Elizabeth M. Caringola
Historic Maryland Newspapers Librarian
Digital Programs and Initiatives
Digital Systems and Stewardship

Wochen-Ausgabe
Der Deutsche Correspondent.
Zweiter Theil.


Historic Maryland Newspapers Project
Presentation for Digital Maryland Conference 2014
March 7, 2014

Elizabeth M. Caringola
Historic Maryland Newspapers Librarian
Digital Programs and Initiatives
Digital Systems and Stewardship

UNIVERSITY OF MARYLAND

UNIVERSITY LIBRARIES
Introduction to the NDNP

• The National Digital Newspaper Program (NDNP) is a joint effort by the National Endowment for the Humanities (NEH) and the Library of Congress (LC) to digitize historic newspapers from every U.S. state and territory.

• The goal is to create “an Internet-based, searchable database of U.S. newspapers with descriptive information and select digitization of historic pages.”

• Each state/territory can be awarded an NDNP grant to digitize newspapers published between 1836 and 1922.

• Newspapers are digitized from a second-generation duplicate of the camera master microfilm.

• During 2-year grant cycle, awardee institutions must deliver 100,000 digitized pages to LC for upload to Chronicling America.
Selecting newspapers for digitization

The NDNP content selection guidelines ensure that relevant titles and suitable microfilm are chosen for digitization.

Content criteria
- Research value
- Geographic representation
- Temporal coverage
- Orphan titles
- Diversity
- Online availability

Microfilm
- Technical quality
- Bibliographic completeness of the microfilm copy
Technical specifications:

Images

NDNP Technical Guidelines for 2012 Awards

- Technical targets
- Master image
  - uncompressed TIFF 6.0
  - 8-bit grayscale
  - 300-400 dpi
- Use images
  - JPEG2000
  - PDF that supports full-text search

Technical specifications: Metadata

NDNP Technical Guidelines for 2012 Awards

- Full and up-to-date Cooperative Online Serials (CONSER) bibliographic record at the title level for the print newspaper
- Issue- and page-level metadata
- Reel metadata

- All metadata is delivered in METS object structure according to an XML template
Optical character recognition (OCR) is captured for every page that is digitized

- ALTO XML schema captures the content and position of printed text
- Allows for full-text search and highlighting of search terms
Summary of digitized content

NDNP Technical Guidelines for 2012 Awards

Page
- Image files: TIFF, JPEG2000, PDF
- XML file that contains OCR

Issue
- XML file that contains issue- and page-level metadata

Reel
- Image files for preservation targets and microfilm targets
- XML file that contains reel technical metadata and metadata for targets

Batch
- A batch manifest lists all reels and issues included in the batch
Historic Maryland Newspapers Project

- UMD Libraries joined the NDNP during the 2012-2014 award period

- To date:
  - 35,916 pages of Maryland newspapers are live on Chronicling America
  - 46,763 pages at LC awaiting ingest
# Titles selected for digitization, 2012-2014

<table>
<thead>
<tr>
<th>Title</th>
<th>Publication location</th>
<th>Years to digitize</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Republican and Baltimore daily clipper</td>
<td>Baltimore, Md.</td>
<td>1844-1846</td>
</tr>
<tr>
<td>Baltimore commercial journal, and Lyford's price-current</td>
<td>Baltimore, Md.</td>
<td>1840-1849</td>
</tr>
<tr>
<td>Baltimore daily commercial</td>
<td>Baltimore, Md.</td>
<td>1865-1867</td>
</tr>
<tr>
<td>Civilian &amp; telegraph</td>
<td>Cumberland, Md.</td>
<td>1859-1875</td>
</tr>
<tr>
<td>The daily exchange</td>
<td>Baltimore, Md.</td>
<td>1858-1861</td>
</tr>
<tr>
<td>Der deutsche Correspondent</td>
<td>Baltimore, Md.</td>
<td>1858-1918</td>
</tr>
<tr>
<td>The pilot and transcript</td>
<td>Baltimore, Md.</td>
<td>1840-1841</td>
</tr>
<tr>
<td>Maryland free press</td>
<td>Hagerstown, Md.</td>
<td>1862-1868</td>
</tr>
<tr>
<td>Title</td>
<td>Publication location</td>
<td>Years to digitize</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>The aegis &amp; intelligencer</td>
<td>Bel Air</td>
<td>1864-1922</td>
</tr>
<tr>
<td>The Baltimore daily news</td>
<td>Baltimore</td>
<td>1885-1892(?)</td>
</tr>
<tr>
<td>Calvert gazette</td>
<td>Prince Frederick</td>
<td>1885-1922</td>
</tr>
<tr>
<td>Calvert journal</td>
<td>Prince Frederick</td>
<td>1867-1922</td>
</tr>
<tr>
<td>Catoctin clarion</td>
<td>Mechanicsville</td>
<td>1871-1922</td>
</tr>
<tr>
<td>The Cecil Democrat</td>
<td>Elkton</td>
<td>1850-1922</td>
</tr>
<tr>
<td>The citizen</td>
<td>Frederick</td>
<td>1895-1922</td>
</tr>
<tr>
<td>The Cumberland daily news</td>
<td>Cumberland</td>
<td>1871-1890</td>
</tr>
<tr>
<td>The daily banner</td>
<td>Cambridge</td>
<td>1902-1922</td>
</tr>
<tr>
<td>Democratic messenger</td>
<td>Snow Hill</td>
<td>1869-1922</td>
</tr>
<tr>
<td>Frederick herald</td>
<td>Frederick</td>
<td>1832-1861</td>
</tr>
<tr>
<td>Frostburg mining journal</td>
<td>Frostburg</td>
<td>1871-1913</td>
</tr>
<tr>
<td>Havre de Grace Republican</td>
<td>Havre de Grace</td>
<td>1881-1922</td>
</tr>
<tr>
<td>The leader</td>
<td>Laurel</td>
<td>1897-1922</td>
</tr>
<tr>
<td>Montgomery County sentinel</td>
<td>Rockville</td>
<td>1856-1922</td>
</tr>
<tr>
<td>The Republican citizen</td>
<td>Frederick</td>
<td>1836-1890</td>
</tr>
<tr>
<td>St. Mary's beacon</td>
<td>Leonard Town</td>
<td>1845-1863</td>
</tr>
<tr>
<td>St. Mary's gazette</td>
<td>Leonard Town</td>
<td>1863-1867</td>
</tr>
<tr>
<td>Saint Mary's beacon</td>
<td>Leonard Town</td>
<td>1867-1922</td>
</tr>
</tbody>
</table>
Why the NDNP?

• Standard for newspaper digitization

• *Chronicling America* is a free, national database, and it openly shares its data

• LC preserves the master TIFF files and microfilm for perpetuity

• Awardees can use the digitized images and metadata for their own projects/repositories
Resources

• National Endowment for the Humanities

• Library of Congress
  – Chronicling America, http://chroniclingamerica.loc.gov/

• Historic Maryland Newspapers Project at UMD Libraries
  – Blogs
    • DigiStew, Division of Digital Systems and Stewardship, http://dssumd.wordpress.com/
    • Special Collections, http://hornbakelibrary.wordpress.com/