



# Aggregating and Accessing the 'Daybooks of History' with the Minnesota Digital Newspaper Hub

2014 Library Technology Conference  
Wednesday, March 19, 2014  
Sarah Quimby [sarah.quimby@mnhs.org](mailto:sarah.quimby@mnhs.org)  
Jane Wong [jane.wong@mnhs.org](mailto:jane.wong@mnhs.org)

# Newspapers: a very recent history

- Current “newspapers” are not paper anymore
  - Most print news publications are created digitally, submitted digitally, and then printed
- Old newspapers are not paper anymore
  - Newspapers published ca.1870- were printed on high-acid newsprint, which degrades quickly
  - Libraries and archives began microfilming newspapers for preservation and recycling the print

# Newspapers: a very recent history

- Microfilm
  - Stable, but not user-friendly
  - Each reel can only be used at one time in one place
- Born-digital publishers' files
  - Portable, but ephemeral
  - Also not user-friendly in their own way
- Millions of pages of news locked down in microfilm, DVD-ROMs, or network drives

# Digital newspapers

- Many organizations have sought to digitize newspapers:
  - Libraries and historical societies
  - Newspaper publishers
  - Commercial businesses
  - Individual enthusiasts
- The result: newspaper titles are scattered across repositories, digitized to different metadata standards, with different search and display protocols and no way to search them all at once

# Minnesota Digital Newspaper Hub

- The solution: aggregation
  - Minnesota Digital Newspaper Hub: <http://www.mnhs.org/newspapers>
  - 109 titles, 28 counties, 1856-1924\*
  - Seeking to become the most comprehensive aggregation of Minnesota's digital newspapers

\*Newspapers still subject to copyright will only be available in the MNHS reading room during the current phase of development

# Minnesota Digital Newspaper Hub

Why we designed it the way we did:

- Accept 2 streams of digital files (born digital and digitized)
- Accept files from multiple partners
- Collate pages of an issue together
- Collate issues of a title together
- Collate “newspapers” together
- Allow browsing of titles, counties, towns, and years
- Allow keyword searching
- <http://www.mnhs.org/newspapers>



# Partners and files

- Microfilm
  - National Digital Newspaper Program (NDNP)
  - Southeastern Libraries Cooperating (SELCO)
- Born-digital publishers' files
- Digitized files
  - Minnesota Digital Library
  - Internet Archive
  - Other organizations

# Technical Standard: NDNP

1. Microfilm:
  - Digital reproduction should be made from a preservation copy of microfilm (P. 6-7)
2. Image: TIFF & Derivatives (Appendix B)
3. Digitized text: OCR (Appendix B)
  - ALTO (Analyzed Layout and Text Object)
4. Metadata: Issue, Reel, Batch (Appendix A & C)
5. Directory Structure: (Appendix D)

**National Digital Newspapers Program (NDNP): Technical Guidelines**

[http://www.loc.gov/ndnp/guidelines/NDNP\\_201416TechNotes.pdf](http://www.loc.gov/ndnp/guidelines/NDNP_201416TechNotes.pdf)



# Digitizing from Microfilm

- All microfilm is not created equal!
  - Camera master, master negative, positive use copy
  - Acetate base, polyester base, vesicular positive, diazo
  - The best images will be obtained from master negative microfilm that has been properly stored
- Digitizing from print is slower and more costly than from microfilm
  - Most agencies will film first, then digitize that

# Born Digital and Digitized Files

- Retain some original digitization characteristics
- Acquire some NDNP characteristics
- Result:
  - Improved access platform
  - Additional digitization standard

# I want to digitize newspapers

## Considerations

- What do you want to digitize?
- How will you provide access?
- What standards do you use?
- How do you plan for preservation and sustainability?

# I want to digitize newspapers!

- All about ACCESS!
  - Access platform options?
    - What digital newspaper standards are used?
  - What is the future of the access platform?
    - What happens to the digital files if the access platform goes away?
  - What if I want to combine my digital newspapers with someone else's? Aggregation!
  - What if I want to get out of the newspaper access business?
    - Who else uses the digital newspaper standards?

# Copyright and partnerships

- Think very carefully about what you want to do with your digital newspapers
  - A preservation archive with limited in-house access?
  - Public access over the web?
  - To make money?
- *and get permission first!*
  - From the copyright owner
  - From the owner of the master negative
  - From the administrators of the access platform

# Sustainability

- Think very carefully about what you want to do with your digital newspapers
  - How do digital newspapers fit your institution's mission?
  - Who is your audience?
  - Who are your partners?
  - How are you going to fund maintenance and growth in the long term?

# The MNHS can help!

Technical standards and digitization projects:

Jane Wong [jane.wong@mnhs.org](mailto:jane.wong@mnhs.org)

Born-digital publisher files:

Sarah Quimby [sarah.quimby@mnhs.org](mailto:sarah.quimby@mnhs.org)

State and local history grants:

David Grabitske [david.grabitske@mnhs.org](mailto:david.grabitske@mnhs.org)

