Collaborative Preservation Scaled to Serve the Texas Press Association

T E X A S  D I G I T A L  N E W S P A P E R  P R O G R A M

Cape Town, South Africa
IFLA News Media Section
August 16, 2015

Ana Krahmer, Digital Newspaper Program Coordinator

ana.krahmer@unt.edu
Overview

• What is TDNP?
• The Texas Press Association Archive
• Other Preservation Initiatives in the U.S.
What is TDNP?

About us
What is TDNP?

• Dedicated to preserving Texas newspapers, from any time or place, for any title.
• Our earliest newspaper is from 1829, and we are hoping to find earlier Texas issues to beat this date.
• We ingest newspapers on a weekly basis, so numbers are always rising.

ana.krahmer@unt.edu
TDNP Statistics as of 8/11/2015

Statistics for Texas Digital Newspaper Program

<table>
<thead>
<tr>
<th>Item Usage</th>
<th>Added Items</th>
<th>More Data</th>
</tr>
</thead>
</table>

9,581,174 Total Uses / 307,334 Total Items (3,217,386 files) / 296,333 Visible / 11,001 Hidden

Usage by Month/Year

<table>
<thead>
<tr>
<th>Year</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>220,271</td>
<td>241,844</td>
<td>238,711</td>
<td>241,919</td>
<td>287,969</td>
<td>207,352</td>
<td>201,568</td>
<td>61,230</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,700,884</td>
</tr>
<tr>
<td>2014</td>
<td>178,865</td>
<td>165,850</td>
<td>179,184</td>
<td>164,216</td>
<td>227,182</td>
<td>237,335</td>
<td>463,178</td>
<td>172,802</td>
<td>193,984</td>
<td>317,866</td>
<td>174,249</td>
<td>191,622</td>
<td>2,666,223</td>
</tr>
</tbody>
</table>
Texas Press Association Archive

Preserving Recent Texas History
The TPA PDF archive represents 661 titles from across Texas, comprising:

- 457 Paid Circulation Newspapers (Legal Publications by Statute)
- 20 African-American Newspapers (all are free publications)
- 49 Latino Newspapers (all are free publications)
- 135 Suburban Community Newspapers (all weeklies)
- 12TB of data from 2010-2014
TPA Archive Partnership

• Collaboration with the Texas Press Association and NewzGroup out of Missouri.

• 12TB of PDF newspapers, prepared for preservation through batch processing of PDF content.

• Range from 2010-August 2014.

• Scheduled transfer of August 2014-August 2015 newspapers from NewzGroup.
TPA Archive Partnership: File Workflow

- QC work performed after batch processing.
- Add initial layer of metadata, pre-OCR.
TPA Archive Partnership: File Workflow
TPA Archive Partnership: File Workflow

[Image of file browser with various file details]
TPA Archive Partnership: File Workflow

- Example of batch metadata for PDF issues
- Applied in XML file to sets divided by year and/or by content changes (managing editor, publisher, masthead, etc.)
- Newspapers uploaded with embargo are marked as “hidden”=TRUE.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <title qualifier="officialTitle">The Monitor (Mabank, Tex.)</title>
  <title qualifier="serialTitle">The Monitor</title>
  <creator qualifier="edi">
    <info>Nearing Editor</info>
  </creator>
  <type>per</type>
  <name>Centrell, Pearl</name>
</metadata>
```
TPA Archive Partnership: Permissions

- Publishers are busy, hard people to catch.
- When they respond, they respond with interest, with the exception of large-city dailies.
- Embargos have ranged from the most recent 6 months to 3 years.
- The Texas Digital Newspaper Program holds a membership in the Texas Press Association.
- Krahmer attends all TPA annual conventions and summer leadership meetings, along with as many regional Press Association meetings as possible.
- Preservation is about establishing communication, trust.

ana.krahmer@unt.edu
Initiatives in the U.S.

Born-Digital News Preservation
The RJI Dodging the Memory Hole

• Founded and sponsored by the Reynolds Journalism Institute and MU Libraries
• Part of the larger JDNA agenda to identify and preserve born digital news content of any kind.
• Leveraging extant state and regional preservation initiatives
• Current focus on building awareness, strengthening the stakeholder network and public-private partnerships
• Knight Foundation-funded JDNA research on evaluating news archives from small newspapers for possible business models
“Beyond-NDNP”

- Originally was a single meeting supported by the Institute of Museum & Library Services to gather NDNP “alumni” together to discuss what the next steps in newspaper preservation would be for their states.
- Discussion surrounds any kind of newspaper preservation.
- Leveraging training, knowledge, and skills attained through the NEH’s NDNP to support state-wide newspaper preservation programs.
- Next meeting: September 16, 2015.
Questions?
Email: Ana.Krahmer@unt.edu
Call: 940-565-3367
Visit: http://texashistory.unt.edu/explore/collections/TDNP/