

# NATIONAL ENDOWMENT FOR THE HUMANITIES



SAMPLE APPLICATION NARRATIVE

## National Digital Newspaper Project

Applicant Institution: Pennsylvania State University Libraries

This sample narrative was prepared using a previous set of application guidelines. Please refer to the guidelines for the 2008 competition when preparing your application. The guidelines are at: <http://www.neh.gov/grants/guidelines/ndnp.html>

# Our Common Wealth: Digital Access to Historic Pennsylvania Newspapers, 1880 – 1922

## Description of the Project

On behalf of the Commonwealth of Pennsylvania, the Pennsylvania State University Libraries requests funding in the amount of \$393,650 to support its participation in the National Endowment for the Humanities' National Digital Newspaper Program coordinated by the Library of Congress. In addition to the amount requested from NEH, the Pennsylvania State University Libraries will provide \$150,146 in cost-share support for the project.

The Pennsylvania State University Libraries has a strong commitment to both preserving and providing on-line access to Pennsylvania newspapers. The Libraries was a regional site of the NEH-funded Pennsylvania Newspaper Project (PaNP) from 1983 to 1990 and is currently serving as the lead institution for the resurrected Pennsylvania Newspaper Project (Phases II and III), which began in 2004 and will conclude in March 2008. In recent years, the Penn State Libraries has endeavored to provide on-line access to numerous historically significant newspapers held in its own collection, including its student newspaper, the *Daily Collegian*, two associated campus papers, and the backfile of *Lancaster Farming*, a national farming weekly. In 2004, the Libraries launched a collaborative venture with cultural institutions across the state to select previously microfilmed titles and create an on-line Pennsylvania Civil War Newspaper archive.

With past experience in completing federally-funded projects and creating and launching large-scale historical newspaper digitization projects, the Penn State University Libraries faculty and staff fully understand the high level of commitment needed to achieve success. To ensure project success, it is essential to possess a knowledgeable and highly skilled staff versed in both preservation microfilm production and in building digital collections, as well as microfilm-to-digital conversion. Furthermore, the staff must have strong management skills with an acute attention to detail. Other key factors to success include having the appropriate technological infrastructure in place to handle the massive electronic storage needs of the project, assembling a core of scholars and cultural heritage leaders to participate in the difficult selection decisions, partnering with experienced and reliable service providers who can meet the required project specifications, and a committed spirit to build strong working relationships with the staff from the Library of Congress to fulfill the requirements of the National Digital Newspaper Program (NDNP).

All deliverables for the project will comply with exact specifications detailed in NDNP *Technical Guidelines for Applicants*. Included with the deliverables will be a listing of the newspapers selected for the project, including information about the title, the geographic and chronological coverage, and its LCCN. A thorough technical analysis of the available microfilm for optimal scanning of the selected titles and a production schedule with agreed-upon timetable for sample and final file delivery to the Library of Congress will be delivered as well. Within six months of the beginning of the project, a sample of the digital assets will be sent to the Library of Congress for their review. The digital assets will include TIFF 6.0 uncompressed, 8-bit grayscale images, OCR text files with bounding-box coordinates, PDFs with hidden text, JPEG200 images, and all metadata (as described in the NDNP *Technical Guidelines*) delivered in a METS object structure according to the prescribed XML batch template structure. By no later than June 2009, the digital assets for twenty-five percent of the total number of pages to be scanned during the life of the project will be delivered to the Library of Congress. The

remaining digital assets will be delivered monthly to the Library of Congress, in conformance with the program's technical guidelines, by the end of the award period. In addition, a summary description will be written on the scope, content, and significance of each digitized newspaper title (a maximum of 500 words) and CONSER records will be updated for each title for inclusion in the LC directory. Finally, a second-generation (2N) duplicate silver negative microfilm (made from the camera master) of the scanned title will be deposited at the Library of Congress.

The Pennsylvania State University Libraries main campus, located at University Park, Pennsylvania, will serve as the headquarters for the Pennsylvania Digital Newspaper Project (PaDNP), *Our Common Wealth: Digital Access to Historic Pennsylvania Newspapers 1880-1922*. Key staff from the Libraries' Digitization and Preservation Department, the Cataloging and Metadata Services Department, along with hired project staff will comprise the PaDNP team. The Digital Libraries Technology unit of the University's Information Technology Services will provide the needed technological support for the Project. The PaDNP Advisory Board, consisting of scholars, historians, and librarians from across the state, has been formed and title selection discussions have already commenced. Newspaper titles with camera master negative microfilm produced by the NEH-funded PaNP phases I, II, and III that are housed at the Pennsylvania State Archives vaults for the State Library of Pennsylvania and master negative holdings of the Free Library of Philadelphia and other cultural repositories will be the source of newspaper negatives from which selections will be made for this project.

While the local loading of digital assets created by this Project are considered outside the scope of this grant application and not reflected in this document, statewide discussion have begun to consider options for future implementation.

The Commonwealth of Pennsylvania looks forward to participating in the NEH/LC National Digital Newspaper Program by contributing significant, historic Pennsylvania newspaper titles to the *Chronicling America: Historic American Newspapers* database.

# Pennsylvania Digital Newspaper Program

## Table of Contents

<b>Description of the Project</b> .....	i
<b>Table of Contents</b> .....	iii
<b>Narrative</b> .....	1
History and Scope of the Project .....	1
Newspaper Publishing in Pennsylvania .....	1
From Colonial Beginnings to the Civil War .....	1
Era of Industrialization and Urbanization .....	2
News Worthy Events: 1880 -1899 .....	3
Progressive Era .....	5
News Worthy Events: 1900 – 1922.....	5
Modern Newspaper History .....	7
Pennsylvania’s Newspaper Repositories on Microfilm and in Paper .....	8
Preservation Microfilming Efforts in the Commonwealth .....	8
Pennsylvania Newspaper Project, Phase I – III .....	8
Other Microfilming Efforts in the Commonwealth .....	9
Newspaper Digitization Efforts in the Commonwealth .....	10
Methodology and Standards .....	12
Access to Master Negative Microfilm, 1880 -1922 .....	12
Title Selection Plan .....	12
Intellectual Content Criteria: Part I .....	13
Refined Selection Criteria .....	13
Identifying Candidates in Urban Areas .....	14
Title Ranking .....	14
Microfilm Evaluation: Part II .....	15
Image Quality Selection Criteria .....	16
Final Title Selection .....	16
Partnerships and Existing Infrastructure .....	17
CONSER Cataloging .....	17
Vendor Specifications and Contracting for Digitization .....	17
Work Plan .....	18
Project Ramp Up: July 2008 – October 2008 .....	18
Selection of Titles: August 2008 – December 2008 .....	18
Evaluation of Microfilm and Metadata Collection: September 2008 – June 2009 .....	19
Scanning and Master Image Formatting by Vendor: October 2008 – April 2010 .....	19
Image Processing and Content Conversion by Vendor: October 2008 – April 2010 .....	19
In-house Quality Control: October 2008 – April 2010 .....	20
Delivery to Library of Congress: June 2009 – April 2010 .....	20
Update MARC Record from the CONSER Database: June 2009 – April 2010 .....	21

Wrap Up: May – June 2010 .....	21
Production Schedule for Project Milestones and Deliverables .....	22
Project Staff .....	23
Conclusion .....	24
References .....	25

## **Budget Form**

### **Appendices**

- A. Master Negative Vault Holdings List
- B. Newspaper Digitization Efforts in the Commonwealth
  - B.1 Penn State Libraries Digitized Collections
  - B.2 Access Pennsylvania Digital Repository
  - B.3 Lancaster County Digitization Project
  - B.4 Pittsburgh Jewish Newspaper Project
  - B.5 Hamlin Memorial Library of Smethport
- C. Letters of Commitment
  - State Library of Pennsylvania
  - Free Library of Philadelphia
- D. Advisory Board
- E. Request for Information
- F. Resumes

## **History of Grants**

### **Letters of Support**

State Library of Pennsylvania  
 Free Library of Philadelphia  
 Carnegie Library of Pittsburgh  
 Pennsylvania Historical & Museum Commission  
 Pennsylvania Newspaper Association  
 The University of Pittsburgh, University Library System  
 Temple University Libraries  
 Susquehanna University, Blough-Weis Library  
 David Hoffman, historian

## Narrative

On behalf of the Commonwealth of Pennsylvania, the Pennsylvania State University Libraries will digitize from previously produced microfilm and make publicly accessible via the Library of Congress' *Chronicling America: Historic American Newspapers* database <http://www.loc.gov/chroniclingamerica/index.html> 100,000 newspaper pages from Pennsylvania. The selected titles will represent a balanced geographic cross-section of the state in English and will cover the time period 1880 - 1922.

## History and Scope of the Project

### Newspaper Publishing in Pennsylvania

#### From colonial beginnings to the Civil War

Newspaper publishing began in the pressrooms of two energetic Philadelphia editors, Andrew Bradford and Benjamin Franklin. In 1719, Bradford issued his *American Weekly Mercury*, the first continuous newspaper in the Middle Colonies and the third in all the colonies. Ten years later, Franklin purchased the nascent *Pennsylvania Gazette* and shaped it into one of the nation's truly engaging sources of information, advice, and criticism in the colonies. The lure of editorship eventually captured other imaginations: Bradford's nephew William began the *Pennsylvania Journal* in 1742; William Goddard's *Pennsylvania Chronicle* was founded in 1767; and on the eve of the Revolution, John Dunlap pulled the first issue of his *Pennsylvania Packet and Daily Advertiser*, the nation's first daily and direct ancestor of today's *Philadelphia Inquirer*.

Between 1740 and 1776 the number of Philadelphia printing houses tripled, and the output served not only Pennsylvania's English-speaking readership, but its German population as well. The German press in America began in 1739 with Christopher Sauer's newspaper, *Der Hoch-Deutsche Pennsylvanische Geschicht-Schreiber*, which under the later name of *Die Germantauer Zeitung*, developed into one of the most influential pre-Revolutionary weeklies in America. The first newspaper in Lancaster, the fortnightly *Gazette*, displayed alternate columns in English and German. William Miller's *Philadelphische Staatsbote*, founded in 1762, enjoyed statewide popularity throughout the German-speaking community. Modern cultural historians note that Pennsylvania was so much a center for the German press and influence in colonial times, that German very nearly became the official language of the province. According to the census of 1790, 110,357 of Pennsylvania's 423,373 residents were of German extraction. Pennsylvania was home to 70.5% of the new nation's German residents.

By 1776, Philadelphia had no fewer than seven newspapers, with editions ranging from 500 to 3,000 copies and circulations as far away as the West Indies. John Scull and Joseph Hall launched their *Pittsburgh Gazette* in 1786, the first newspaper west of the Appalachian Mountains. It was followed by other western papers: the *Uniontown Fayette Gazette* in 1798, the *Washington Western Telegraph* in 1795, the *Bedford Gazette* in 1805, and Meadville's *Crawford Weekly Messenger* in 1805. A new journalistic phenomenon, the penny newspaper, arose during this period, facilitating the spread of information to a far wider audience than the expensive papers had reached. By 1840, Pennsylvania led all other states in the number of newspapers published. Philadelphia alone had 52 weekly and 15 daily papers published during the 1850s. The total daily product of Pennsylvania's newspaper press had climbed past 200 titles by 1860.

## Era of industrialization and urbanization

The news industry expanded greatly during the Civil War years and continued to grow until the close of the century. For instance, by 1880, Pittsburgh had ten daily newspapers, including the new *Evening Telegraph*, the *Post*, and the *Chronicle Telegraph*. Harrisburg boasted three dailies, the *Telegraph*, the *Patriot*, and the *Harrisburg Independent*. The *Scranton Morning Republican* of 1867 was the first postwar daily outside Philadelphia. The *Philadelphia North American*, established in 1839, became one of the truly important progressive newspapers of the eastern United States, and the *Evening Bulletin*, which began as *Cumming's Evening Bulletin* in 1847, evolved into one of the largest evening newspapers in the country. The famous *Williamsport Grit*, which started in 1882 as a rustic week-end filled with ethnic "grit" and humor, stood as the nation's most widely distributed country and small-town weekly and produced both state and national editions.

Prior to the start of World War I, the African American population in Pennsylvania grew slowly but, migration from the South intensified as World War I curtailed European immigration (PHMC 2007). The *Pittsburgh Courier* founded in 1907, the *Christian Recorder*, and the *Philadelphia Tribune* at 120 years old, are all among the oldest and most influential African-American newspapers in the country. The *Pittsburgh Courier* was the first African-American newspaper to publish both national and local editions; and at one time, had 14 editions in circulation.

In the 19th century, a new American culture was developing from a mix of new immigrants from a variety of ethnic and religious groups. The newspapers of Pennsylvania document the emerging nationalism, the spirit of discovery, and the concept of manifest destiny. While the Germans are the most prominent European ethnic group to leave a record of their settlement through the medium of newspapers in Pennsylvania during the late 18th and early 19th centuries, the new immigrants from Italy, Lithuania, Slovakia, Ukraine, and Sweden have a newspaper publishing legacy from the late 19th to the early 20th centuries. With two million Poles immigrating to the United States between 1850 and 1924, they were one of the largest European ethnic groups to settle in Pennsylvania. No other state had as many foreign-language newspapers: at least 60 were being published at the end of the century.

The newcomers found work in the expanding northeastern anthracite mining area, the western Pennsylvania bituminous coal region, and in Pittsburgh with its heavy industry for iron and steel; all of which attracted unskilled male immigrants, particularly those from southern and eastern Europe. By 1920, Pittsburgh had about 200,000 Poles and Philadelphia had at least 50,000. Another 75,000 had settled in the anthracite cities of Scranton, Wilkes-Barre, and Hazleton, as well as the smaller towns of Shenandoah, Mount Carmel, and Shamokin. The cities of Reading, Conshohocken, and Erie also had prominent Polish populations. Linking it all together was the Pennsylvania Railroad. With 30,000 employees and \$400 million in capital, it was the world's largest corporation in 1880.

During the second half of the nineteenth century, the population of Pennsylvania increased from 2.3 million in 1850 to 6.3 million in 1900 (Licht 2002). Consequently, between 1880 and 1920, 11 of the 100 largest cities in the United States were located in Pennsylvania (Gibson 1998). The population of Pittsburgh, for example, grew seven-fold (Licht 2002). This growth was reflected in increased newspaper circulation. From 1881 to 1921, the Republican oriented daily newspapers grew from a combined total circulation of 39,000 to 237,403. The *Post*, which claimed to be the only Democratic newspaper in western Pennsylvania, saw its circulation increase from 8,500 to 67,696 during this 41 year period. The *Press*, which was established in 1884 as an Independent newspaper, bragged that it was the paper of the masses with a circulation of 118,372 in 1921.

**Daily Circulation of Major, English Newspapers in Pittsburgh 1881-1921**

Title	Political orientation	Circulation			
		1881	1891	1911	1921
Dispatch	Republican/ Independent Republican	12,000	31,828	65,323	55,416
Commercial Gazette	Republican	14,000	25,026		
Post	Democratic	8,500	18,075	55,009	67,696
Press	Independent		38,790	100,804	118,372
Telegraph	Republican	9,000			
Sun	Democratic/ Independent			64,876	75,971
Chronicle Telegraph	Independent/ Republican		22,000	77,286	93,180
Leader	Independent	13,217	23,506	75,000	81,202
Gazette Times	Republican			81,415	88,807
Times	Republican	4,000	35,045		
Evening Chronicle	Independent	7,200			
<b>Total daily newspaper circulation</b>		<b>67,917</b>	<b>194,270</b>	<b>519,713</b>	<b>580,644</b>
		Population U.S. Census			
		1881	1891	1911	1921
		156,381	238,617	533,905	588,193

Source: N. W. Ayer & Son (1881, 1891, 1911, 1921). American Newspaper Annual. Philadelphia, PA

Industrialization and urbanization were the overarching themes in Pennsylvania newspapers from 1880 to 1922. Spanning both the height of the Gilded Age (1881-1897) and the Progressive Era (1897-1921), this timeframe included industrial capitalist development, the rise of the corporation, the boom-and-bust cycles of the economy, tensions between capital and labor, ethnic and racial hostilities, the growth of cities, and urban machine politics and corruption (Licht 2002).

Conflict between labor and the captains of industry marked the decades between 1880 and 1922. Miners and steel workers faced physical danger, economic uncertainty, and exploitative working conditions. In company towns, miners and their families were in constant debt to the company store and in danger of eviction from their company-owned homes. Many children labored in the anthracite industry as breaker boys. Between 1902 and 1920, mine accident deaths occurred on an average of 525 per year. As a consequence, Pennsylvanians played an important role in the development of the labor movement, and the Commonwealth was the site of some of the largest strikes in the history of American labor (PHMC 2007). Workers organized to fight long working hours, low wages, and dangerous conditions. A search of *Chronicling America: Historic American Newspapers* found 33 labor related newspapers published in Pennsylvania during this era.

### ***News worthy events: 1880 – 1899***

1881 November 15

The American Federation of Labor was organized in Pittsburgh with Samuel Gompers as a prime mover. It continues as one of the strongest unions in the world as part of the AFL-CIO (Reynolds 1999).



1882 March 6

The Homestead mill of the Pittsburgh Bessemer Steel Company, owned by Andrew Carnegie, had its first strike when mill workers refused to sign "yellow dog" contracts; violence followed (Lorant 1999).

1882 June 1

A nationwide strike resulted from a general stoppage of work called by the Amalgamated Association of Iron and Steel Workers in a wage dispute (Lorant 1999).

1882 - Celebration and reenactment held in Chester and Philadelphia in honor of the 200<sup>th</sup> anniversary of the landing of William Penn (March 1915).

1884 – The Punxy Groundhog Club started a “mock pagan” ritual of weather forecasting (Reynolds 1999).

1885 - Elizabeth Cochrane (1867-1922), the Nellie Bly of the "Around the world in 72 days" fame began her newspaper career as a reporter for the *Pittsburgh Dispatch*. She investigated the factories and public institutions of the city and reported about them (Lorant 1999).

1887 – Centennial of the adoption of the U. S. Constitution was celebrated in Philadelphia by a great industrial parade one day and a military parade on the next (March 1915).

1887 – An amendment to the state constitution was proposed by Prohibitionists. It was defeated by a large majority, the people in the largest cities not being ready for such a measure (March 1915).

1888 March 11

The Great White Hurricane, a blizzard of snow over the eastern part of the country, cut off all communication from Philadelphia, destroyed shipping, and blocked railroads. Workmen who attempted to clear the drifts from the tracks were frozen to death (March 1915).

1889 May 31

The Great Johnstown Flood destroyed the city when heavy rain caused a dam to break above the city at the South Fork Fishing and Hunting Club whose members included wealthy industrialists, bankers, and attorneys. Between two and three thousand people perished. “Great masses of wreckage and bodies of human beings were piled up against a stone bridge of the Pennsylvania Railroad below the city. This caught on fire and its lurid glare added to the horror of the scene. The magnitude of the disaster aroused the sympathies of the whole world and people everywhere sent assistance to the flood sufferers” (March 1915).

1889 September 2

Labor Day celebrated as a legal holiday for the first time in Pennsylvania (Walther 1925).

1892 July 6

Homestead Strike, armed private police from the Pinkerton National Detective Agency hired by Henry Frick, general manager of the Homestead steelworks owned by Andrew Carnegie, battled armed Amalgamated Association of Iron, Steel & Tin Workers. National Guard troops were ordered to Homestead by the governor to prevent further violence; the strike ended soon thereafter (Licht 2002).

1893 March 20

Ground was broken for the first Ferris wheel in Pittsburgh, invented by George Washington Ferris of the city (Lorant 1999).

1894 – The state capitol burned and many valuable papers and documents were lost. “After the destruction of the building, agitation arose in favor of moving the capitol elsewhere. It was finally determined, however, to rebuild upon the old site. The agitation had at least one good effect, for it awoke the people of Harrisburg from their lethargy and made them introduce civic reforms which placed their city among the most progressive of the state” (March 1915).

1895 – To discourage child labor, the General Assembly enacted a Compulsory Education Act mandating that children between 8 and 13 years of age attend school for at least four months per year (PHMC 2007).

1897 September 10

In the Massacre at Lattimer Mines, unarmed striking miners were fired on by a posse organized by the Luzerne County Sheriff at the behest of miner owners (PHMC 2007)

1898 – Willa Cather, novelist, joined the staff of the *Pittsburgh Dispatch*. Two years later she became reporter for the *Pittsburgh Leader*, the city’s leading newspaper (Lorant 1999).

1898 – Using lantern slides at lectures and the newspapers, Mira Lloyd Dock and J. Horace McFarland began to advocate for the “City Beautiful” movement to clean up industrialized Harrisburg (Contosta 2002).

## **Progressive Era**

Progressive reformers recognized the serious political, environmental and social problems resulting from the rapid industrialization and urbanization that occurred during the last half of the 19<sup>th</sup> century. They pointed to political corruption, business monopolies, the erosion of democracy in general, the exploitation of workers, and disease and poverty among the urban population (Contosta 2002). At the beginning of the 20<sup>th</sup> century, Pennsylvania’s reformers had limited success in cleaning up entrenched political machines in the state’s largest cities and in the state capital. Although they had some success in pushing through social reforms, labor strife continued.

### ***News worthy events: 1900 – 1922***

1901 February 1

United States Steel was incorporated; it was the first billion-dollar corporation in American history (Contosta 2002).

1901 April 17

First automobile accident in the city of Pittsburgh was reported (Lorant 1999).

1901 July 14

The Amalgamated Association of Iron, Steel and Tin Workers called a general strike against the U.S. Steel Corporation subsidiaries, the first steelworker’s strike since 1892 (Lorant 1999).

1901 August 5

Pittsburgh newspapers and the general public demanded construction of a water filter system as hospitals became crowded with 266 victims of typhoid fever (Lorant 1999).

1902 -1903 – Anthracite coal miner strike, in which President Theodore Roosevelt intervened, set the pattern for non-violent arbitration in labor relations (PHMC 2007).

1903 October 13

The first of baseball's modern World Series ended in Pittsburgh before 7455 persons at Exposition Park; the Pirates were defeated by Boston, 4 to 3, and lost the Series three games to five (Lorant 1999).

1904 – Ida Tarbell, a native of Pennsylvania's oil region, published her *History of the Standard Oil Company*. Labeled a muckraking account, it revealed the illegal means used by John D. Rockefeller to monopolize the early oil industry (PHMC 2007).

1905 – Mounted state police were created to patrol the rural districts and other unprotected places. They were used in protecting life and property in times of strikes, and in capturing criminals out of the jurisdiction of the ordinary police. It was the first uniformed police organization of its kind in the United States (March 1915).

1905 – H.J. Heinz Co. led a movement for model factories based on the principle that workers deserve clean, pleasant work conditions with some chance for self-improvement (PHMC 2007).

1905 June 19

The first all-motion-picture theater in the world was opened on 433-35 Smithfield Street in Pittsburgh by John P. Harris and Harry Davis. The term "nickelodeon" was coined there (PHMC 2007).

1906 October 4

The new state capitol was dedicated. Citizens were led to believe it was constructed within its \$4 million budget. They soon learned that overpriced decorations and furnishings, as well as kickbacks and graft had inflated the cost to \$13 million (Contosta 2002).

1907 – First statewide farm products show was held in Harrisburg (PMHC 2007)

1907 – Pittsburgh's first Bureau of Smoke Control was instituted. In the national press, Pittsburgh was the prime example of a dirty and polluted industrial city (Contosta 2002).

1909 July 15

Rioting workers of the Pressed Steel Car Company, near McKees Rocks, clashed with guards; Coal and Iron Police and State Constabulary were called out; many strikers were wounded by volleys of buckshot fired to halt their charges; State Police were ordered to "shoot to kill" if attacked by strikers (Lorant 1999).

1911 January 1

The first old age retirement plan for United States Steel employees was created with a fund of \$12,000,000 of which \$4,000,000 had been provided by Andrew Carnegie ten years before (Lorant 1999).

1912 October 19

The Pitt Panthers were defeated, 45 to 8, by Jim Thorpe and the Carlisle Indians from the Carlisle Indian Industrial School which was established in 1879 by the federal government and social reformers to transform Native Americans into self-supporting "Americans" (Lorant 1999).

1913 – The semi centennial celebration of the anniversary of the battle of Gettysburg was organized as a peace jubilee. The veterans of the North and South met as friends in a great encampment upon the Gettysburg field where they had fought. They came from every state in the Union as the guests of Pennsylvania and all of their expenses within the state were paid out of appropriations made by the legislature (March 1915).

1913 – The General Assembly approved a women’s suffrage amendment to the state’s Constitution but Pennsylvania’s male voters rejected the amendment (PHMC 2007).

1913 December 1

The world’s first “drive-in gas station” opened in Pittsburgh (PHMC 2007).

1917 -1918 – Pennsylvania furnished nearly 300,000 men for World War I. This was almost one out of every twelve of the total number. Pennsylvania’s shipyards, mills, and factories provided a large part of the war materials for the nation (PHMC 2007).

1918 April 1

Daylight-saving time, conceived and promoted by Robert Garland, of Pittsburgh, went into effect (Lorant 1999).

1918 June-July

Violent racial incidents occurred in Philadelphia as the black population increased during World War I (Contosta 2002).

1918 October 5

An influenza epidemic, starting at the Cantonment Hospital in Point Breeze, victimized 23,268 Pittsburgh residents. Of the total, 1374 died of lobar pneumonia and 678 of bronchopneumonia (Lorant 1999).

1920 November 2

The first commercial broadcast station in the world was KDKA in Pittsburgh, which started daily schedule broadcasting. The election returns were its first scheduled broadcast. The first church service broadcast by radio occurred on KDKA a year later, and the first public address by radio was made by Herbert Hoover at the Duquesne Club in Pittsburgh in 1921 (PHMC 2007).

1922 April

Miners in Windber went on strike to gain union recognition, preserve their pay rates, and have the coal that they dug weighed accurately. It was the first national strike by both anthracite and bituminous miners (WITF 2003).

## **Modern newspaper history**

Patterns of newspaper growth have changed drastically in the twentieth century. Economic pressures forced many newspapers either into bankruptcy or consolidation. Even Pittsburgh and Philadelphia have been reduced to two major dailies; the *Pittsburgh Post-Gazette*, the *Pittsburgh Tribune-Review*, and the *Philadelphia Daily News* and the *Philadelphia Inquirer*, respectively. In Harrisburg, two once independent newspapers – the morning *Patriot* and the afternoon edition of the *Evening News* – combined to form the *Patriot-News*. Yet while urban dailies are disappearing, the newspaper medium is persisting in a form quite different from the large metropolitan press. The spectacular growth of suburban newspapers seems to suggest a regeneration of the old country tabloid. These slim, ephemeral, neighborhood papers record valuable “local” events and opinions often left unnoticed in the big city newsrooms.

The total number of Pennsylvania newspapers currently published remains quite high. The 2005 *Pennsylvania Newspaper Directory* for the Pennsylvania Newspaper Association lists 199 weekly newspapers, 83 daily newspapers, 2 biweekly newspapers, and 4 monthly newspapers. Since the mid-1990s however, many of the major city and town papers across the Commonwealth are now featured as online publications. One hundred twelve Pennsylvania newspapers appear as online versions. For a

current listing of online titles, see: <http://www.usnpl.com/panews.php/>. Despite shifts in editorial principles and distribution practices, Pennsylvania news journalism continues to enjoy an excellent reputation for its diversity, quality and comprehensiveness.

## **Pennsylvania's Newspaper Repositories on Microfilm and in Paper**

As one of the largest research libraries in the Commonwealth, the State Library of Pennsylvania has holdings in almost every area of human concern in its massive collection, including newspapers. The Library has been in existence since 1745 when Benjamin Franklin purchased the first library materials (including newspapers!) for the General Assembly. Today, the collection of Pennsylvania newspapers held by the Library is one of the largest in existence. The collection is comprised of 1,709 titles, represented in 6,510 historic volumes in paper and 105,350 microfilm reels. In 2006 the Library re-defined its collection development mission whereby designating the State Library as the repository of choice for all (historical and current) Pennsylvania newspapers in paper format in perpetuity.

The earliest Pennsylvania newspaper held by the Library is the December 22, 1719 issue of the *American Weekly Mercury*. This newspaper is reputed to be the first newspaper published in the middle colonies, and the fourth, in order of time, published in America. It is preserved on microfilm. The earliest Pennsylvania newspaper held in both film and paper formats is the April 25, 1771 issue of the *Pennsylvania Journal and Weekly Advertiser*.

The State Library receives 128 current newspapers, including all the major dailies published in Pennsylvania and at least one newspaper from each of Pennsylvania's 67 counties. The News Sources page on the Library's website provides access to prominent news sources available on the Internet, as well as access to indexes such as Infotrac and subscription databases such as ProQuest and the *Philadelphia Inquirer*.

In addition to the State Library's massive Pennsylvania newspaper collection, holdings of historical newspapers in both microfilm and paper format are located at numerous other cultural repositories throughout the state, such as the Carnegie Library of Pittsburgh, where an extensive collection of local historical newspapers is held. Typically, these collections are local, regional, or even neighborhood-specific in scope and frequently duplicate the holdings found at the State Library. The Free Library of Philadelphia is the exception. A recent survey of their master microfilm holdings revealed that significant runs of many Philadelphia titles are held here. See the *Master Negative Vault Holdings List* in Appendix A.

## **Preservation Microfilming Efforts in the Commonwealth**

### **Pennsylvania Newspaper Project, Phase I – III**

The notion of preserving Pennsylvania's documentary record is not a new concept for the Commonwealth. Years before the *Pennsylvania Preservation Plan* was conceived and drafted in 1996, individual cultural repositories – libraries, museums, historical organizations, archives, and other memory historical organizations – assumed the monumental task of being good stewards of the state's resources through the efforts of housing, reformatting, stabilization and conservation. While the *Pennsylvania Preservation Plan* remains the foundation of statewide planning, it is yet to be fully realized. Active stewardship and cooperative action to not only preserve collections, but to create freely-available on-line

digital access to collections has been a continuing activity at many individual institutions and among regional cooperatives and consortia.

Within the past ten years, there has been a noticeable increase in individual institutions awareness of the preservation needs of their own collections. This new found awareness led many to initiate new preservation programs, while others have undertaken bold action to safeguard collections at risk through reformatting and republishing on the web. Of all the programs and activities launched, perhaps the most notable statewide cooperative in sheer size and scale was the Pennsylvania Newspaper Project (PaNP). Funded initially in 1983 by the National Endowment for the Humanities as a planning grant for part of the United States Newspaper Program, the PaNP developed into a viable funded project and continued its mission to locate, inventory, catalog, and microfilm Pennsylvania newspaper titles through 1990. The State Library served as the headquarters for both project administration and general oversight. In addition, regional cataloging centers were established at four partner institutions: The Pennsylvania State University Libraries, The University of Pittsburgh Libraries, the Historical Society of Pennsylvania, and the State Library. At the conclusion of the PaNP Phase I in 1990, 7,921 newspaper titles had been located from the 67 counties in the state. Holdings for each title were inventoried, titles cataloged in OCLC, and 3,203,114 pages (represented in 270 titles) were reformatted to microfilm.

In 2004, after a fourteen-year hiatus, the Pennsylvania Newspaper Project, Phase II was revived with funding from NEH to continue the reformatting work from 1990. A continuation grant was sought and awarded from NEH in 2006 to continue filming additional titles (PaNP, Phase III). The Penn State University Libraries continues in its role to serve as the lead institution for Phase II and Phase III of the PaNP and will conclude in March 2008. The funding from NEH for these two phases provided the Commonwealth with the needed financial support to reformat previously identified newspapers for microfilming, and to respond to local community nominations (pleas) to preserve their newspaper publishing histories. At the conclusion of the current project in 2008, over 150 newspaper titles, represented in 755,000 pages, will be converted to preservation-quality microfilm. All generations of microfilm produced by the PaNP (Phases I, II, and III) are held by the State Library in perpetuity for the Commonwealth and available via Interlibrary Loan access. In addition, the PaNP OCLC symbol continues to be maintained by the State Library cataloging department.

### **Other microfilming efforts in the Commonwealth**

The need to preserve the state's publishing heritage is highly desirable among many institutions. Some cultural repositories have found local funding to support preservation microfilming activities while others have sought and received funding through LSTA grants or local funding agencies. As early as the mid-1950s and early 1970s the Pennsylvania State University Libraries microfilmed newspaper titles not considered attractive to commercial microfilm producers, which for the most part were titles from Centre County. As funding permitted, other newspaper titles were reformatted to film including Penn State's student newspaper, the *Daily Collegian*. At the Free Library of Philadelphia, local Philadelphia city papers were preserved on microfilm. Within the last seven years, the Penn State University Libraries has had the opportunity to partner with local area historical societies to fill gaps in film holdings of several local newspapers including the *Centre Democrat*, the *Democratic Watchman* (Centre County titles) and Clearfield County's *Raftsmen's Journal*. Additionally, plans were carried out to partner with the Penns Valley Historical Museum in Centre County, PA to microfilm a major backfile of the defunct title the *Millheim Journal*, which had recently become publicly accessible after the paper backfile was transferred to the local historical society.

Numerous historical societies followed suit to preserve their local newspaper heritage. One such local effort was the work of area citizens and the historical society of Freeland, PA. Thirty-five years' worth of several early Freeland newspapers owned by MMI Preparatory School were filmed during spring 2003

after a successful fund-raising campaign. The reformatted papers were issues of the *Freeland Tribune* and the *Semi-Weekly Progress*, published from the late 1880s to 1921. Another example of a local attempt to preserve newspapers was the Huntingdon County Historical Society's effort to locate all existing issues of the *Flea-Bite* and the *Watchman* from Newton Hamilton, PA. As of fall 2002, Society members tracked down all known single issues from private citizens from the Newton Hamilton area.

The Phoenixville Historical Society in Chester County preserved their local heritage by filming the local papers while the State Library purchased copies of those reels as they were issued. The Old Line Museum in Delta, PA (York County) managed to secure private funding to preserve their local papers which were located at the State Library. The Butler County Public Library took steps to microfilm their local papers and borrowed originals from the State Library for inclusion in their project.

Preserving historical newspaper collections on microfilm, and securing the funding to do so, remains one of the highest priorities held by many cultural institutions statewide.

### **Newspaper Digitization Efforts in the Commonwealth**

In the last several years, the ability to digitize back issues of microfilmed newspaper pages and mount scanned images as searchable page or article-level text has become viable, and highly desirable. In July 2004, the Penn State University Libraries launched its first on-line newspaper repository – the *Historical Digital Collegian* archive. Using Olive Software's ActivePaper Archive, the Libraries was able to provide date and keyword searching at the article level for Penn State's student newspaper. Today the entire allowable backfile covering 1887 to 1987, per the agreement with the publisher, is available for use by researchers, students and scholars worldwide. See:

<http://www.libraries.psu.edu/digital/newspapers/historicalcollegian/>

Since the launch of the *Collegian* in 2004, the University Libraries has mounted additional on-line newspaper repositories. Included among these were the Penn State campus student newspapers from the Behrend College, the *Behrend Beacon*, and the *Hazleton Collegian* from the Penn State Hazleton Campus. Also added for public access was the *Lancaster Farming*, 1955 to 1983. *Lancaster Farming* is considered the third largest weekly agricultural publication in the United States and the largest agricultural publication being published on the east coast. *Lancaster Farming* is one of Pennsylvania's most popular farming newspapers.

Also launched in 2004 was the *Pennsylvania Civil War Newspaper* archive, a statewide collaborative digitization effort. Previously microfilmed newspapers owned by the State Library, the Lancaster Historical Society and the Penn State University Libraries were selected among available titles that represented a cross-section of the state's publishing record during the pivotal years before, during, and after the United States Civil War, roughly from 1840 to 1870. To date seventeen titles from ten Pennsylvania cities and towns are available on-line. In 2006/07, an additional twenty-three newspaper titles were scanned from microfilm and processed for inclusion at the site. As of this writing, 324,467 newspaper pages have been digitized by the University Libraries.

For more information about the digitized newspaper collections at the Penn State University Libraries, see: <http://apps.libraries.psu.edu/digital/index.cfm/> See Appendix B.1.

In 2005, The Office of Commonwealth Libraries launched a statewide digitization service known as the *Access Pennsylvania Digital Repository*, see: <http://www.accesspadigital.org/> Managed by the Health Sciences Libraries Consortium (HSLC)/Access PA on behalf of the Office of Commonwealth Libraries,

and using the CONTENTdm software product licensed through OCLC, the service permits the storage and retrieval of digitized collections created by libraries and other entities located throughout Pennsylvania. The primary focus of the service is to build digital collections by, for, and about Pennsylvania for the citizens of Pennsylvania. In cooperation with other Pennsylvania agencies, including the Office of Commonwealth Libraries, PALINET, and PALCI, the goal of HSLC/Access PA is to work toward a unified approach to digitization in Pennsylvania. Participation in the Digital Repository is open to all Pennsylvania libraries who participate in the *Access Pennsylvania Database*. See Appendix B.2.

As a result of this new statewide service, many libraries sought Library Services and Technology Act (LSTA) grants to digitize Pennsylvania-related collections, including local historical newspaper resources. In the two year period from 2006 to 2007, twenty separate LSTA grant applications were received to digitize newspaper holdings. While not all of these grants were awarded, many did receive planning support or outright funds to digitize newspaper collections.

Newspaper collections currently available through the *Access Pennsylvania Digital Repository* are:

- 1) *Ambler Gazette*, 1897 – 1923 (by the Wissahickon Valley Public Library)
- 2) *Huntingdon Gazette*, 1806 – 1812, 1815 – 1839 (by Juniata College Library)
- 3) *Huntingdon Journal*, 1838 – 1880 (by Juniata College Library)
- 4) *The Campus of Allegheny College*, 1948 – 1957 (by Allegheny College Library)

In addition to these runs, other newspaper-related collections can be found on the *Access Pennsylvania Digital Repository* site. Three noteworthy collections digitized by the State Library were the *Red Man*, the newspaper from the Indian School in Carlisle, PA, the *Pennsylvania Necrology Scrapbook* collection and the *Harrisburg Newspaper Index* covering 1799 – 1827. The *Newspaper Index* is a searchable historical index and contains the personal names, death and birth dates, and marriage announcements from four early Pennsylvania newspapers – the *Dauphin Guardian*, the *Oracle of Dauphin*, the *Harrisburg Chronicle* and the *Harrisburg Republican*. The *Pennsylvania Necrology Scrapbook* was created from microfilmed scrapbooks of obituaries clipped from various Pennsylvania newspapers dating from October 16, 1891 to March 3, 1904 and includes many Civil War veterans. For more information on the *Access Pennsylvania Digital Repository* see:

[http://www.accesspadigital.org/pdf/access\\_pa\\_dig\\_rep\\_guidelines.pdf](http://www.accesspadigital.org/pdf/access_pa_dig_rep_guidelines.pdf)

Other significant newspaper digitization activities in the state have happened through the efforts of regional cooperatives or a single library's desire to mount their local papers. One such cooperative leading the digitization effort in the central region of Pennsylvania is the Lancaster County Digitization Project (LCDP). See Appendix B.3. With ten participating institutions (academic libraries, historical societies, and museums) with an interest in digitizing Lancaster County documents, the group has successfully mounted the backfile of the *Columbia Spy*, 1830 – 1889, the *Lancaster Journal*, 1796-1816, and the student newspaper, the *College Reporter*, 1915 – 2001 (from Franklin & Marshall University). For more information on the LCDP see:

<http://edisk.fandm.edu/christopher.raab/lcdp.html>

In the western part of the state, newspaper digitization efforts include the Pittsburgh Jewish Newspaper Project, hosted by Carnegie Mellon University Libraries. See Appendix B.4. Utilizing collections from The Rauh Jewish Archives at the Senator John Heinz History Center and The Jewish Chronicle of Pittsburgh, the project was underwritten in part by the United Jewish Federation of Pittsburgh. As a result, an on-line resource containing *The Jewish Criterion* (1895-1962), *The American Jewish Outlook*



(1934-1962), and *The Jewish Chronicle* (1962-Present) is now available. This project serves as both an on-line reference source and a digitized historical documentation of the Jewish community of Pittsburgh and its outlying areas. For more information on this Pittsburgh Jewish Newspaper Project see: <http://pjn.library.cmu.edu/>

The Hamlin Memorial Library of Smethport is an example of a single library's attempt to digitize their local historic newspapers with assistance from a commercial vendor. See Appendix B.5. The on-line repository at the Library of Smethport contains over 53,000 historical newspaper pages representing sixteen titles from McKean County, Pennsylvania. For more information see: <http://smethport.newspaperarchive.com/>

## **Methodology and Standards**

### **Access to Master Negative Microfilm, 1880 -1922**

Due to the diverse and independent nature of microfilming programs in the state, master negative microfilm holdings exist in numerous locations, including libraries, cultural repositories and national commercial vendors. With the vast microfilm holding held by the State Library and those held by the Free Library of Philadelphia, the PaDNP will rely on these holding locations to identify titles for consideration by the PaDNP Advisory Board as candidates for the Project. Both institutions are committed to partnering with the Penn State University Libraries to lend camera master negatives for technical evaluation and scanning. Letters of commitment were received from both institutions giving PaDNP access to their master negative collections. See Appendix C.

### **Title Selection Plan**

Title selection will be limited to newspapers published in Pennsylvania in English between 1880 and 1922. A two part title selection plan will be implemented to ensure the best candidates for digitization. Part one will evaluate the intellectual content and completeness of available titles in the targeted time frame, while part two will consist of analyzing the image quality of a given set of microfilm. The Project Manager and the Microfilm Project Assistant, who will be familiar with the current guidelines for preservation microfilming of newspaper for the United States Newspaper Program, will be responsible for this aspect of the work plan. An Advisory Board comprised of scholars, librarians, archivists, genealogists, and historians has been formed. Each member of this esteemed panel of experts has already expressed a commitment to serve and will constitute the statewide selection committee for the PaDNP. See Advisory Board members in Appendix D.

Once PaDNP is awarded funding, Project staff will turn to the *Master Negative Vault Holdings List* (see Appendix A) to begin the selection planning process. As an on-going effort, the Project staff will endeavor to locate other caches of camera negative film holdings that may exist elsewhere.

First generation camera negative microfilm will be the primary material type sought for the title selection phase of the Project and will be used as the source for film duplication and in digitization. The original paper copy, when available, will be used only if needed to fill gaps in the collection.

## Intellectual content criteria: Part I

In the first part of the title selection plan, the Advisory Board will evaluate and rank titles based on the intellectual content selection criteria set forth in the NEH/NDNP guidelines. This includes research value, geographic representation, and temporal coverage.

- The titles with the highest *research value* will significantly reflect the political, economic and cultural history of the Commonwealth. The titles will be of statewide or regional influence and important as a source of information about a variety of ethnic, racial, political, economic, religious, or other special audiences or interest groups. In addition, the titles will be known as the “paper of record” at the state or county level and contain published legal notices, news of state and regional governmental affairs, and announcements of community news and events, including births, deaths, and marriages.
- The title’s *geographic representation* will be indicated by broad geographic coverage at the state or at least multi-county level of a majority of the population.
- The title’s *temporal coverage* will have a long continuous run that includes the 1880 to 1922 targeted time frame, but may include runs from 1836 to 1922 for future digitization as NDNP increases its chronological span.

### *Refined selection criteria*

If more than one title meets historical, geographic, and temporal criteria, the Advisory Board will consider four additional factors in making the final selections.

- *Bibliographic completeness* of the microfilm copy within the 1880-1922 time frame is of utmost importance. A nearly complete microfilm file, however, can be enhanced through the scanning of paper copies when available. Important titles with only scattered issues on microfilm will not be considered.
- *Orphan titles*, which are high research value titles that have ceased publication and lack proactive owners who are likely to invest in future digital conversion, will receive special consideration.
- *Diversity* will be considered as a means to balance content with information that documents a significant minority community at the state or region level during the targeted time period (1880-1922).
- *Duplication* of other institutional efforts will be avoided; newspaper titles that have already been converted to digital form in an open access environment or as a commercial product will not be considered.

Project staff will research titles and provide the Advisory Board with information about each title including but not limited to: the OCLC bibliographic record, contents and news coverage, a list of historical events, editorial stance, political orientation, demographics of the readership, circulation statistics, geographic representation, and temporal coverage.

Title information and the ranking form will be made available to the Advisory Board through a password protected website. In addition, a listserv will be hosted and maintained by the Pennsylvania Newspaper

Association for communication and discussion between Project staff and the Advisory Board. Project staff will tabulate the forms and provide the Advisory Board with a ranked list of titles.

### ***Identifying candidates in urban areas***

A search of the *Chronicling America: Historic American Newspapers* database for English language newspapers published in Pennsylvania between 1880 and 1920 for all frequencies yielded 1,689 newspaper titles captured on service copy microfilm. In order to identify titles of statewide or regional influence with the broadest geographic representation, the U.S. Census Data for 1880-1920 was used to determine which Pennsylvania cities had the largest populations. Results from the census data revealed that 12 Pennsylvania cities were ranked among the 100 largest urban places in the United States.

Rank by Population of Pennsylvania's 12 Largest Urban Places  
Out of 100 in the United States, 1880 – 1920

<b>City</b>	<b>1880</b>	<b>1890</b>	<b>1900</b>	<b>1910</b>	<b>1920</b>
Allegheny	23	28	27	(A)*	
Allentown					95
Altoona			96		
Erie	70	73	73	85	75
Harrisburg	60	75	77	88	92
Johnstown				98	
Lancaster	77	94	90		
Philadelphia	2	3	3	3	3
Pittsburgh	12	13	11	8	9
Reading	41	48	50	55	64
Scranton	39	39	38	38	47
Wilkes-Barre	79	80	75	83	94

(A)\* Annexed by Pittsburgh

Source: U.S. Bureau of the Census

<http://www.census.gov/population/documentation/twps0027/tab01.txt>

Based on these findings, The *Chronicling America: Historic American Newspapers* database was then used to generate a list of titles available as service copy microfilm published in each of these 12 cities. Using this method, the number of candidates for digitization was narrowed down to 418 titles from 1,689. Finally, institutional vault listings were used to find the holder of the master negative microfilm for each of the titles known to exist on service copy film. This resulted in identifying 48 titles representing 700 years totaling approximately 1.6 million pages. From this listing we will select 100,000 pages. See the *Master Negative Vault Holdings List* in Appendix A. Once PaDNP is awarded funding, Project staff will continue to investigate caches of camera negative film holdings that may exist elsewhere.

### ***Title ranking***

The intellectual content priority ranking procedure that PaDNP will employ for ranking titles will be modeled after the University of North Texas (UNT) 2006 NEH/NDNP Grant Application. Using the research information provided to them, the PaDNP Advisory Board will rank their selections using a priority scale. The UNT priority scale features the following ranks:

**Highest priority for digitization** - Newspaper meets all intellectual selection criteria and the lack of inclusion of title would constitute a serious gap in the project. Inclusion furthers scholarship.

**High priority**– Newspaper meets most intellectual selection criteria. Lack of inclusion of title would be noticed, yet other titles may suffice.

**Medium priority** – Newspaper meets some intellectual selection criteria. Lack of inclusion of title may not be noticed because other titles provide similar coverage.

**Low priority** – Newspaper meets few intellectual selection criteria. Inclusion of title may fill some gaps, but generally lack of inclusion would not affect scholarship.

**Not recommended** – Newspaper does not meet intellectual selection criteria

The foremost reason for selection of an individual title for digitization is that the newspaper significantly reflects the political, economic and cultural history of the Commonwealth. Preference will be given to titles that are recognized as the “paper of record” at the state or county level.

In addition, to achieve the highest priority for temporal coverage, the master negative microfilm files must be complete within the 1880 to 1922 time frame. A nearly complete microfilm file can be enhanced through the scanning of paper copies when available. Important titles with only scattered issues on microfilm will not be candidates for digitization. During the intellectual content selection process, this information will be obtained from the OCLC holdings records. Holding records will be verified later during frame-by-frame collation of the master negative microfilm.

## **Microfilm evaluation: Part II**

Newspaper titles meeting the intellectual content selection criteria will need to be of sufficient image quality on microfilm to yield satisfactory digital images and automated text conversion output. For the second part of the title selection plan, the camera master negative microfilm (1N) for highly ranked titles, without restrictions that interfere with the goals of the program, will be borrowed from collaborating repositories. Following the University of Kentucky’s NDNP film evaluation protocol and procedures, Project staff will perform a technical analysis of the 1N film for adherence to standards for image quality, which includes resolution, reduction ratio, density, and focus.

During the frame-by-frame collation and microfilm evaluation process, Project staff will record structural metadata elements in an electronic format that will include title-level, issue/edition-level, and page-level elements. Project staff will also capture technical metadata concerning the quality characteristics of the first-generation camera negative microfilm (1N) that will be used to create a duplicate second-generation negative (2N). The Project Coordinator and Project Manager will be responsible for adding the administrative metadata.

Note: Penn State University Libraries staff attended “Meta Morphosis: The University of Kentucky Film-to-Digital Institute,” in Lexington, KY, September 12-13, 2007 and have held subsequent conversations with the University of Kentucky NDNP staff to adopt their electronic microfilm evaluation form (for metadata and frame-by-frame collation data) for the PaDNP.

### ***Image quality selection criteria***

The technical factors that affect the success of microfilm scanning and optical character recognition (OCR) will be considered during the title selection process. They are as follows:

1. The quality of original text and microfilm capture. Project staff will use a light box and loupe to inspect the microfilm for page curvature, gutter shadows, out of focus pages, and skew. Preference in selection will be given to titles on higher quality microfilm.
2. The reduction ratio used when microfilming the original newspaper. Preference in selection will be given to titles with a reduction ratio of 20x or lower.
3. The resolution of the camera master negative microfilm. Project staff will use a 100x power microscope to read resolution test patterns on the camera master negative microfilm. For film without resolution test charts, resolution will be estimated by comparison to film with resolution test charts and the original material if available. Preference in selection will be given to film with resolution test patterns readable at 5.0 or higher.
4. Variations in density within images and between exposures. Project staff will use a densitometer to take 10 maximum density readings on the camera master negative microfilm. Preference in selection will be given to film with a narrow density range (e.g. .90-1.20) and with variations in density readings of no more than 0.2 within an image and between exposures.

The physical condition of the film will be assessed relative to the film type. Frame-by-frame collation will be performed to verify the completeness of the run on film. An effort will be made to deliver as complete a title run as possible by locating and substituting a limited number of scanned images from paper. For a camera master negative that is questionable, with respect to any technical image quality criteria, sample digital images will be tested for usable OCR confidence levels to determine suitability for selection.

### **Final title selection**

Titles which meet the NDNP standards for both intellectual content and image quality will be selected for digitization. After title selection is confirmed by the Advisory Board, the Project Manager will write a 500 word summary description of the scope, content, history, and significance of each title selected for digitization based on the information previously provided to the Advisory Board.

Project staff will send selected camera master negatives to a reputable RLG Preservation Microfilming compliant-vendor for the production of a clean second-generation duplicate silver negative (2N) to be used for scanning. Following RLG Preservation Microfilming guidelines the vendor will inspect the 2N for resolution, maximum density (10 readings per reel) and minimum density. At the end of the award period, Project staff will ship bar-coded second-generation print negative (2N) film to the Library of Congress for deposit.

A single service copy (3P) duplicate for all PaDNP titles will be created for deposit at the Penn State University Libraries. This film copy will be used for cataloging purposes and CONSER authentication. The purchase of the service copy microfilm is outside the scope of this grant application and is not reflected in the PaDNP budget.

## **Partnerships and Existing Infrastructure**

Strong collaborative partnerships and a robust technological infrastructure to handle the massive electronic manipulation and storage needs of the Project are critical to fulfilling the requirement of a National Digital Newspaper Program. The Penn State University Libraries currently enjoys an enthusiastic partnership with the State Library and with the Free Library of Philadelphia. All three institutions meet regularly along with the Carnegie Public Library as the four-named State Resource Centers for the Commonwealth. Additionally, staff from each of these institutions currently serve on the PaNP, Phase III Advisory Board and have agreed to serve on the PaDNP Advisory Board. Preserving the Commonwealth's cultural heritage, resource sharing, and building strong collections (digital or traditional) are all common goals of these institutions.

The technical infrastructure found at the Penn State University Libraries is managed and maintained by Digital Library Technologies (DLT), a unit of the University's Information Technology Services division. DLT provides technical leadership in research and development of digital library initiatives, explores and evaluates suitability of new technologies to support the University Libraries' programs and services, and supports workstation and network services for the University Libraries. DLT is also responsible for implementing and maintaining the *Library Information Access System (LIAS)*, the primary tool for accessing the University Libraries resources and services. Named Project staff from the Libraries' Digitization and Preservation Department (the Project Coordinator and the Digital Project Coordinator) will work closely with DLT staff to secure and maintain the needed internal server space for file manipulation and validation for the PaDNP. DLT fully supports the technological needs of the PaDNP work plan.

Perhaps the most important partnership forged by the Project staff will be with the reformatting vendor selected for the Project. The Penn State University Libraries Digitization and Preservation Department staff have years of relevant experience working with service providers to reformat collections, including preservation microfilming, paper to digital, microfilm to digital, and paper to digital to film conversion.

## **CONSER Cataloging**

In 2007 the Cataloging and Metadata Department at the Penn State University Libraries achieved membership in CONSER (Cooperative On-Line Serials). Lead by serials cataloging librarian, Christopher Walker, the University Libraries has positioned itself among the elite consortium of government and academic libraries (fewer than 50 members) that provide high quality cataloging records for the national bibliographic database. Penn State is the first library to receive full membership since 2002, and is one of only three academic institutions in Pennsylvania that have achieved this status. The University Libraries' Cataloging and Metadata Services Department fully supports the bibliographic and CONSER cataloging needs of the PaDNP work plan.

## **Vendor Specifications and Contracting for Digitization**

As per the Pennsylvania State University guidelines, a formal bidding process will be conducted to select an acceptable vendor(s). A Request for Proposals (RFP) will be prepared in adherence with the *NDNP Technical Guidelines for Applicants* and competitive bids will be sought on the costs of microfilm duplication from the archival master negative, technical analysis and metadata concerning the quality characteristics of the second generation duplicate silver negative microfilm (2N), scanning the 2N microfilm to generate high-quality digital images, image processing and content conversion, quality

assessment of all page images, METS/ALTO, JPEG 2000, and PDF formats with all required metadata tagging, technical validation of digital objects, and shipping and handling of deliverables to PaDNP. A contract will be written with the successful vendor(s) upon review and careful consideration of all received request for proposal responses. The Project Coordinator will oversee all aspects of the contract with the vendor(s). The successful vendor(s) will identify a designated contact person with whom the Project Coordinator may communicate for the duration of the project

Note: A Request for Information (RFI) for the PaDNP was requested from OCLC Preservation Services Center, Bethlehem, PA and the information and pricing provided was used in calculating the budget documentation for the grant application. See the OCLC Preservation Services Center documentation in Appendix E.

## **Work Plan: July 2008 – June 2010**

### **Project Ramp Up: July 2008 – October 2008**

- Hire Project Manager
- Hire and train Microfilm Project Assistant to perform microfilm evaluation and collation
- Establish offices for Project Manager, Microfilm Project Assistant, and Digital Project Assistant at Penn State University Libraries' new off-campus facility. Note: The plan for this office space is in progress as the University Libraries moves to occupy this new facility.
- Set up microfilm evaluation lab at the same facility
- Send two representatives to meeting of NEH/LC NDNP award recipients
- Host a technical representative from the Library of Congress for a site visit and technical consultation
- Contact Advisory Board to announce grant award and to confirm their participation
- Set up the listserv to facilitate communication between staff and Advisory Board
- Develop website with access for Advisory Board during title selection process
- Post comprehensive title list, title information, and ranking forms on website
- Write Request for Proposals for film duplication and digital conversion; along with the University's Purchasing Services agent seek bids from qualified vendors
- Contract with vendor to produce deliverables as defined in the NDNP Technical Guidelines
- Hire and train Digital Project Assistant to perform in-house quality control of digital images

### **Selection of Titles: August 2008 – December 2008**

- Place a statewide call to cultural repositories to locate additional master negative newspaper holdings
- Schedule first meeting of Advisory Board for August 2008
  - Review the *Master Negative Vault Holdings List* (see Appendix A) and title information and any additional found holdings
  - Review NDNP intellectual content criteria and ranking procedure
  - Select title with at least 5,000 pages for test sample as per NDNP intellectual content criteria
  - Demonstrate use of website for ranking remaining titles
- Set October 2008 deadline for completing ranking forms

- Tabulate results of title ranking
- Discuss title rankings with Advisory Board through listserv
- Create prioritized list of titles for microfilm evaluation

### **Evaluation of Microfilm and Metadata Collection: September 2008 – June 2009**

- Develop a project tracking mechanism for film and files
- Develop an electronic format for recording descriptive, technical, structural, and administrative metadata based on the University of Kentucky's seven part electronic evaluation form
- Request archival master negative (1N) microfilm of selected titles from repository(ies)
- Inspect 1N microfilm for physical condition (i.e., film base) and for technical quality image indicators (i.e., resolution, density, reduction ratio, focus) in accordance with USNP preservation microfilming guidelines and *NDNP Technical Guidelines*; record technical metadata in the electronic form as required in Appendix A of the *NDNP Technical Guidelines for Applicants*;
- Collate 1N microfilm frame-by-frame; record descriptive and structural metadata in the electronic form as required in Appendix A of the *NDNP Technical Guidelines*
- If necessary, search for, borrow, and scan original paper copies of missing issues and/or pages
- Send 1N microfilm that meets image quality indicators to vendor for duplication along with associated metadata
- After duplication, return 1N microfilm to repository(ies)
- Return failed 1N microfilm to repository(ies)
- Repeat cycle through list of prioritized titles until 100,000 pages of quality microfilm are identified
- Write summary description for each digitized newspaper title

### **Scanning and Master Image Formatting by Vendor: October 2008 – April 2010**

- Produce a second generation duplicate silver negative (2N) microfilm from each 1N reel
- Produce a positive service copy generation (3P) for the Penn State University Libraries to be used for cataloging purposes. Note: The creation of the 3P film copy will not be charged to the PaDNP grant.
- Inspect 2N for resolution, maximum density (10 readings per reel), and minimum density
- Return 1N to PaDNP
- Scan the 2N and create the master digital images according to *NDNP Technical Guidelines*
- Send the 2N to PaDNP
- Each newspaper digitized for NDNP will be supported by coherent metadata to provide intellectual access and support navigation of the structure of the publication. The tables in Appendices A, B & C: of the *NDNP Technical Guidelines for Applicants* will be used to identify the elements appropriate for metadata collection and creation.
- Perform quality control of digital page images and associated metadata

### **Image Processing and Content Conversion by Vendor: October 2008 – April 2010**

- Process digital scans to produce derivative images and OCR files that conform to the *NDNP Technical Guidelines*



- Create METS objects with descriptive, technical, and structural metadata
- Divide data into batches of approximately 500GB
- Validate data with an NDNP program-specific software application to ensure technical conformance with the digital object profiles and specifications
- Load validated batch onto durable external hard drives and deliver to PaDNP for quality inspection and validation

### **In-house Quality Control: October 2008 – April 2010**

- Upon receipt of delivered media, run batch integrity verification using LC validation processor according to the technical validation of digital objects section of the *NDNP Technical Guidelines*
- Review directory structure, image files, OCR, descriptive, technical, and structural metadata for quality and adherence to *NDNP Technical Guidelines*
  - Evaluate OCR by viewing a sample of pages in the Digital Validation Viewer and by examining PDF hidden text; calculate word-level accuracy for an article as a percent of total words
  - Verify number of pages per issue against collation sheets
  - Check all images in thumbnail view to ensure no errors occurred in file transfer; view a sample of 10-20 percent (depending on volume) of TIFF, JPEG2000 and PDF files in the LC Digital Validation Viewer for quality
  - If image appears to have defects, check the microfilm negative to determine source of problem
  - Check metadata in the LC Digital Validation Viewer and by reviewing the raw XML data. Give special attention to issues with title changes, missing pages, or other anomalies
- Communicate with production vendor(s) to correct problems with quality or validation errors
- If corrections are needed, run validation of modified batch
- Send validated files to LC according to the specifications outlined in the Delivery of Digital Assets section of the *NDNP Technical Guidelines*; delivery of the digital assets will be via tracked shipments of durable external hard drives (Firewire-enabled)

### **Delivery to Library of Congress: June 2009 – April 2010**

- Deliver 25,000 pages as digital assets\* to LC on or before June 1, 2009
- Deliver 6,818 pages per month as prescribed digital assets\* to LC from June 2009 to April 2010 for a total of 75,000 pages
- Deliver title essays as the digital files are delivered to LC

\*Digital assets will include TIFF 6.0 uncompressed, 8-bit grayscale images, OCR text files with bounding-box coordinates, PDFs with hidden text, JPEG200 images, and all metadata delivered in a METS object structure according to the prescribed XML batch template structure as described in the *NDNP Technical Guidelines*.

**Update MARC Record from the CONSER Database: June 2009 – April 2010**

- For each PaDNP title, update the existing MARC record to the CONSER database
  - Records will be in MARC 21 Communications format
  - LCCN will be used as the basis for the title identifier; all LCCNs provided in metadata will be normalized to MARC 21 standard.
- Bibliographic records will be updated to include links to digital newspapers in the national newspaper directory.
- Deliver bibliographic copy with the project data to LC

**Wrap Up: May – June 2010**

- Apply barcodes provided by Library of Congress to second generation duplicate silver negative (2N) microfilm reels
- Send the barcoded 2N microfilm holdings of Pennsylvania newspapers to LC
- Send final report to NEH



## Project Staff

PaDNP staffing will consist of a Project Manager, a Microfilming Project Assistant, and a Digital Project Assistant, all of which will be hired staff exclusively for the project. In addition, five staff from the Penn State University Libraries will provide specialized expertise as a cost share contribution. All project staff will be headquartered at the University Park campus of The Pennsylvania State University. See below for project staff a summary of duties and qualifications. See Appendix F for resumes.

**L. Suzanne (Sue) Kellerman**, Full-Librarian, The Judith O. Sieg Chair for Preservation, Head of the Digitization and Preservation Department at the University Libraries at The Penn State University will commit 25% of her time for the two-year duration of the Project as cost share. Ms. Kellerman will serve as the Principal Investigator and Project Coordinator and will be responsible for project administration and budget oversight. Other activities include working with the University's Purchasing Services to prepare and finalize contract agreements, communicate and serve as the liaison with vendors, repositories, Project staff, and to technology support units. Ms. Kellerman will be responsible for hiring and supervising the Project Manager and supervision of the Digital Projects Coordinator. Along with the Project Manager and the Digital Projects Coordinator, Ms. Kellerman will oversee the operations and management of workflow procedures for the duration of the grant to ensure NDNP compliance. Along with the Project Manager, Ms. Kellerman will convene the PaDNP Advisory Board for title selection and prioritized rankings, openly share information and facilitate publicity for the Project. In her role as head of the Digitization and Preservation Department, Ms. Kellerman is responsible for developing policies and programs to preserve traditional and digital assets of the Libraries 4.7 million volume collection. She has responsibility for preservation program activities in conservation, collections care, commercial binding, preservation microfilming, mass deacidification, disaster recovery and response, and digital library production initiatives that include digital conversion, e-publishing and preservation of electronic assets.

**Larry Wentzel**, Digital Projects Coordinator, Digitization and Preservation Department at the University Libraries at The Penn State University will commit 15% of his time in Year 1 and 25% of his time in Year 2 as cost share. Mr. Wentzel will train the Digital Project Assistant, will assist in managing the operations to assess the quality of digital objects, serve as the primary liaison with the Digital Library Technologies unit and assist in managing workflow procedures for the duration of the grant to ensure NDNP compliance. Mr. Wentzel has an extensive technical background in digitization and planning digitization products.

**Cheri Banks**, Staff Assistant, Digitization and Preservation Department at the University Libraries at The Penn State University will commit 10% of her time for the two-year duration of the Project as cost share. Ms. Banks will assist in ordering and receiving supplies and will handle the financial accounting and budgetary reporting associated with the grant.

**Kevin Clair**, Metadata Librarian, Cataloging and Metadata Services at the University Libraries at The Penn State University will commit 5% of his time for the two-year duration of the Project as cost share. Mr. Clair will serve as the metadata knowledge expert for the Project team. Mr. Clair has experience with METS having worked on University Libraries' Making of America, II digitization project.

**Christopher Walker**, Serials Cataloging Librarian, Cataloging and Metadata Services at the University Libraries at The Penn State University will commit 5% of his time in Year 2 of the Project as cost share. Mr. Walker will be responsible for updating the existing bibliographic records for each of the PaDNP titles to CONSER-level records as outlined in the work plan section of the grant application. He will also

be responsible for cataloging on OCLC and the University Libraries' LIAS catalog. Mr. Walker spearheaded the effort for the University Libraries to achieve membership in CONSER.

**Project Manager** will devote 100% of his/her time for the two-year duration of the Project and will assist the Project Coordinator in managing operations and workflow procedures, convening the PaDNP Advisory Board for title selection and prioritized rankings of the titles, oversee the microfilm evaluation and metadata collection and facilitate publicity for the Project. Other responsibilities will include hiring and training the Microfilm Project Assistant and the Digital Project Assistant in the technical inspection of microfilm in adherence with ISO specifications and national guidelines and best practices for preservation microfilm and in conformance with the *NDNP Technical Guidelines*. The Project Manager will prepare the 500 word summary description of the scope, content, and significance of each title selected for digitization. Additionally, this position will collaborate with repositories to borrow camera negative film, oversee the tracking and shipping of microfilm and external hard drives between repositories, vendors, PaDNP and LC, maintain project records, compile statistics, prepare reports and coordinate Project website design and listserv maintenance. Strong managerial skills, project management experience, supervisory experience, relevant practical experience in preservation microfilm, knowledge of technical standards and best practices for preservation microfilm production and digitization, and knowledge of library operations including an understanding of bibliographic control methods will be requirements for the position. This position will report to the Project Coordinator.

**Microfilm Project Assistant** will devote 100% of his/her time at 40 hours/week for Year 1 of the Project. This position will research and compile title information, perform technical inspection of microfilm in adherence with ISO specifications and national guidelines and best practices for preservation microfilm, perform frame-by-frame collation, and collect metadata. The Microfilm Project Assistant will be responsible for tracking of the film evaluation in electronic form in conformance with the *NDNP Technical Guidelines* and will apply barcodes to the second-generation microfilm produced for deposit at the Library of Congress. This Assistant will also prepare and track microfilm shipments between repositories, vendors, PaDNP and LC. Project completion experience, an acute attention to detail and familiarity with photography or microfilm along with knowledge of library operations and an understanding of bibliographic control methods will be requirements for the position. This position will report to the Project Manager.

**Digital Project Assistant** will devote 100% of his/her time for the two-year duration of the Project; 20 hours/week in Year 1 and 40 hours/week in Year 2. Primary duties will include assisting with the technical inspection of microfilm in adherence with ISO specifications and national guidelines and best practices for preservation microfilm, perform collation and collect metadata. The Digital Project Assistant will also perform quality inspection of digital assets and assist in sending and tracking shipments as outlined in the *NDNP Technical Guidelines*. Project completion experience, an acute attention to detail, relevant practical experience in digitization, preferably film to digital conversion, computer application and system support analysis will be requirements for the position. This position will report to the Project Manager.

## Conclusion

Since 1745, The Commonwealth of Pennsylvania has been actively collecting and preserving our common wealth of newspaper publishing history. Through the preservation microfilming activities from the 1950s, 1970s and the NEH-funded PaNP projects of today, scores of newspapers have been preserved on film. While modest film-to-digital conversion projects have sprung up across the state in recent years, mass digitization efforts have yet to be realized until now. With this grant, the Commonwealth eagerly

looks forward to participating in the National Digital Newspaper Program by contributing Pennsylvania titles to the Library of Congress' *Chronicling America: Historic American Newspapers*.

## References

Ayer N. W. & Son. 1881, 1891, 1911, 1921. *American newspaper annual*. Philadelphia: N. W. Ayer & Son.

Contosta, D. R. 2002. Reforming the Commonwealth: 1900-1950. In *Pennsylvania: A history of the Commonwealth*, ed. Randall M. Miller and William Pencak, 258-316. Pennsylvania State University Press and Pennsylvania Historical and Museum Commission.

Gibson, C. 1998. *Population of the 100 largest cities and other urban places in the United States: 1790 to 1990*. Washington D. C.: U.S. Bureau of the Census, Population Division.  
<http://www.census.gov/population/www/documentation/twps0027.htm>.

Library of Congress. 2007. *Chronicling America: Historic American Newspapers*.  
<http://www.loc.gov/chroniclingamerica/>

Licht, W. 2002. Civil Wars: 1850-1900. In *Pennsylvania: A history of the Commonwealth*, ed. Randall M. Miller and William Pencak, 203-256. Pennsylvania State University Press and Pennsylvania Historical and Museum Commission.

Lorant, S. 1999. *Pittsburgh: The Story of an American city*. Rowman & Littlefield Publishing Group.  
[http://digital.library.pitt.edu/cgi-bin/chronology/chronology\\_driver.pl?searchtype=dbrowse&year=1880&year2=1889](http://digital.library.pitt.edu/cgi-bin/chronology/chronology_driver.pl?searchtype=dbrowse&year=1880&year2=1889)

March, T. S. 1915. *A history of Pennsylvania*. American Book Company  
<http://collection1.libraries.psu.edu/u/?digitalbks2,10834>

Pennsylvania Historical and Museum Commission. 2007. 1861-1945: Era of industrial ascendancy. In *Outline of Pennsylvania history*. <http://www.phmc.state.pa.us/bah/pahist/industry.asp?secid=31>

Reynolds, P. M. 1999. *Pennsylvania firsts: The famous, infamous, and quirky of the Keystone State*. Philadelphia: Camino Books, Inc.

Walther, R. J. 1925. *Happenings in ye olde Philadelphia 1680-1900*. Philadelphia: Walther Printing House. <http://www.ushistory.org/philadelphia/timeline/index.html>

WITF, Inc. 2003. Windber Strike of 1922-1923. *Behind the marker*.  
<http://www.explorepahistory.com/hmarker.php?markerId=938>