

Multiple options for supporting online newspaper collections

CONTENTdm offers a range of solutions for newspapers

CONTENTdm® Digital Collection Management Software offers a range of solutions for your historical newspapers, including support of the National Digital Newspaper Program (NDNP) data format. With CONTENTdm, you can turn your historical newspapers into fully searchable online resources available from any Web browser.

Select the best option for your newspaper digitization project

CONTENTdm offers a range of solutions for your institution's historical newspapers, from a basic "do-it-yourself" solution to an integrated turnkey solution. We will work with you to select what is most appropriate for your institution, the size of your project, the digital file format you prefer and your budget. Additionally, since the CONTENTdm software supports emerging newspaper digitization specifications—including the National Digital Newspaper Program (NDNP) format and METS/ALTO format—CONTENTdm is the best choice if your institution is pursuing grant funding.

Using CONTENTdm on your desktop, you may process your newspapers for full searchability and for high-resolution, rapid display. Alternatively, you can let us do the work for you. We will scan your paper or microfilm and return a completely built CONTENTdm newspaper collection accessible via the Web, along with your archival source files. We can even install the collection on your CONTENTdm server.

No matter which solution you choose, you'll be able to provide convenient, 24/7 access to your digital newspaper collections through customizable Web interfaces. Also, with CONTENTdm, you'll provide end users with a number of searching features including full-text searching, searching across multiple digital collections of all item types and highlighted search words in both images and text.

Whether you choose to process your newspapers in-house or outsource the project, we recommend you use JPEG2000 for display images so that your users can view a high-resolution image of the newspaper pages without the slow load time of a full-resolution TIFF file.

Available features:

- Import of the NDNP data format
- JPEG2000 for large format pages
- Printable PDF versions
- An electronic article clipper
- Customizable interfaces
- Full-article segmentation highlighting
- Persistent reference URL to individual newspaper pages
- Full-text searching
- Searching across multiple collections of all media types, including photos, audio and video
- Highlighted search words



End users can view a newspaper page, and then select an individual article for closer viewing. They also can select and extract the article to save locally, print or e-mail.

The San Francisco Call, courtesy of The University of California, Riverside.

Multiple options for supporting online newspaper collections
CONTENTdm offers a range of solutions for newspapers

In-house newspaper processing

Among the desktop publishing options, you have a number of choices for processing your historical newspaper collections and showcasing them on the Web.

One option is to import your scanned newspaper files into CONTENTdm as compound objects. If you already have full-text transcripts of the newspapers, your users will see search terms highlighted in the text window, side-by-side with the page image.

Alternatively, using the CONTENTdm Project Client's integrated OCR (optical character recognition) option, you can produce searchable transcripts in one step. The end user will see the retrieved term highlighted not only in the text window, but also within the page image.

In either case, if you want to make printable PDFs available to end users, you can choose to generate a PDF of the entire document for printing during the import process.

If your newspapers were "born digital", i.e., produced originally as PDF files, you may import them as simple items, taking advantage of several powerful CONTENTdm features:

- Automatic text extraction for cross-document and cross-collection searchability with term highlighting
- Page-level indexing for rapid retrieval and display

View newspapers on the Web with CONTENTdm

To view several newspapers that are available on the Web and through CONTENTdm, please visit: <http://newspapers.contentdm.oclc.org/>. If you are interested in seeing CONTENTdm's article segmentation feature, please view *The San Francisco Call*.

Integrated turnkey processing

The Integrated turnkey option is a professional newspaper processing solution. OCLC digitizes your newspapers (paper or microfilm) and processes them to your specifications.

Once processed, we can build and install the CONTENTdm collection for you, providing convenient online access for your users. All of the features enjoyed with the desktop publishing option are included when we process the newspapers for you—full-resolution images are returned to you in addition to the fully built collection with JPEG2000 display images and thumbnails. Full-text searching is available across all of your newspapers, as well as across all collections published on your CONTENTdm server.

In addition, with the turnkey option, you also can choose to have article segmentation data generated. If you do, your users will be able to highlight entire articles. Whether they have searched on term that is highlighted within the page image, or just moved the mouse over a segmented article, the whole article is highlighted and can be extracted with one click. The title information is also included in the 'clipping'.

For more information

Contact OCLC at 1-800-848-5878, ext. 4301
or e-mail contentdm@oclc.org.

www.oclc.org/contentdm