checkbox"/>	Colorado Daily Chieftain (Electronic Resource) Internet; F771 .C653	
11	<input type="checkbox"/>	Colorado Journal (Denver, Colo. : Weekly) Internet; F771 .C654	
12	<input type="checkbox"/>	Colorado Miner (Georgetown, Colo.) Internet; F771 .C655	

Result page: 1 2 3 4 5 Next

SKYLINE

Record 2 of 52
Record: [Back](#) [Next](#)

Title: [The bayfield blade \(electronic resource\)](#).

Publisher: Bayfield, La Plata County, Colo. : D.L. Eggen.

Click on the following to:

[Locate in title list](#) [View full record](#) [View full record](#) [View full record](#)

LOCATION	CALL #	STATUS
Internet	F771 .B29	ONLINE

Period title: [Bayfield blade](#)

Call #: F771 .B29

Location: Internet

Frequency: Weekly

Pub history: Print published 1905-1925

Series: [Colorado's historic newspaper collection](#)

Note: Digitized by Colorado Digitization Program in collaboration with the Colorado State Library and the Colorado Historical Society.
Note from the click down menu of opening screen of Colorado historical newspaper collection (viewed July 14, 2004):
Mode of access: World Wide Web.

Subject: [Newspapers -- Colorado -- La Plata County](#)
[Colorado -- History -- Sources](#)
[Colorado -- History -- Local -- Sources](#)

Other contrib: [Colorado Digitization Program](#)
[Colorado State Library](#)
[Colorado Historical Society](#)

OCLC #: 35344925

Record 2 of 52
Record: [Back](#) [Next](#)

Autumn Library, 1500 Lawrence St., Denver, Colorado 80202
Circulation Desk: 303.556.2639 | Reference Desk: 303.556.2365

Colorado State University Libraries
 Catalog (Range) | Databases | Journals | Media | Search | Home | About Us | Web Search

Databases by Subject (newspaper and newswire services)

More Information | Help & Support | Full Text | New Database | Off Source | Partial Full Text | All Off-Campus Access | All Full-text (page 2 of 2)

[Return to Main Page](#)

Database	Features	Coverage	Notes/Warnings
Academic OneFile		Dates: -- current Updated: varies	<ul style="list-style-type: none"> 1) Do not use the browser Back button -- use the system-provided navigation. 2) To avoid spam pages click here. When using Company Profiles within the Business portion of Academic OneFile News, please see this link.
Business Database		Dates: 1990 - present Updated: weekly	<ul style="list-style-type: none"> Visit the OCLC System Alerts page to get the most up-to-date information on any systems problems. click here.
Business First			<ul style="list-style-type: none"> 1) Users with older browsers or slower modems may use the Text Only version. 2) Users report results click here. Please note the minimum browser requirements: * Netscape 7.x, or * Microsoft Internet Explorer click here.
Business Industry		Dates: 1994 - present Updated: daily	<ul style="list-style-type: none"> Visit the OCLC System Alerts page to get the most up-to-date information on any systems problems. click here.
Cofeagan		Dates: 1990 - present	
Colorado's Historic Newspaper Collection			
EBSCO Regional Business News			<ul style="list-style-type: none"> 1) Users with older browsers or slower modems may use the Text Only version. 2) Users report results click here. Please note the minimum browser requirements: * Netscape 7.x, or * Microsoft Internet Explorer click here.
Ethnic NewsWatch		Dates: 1970 - present Updated: monthly	
Eisenfax		Dates: 1997 - present Updated: varies	<ul style="list-style-type: none"> This database is currently not compatible with Netscape version 4.7 or older.
Historical Newspapers: The New York Times		Dates: Sept. 18, 1851 - ongoing	<ul style="list-style-type: none"> For Netscape Browser, "Enable all cookies" or "Enable cookies for the originating Web site only". Full text coverage is for the Eastern Edition of the NYT which the publisher considers the paper of record.

Douglas County LIBRARIES

Search

Home | About Us | Board of Trustees | Foundation | History | Hours & Locations | Help | Search | Terms

[Home](#) | [About Us](#) | [Board of Trustees](#) | [Foundation](#) | [History](#) | [Hours & Locations](#) | [Help](#) | [Search](#) | [Terms](#)

Magazine & Newspaper Links

Local & Regional

Magazines & Newspapers-Local & Regional Resources

Magazines & Newspaper Databases
 Connect to our Research Sources

Colorado's Historic Newspapers Collection
 View Colorado's historic newspapers.

Denver Business Journal
 Top business journal for the Denver metro area covers all the industries.

Denver Post
 Local and national news as well as classified ads! Find apartments, homes, cars, jobs, news, entertainment and more.

Douglas County Daily News Press
 Read the newspaper of Douglas County, or search their archives. You can also browse the main articles of the paper from 1996 to now!

Rocky Mountain News
 More local news and classified ads! Find apartments, homes, cars, jobs, news, entertainment and more.

Westward
 Find out about the local music, art and entertainment scenes and read hard hitting news stories of local importance.

205

Colorado Virtual Library Presented by: Colorado State Library

Welcome to Colorado Virtual Library (CVL)
<http://www.coloradovirtuallibrary.org>
 CVL provides access to information resources of Colorado libraries.

Search Colorado Libraries

Search Type: Keyword Author (last, first) Title
 Search Group: [Quick Search](#) | [Create a Group](#)
 Search Text:

Use the Quick Search Group to search Colorado's largest libraries.

Search Colorado Libraries by

Region: 
 Type/Kind of Library:
 K-12
 Public
 Academic
 Other

Click on a region of the map to search a group of libraries.

MyCVL Login:

Name:
 Password:
[Forgot a MyCVL?](#)
[Create a New Account](#)

Latest by Telegraph.
 The Hall Road Occupied.
 The Last Rifle Drive.
 First Train Through From Cheyenne, Horsham!!!
 Glory Badolgerum!!!
The Daily Register-Call.
 June 23, 1878 p. 2

This week in Colorado History
 Click on the image for the full story from Colorado's Historic Newspaper Collection.

CVL Resources

- [Info](#)
- [Teachers](#)
- [Digital Collections](#)
- [Identification Links](#)
- [Historical Newspapers](#)
- [Donation Info](#)

AskColorado - Anytime
 A series of Colorado Libraries
 Chat live with a librarian anytime, 24 hours a day, 7 days a week except major holidays.

Colorado's Historic Newspaper Collection
 Search and read Colorado newspapers from 1859-1923.

[Contact Us](#) | [About Us](#) | [FAQ/Help](#) | [Colorado Libraries Collaborate](#)

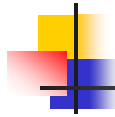
Collection Development

- General policies:
 - Historic newspapers published in Colorado
 - Start with earliest available published issue of a title and move forward chronologically
 - Currently driven by donations
 - Target fundraising efforts to counties with no papers online
 - Focus on issues published prior to 1923 due to copyright concerns
 - Use master negatives only unless special circumstances



Collection Development

- Unintended positive consequences
 - Easier to determine missing issues
 - Has helped to identify existing paper copies for filming



Copyright Concerns

- Format owners
 - E.g. Owner of the microfilm negative of newspaper
- Content owners
 - Newspaper owner
 - Heirs of family, commercial firm, purchaser of newspaper
 - Writers
 - Photographers
 - Newswires



Usage

- First year of operation - 2004
 - 1.3 million views of articles, ads, pages
- Second year - 2005
 - 3.5 million views
- Third year - 2006 to date
 - 1.5 million views
 - 3,500 unique visitors monthly
 - Average visit length – over 25 minutes



User Survey Results

- 52% over age of 60
- 45% between age 40 and 60
- 71% live in Colorado
- 44% use it for genealogy/family history
- 34% use it for general and/or historical research



Funding

- Project started with 2 grants to CDP, CSL & CHS
 - LSTA - \$120,000 & IMLS - \$249,232
 - Software, hardware, & first 96,000 pages digitized
 - Focus groups with educators, librarians, researchers
 - Creation of K12 lesson plans using primary source materials
 - Marketing flyers



Funding

- Donations of \$425,000+
 - \$1.25 digitizes and indexes 1 page
 - \$400 for each newspaper title "profile"
 - Donors can provide general support or help fund specific newspapers
 - Estimate 2.25 million pages available to digitize < 1923



Future Plans

- Add more newspapers monthly (as funding permits)
- Add more featured groups of articles
 - E.g. historic events, important people
- Add history of newspapers
- Add Colorado timeline
- Assess feasibility of adding keywords and/or subject headings to articles
- Upgrade to OAI Harvestable version of Olive
- Make CHNC Z39.50 searchable



Colorado's Historic Newspaper Collection



Brenda Bailey-Hainer
CHNC Program Director
Colorado State Library
bailey_b@cde.state.co.us
303.866.6907

Steve Walker
CHNC Systems Administrator
sysadmin@aclin.org
303.751.6277 ext. 118

www.ColoradoHistoricNewspapers.org