

Research from Many Angles: Evaluating Usage and Impact of Digital Newspapers in the  
Texas Digital Newspaper Program

Ana Kraemer

Mark Phillips

University of North Texas Libraries

## Abstract

The Texas Digital Newspaper Program has partnered with communities and institutions from across Texas to digitize, preserve, and make publicly available over 1.3 million pages of newspapers, representing over 70 counties and over 500 titles. Such a large accumulation of content has allowed University of North Texas Libraries to study how these newspapers are being used and contributed by the public through multiple venues, including search term analysis, geographical visualizations, surveys, and of course, usability analysis. The user groups who utilize the Texas newspapers range from university researchers, to K-12 students and educators, to lay historians, to genealogists. Each of these groups represents a significant percentage of the people who regularly work with the digital Texas newspapers and, therefore, UNT Libraries strives to serve and to better understand these populations with as many methods as possible targeted to as many relevant user groups as possible. This paper will examine how these research approaches help build understanding of the impact of the Digital Newspaper Program on the public, and it will exemplify cases in which UNT Libraries have utilized specific research results to improve access, presentation, and the public profile of its digital newspaper content and of the Program overall.

## **What is TDNP, and Who Does it Serve?**

The Texas Digital Newspaper Program (TDNP), hosted out of University of North Texas Libraries, digitally preserves and makes accessible newspapers from any time period, from any Texas community. TDNP partners with newspaper publishers, private foundations, public libraries, museums, historical societies, genealogy societies, small and large universities, and individual collectors to locate and digitize newspapers. At this time, TDNP hosts over 1.5 million pages of Texas newspapers.

The digital newspaper collection is a valued and valuable asset to UNT Libraries, and as such, we regularly assess the collection's impact, through growth by county, by pages of newspapers digitized, by collaborative partnerships, and by titles. This regular assessment allows us over time to analyze newspaper issues digitized, numbers of titles digitized, counties represented via newspapers, numbers of communities participating in TDNP, and types of groups who utilize the newspaper collection. We use this data for multiple research purposes in addition for general library statistics-gathering.

Understanding the usage and growth patterns of digital library collections in comparison to traditional collections is of prime importance to UNT Libraries. By analyzing this usage data, UNT Libraries can plan to focus resources, in the form of effort and funding, toward library initiatives that will reap the highest return on investment, and which stand to reach the widest possible set of users.

User groups who work with the Texas Digital Newspaper Program include, but are not limited to, students and educators in K-12 and higher education; undergraduate and graduate academic researchers; lay-researchers; lifelong learners; family historians; publishers; public and academic librarians; and civic administrators. A mixed-methods study launched in 2011 by Kathleen Murray and Dreanna Belden explored the impact on

users of the collective set of primary source objects available on The Portal to Texas History, and many respondents noted that the digitized newspapers, in particular, were incredibly important to them. This survey received 573 responses over the course of 4 months. Of these 573 respondents, 36% self-identified as genealogists, 19% as lifelong learners, 19% as historians, 6% as librarians, 5% as students, and 15% as “other” (p. 14). Approximately 20% of the optional comments within this survey cited the digital newspaper collection as being especially valuable.

UNT Libraries is also heavily involved in K-12 education through projects like National History Day, Teaching with Primary Sources, and Resources 4 Educators. Tara Carlisle, Collections Development Librarian, constantly works with educators across Texas to promote use of the Texas newspapers available on the Portal within the K-12 classroom. Texas History is taught in 4th and 7th grades in Texas, and the Texas Digital Newspaper Program collection hosts 79,604 pre-1920 Texas newspapers, many of which have been utilized in *Resources 4 Educators’* “Newspaper Narratives,” a set of primary source classroom lesson plans for teaching K-12 students how to utilize Texas newspapers in their research

(<http://education.texashistory.unt.edu/lessons/newspapers/index.html>). According to Carlisle, “Overall, the usage follows the school year calendar, but according to Google Analytics it appears that March through April and October through November have the highest usage” (2014). This usage peak is visible on the TDNP statistics page (<http://texashistory.unt.edu/explore/collections/TDNP/stats/>).

## **Data Gathered about TDNP**

The data we gather to study the impact of the Texas Digital Newspaper Program is both quantitative and qualitative, and we are thus able to use it for multiple research and representation purposes. Numerical data inform us about the breadth of digital newspaper coverage over miles and years and about digital newspaper users over time, while subjective data, such as survey results, reference questions, and case study analysis, provide depth in understanding usage cases and impact on our audience. Quantitative data elucidate patterns over time and newspaper usage over time, and we can leverage this information to judge how we might draw more users to, and generate more data for, the Texas Digital Newspaper Program through analysis of trends. Qualitative information provides us with observational data that is credible, transferable, and dependable (Hughes and Hayoe 2008, p. 78). All data types combined inform how the Digital Newspaper team can move forward with future newspaper preservation projects.

UNT Libraries leverages this combination of data for multiple purposes, ranging from future newspaper digitization project planning to preparing grant applications to annual planning for student staffing. This combination of data also helps TDNP because it shows where gaps in the collection exist (quantitatively), sometimes even how those gaps affect people (qualitatively), and allows the TDNP team to focus on certain areas or collections over others to overcome inconsistencies illuminated by the collection data analysis. The ultimate goal for gathering this type of data is to benefit users of the Texas Digital Newspaper Program toward increasing the total number of newspapers that are digitally preserved and publicly available on The Portal to Texas History.

## Generating the Quantitative Dataset

To illustrate the makeup and characteristics of the Texas Digital Newspaper Program collection in The Portal to Texas History, we were able to generate a dataset utilizing a series of python scripts by targeting different metadata elements across all metadata records in the TDNP newspaper collection. Our findings below highlight some of the patterns in newspaper data acquisition, as well as characteristics of the geographic and temporal makeup of the collection.

The descriptive metadata for the Texas Digital Newspaper Program (<http://texashistory.unt.edu/explore/collections/TDNP/>) collection is available via the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) (<http://www.openarchives.org/pmh/>) in a number of metadata formats including the required OAI\_DC format ([http://texashistory.unt.edu/explore/collections/TDNP/oai/?verb=ListRecords&metadataPrefix=oai\\_dc](http://texashistory.unt.edu/explore/collections/TDNP/oai/?verb=ListRecords&metadataPrefix=oai_dc)), an RDF/XML format, and finally--the format used locally by the system--the UNTL metadata format. The OAI-PMH repository for the TDNP collection is an example of one of the Application Programming Interfaces (APIs) that UNT Libraries has made freely accessible for The Portal to Texas History. Full information on this and other APIs offered for this collection are available at the API page for the collection (<http://texashistory.unt.edu/explore/collections/TDNP/api/>). The metadata was gathered using an Open Source, Python-based metadata harvester called **pyoaiharvester** (<https://github.com/vphill/pyoaiharvester>) from the TDNP OAI Repository endpoint (<http://texashistory.unt.edu/explore/collections/TDNP/oai/>). This harvester allowed us to download all of the metadata records in the collection as of the writing of the paper; we

have, however, removed data from 2014 since this represents an incomplete year. At this time, all years combined total 165,298 metadata records in the collected dataset.

Once the metadata records were harvested, Mark Phillips prepared a Python script to parse the UNTL metadata records and to extract information relevant to the analysis of the collection for the purposes of this article. The full extraction code and information about obtaining the dataset are freely available via Phillips' GitHub account.

([https://github.com/vphill/untl\\_metadata\\_extraction/blob/master/tdnp\\_dataset.py](https://github.com/vphill/untl_metadata_extraction/blob/master/tdnp_dataset.py))

This extraction script uses the output of the **pyoaiharvester** tool as input, and it returns a tab-delimited file that displays one newspaper issue per row. The fields in the file are the following:

**Table 1: Fields in TDNP dataset.**

Field	Field Description	Example Data
ARK	ARK Identifier for issue	ark:/67531/metaph16320
Partner	Contributing Partner Code	BDPL
Year Online	Year issue went online (prefixed with "od:")	od:2006
Year	Year of newspaper issue	1934
Decade	Decade of newspaper issue	1930
County	County of newspaper issue	Palo Pinto County
Community	Community of newspaper issue	Mineral Wells
Title	Title of newspaper issue	The Tattler

All values in the Partner field can be resolved from the controlled vocabulary

(<http://digital2.library.unt.edu/vocabularies/institutions/>).

The resulting tab-separated value dataset had 165,298 rows of data with one row representing a newspaper issue. We were then able to utilize this information to aggregate and analyze the TDNP collection. We arranged this according to a variety of data points to track different characteristics of the TDNP collection, including issues added per year, titles added per year, number of counties and communities added per year, and number of new counties and communities added per year, along with different geographic heatmap representations of digitized newspapers across Texas.

## **Datasets and Analysis**

When the Texas Digital Newspaper Program first began in 2006, the digital newspaper staff was comprised of one dedicated full-time employee and the lab manager who oversaw all digitization activities for The Portal to Texas History. In 2007, University of North Texas Libraries was selected to serve as the lead institution in Texas for the National Digital Newspaper Program. In 2009, the NEH awarded NDNP-Oklahoma to the Oklahoma Historical Society in collaboration with UNT Libraries, which enabled UNT to fund one more full-time employee whose primary duties would be NDNP-Oklahoma newspaper digitization, and whose secondary duties would include assisting with the Texas Digital Newspaper Program. When it became evident in 2010 that TDNP was going to grow very large very quickly, the digital newspaper staff expanded to four full-time employees, with plans to hire a supervisor for all four. The 2010 acquisition of a Mekele Mach V microfilm scanner facilitated the development of an efficient in-house newspaper digitization workflow. This team was to coordinate all newspaper digitization activities. The rise in numbers after 2009 represents the addition of the newspapers scanned in-house, as well as PDF e-print editions, which we also began to preserve in 2010 via The Portal to Texas History. In 2013,



we added in-house physical newspaper scanning to our workflow, and we project an increase in content acquisition through this.

**Numbers over Time**

Our aim is to preserve any and all newspapers, from any media format (PDF, physical pages, or microfilm), from any year, or from any location in Texas. Certain trends are easy to explain based upon events happening with the Digital Libraries at UNT. For example, the jump in unique titles added in 2011 (see Figures 1 & 2 and Tables 2 & 3) indicates the consolidation of staff into a single, unified newspaper team on which the work of one employee, starting in 2011, is dedicated for approximately 30% of time to gather more newspaper content. The title decrease in 2013, analyzed in light of 2013 having been the year when the highest number of issues had been added in the program’s history, illustrates one of the newspaper team’s goals for 2013: to add complete, or close-to-complete, newspaper runs to TDNP.

**Table 2: Number of issues added per year (n=165,149)**

<b>Year</b>	<b># of Issues</b>	<b>Cummulative #</b>	<b>% of Issues</b>	<b>Cummulative %</b>
2006	54	54	0.03%	0.03%
2007	0	0	0%	0.03%
2008	44	98	0.03%	0.06%
2009	7,263	7,361	4.40%	4.46%
2010	44,788	52,149	27.10%	31.56%
2011	32,626	84,775	19.74%	51.30%
2012	30,836	115,611	18.65%	69.95%
2013	49,538	165,149	29.97%	99.92%

Figure 1: Numbers of newspaper issues added to the Portal each year.

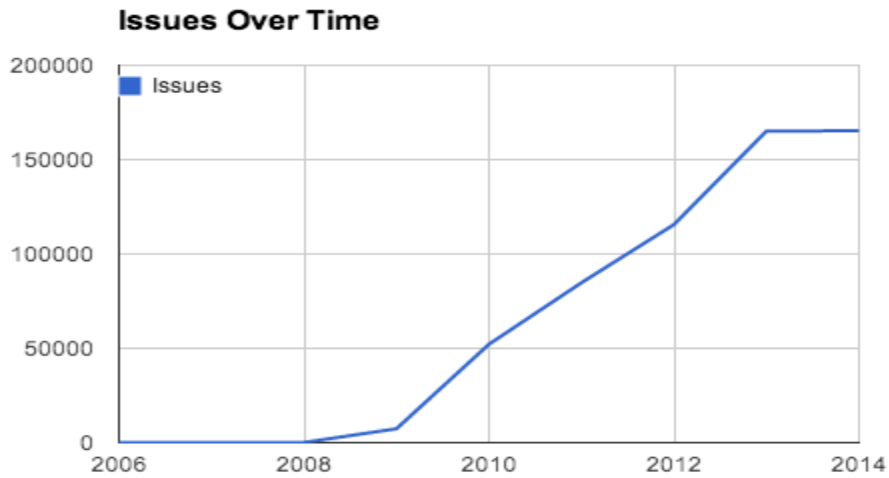
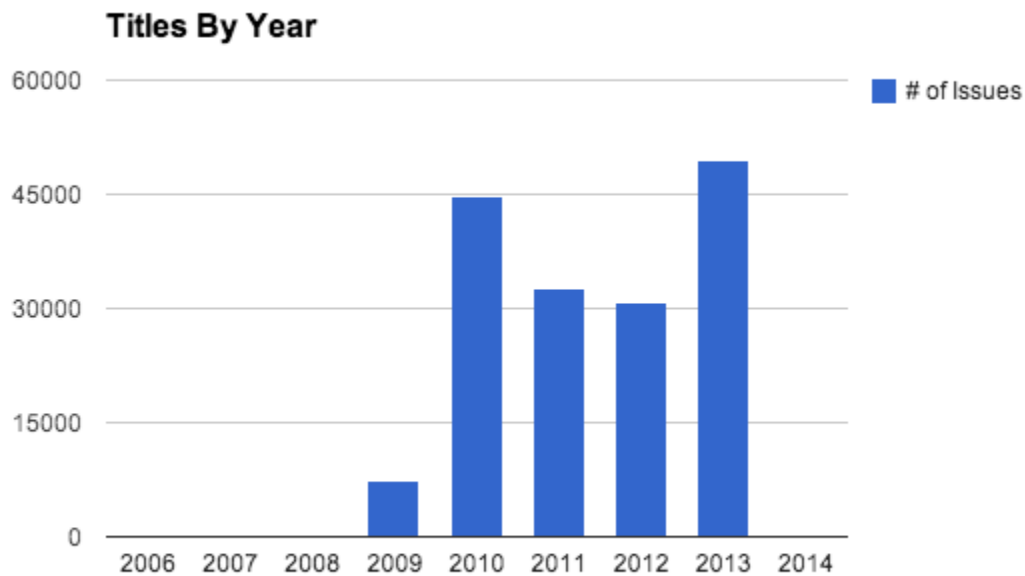


Table 3: Number of unique titles added per year (n=619)

Year	# of Titles Added	# of New Titles Added	Cumulative # of Titles
2006	2	2	2
2007	0	0	2
2008	2	2	4
2009	26	25	29
2010	121	114	143
2011	269	252	395
2012	163	145	540
2013	98	79	619

Figure 2: Number of titles added to the Portal each year.



### ***Numbers by Geographic Location***

There are 254 counties in the State of Texas. Though challenging, it is a goal of TDNP to have publications representing as many communities and counties as possible, and we are constantly working towards that goal. The dataset used for this paper only lists the first county or community for an issue of newspaper. Because much of the data collection depends on strong communication and extensive, state-wide travel, the number of counties added will generally remain consistent and proportional to the number of employee hours dedicated to communication and travel activities. In other words, the number of counties will consistently be proportional to what is humanly possible, in terms of the amount of employee time allocated for travel across Texas.

Using the place field in the UNTL metadata system to distinguish between counties and communities, we can report numbers of counties added to the Portal over the course of

each year (see Table 4). As explained earlier, this data also allows us to create acquisition goals for the short-term and the long-term along with goals for staff and student employment budgeting.

County tables and maps that represent annual additions like those below help us plan for areas of collection interest for the succeeding year. For example, in comparing the cumulative map with the 2013 map, we have determined that 2014 will be a year in which we strive to acquire more West and South Texas content, and we have thus initiated contact with prospective partner institutions about newspaper digitization. Finally, it is important to note that while we would love to add a newspaper from each of the 254 counties in Texas, some counties have never published newspapers. Therefore, we can only strive to digitize as many counties as possible, even though we may never reach 254 counties.

**Table #4: Number of counties added per year, as represented by newspapers published within a given county (n=109)**

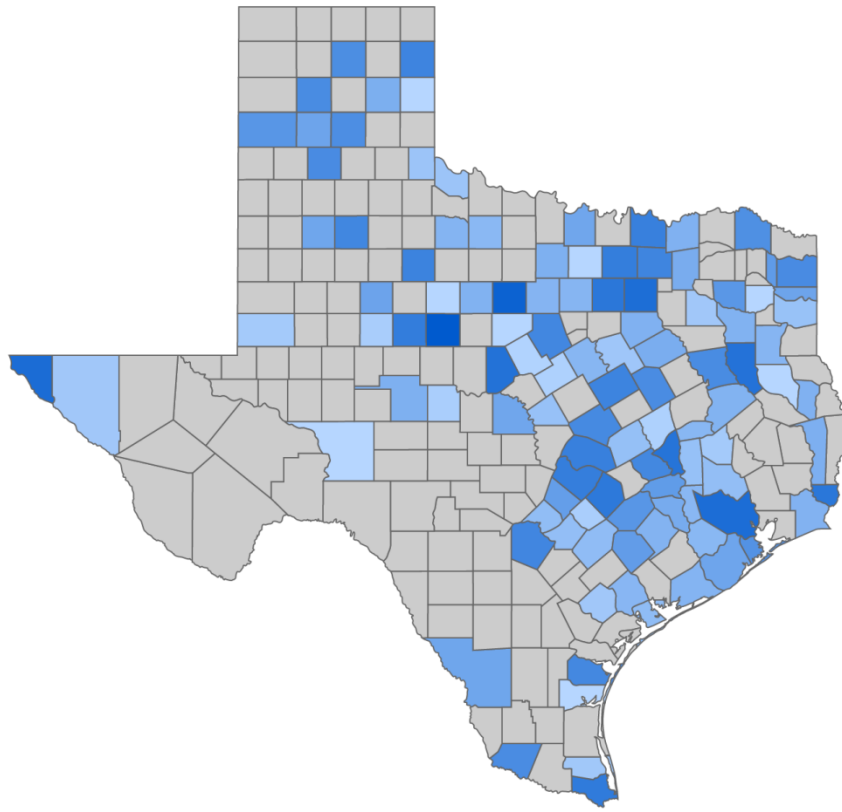
<b>Year</b>	<b>Number of Counties Added</b>	<b>Number of New Counties Added</b>	<b>Cumulative Number of Counties</b>
2006	2	2	2
2007	0	0	2
2008	2	2	4
2009	16	14	18
2010	32	20	38
2011	60	34	72
2012	57	23	95
2013	42	14	109

The geographical heatmaps demonstrate historical migration and movement patterns of the western frontier in Texas and the United States. Some of the first content we added to the Texas Digital Newspaper Program was from older newspaper titles that had been published in Texas for fifty to one hundred years; not coincidentally, these newspapers were published in counties that housed the earliest Texas settlements, including the now large cities of Austin, Dallas, San Antonio, Houston, and El Paso.

Finally, we can explain remarkably high additions of issues added during a year through simple institutional memory: in 2011, UNT Libraries added a large number of issues of the daily newspaper, *The El Paso Morning Times*, for the decade of 1910-1920. This was the only newspaper title in Texas to report on the Mexican Revolution in both Spanish and English, and we collaborated with University of Texas at El Paso to make this openly available to researchers. This influx is also apparent in the table documenting the numbers of issues by decade.

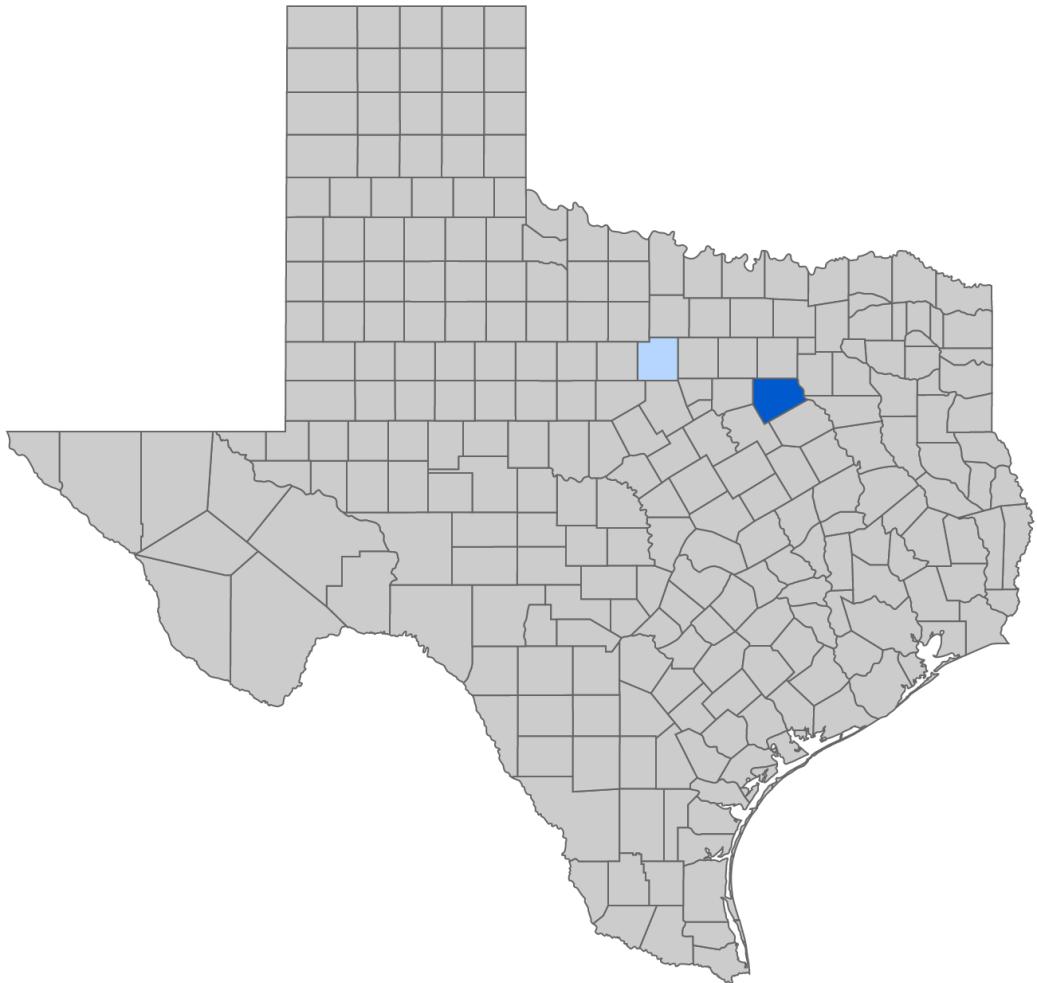
Heatmap visualizations of newspaper coverage in the Texas Digital Newspaper Program provide us with an easy way to target our communication and collecting strategies for each cycle of state library and private foundation grants. For example, in 2013, Portal staff made frequent visits to eastern and south-eastern Texas, which has resulted in an increase in coastal newspaper digitization projects representing communities like Port Aransas and Orange. We also observe additions of newspaper submissions from counties adjacent to grant-receiving counties from the previous year. Libraries in neighboring counties often report to us that they are motivated to submit newspapers after speaking to the previous awardee library about their newspaper digitization experience.

Figure 3: Counties currently represented in the TDNP Collection, including 2014.



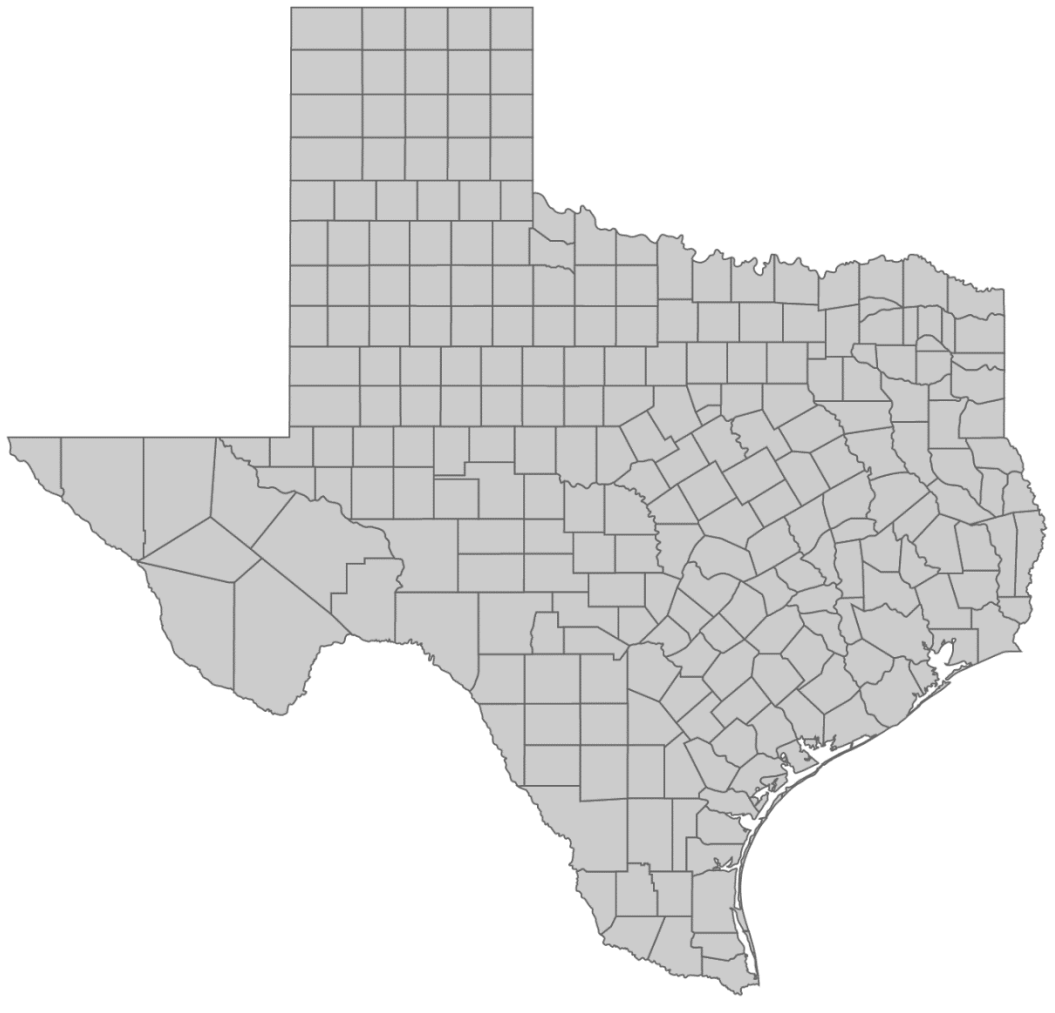
**Figure 4:**  
Counties added  
in 2006.

In 2006, we initiated our first newspaper digitization projects. We selected newspapers from communities close to the Denton area, from which we could easily gather pre-1923 content. These projects represented Ellis County and Palo Pinto County.



**Figure 5:**  
**Counties added**  
**in 2007.**

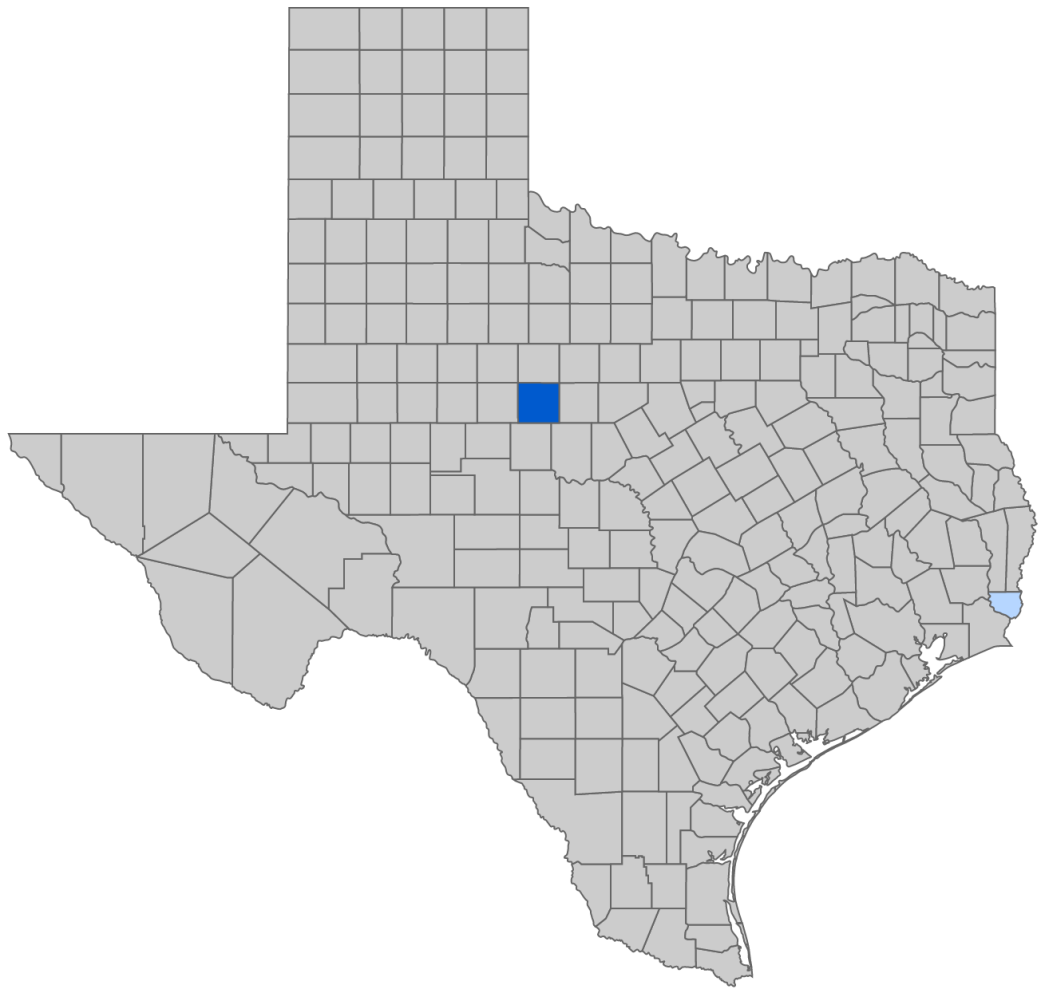
Because this was the year we began the National Digital Newspaper Program for Texas, we focused our energies on this grant project, which would lead to a large content ingest later.





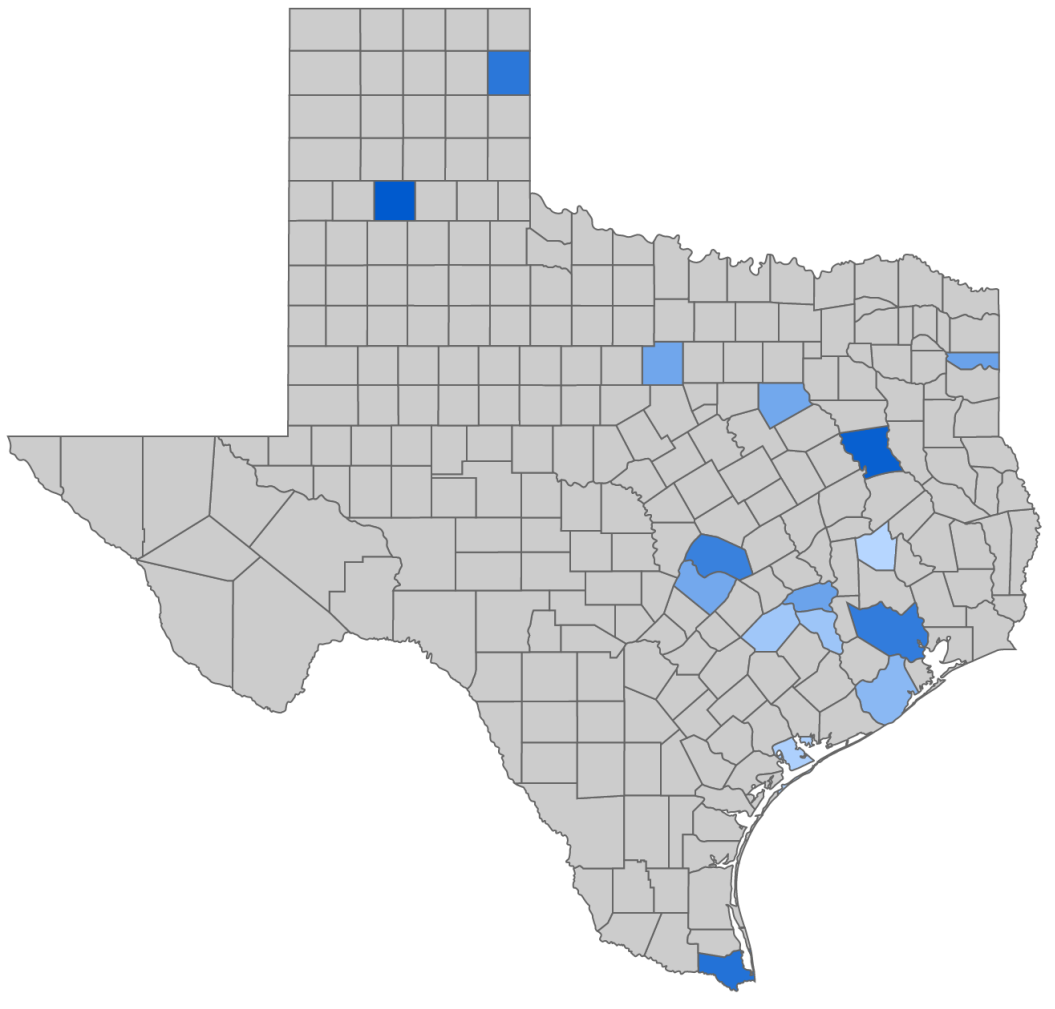
**Figure 6:**  
**Counties added**  
**in 2008.**

We continued pilot projects with two additional partner institutions, representing newspapers from coastal and central Texas, in 2008.



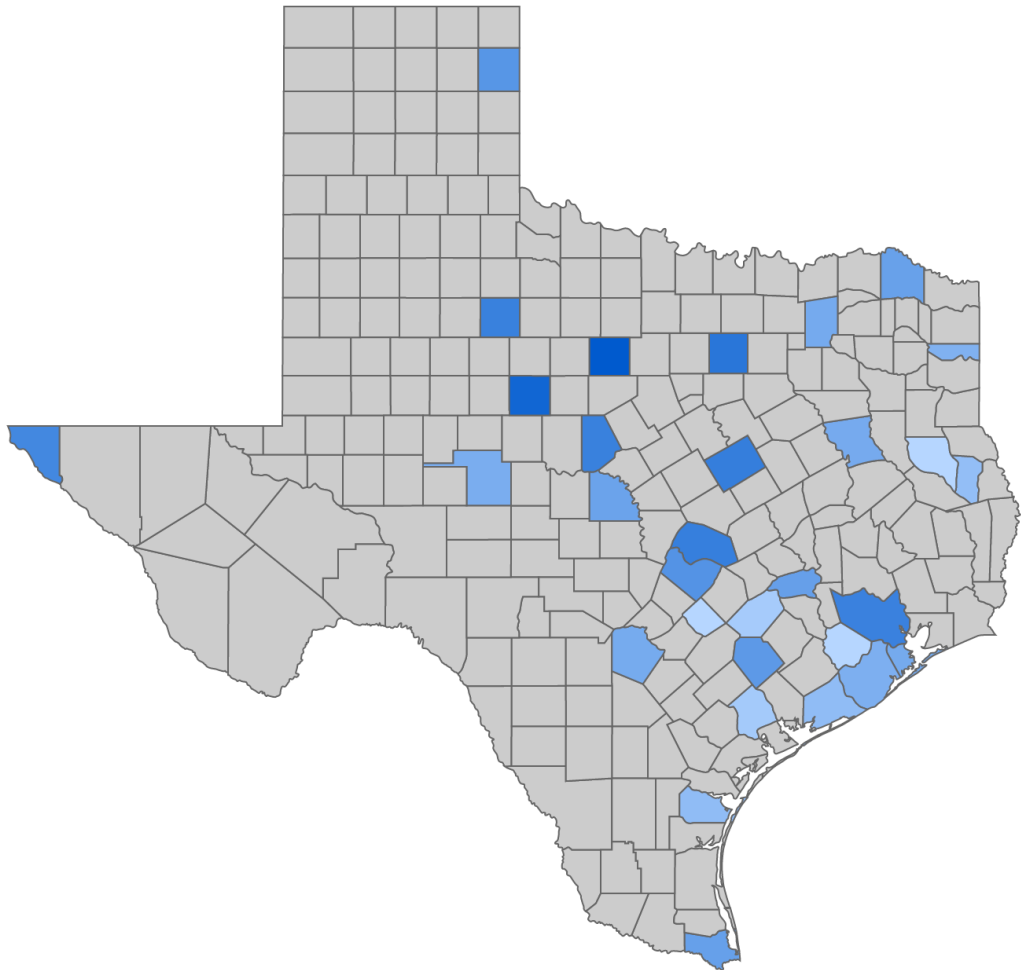
**Figure 7:**  
**Counties added**  
**in 2009.**

Two things happened to increase our county numbers in 2009: 1) We uploaded the first set of newspapers from the Early Texas Newspaper Project, through our partnership with the UT Dolph Brisco Center for American History; 2) We first partnered with the Tocker Foundation, a foundation who supports newspaper digitization for public libraries serving populations below 12,000.



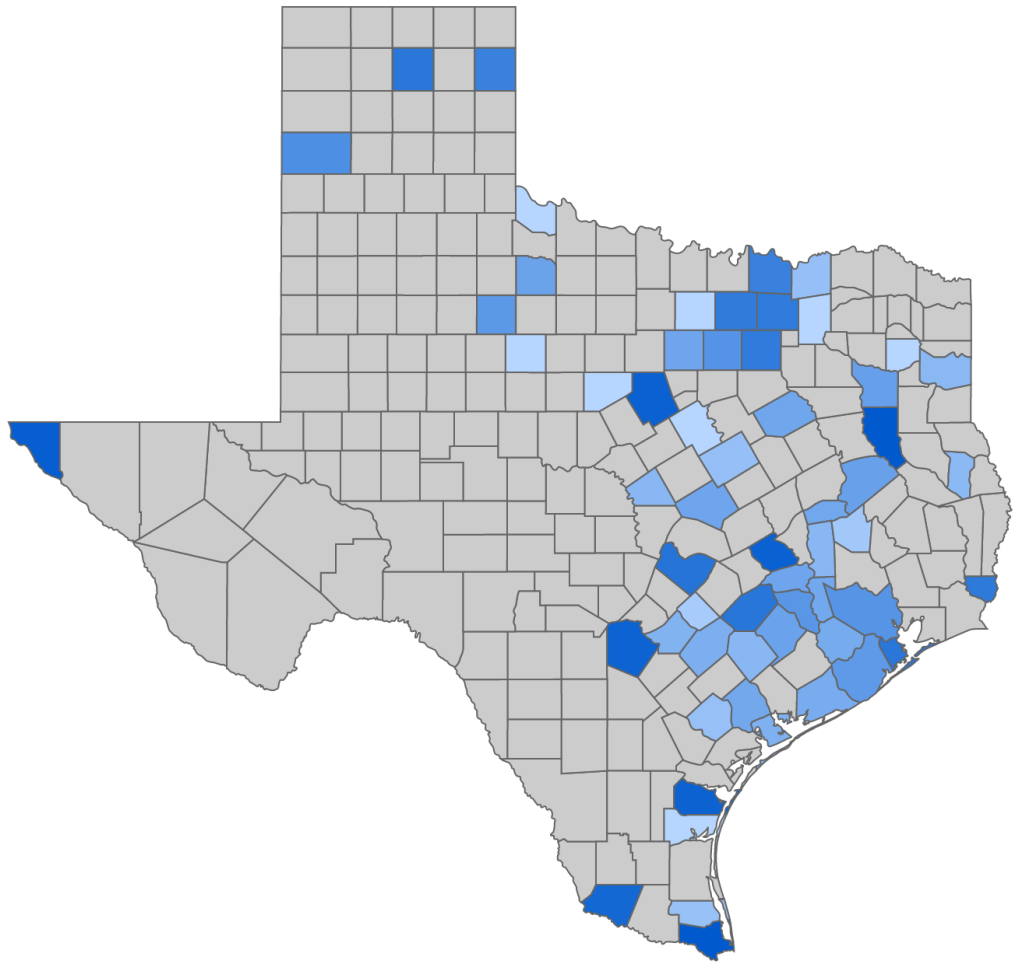
**Figure 8:**  
**Counties added**  
**in 2010.**

Gradually, more counties were added as we added more NDNP-Texas newspapers and partnered with more institutions. 2010 saw the addition of content from what is still our western-most partner, from El Paso, Texas.



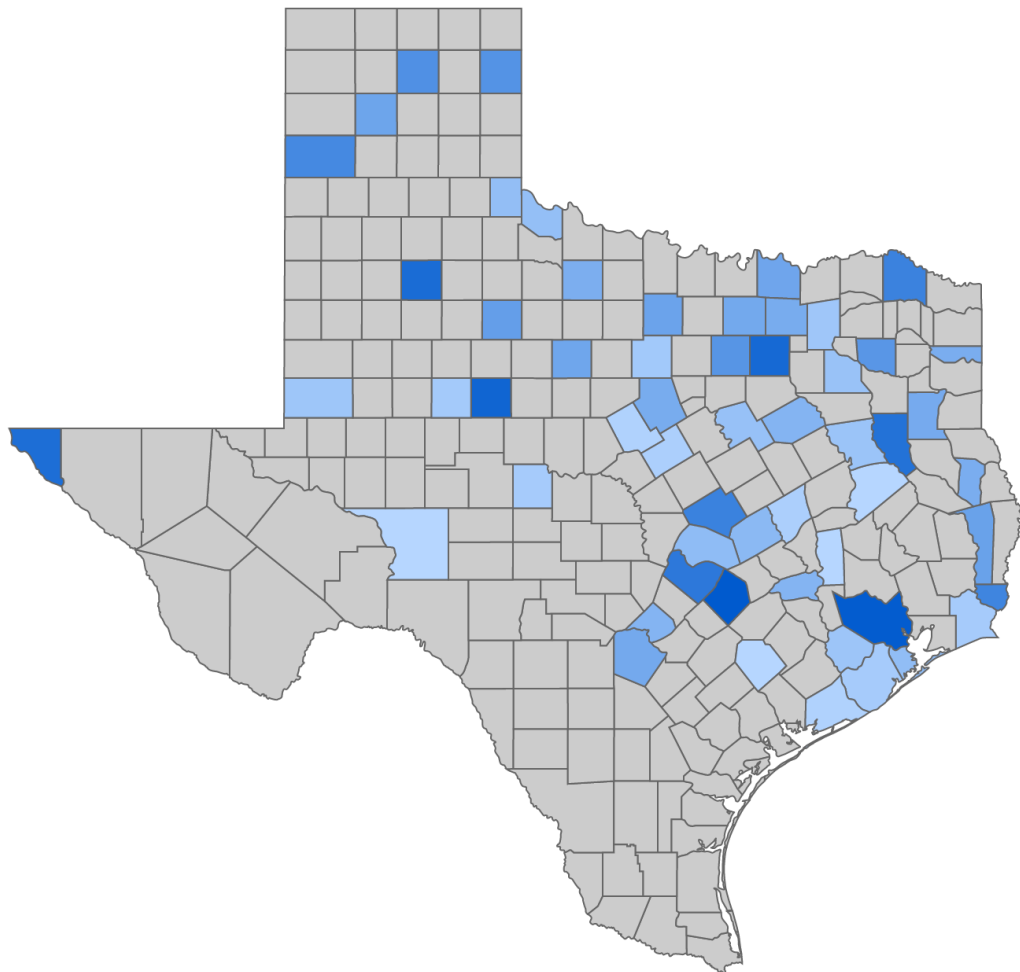
**Figure 9:**  
**Counties added**  
**in 2011.**

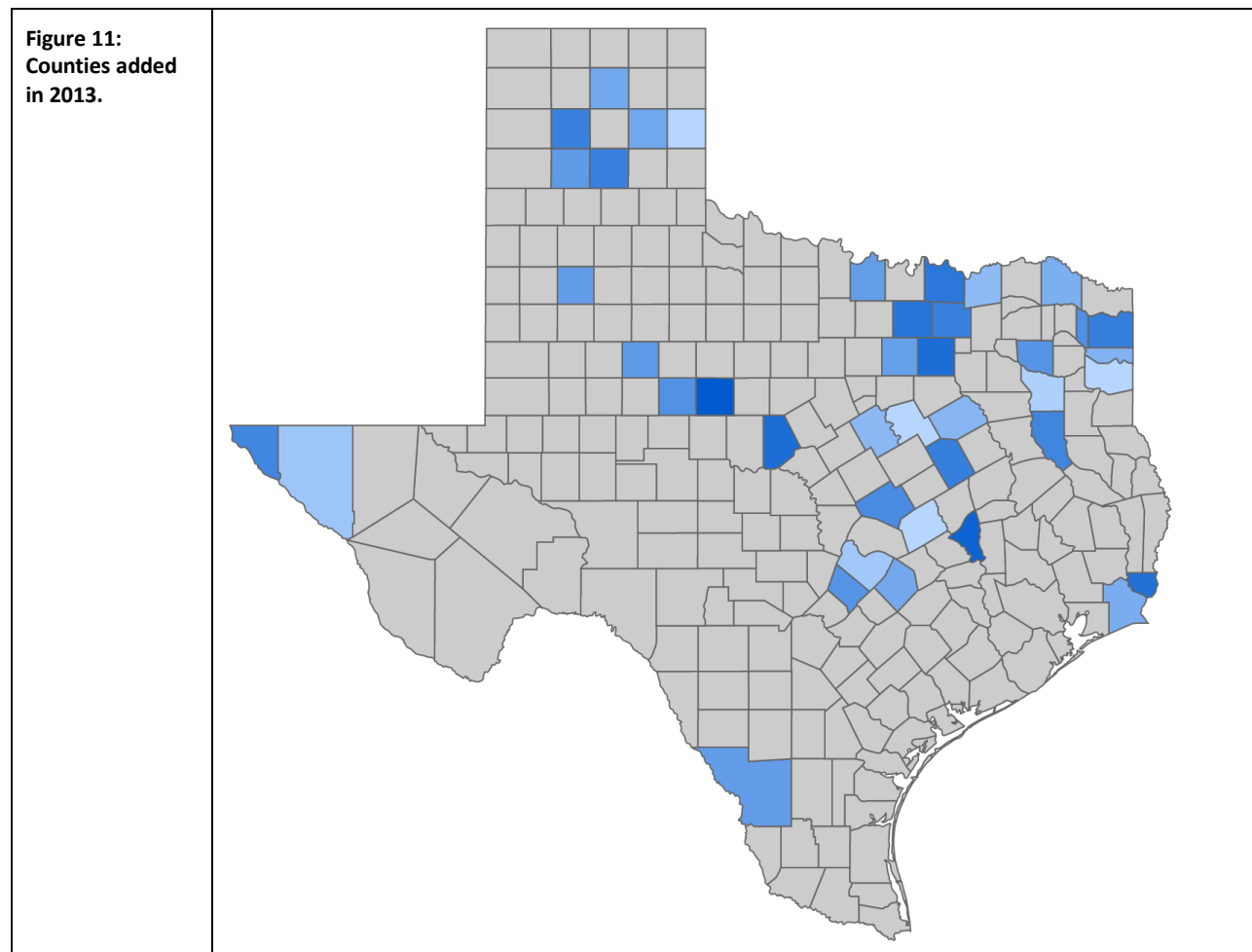
To support the increase in newspaper digitization projects, by 2011 the Digital Newspaper Team had grown into an individual unit, comprised of 5 full-time staff and 2 student assistants.



**Figure 10:**  
**Counties added**  
**in 2012.**

From 2011 and onward, we have continued to work with partner institutions to digitize an increasing amount of content, toward our long-term ultimate goal of representing all counties in Texas that have newspapers.





The UNTL metadata schema lets us identify newspapers from largest geographic region down to the smallest: the community level. As a result, we can document the numbers of communities that are represented by the digital newspapers, as depicted by their geographic coverage fields. However, many Texas newspapers represent entire counties, and we are careful to broaden the coverage=placeName entry to county so that we do not misrepresent the original coverage of the newspaper during its publication.

**Table 5 Number of communities added per year (n=142)**

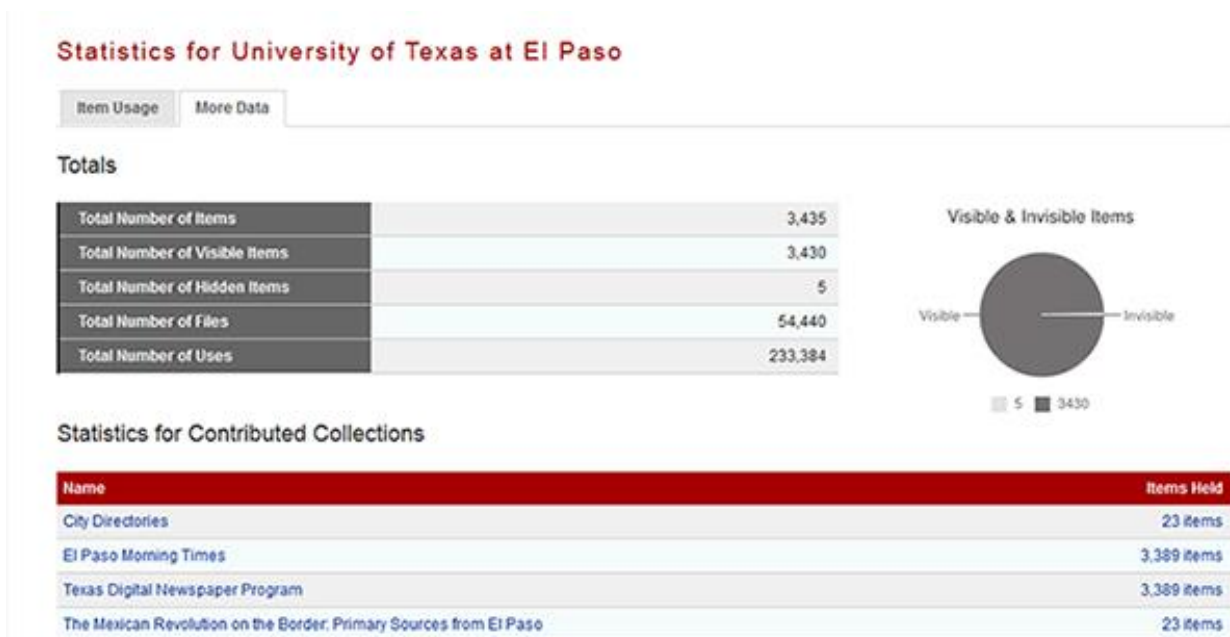
<b>Year</b>	<b># of Communities Added</b>	<b># of New Communities Added</b>	<b>Cumulative # of Communities</b>
2006	2	2	2
2007	0	0	2
2008	2	2	4
2009	17	16	20
2010	35	22	42
2011	62	42	84
2012	66	37	121
2013	47	21	142

## **Partners**

One of the founding goals of The Portal to Texas History is to unite content providers across Texas with UNT Libraries' capacity for digitization services, digital preservation infrastructure, and open data access. Groups who partner with UNT Libraries to digitize their newspapers frequently cite the lack of fees for continuing access or continued sustainability to their communities as serving as a deciding factor in their decision to digitize with UNT Libraries. One library director whose group had received a grant from the Tocker Foundation to digitize and preserve their newspaper in the Texas Digital Newspaper Program stated, "The newspaper digitization project places our library in position to reach out toward the future. By taking a piece of the past and bringing it with us, we are sure to grow and learn, appreciate and respect what was, what is and what will be." A large portion of partners who contribute materials to the Texas Digital Newspaper Program are small community public libraries who cannot not necessarily afford continuing sustainability fees.

All groups who contribute newspapers to the Texas Digital Newspaper Program are assigned two organizational units to accompany each of their newspaper issues: a “Partner” and a “Collection.” We define “Partners” as those groups who have organized funding for the digitization project. We define “Collection” as a contextual whole that represents newspaper dataset; “Collection” could represent newspapers from multiple towns within one county, or it could represent a series of newspapers and photographs that all come from the same city. Many partners, like the University of Texas at El Paso, also include non-newspaper content on The Portal to Texas History, and we can combine their data under the umbrella of their institutional biography while still separately measuring the number of different data types that exist in other collections under their institution.

Figure 12: Image displaying example Partner statistics and contributed collections.



For example, University of Texas at El Paso has placed 3,389 newspaper issues on the Portal, along with 46 other objects that exist in different Collections, including a City Directories collection and a collection containing primary sources from the Mexican



Revolution. All newspapers are also always included in the Texas Digital Newspaper Program collection.

Data about the numbers of new partners added each year directly feeds into content acquisition goals for the long-term and short-term. Because most newspaper digitization projects for the Portal are funded according to a grant cycle schedule, as previously mentioned, we use this information to seek out prospective new partners who might be interested in applying for grants to fund digitization of their newspaper content.

**Table 6: Number of partners added per year (n=48)**

<b>Year</b>	<b># of Partners Added</b>	<b># of New Partners Added</b>	<b>Cumulative # of Partners</b>
2006	2	2	2
2007	0	0	2
2008	2	2	4
2009	6	5	9
2010	10	6	15
2011	16	12	27
2012	20	10	37
2013	24	10	47
2014	3	1	48

Often, we see that partners from different types of institutions have distinctive patterns in newspaper content holdings. For example, genealogical/historical societies and museums will have a variety of regional newspaper titles that range in date from very early to mid-twentieth century, but they rarely have complete newspaper runs. Academic libraries very often have a combination of regional newspapers as well as a complete run of their

student newspaper. Public libraries frequently either own an entire run of their community's newspaper, or have direct access to the newspaper publishing offices.

**Table 7: Number of TDNP partners by type of represented institutions (n=48)**

<b>Partner Type</b>	<b># of Partners by Type</b>	<b># of Issues</b>
Public Libraries	27	66,320
Academic Libraries & Archives*	13	41,308
Genealogical/Historical Societies	4	5,065
Museums	2	2,544

\* "UNT Archives" and "UNT Libraries" are two partner institutions, thus actually totaling to 48 partners. This table indicates content whose digitization was funded by external partners. Content funded internally by UNT totals to 49,805 and has been removed from this table.

Data about numbers of issues contributed by each partner gives the TDNP team a good starting point for encouraging partners to digitize more newspaper content and to apply for newspaper digitization grants from local foundations and state agencies. This data is easily viewable from the Texas Digital Newspaper Program collection statistics page (<http://texashistory.unt.edu/explore/collections/TDNP/stats/>). Current partners report that they use their collection and institution statistics pages to determine how much more of their newspaper content they might digitize, while prospective partners utilize these pages to see which groups in their region have already contributed newspaper content to TDNP and as a deciding factor in whether they also want to contribute their newspapers. For example, multiple institutions in the panhandle region of Texas contributed one after another, after communicating with each other about the process and after using the Newspaper Program statistics pages.

**Table 8: All TDNP partners and numbers of issues each has contributed (Continued on next page.)**

<b>Partner</b>	<b># of Issues</b>
Abilene Christian University Library	3,461
Abilene Public Library	9,043
Atlanta Public Library	2,777
Bartlett Activities Center and the Historical Society of Bartlett	3,355
Bastrop Public Library	4,824
Boyce Ditto Public Library	1
Breckenridge Public Library	11,914
Casey Memorial Library	1,893
Collin County Genealogical Society	1,568
Crosby County Public Library	2,514
Deaf Smith County Library	1,103
Denton Public Library	90
Ferris Public Library	155
Gibbs Memorial Library	1,888
Grayson County Frontier Village	2,543
Hardin-Simmons University Library	2,553
Harrie P. Woodson Memorial Library	2,381
Hemphill County Library	2,817
Heritage House Museum	1
Honey Grove Preservation League	141
Howard Payne University Library	6,474
Hutchinson County Library, Borger Branch	1,762
Lamar State College – Orange	5,196

McMurry University Library	1,858
Meridian Public Library	101
Mineola Memorial Library	1,005
Mineral Wells Heritage Association	1
Montgomery County Memorial Library	49
Rice University Woodson Research Center	2,710
Richard S. and Leah Morris Memorial Library	1,702
Rio Grande City Public Library	1,888
Sachse Public Library	408
Schulenburg Public Library	1,095
Singleary Memorial Library	4,464
Smith Public Library	1,529
South Texas College of Law	277
Stella Hill Memorial Library	1,984
Stonewall County Library	2,982
Sweetwater/Nolan County City-County Library	3,986
Swisher County Library	1,965
Tarleton State University	2,431
Texas Wesleyan University	703
The Dolph Briscoe Center for American History	13,488
UNT Archives	391
UNT Libraries	49,414
University of Dallas	245
University of Texas at El Paso	3,389
University of Texas at San Antonio Libraries Special Collections	381

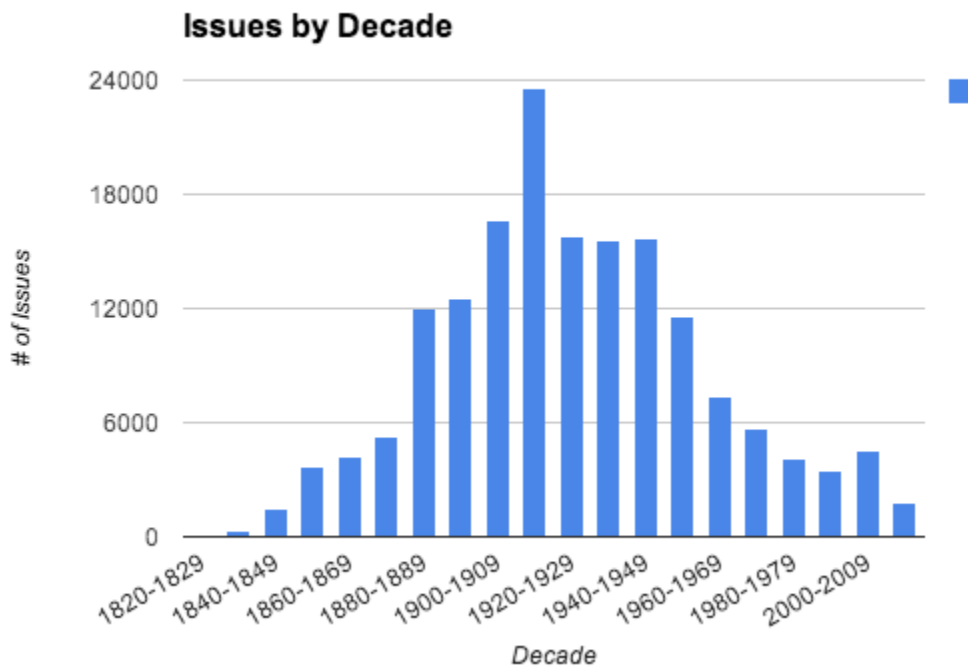
As the Texas Digital Newspaper Program was first beginning, one of its greatest external supporters was the University of Texas Dolph Briscoe Center for American History. UNT Libraries partnered with this group initially on the National Digital Newspaper Program grant, and then continued the partnership through a project called *The Early Texas Newspapers Project* (<http://texashistory.unt.edu/explore/collections/ETXNP/>), funded by a Texas State Library and Archives Commission grant program, TexTreasures. This project was awarded and then renewed over the course four cycles, and this partnership digitized nearly sixty-thousand pages of Texas newspapers, representing early Colonial Texas up to 1923. The titles contributed by UT’s Dolph Briscoe Center for American History represent most of the earliest Texas newspapers available on The Portal to Texas History.

**Table 9: Number of issues in each representative decade (Continued on next page.)**

<b>Decade</b>	<b># of Issues</b>
1820-1829	2
1830-1839	302
1840-1849	1,439
1850-1859	3,665
1860-1869	4,185
1870-1879	5,215
1880-1889	11,963
1890-1899	12,546
1900-1909	16,668
1910-1919	23,619
1920-1929	15,806
1930-1939	15,588

1940-1949	15,691
1950-1959	11,567
1960-1969	7,408
1970-1979	5,679
1980-1979	4,148
1990-1999	3,521
2000-2009	4,492
2010-2019	1,794

Figure 13: Graph of number of issues in each representative decade.



## Conclusion

The goal of The Portal to Texas History is to unify primary source objects in one digital preservation and access interface, and the Texas Digital Newspaper Program comprises a large portion of this interface. The APIs associated with the TDNP collection have given us an excellent opportunity to arrange the data in a meaningful way that we can re-appropriate for multiple research purposes. All data we gather, and have gathered for the purposes of this paper, can be single-sourced, recoded, and repurposed for different analysis projects, thus we have made this dataset freely available

(<http://digital.library.unt.edu/ark:/67531/metadc275016/>). This supports open access and University of North Texas' open access initiative (<https://openaccess.unt.edu/resources>).

UNT Libraries uses this type of data to support further Texas newspaper preservation through digitization and to assist groups who might not have the technological infrastructure or the background in digital preservation to digitally preserve their newspapers in an affordable way, thus combining expensive technical infrastructure with priceless newspaper content.

## References

Belden, D. & Murray, K. (2012). The Impact of digitizing resources: Content partner and user perspectives. *Digital Frontiers 2012 Conference*. Denton, TX: UNT Digital Library.

Carlisle, T. (2012). Resources for educators. *The Portal to Texas History*. Retrieved from <http://education.texashistory.unt.edu/lessons/newspapers/index.html>

Hughes, M. & Hayhoe, G. (2008). *A Research primer for technical communication: Methods, exemplars, and analyses*. New York, NY: Routledge.

Open Archives Initiative. (2008). Protocol for Metadata Harvesting (Version 2.0)  
Available from <http://www.openarchives.org/pmh/>

Phillips, M. (2014). UNTL Metadata Extraction Script (Version 1.0) [Software]. Available from  
[https://github.com/vphill/untl\\_metadata\\_extraction/blob/master/tdnp\\_dataset.py](https://github.com/vphill/untl_metadata_extraction/blob/master/tdnp_dataset.py)

University of North Texas Libraries. (2014). *The Portal to Texas History*. Retrieved from <http://texashistory.unt.edu>.