

7-2013

# The Vermont Digital Newspaper Project and the National Digital Newspaper Program: Cooperative Efforts in Long-Term Digital Newspaper Access and Preservation

Tom McMurdo

*Vermont Department of Libraries*, [thomas.mcmurdo@state.vt.us](mailto:thomas.mcmurdo@state.vt.us)

Birdie MacLennan

*University of Vermont*, [bmaclenn@uvm.edu](mailto:bmaclenn@uvm.edu)

Follow this and additional works at: <http://scholarworks.uvm.edu/libfacpub>

 Part of the [Library and Information Science Commons](#), and the [United States History Commons](#)

---

## Recommended Citation

Tom McMurdo and Birdie MacLennan, "The Vermont Digital Newspaper Project and the National Digital Newspaper Program," *Library Resources & Technical Services* 53, no.3 (July 2013): 148-163.

This Article is brought to you for free and open access by the University Libraries at ScholarWorks @ UVM. It has been accepted for inclusion in University Libraries Faculty and Staff Publications by an authorized administrator of ScholarWorks @ UVM. For more information, please contact [donna.omalley@uvm.edu](mailto:donna.omalley@uvm.edu).

# Notes on Operations

## The Vermont Digital Newspaper Project and the National Digital Newspaper Program

### Cooperative Efforts in Long-Term Digital Newspaper Access and Preservation

Thomas McMurdo and Birdie MacLennan

**Thomas McMurdo** (thomas.mcmurdo@state.vt.us) is Collections and Digital Initiatives Librarian, Vermont Department of Libraries, and **Birdie MacLennan** (bmacleonn@uvm.edu) is Director and Principal Investigator, Vermont Digital Newspaper Project, and Director, Resource Description and Analysis Services, Bailey/Howe Library, University of Vermont.

Submitted October 24, 2012; tentatively accepted November 21, 2012, pending modest revision; revision submitted December 10, 2012, and accepted for publication.

The Vermont Digital Newspaper Project (VTDNP) is made possible with funding from the National Endowment for the Humanities and with technical support from the Library of Congress. The authors wish to thank VTDNP state institutional partners at the University of Vermont, the Vermont Department of Libraries, the Ilsley Public Library, the Vermont Historical Society, and the Vermont State Archives and Records Administration for their support, encouragement and advice. We are especially grateful to Project Management Group members and administrative consultants, past and present, including Sally Blanchard, Prudence Doherty, Paul Donovan, Chris Kirby, Lorraine Lanisus, Martha Reid, and Mara Saule.

We are also grateful to our state Advisory Committee (<http://library.uvm.edu/vtnp/advisors.html>) for their advice and enthusiasm, notably in guiding our title selections. We also thank LRTS reviewers and editors for their insightful comments and advice on this paper.

*The Vermont Digital Newspaper Project (VTDNP) is a state partner in the National Digital Newspaper Program (NDNP). Developed by the National Endowment for the Humanities (NEH) and the Library of Congress (LC), the NDNP is a long-term, national effort to build a freely accessible, searchable Internet database of historical US newspapers. NEH provides funding to state projects to select and digitize historic newspapers published between 1836 and 1922. LC provides the technical support and framework for preservation digitization. Digitized newspapers are archived by LC and made freely available through the website *Chronicling America: Historic American Newspapers*. Vermont joined the NDNP in July 2010, when the University of Vermont Libraries were awarded NEH funding to embark collaboratively with state partners—including the Vermont Department of Libraries, the Ilsley Public Library of Middlebury, and the Vermont Historical Society—on the Vermont Digital Newspaper Project. Institutional partnerships and the engagement of committed individuals serve as a foundation to the VTDNP and provide an avenue to expand statewide infrastructures to accommodate large-scale microfilm-to-digital conversion and preservation efforts. Through collaboration and outreach, project partners select and digitize historical newspapers from microfilm and promote *Chronicling America*, a tool for discovery of these primary historical resources.*

The National Digital Newspaper Program (NDNP) was launched in spring 2004 as a partnership between the National Endowment for the Humanities (NEH) and the Library of Congress (LC) with the goal of developing a unique digital resource to provide enhanced access to United States newspapers as primary resource materials for the study of American history and culture.<sup>1</sup> Under the terms of the partnership, NEH awards funding to state projects to select and digitize historically significant newspapers published from 1836 to 1922, primarily by converting them from microfilm. The digitized titles are aggregated and permanently maintained by LC where they are freely and universally accessible on the Internet in a searchable database, *Chronicling America: Historic American Newspapers* (<http://chroniclingamerica.loc.gov>). The period from 1836 to 1922 was chosen to complement other digital newspaper resources that cover earlier periods of US history. Many newspapers published after 1923 are not in the public domain and therefore are not under consideration as candidates for inclusion

in the NDNP.<sup>2</sup> The NEH aims, over a twenty-year period, to have every state and US territory represented in *Chronicling America*.<sup>3</sup>

Vermont joined a growing roster of states participating in the NDNP in July 2010 when NEH awarded the University of Vermont Libraries a two-year grant to embark collaboratively with partnering institutions<sup>4</sup>—including the Vermont Department of Libraries, the Ilsley Public Library in Middlebury, and the Vermont Historical Society—in the Vermont Digital Newspaper Project (VTDNP). The project's plan, consistent with NDNP guidelines,<sup>5</sup> includes selecting, digitizing, and making available to LC approximately 100,000 pages of Vermont newspapers, published between 1836 and 1922, from the microfilm collections of the Vermont Department of Libraries and the University of Vermont. The project builds on work of the NEH-funded Vermont Newspaper Project which (<http://library.uvm.edu/vtnp/vtnp.html>), from 1997 to 2001, under the auspices of the United States Newspaper Program (USNP), identified, cataloged, and selectively microfilmed approximately 1,000 historical Vermont newspaper titles from 108 state repositories—including libraries, historical societies, museums, and court houses—to contribute nearly 3,000 unique local holdings records to the Cooperative Online Serials Program (CONSER) database in the Online Computer Library Center (OCLC).

This paper offers an overview of the USNP and the NDNP through the literature, and considers the issues, challenges, and benefits associated with developing support and infrastructure for a statewide preservation digital newspaper program. Methods for microfilm-to-digital conversion are explored and charted, specifically in relation to the technical requirements and production framework of the NDNP two-year grant cycle.<sup>6</sup> The processes and procedures outlined in

this article serve to inform colleagues at institutions that are considering ways to build collaborative, large-scale projects—notably around microfilm-to-digital conversion—as a means of giving new life and access to historical documents in the humanities, and preserving these documents in digital form for generations to come.

### Literature review

The National Digital Newspaper Program (NDNP) builds on the work of the United States Newspaper Program (USNP),<sup>7</sup> a cooperative national effort funded by NEH with technical assistance from LC. Over nearly three decades, from 1982 to 2010, NEH and LC worked with projects in all fifty states and US territories to locate, catalog, and selectively preserve on microfilm American newspapers from 1690 to the present.<sup>8</sup> The USNP revealed the extent and variety of newspaper holdings in the United States and resulted in the creation of a national database of newspaper titles and holdings on OCLC.<sup>9</sup> These records are central to identifying the selection of titles to be digitized under the NDNP. The widespread engagement and diverse collaborative partnerships of the USNP over nearly three decades has resulted in a substantial body of literature.

Connell's *USNP Bibliography of Newspaper and Periodical* (1983 to 1998) includes more than 200 citations that chronicle the diverse perspectives and experiences of USNP engagement at both the state and national levels.<sup>10</sup> Robert Harriman, LC's technical advisor to the USNP, provides background and context for the start-up of the USNP, including the focus on gaining bibliographic control of existing newspapers collections, the unique experiences of USNP participants working in the field to locate, catalog and inventory previously uncollected newspaper collections, and the preservation

challenges in working with fragile or disintegrating collections. He also discusses the adaptation of cataloging standards for newspaper bibliographic and holdings records in the CONSER/OCLC serials database.<sup>11</sup> In a subsequent article, Harriman reviews the structure of cooperative relationships that formed the basis of the USNP and provided a framework for the expansion of the CONSER/USNP database to encompass newspapers using a “master record convention” permitting catalogers to create a single bibliographic record for the newspaper as published (i.e., newsprint) and to attach local data records reflecting institutional holdings for each format (i.e., print or microfilm) for which associated issues are held.<sup>12</sup> The article also discusses efforts to inventory and evaluate existing microfilm for completeness and quality while recognizing the limitations of the USNP. For example, the structure and format of the local data (or holdings) record “prohibited the entry of any meaningful descriptive information . . . publication details about the microfilm producer and descriptive information necessary to determine film characteristics are not recorded.”<sup>13</sup> Input of coded information to denote whether film is positive or negative, service copy or master, etc. has been optional, rather than required. Thus, “we are lacking information about the existence of preservation master negative film produced prior to [the USNP].”<sup>14</sup> Nonetheless, Harriman realized that “the [USNP] is being asked to serve both as a model for cooperative preservation programs and as a laboratory to test methods and procedures crucial to the success of such programs . . . the expected result . . . should be a solid foundation of bibliographic information about newspapers on which to build.”<sup>15</sup>

Indeed, some twenty years later, in describing the start-up phase of the NDNP, Mark Sweeney acknowledges the distributed cataloging and

preservation microfilm work of the USNP and summarizes its accomplishments: "About \$55 million in grants to institutions representing all fifty states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands resulted in the bibliographic description of over 140,000 newspaper titles and close to one million local holdings records. In addition, over sixty-five million newspaper pages were preserved on microfilm."<sup>16</sup> Sweeney goes on to describe the NDNP initiative, modeled on a cooperative, phased development approach (as was the USNP), in the creation of *Chronicling America*, "a freely available national digital resource," which provides two different data sets to support use of newspapers: (1) a newspaper directory with descriptive records about American newspapers (from 1690 to the present) and where holdings are located, and (2) a growing collection of keyword searchable, digitally converted public domain newspaper pages selected and contributed by participating state projects.<sup>17</sup>

By the time funded activity under the USNP was completed in 2010, newspaper digitization initiatives in the US were a decade in the making. The Utah Digital Newspapers (UDN) Program emerged in 2001 when the University of Utah's Marriott Library was awarded a Library Services and Technology (LSTA) grant "to research and demonstrate a newspaper digitization project."<sup>18</sup> By the end of 2002, the Marriott Library had digitized 30,000 pages from thirty years of three weekly Utah newspapers and launched the Utah Digital Newspapers website (<http://digitalnewspapers.org>). The success of this first grant led to subsequent funding from both the LSTA and the Institute for Museum and Library Services (IMLS), and the formulation of partnerships with other state sites, including Brigham Young University and Utah State University, for the cooperative development of a distributed, aggregated, statewide

digital newspaper collection, universally accessible through the Internet. Arlitsch, Yapp, and Edge provide background and a detailed overview of early newspaper digitization efforts and the start-up history of the UDN Program.<sup>19</sup> A subsequent article by Arlitsch and Herbert discusses equipment and procedural considerations for newspaper digitization from microfilm and paper—including an overview of Optical Character Recognition (OCR) technologies and assessment models in producing searchable full-text from digital images.<sup>20</sup> Between 2002 and 2005, Utah's pioneering efforts in newspaper digitization resulted in the availability of approximately 400,000 pages of searchable newspaper content served on the web to a global audience. In two successive articles published in 2008, Herbert and Estlund further discuss UDN progress and subsequent funding, including engagement with the NDNP, as well as ongoing considerations for scanning, article-level zoning, and OCR software.<sup>21</sup>

In 2005, one year after the NEH announcement of the formation of the NDNP, the Marriott Library at the University of Utah was awarded NEH funding, along with five other state institutions in California, Florida, Kentucky, New York, and Virginia, to launch the start-up of the NDNP.<sup>22</sup>

In May 2006, Utah hosted the annual conference of the Newspapers Section of the International Federation of Library Associations (IFLA), along with the annual meeting of the NDNP by LC and NEH.<sup>23</sup> The proceedings of the conference, under the editorship of Hartmut Walravens, provide an international scope into newspaper digitization initiatives.<sup>24</sup> NDNP perspectives figure prominently. Helen Aguëra of NEH, along with Mark Sweeney, Ray Murray, and George Schlukbier of LC, presented "The U.S. National Digital Newspaper Program—Thinking Ahead, Designing Now," a four-part segment that documents respectively: (1) program

development, (2) preservation planning, (3) technical specifications, and (4) repository development.<sup>25</sup> Newspaper digitization outlooks from NDNP start-up participants in Utah, California and Kentucky are also included in the conference proceedings.<sup>26</sup>

As of 2012, there are thirty-two state projects participating in the NDNP, along with a cumulative body of literature that includes diverse perspectives on newspaper digitization, and user experiences that document the significance of universally accessible digitally archived newspapers on research and education. LC has made available a website of "Presentations and Publications," including works by NDNP developers and implementers at NEH and LC.<sup>27</sup> The NDNP "Extras" site showcases content in *Chronicling America* and provides examples of educational tutorials, search techniques, and teaching resources produced by state projects, as well as how some states are reusing digitized historic newspapers and associated metadata records in unique ways.<sup>28</sup>

Several state projects have published articles about newspaper digitization experiences. Terry's 2009 article, "The Digitization of Historic Newspapers on Microfilm: The Kentucky Experience,"<sup>29</sup> is an instructive, step-by-step description of the film-to-digital process at the University of Kentucky, home to Kentucky's NDNP Project. Newspapers published between 1836 and 1922, the NDNP timeframe, are notoriously irregular, often changing names, publication days, and ownership without warning. Combine the ravages of time and idiosyncrasies of "prestandard" filming practices on vulnerable acetate negatives with publishing irregularities, and the problems associated with digitization are compounded. Terry's article offers some practical ways to address these inconsistencies, emphasizing a clear and thorough process for collating the reels. Collation requires a

frame-by-frame inspection of every reel, recording issue dates, titles, and page order. Terry walks readers through this process.

The University of Kentucky has been a leader in providing instruction for prospective and existing NDNP state projects. Their Meta | Morphosis Film-to-Digital Institute, offered from 2006 to 2009, included two days of intensive instruction for film-to-digital conversion, giving new and potential NDNP participants the opportunity to see a full scale state program in action, and to discuss issues with experienced staff. As of this writing, Kentucky is re-tooling the Meta | Morphosis Institute for the online environment, developing a series of interactive tutorials and podcasts.<sup>30</sup> The University of Kentucky has also produced an online video, featuring their in-house operations for newspaper digitization.<sup>31</sup>

Robinson's 2010 article, "The Evolution of Newspaper Digitization at the Washington State Library"<sup>32</sup> recounts the Library's two-year pilot project, from 2005 to 2007, to complete newspaper digitization in-house for approximately 11,000 pages of historical newspapers. Robinson's article provides insight into the issues and challenges facing institutions attempting to complete large-scale digitization projects in-house. Relatively few NDNP state projects choose to do in-house digitization. This is due to high start-up costs for equipment, the need for extra staff, and scanning expertise. Digitization vendors can often do the work faster than homegrown projects, and with virtually no upfront expense. Conversely, the advantages to doing digitization work in-house are the ability to maintain control of digitization each step of the way, building expertise within one's organization, and, significantly, retaining purchased scanning equipment.

In 2007, the Washington State Library staff opted to outsource the scanning, metadata and OCR production as a means to increase efficiency

and output. Robinson makes it plain that the experiences gleaned from the pilot enhanced the expertise and understanding of newspaper digitization processes at the Washington State Library and positioned them for an NEH award in 2008 to launch their participation in the NDNP.

The planning overview and preparation for developing an NDNP initiative in Illinois is described in Scott's 2008 article, "The Illinois Digital Newspaper Project."<sup>33</sup> Scott places the history of Illinois and the Illinois press in the context of the 1860–1922 time-frame of the 2009 NDNP grant cycle, and outlines a "plan of action." The article is useful as a checklist for prospective projects, and provides readers with historical context for the significance of newspapers in documenting local experience of regional, national and international events.

NDNP guidelines and resources for digitization are available from LC.<sup>34</sup> Technical Guidelines for Applicants provide a full overview of all requirements and deliverables for NDNP participants.<sup>35</sup> It is "must reading" for those involved in, or considering embarking on, an NDNP project. Everything from selection guidelines, scanning and OCR specifications, and requirements for metadata, file structure and naming conventions is included, along with a comprehensive listing of associated resources.

### **Methodologies: the Vermont Experience**

Preservation reformatting of newspapers on microfilm has been an active concern in Vermont for over fifty years. The Vermont Public Records Division (Public Records) filmed newspapers in conjunction with the Vermont Civil War Centennial Committee in the early 1960s. Public Records undertook a massive newspaper microfilming enterprise in the late 1970s that extended into the early 2000s. In 2008,

Public Records was absorbed by the Vermont State Records and Archives (VSARA).<sup>36</sup> The Vermont Department of Libraries, as the central state newspaper repository, in partnership with the VSARA, now holds a collection of several million pages of newspapers on master negative microfilm. Approximately 850,000 of these pages fit within the NDNP publication timeline of 1836–1922.

Vermont was the fiftieth and final state to join the United States Newspaper Program (USNP), doing so in 1997 as the Vermont Newspaper Project (VTNP), a collaborative endeavor between the University of Vermont and the Vermont Department of Libraries. VTNP staff conducted a statewide survey to identify existing state newspapers, and subsequently traveled throughout Vermont's fourteen counties to inventory, catalog and consolidate holdings for collections of newspapers, both in print and on microfilm. VTNP staff cataloged, according to CONSER standards, approximately 1,000 unique titles with nearly 3,000 holdings (or "ldr"—local data records) in 108 Vermont repositories. Catalog records and associated holdings data for each repository were added to the CONSER database on OCLC and also maintained in a local database. The VTNP initially produced a printed Union List of Vermont Newspapers that was distributed to all participating repository institutions.<sup>37</sup> The database (<http://vtnp.uvm.edu>) was developed into an ExLibris Voyager interface, and made freely available on the Internet.

The VTNP database, or "catalog," as it is also designated, provides the framework for identifying titles for digitization. By generating reports from the catalog's bibliographic and holdings data, it was possible to learn what newspapers were published from 1836 to 1922, the period under consideration by the NDNP. The holdings records reveal what issues survived and in what format(s) they were

preserved (i.e., print and/or microfilm). The bibliographic data helps contextualize newspapers by indicating title changes, frequency, geographic locations, period of publication, and other newspapers that were competing within the given timeframe. Since NDNP state projects are limited to digitizing 100,000 pages per two-year grant cycle, this information is essential to developing title lists and page counts and to making informed selection decisions.

With NEH support, the USNP provided funds for microfilming newspapers. Some important collections and runs of issues were discovered during USNP canvassing that substantially contributed to the available microfilmed record of newspapers.

The USNP-Vermont Newspaper Project produced approximately 170,000 pages of film. While the VTNP uncovered a variety of newly identified titles, with runs of varying length that had not previously been filmed during the decades of preservation microfilming that preceded the advent of the USNP in Vermont, much of the filming effort focused on short runs or individual issues of titles that fill gaps in existing microfilm holdings. The VTNP also filmed several titles that were published after 1922 and are therefore outside the scope of the NDNP. As the Vermont Digital Newspaper Project (VTDNP) planning efforts began in 2009, we identified approximately 500 titles in the VTNP database on microfilm that were published between 1836 and 1922. While these titles formed the genesis of content from which to build a digital project, it would turn out that a relatively small proportion of titles to be selected for digitization under the VTDNP would come from USNP-produced microfilm. The establishment of partnerships and cooperative working agreements to build a statewide infrastructure in support of newspaper digitization would prove to be a critical element in developing the grant proposal, and essential

to successful outcomes for Vermont engagement with the NDNP.

### Project Genesis and Organization

From the moment that Vermont newspaper titles were made available online through OCLC's WorldCat and through the VTNP online catalog, Vermont newspaper enthusiasts from around the country began to inquire if the full-text to these titles could also be found online. Before the advent of the NDNP, these users were referred to their local repository for retrieval of microfilm. The primary impetus for Vermont engagement with the NDNP began with Chris Kirby, Adult Services and Technology Librarian at the Ilsley Public Library in Middlebury. In the summer of 2009, Kirby researched NDNP requirements for film-to-digital conversion and prepared a draft grant proposal giving a procedural compass to the project. The initial draft proposal provided a framework for representatives from diverse state institutions to come together in the fall of 2009 to begin discussions to lay the groundwork for building statewide support and infrastructure to launch an NDNP initiative. Included in the discussions was an assessment of Vermont cultural organizations with digitization experiences that might lend support to a statewide newspaper digitization effort, including the University of Vermont's Center for Digital Initiatives (CDI; <http://cdi.uvm.edu>), the digital collections of Middlebury College (<http://middarchive.middlebury.edu>), and the Online Digital Archive of the Vermont Folklife Center ([www.vermontfolklifecenter.org/archive](http://www.vermontfolklifecenter.org/archive)).

Formed in the context of these ongoing digitization efforts, the proposed Vermont Digital Newspaper Project would be the largest digital project to emerge in Vermont and, as such, was envisioned to foster collaboration with colleagues from around

the state in a shared set of goals to further hone digitization skills in-state, while addressing public demand for universal access to a searchable, full-text Vermont newspaper archive for citizens, historians and researchers. The USNP-VTNP experience had provided a framework for cooperative working relationships between state institutions and formed the basis for renewed partnerships in developing the Vermont NDNP initiative. As discussions advanced in the fall of 2009, a cohesive planning group was formed with representatives from the University of Vermont, the Vermont Department of Libraries, the Ilsley Public Library in Middlebury, and the Vermont Historical Society, to further define staffing and budget needs for completing the grant proposal. The University of Vermont emerged as the lead/application institution and host site. The Vermont Department of Libraries, as the central newspaper repository for the state, engaged staff support for research and development, as well as a collaborative commitment to enable access to preservation microfilm for the project. The Ilsley Public Library engaged staff support for ongoing research, development, and outreach in the public library sector. The Vermont Historical Society offered support in an advisory and outreach capacity. When the project was awarded funding in July 2010, the planning group evolved into an implementation group, and ultimately, into a cooperative, inter-institutional, six-member project management group.

The first six months of operation focused on organization of infrastructure. The University of Vermont established communication protocols using Basecamp (<http://basecamp.com>) browser-based project management software for online collaboration. Basecamp has proved to be an effective platform that includes messaging, file-sharing, "to-do" lists, calendar functions, and other useful features for diverse partners collaborating in

disparate geographical locations.

Other start-up activities included: the formation of a search committee to hire a project librarian to oversee production and manage the day-to-day activities of the project; the establishment of work space, including the procurement of equipment for microfilm inspection (densitometer, microscope, light table) and for digital workflows (two Dell optiplex workstations, Scan-Pro2000 microfilm reader/printer, ten 1 TB hard drives and shipping cases). The project director also initiated a Request for Proposals (RFP) to solicit bids for NDNP-compliant digitization services, including OCR creation, and provision of related metadata.

With the organizational structure in place in January 2011, Tom McMurdo was hired as project librarian. His prior experience with the NDNP's California Digital Newspaper Collection (<http://cdnc.ucr.edu/cdnc>) and the USNP's California Newspaper Project (<http://cnp.ucr.edu>) translated well to the startup of digitization activities for the VTDNP.

### Newspaper Selection Factors

The VTDNP title selection process consisted of three main factors. The first was the formation of an Advisory Committee of scholars, historians, journalists, teachers, librarians, and archivists. NDNP grant application guidelines stipulate the articulation of a plan to involve an advisory board in the selection process.<sup>38</sup> The second was to refine the long and unwieldy alphabetical lists of approximately 500 titles published between 1836 and 1922 that the project planning group had identified and to confirm the availability of corresponding master microfilm negatives. Finally, once the titles on master negatives were ascertained and compiled into coherent title families arranged by county, the Advisory Committee would be engaged to review the list of titles and to rank them based on NDNP

selection criteria.<sup>39</sup> Titles for digitization would be selected based on compiled ranking results.

A provisional Advisory Committee was formed in the planning stages of the grant proposal. Once the VTDNP was funded, Advisory Committee members<sup>40</sup> confirmed their commitment to the project and their engagement in the title selection process. Lorraine Lanius, on behalf of the Project Planning Group, led efforts to compile an Advisory Committee Briefing Book to provide general background information about the NDNP and the VTDNP to guide Advisory Committee members in the selection of newspapers to be digitized and to suggest other areas of engagement for committee members.<sup>41</sup> Reference to a timeline of historical events in Vermont was also provided.<sup>42</sup> An Advisory Committee "project" space was established on Basecamp in the fall of 2010 where the Project Management Group (previously the Planning Group) and Advisory Committee members could exchange correspondence and post documents.

During the planning phase of the VTDNP, an alphabetical list of more than 500 titles published between 1836 and 1922 available in microfilm format was generated from the VTNP database. The title list was subsequently sorted geographically by each of Vermont's fourteen counties and further culled to identify significant titles with long runs in each of the counties. However, newspapers are serials and in the nineteenth and early twentieth century they often changed name, absorbed competing titles, and were themselves absorbed, or split into daily and weekly editions, often without warning or apparent reason. It soon became apparent that the alphabetical lists would be better classified as title families—composed of directly successive titles making up a newspaper "family." By consulting geographical title lists developed by the Project Planning Group and

bibliographic records in the VTNP catalog, with their corresponding preceding and succeeding titles (as identified in MARC record fields 780 and 785), the project librarian grouped logical title families together and, with further research, developed a unique list of fifty-seven title families, representing newspapers from Vermont's fourteen counties.

The VTDNP endeavored to create a system whereby the advisors could rank titles in a manner consistent with the selection criteria outlined by the NDNP.<sup>43</sup> With an estimate of approximately 850,000 pages of newspapers on master microfilm negatives identified within the scope of the grant, and a high level of public interest in project outcomes, it was important to have a clear, transparent process in place for selection of the 100,000 pages to be digitized. A clearly defined, quantifiable selection process was essential to anticipating public concerns. To accomplish this, the VTDNP developed a title ranking form modeled on a form produced by the Pennsylvania NDNP project in their selection process and distributed the form to the advisory committee with their briefing handbook.<sup>44</sup>

The ranking form provides information in two sections: the "Title" area provides title and place of publication information, dates of publication and dates of available master negatives, and an estimated page count for those negatives. The "Ranking Criteria" section enables advisers to rank each title based on a scale of 1 through 5 (5 being the highest), as applied to three NDNP criteria: Research Value, Geographic Coverage, and Temporal Coverage. This ranking system yields a score between 3 and 15. Advisers can augment this ranking by adding a point if the title is a complete run, an orphan title, or expands the diversity of selections. "Complete run," means that the negatives available are a full collection of the title and do not have significant gaps. "Orphan title" means

**Table 1.** VTDNP title ranking form segment for Addison County titles

Vermont Digital Newspaper Project						Ranking Form						
Phase 1, 2010–2012						Ranking Criteria						
Addison County						Ranking Criteria						
Title or Title Family	Place of Publication	Digitization Candidates		Master negative date range	Estimated page count	Research Value 5 = highest priority	Geographic Coverage 5 = highest priority	Temporal Coverage 5 = highest priority	Complete	Orphan	Diversity	Total Score
		Title begin date	Title end date									
<i>Bridport Sun</i>	Bridport	1900	19uu	1901–1915	3,000	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	
<i>Bristol Herald</i>	Bristol	1879	1959	1880–1922	6,500	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	
<i>People’s Press/Galaxy/ Register</i>	Middlebury	1836	1937	1837–1922	18,000	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	
<i>City News/Vermontier/ Citizen/Enterprise</i>	Vergennes	1837	1982	1838–1922	25,000	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	
<i>Orwell Citizen</i>	Vergennes	1900	19uu	1901–1922	5,400	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	
<i>Middlebury Record</i>	Vergennes	1901	19uu	1901–1922	4,000	1 2 3 4 5	1 2 3 4 5	1 2 3 4 5	+1	+1	+1	

that the title has no direct descendent still in publication. “Diversity” means titles that may be of particular value to researchers because they are from either a significant minority community, or contain diverse or distinctive social or political perspectives. In Vermont, this includes abolitionist or antislavery newspapers, ethnic newspapers in Italian and French, and Democratic Party papers, in contrast to the dominant Whig and Republican titles in the mid- to late nineteenth century.

In April 2011, the University of Vermont hosted a day-long meeting where project planners and advisers met in person. Skype’s online conferencing capability was used to establish interactive connections with Advisory Committee members who could not be physically present. During this meeting, advisers familiarized themselves with project background and goals, as well as the selection process and ranking terminology, and came to a procedural understanding for ranking titles using the same rubric. The meeting also offered the occasion for lively group discussion of individual titles. The strength of the VTDNP Advisory Committee is that its members bring to the table a blend of

diverse newspaper experiences and perspectives with strong regional and statewide knowledge of Vermont history. Meeting attendees came away with valuable insights into the newspaper publishing history in Vermont and appreciation for titles under consideration for digitization.

Due to the strict production timetable for deliverables defined by the grant cycle, the Project Management Group proposed three titles to the Advisory Committee to launch digitization: the *Vermont Farmer*, *Rutland Herald*, and *Burlington Free Press*. The *Vermont Farmer* was chosen as a start-up sample reel for its size and format on a compact reel and as a title that offers perspectives into agricultural practices and methods in northeast Vermont during the 1870s.<sup>45</sup> The publisher was well known as an innovative plant breeder and orchardist. The title is not solely an agricultural journal, as it also contains local news, obituaries, and advertisements, and is rightfully called a “newspaper.” The Burlington Free Press and Rutland Herald are well-known as the two longest-running and historically significant titles in the state. These titles were universally approved by the Advisory Committee.

With production underway for

the three titles, Advisory Committee members selected and ranked additional titles over the course of the next month. Advisers with stronger regional knowledge tended to rank newspapers with which they were most familiar. Those with broad knowledge of historical Vermont newspapers tempered this regionalism by ranking titles that encompass a variety of geographical areas and times within the grant period. Overall, the Advisory Committee provided a well-balanced selection of newspapers. The project librarian subsequently compiled rankings. In June 2011, the Advisory Committee agreed on a list of twelve title families comprising thirty-seven unique bibliographic titles from ten of Vermont’s fourteen counties (see appendix).

## Digitization: Planning and Execution

### Planning

The NDNP requires that state projects work with existing master negative microfilm collections as source material because they yield the highest quality digital images of newspapers. It is important to start with the

best available microfilm negative to achieve NDNP standards, which are intended as “best practices” for digital preservation.<sup>46</sup> This requirement puts the charge on NDNP state projects to locate preservation microfilm negatives. Unlike public-use films, or service copies, of which there may be many, there is generally only one master negative, that is, the film that went through the camera at the time of filming. Subsequent generations of film of a given reel are derived from the negative. Availability and condition of master negatives for newspapers is often contingent on preservation archiving policies in a given state, and may vary.

Newspaper pages are to be scanned in 8-bit grayscale, as opposed to bi-tonal pure black or white, at 300–400 dpi.<sup>47</sup> This 300–400 dpi is relative to the original page, meaning that the scan, enlarged to the size of the original newspaper would contain 300–400 dots per inch. For each newspaper page, the NDNP requires delivery of an uncompressed TIFF 6.0, and a JPEG2000 and PDF derived from the original TIFF. Also required is “structural metadata to relate pages to title, date, and edition, to sequence pages within issue or section and to identify associated image and OCR files.”<sup>48</sup>

Submitted XML files at the reel, issue, and page level fit a delineated hierarchical structure that enables users to access newspapers by title, date, issue, and page. Attention to detail in the gathering of this data leads to accurate results for those searching newspapers.

OCR files with word-bounding boxes and columns at the page level are required for NDNP work.<sup>49</sup> OCR files are XML-encoded representations of the text as interpreted by OCR software, such as ABBYY FineReader.<sup>50</sup> This software interprets the scanned TIFF images to generate files of text. OCR text is the basis of keyword searching. When a user enters a keyword into a *Chronicling America*

search, the results are derived from the OCR files. OCR text is located, word by word, to areas on the image, so that a keyword can be highlighted on the page. The quality of image scans directly impacts the accuracy of OCR interpretation, and therefore has a direct effect on the usability of the resource. A common problem with OCR interpretation of poorly scanned pages is letter “blocking.” A newspaper page that has shortcomings in the scan may cause certain letters to “block.” This is similar to when too much ink is applied on a hand pressed page, and all of the white space inside the formation of letters is filled in. In poor scans, the characters on a page may be eye-readable, but the OCR interpreter may not be able to discern the difference between an “a,” “c,” “e,” or “o,” for example. A human reading a page that has a relatively poor scan of the words “Ethan Allen” may be able to easily tell the letters, but a computer may very easily read the words as “Ethon Allon,” or a similar variant. A researcher looking for Ethan Allen would not receive a “hit” for this occurrence of the search terms. This is a primary factor in why it is important to use master negatives when scanning newspapers: the closer one can get to the original, the higher fidelity one will achieve in scanned images.

LC recognizes that many non-NDNP digitization projects will not fully implement the NDNP specifications,<sup>51</sup> but it is fundamental to image quality to use master negatives for scanning. Some newspaper digitization projects, in both commercial and public sectors, do not always use master negatives, instead using at-hand positive service copies as the source material. Scanning from public-use positive reels yields mixed results. Such film often has scratch marks, tape repairs, and other kinds of damage from public use that decreases the quality of scans made from those reels. Any damage to the film will naturally be visible in the scans.

While most libraries do an admirable job of maintaining their microfilm collections for public use, service microfilm is naturally subjected to conditions that are unfavorable for preservation. A large percentage of film still in use is acetate.<sup>52</sup> Acetate is an early generation plastic used into the 1980s, when polyester-based film stock came into widespread use. Acetate microfilm is best stored in constant temperatures, ideally around 35 degrees Fahrenheit, with relative humidity below 30 percent.<sup>53</sup> Acetate film stored in ideal conditions from the onset can go decades with little or no degradation.<sup>54</sup> Polyester film is more resistant to unfavorable storage conditions, and its life is extended from proper storage as well. Properly stored polyester film has an estimated lifespan of more than 500 years.<sup>55</sup>

Ideal storage conditions are plainly impossible to achieve for film in public use areas. Increases in either temperature or relative humidity can dramatically shorten the life of acetate microfilm by bringing on vinegar syndrome. Fluctuations in these storage factors can also accelerate vinegar syndrome.<sup>56</sup> Vinegar syndrome is so named for the strong vinegar odor that the chemical breakdown of acetate film base produces as it ages. This smell intensifies as vinegar syndrome progresses. Beyond the smell, the separation of the film layer base with the gelatin is where the real problem lies. As vinegar syndrome advances and this separation intensifies, it can make film reels unreadable.

For a relatively small number of reels, it would make sense to simply duplicate these reels on modern polyester stock to create a duplicate master negative, but the scope of microfilming preservation efforts in the mid- to late twentieth century is staggering. Millions of pages of newspapers were microfilmed from the 1950s to the 1990s in the US. The vast majority of these film masters are acetate, but duplicating such large collections on

new polyester negatives is cost prohibitive. Today, there is little room in most library budgets for such an undertaking. Indeed, in the last ten to fifteen years many public microfilming operations have closed, and the master negatives dispersed, often ending up in the hands of commercial companies.<sup>57</sup>

Most of this twentieth century film is commonly referred to as “prestandards” film. That is, the film was produced according to local practices, and not with an overarching set of standards in mind.<sup>58</sup> Without such standards, the organization and quality of this film can vary widely. Nevertheless, the vast majority of such films can be used to successfully create images that meet NDNP specifications. Much of this prestandards film has not been stored under ideal conditions and has varying degrees of vinegar syndrome. This aging resource is ripe for digital preservation.

The Vermont Digital Newspaper Project Planning Group initially projected the use of master negatives produced under the USNP Preservation Microfilming Guidelines for the Vermont Newspaper Project.<sup>59</sup> These master negative microfilms are on polyester stock and stored in archival conditions, so there was an expectation that high-quality, easy to use film would be readily available.

As the VTDNP prepared to move forward with title selection, it became clear that the longest runs for the most historically significant titles were filmed in Vermont before the USNP effort on “prestandards” microfilm. Through the VTDNP partnership with the Vermont Department of Libraries (VTLIB), notably through Paul Donovan, Law and Documents Librarian at VTLIB, an agreement was arranged between the Vermont State Archives and Records Administration (VSARA) and the VTDNP, to permit borrowing master negatives on an as-needed basis for the duration of the project.

VTLIB stores the master negatives

generated by the Vermont Public Records Division microfilm program in the vault of the VSARA. Through this cooperative arrangement, Archivist Scott Reilly provided the VTDNP with a list, a 400-plus page document, which provides a reel-by-reel inventory of the master negatives held in the vault. Comparing this list of available masters with title lists generated during the VTDNP planning phase showed all of the titles that fit in the grant period of 1836–1922. Synthesizing the available reels with the title lists yielded an inventory of approximately 850,000 estimated pages of newspapers in significant runs. This abundance of film ensured that the VTDNP would have an ample supply of titles from which to select for the grant period.

### Execution

In January 2011, the project director issued an RFP to solicit bids for NDNP compliant digitization services, including OCR creation, and provision of related metadata. The RFP stipulated that vendors convert a sample test reel according to NDNP specifications. A sample reel of the *Vermont Farmer* was duplicated and distributed to vendors who expressed interest in offering proposals. The VTDNP received four competitive proposals, with complete digital conversions. The digital conversions enabled us to evaluate the quality of each vendor’s work by comparing sample output for page images, XML and OCR files side by side. Vendor selection criteria outlined in the RFP was used to evaluate proposals. Evaluative areas included: experience with similar digitization projects; quality and completeness of the proposal and sample conversion images and metadata; organizational capacity and commitment to serve a long-term, high-volume, deadline-driven digitization project; availability of skilled and reliable project personnel to carry out work and communicate progress; and,

of course, overall cost. After close evaluation of each bid, iArchives of London, Utah was selected as VTDNP’s digitization vendor.

iArchives is an experienced newspaper digitization partner, evidenced by their work with several other NDNP state projects. A list of iArchives NDNP partners is included on their website ([www.iarchives.com/demo-loc.shtml](http://www.iarchives.com/demo-loc.shtml)). Shortly after the digitization contract was signed, two representatives from iArchives came to Vermont for a series of meetings that established procedural workflows, production timetables, and scheduling for ongoing communication. Their familiarity with NDNP specifications enabled production to begin immediately. However, VTDNP personnel had to provide them with content to get underway.

NDNP metadata, that is, data about the newspaper pages, issues, and reels of film, must be generated from a thorough review of each reel of film. Since archival masters and printing masters should not be fed through a microfilm viewer, it is necessary to match a positive, public-use set to the masters for the purpose of review. “Thorough review” for the NDNP involves inspecting every frame on every reel to create an exact metadata record of what is on each reel. Due to the high volume of pages, or frames, to be scanned, thousands of pages per monthly batch of deliverables, this process is very labor-intensive and requires a great deal of lead time and planning.

The initial selection of three titles enabled the VTDNP to launch digitization workflows with iArchives with little delay. Master negatives were sent to a film duplication vendor, Archival Microfilming Services of Connecticut, to create duplicate negatives for digitization. Upon completion of the duplication, production began. The frame by frame review for the first 10,000 page batch was completed single-handedly by the project librarian and

shipped to iArchives on May 17, 2011.

It became quickly apparent that frame by frame inventory of film for a production schedule of 8,000 to 10,000 pages per month conducted by one person, as it was for the first batch, would not be sustainable. The grant proposal's staffing plan provided for a half-time, NEH-funded digital support specialist. While a search effort to fill that position was in progress, it had not been completed as production was starting. To bolster production start-up, the project director allocated three cataloging/metadata specialists (combined 0.3 FTE) from UVM's Resource Description and Analysis Services unit to contribute to frame by frame inspection and metadata collation. This support was vital to launching production. The hiring of the digital support specialist (0.5 FTE) in July 2011 brought the project to a full staffing complement, well-positioned to meet production goals of approximately 10,000 pages per month over the remaining year of the grant.

Nineteenth century newspapers, with all of their oddities, mixed with the sometimes idiosyncratic nature of prestandards microfilm, make for many diverse and unusual reels of microfilm. The frame by frame inventory of these reels demands attention to detail and requires a learning curve to understand what information is important to record, and what is of lesser importance. The decisions made during the inventory can profoundly affect the end user experience. In this process, information about each reel is recorded, including reel dates, titles on a reel and their corresponding LCCN (Library of Congress Control Number, or field 010 in catalog records), issue dates, supplements, pages in each issue, duplicate images or issues, the number of targets, incorrect dates, other printing errors, and volume and issue numbers for each issue. Many irregularities exist, such as undated supplements filmed in sequence with dated issues, temporary title changes,

issues of completely unrelated titles filmed with another title, partially filmed issues, or damaged issues. Each person completing the frame-by-frame film inventory must make case by case decisions about these frames on the film.

These inventory files are the basis of the XML files that are associated with the scanned newspaper images. To understand their function in file structure and information retrieval, it is helpful to think of XML files as pointing toward a given title, issue, or page in an issue. If mistakes are made in the frame by frame inventory, the XML files generated from that data will not point to the correct titles, issues, or pages. This frame by frame inventory has a critical role in the usefulness of the images created.

The VTDNP delivers raw metadata to iArchives along with a duplicate negative for each reel. iArchives scans the duplicate negative reels to generate TIFFs. From those TIFFs, they create a PDF and a JPEG2000 file. These PDF and JPEG2000 derivative files are compressed versions of the TIFF. Since they are compressed, they are much easier to deliver to online users. TIFFs are simply too large and resource-intensive to be delivered in a timely manner. The functionality of JPEG2000 is excellent. While they are compressed files, they are "lossless," and the clarity they deliver to end users is adequate.<sup>60</sup> The PDF files give end users a choice of formats in which to view or save an image.

iArchives also generates the XML data associated with the image files using the raw metadata generated by the VTDNP production team. The Metadata Object Description schema, or MODS, is used as the NDNP standard for descriptive metadata. iArchives converts the raw VTDNP metadata into MODS, inside the Metadata Encoding and Transmission Standard, or METS, wrapper. The newspaper metadata files and the images are hierarchically indexed in this highest

level file. Within each batch file, there are reel files. Within each reel file, there are issue files, and in each issue file, there are page files. For each of these levels, there are corresponding XML "pointer" files to the appropriate image or group of images.

iArchives also generates OCR files for each page. These OCR files correspond to zones on a page image, with a computer-generated set of words that match the page. The OCR files are the basis for keyword searches of the newspapers on *Chronicling America*. A keyword search of *Chronicling America* yields pages with the search term(s) highlighted. Terms are derived from the OCR files which match the terms to the page images and the location of the words on the pages. OCR output is continually improving with successive generations of software, but still has its limits. Nineteenth century, hand-pressed, small-type newspapers, sometimes with damage from a hundred-plus years of storage, filmed without clear standards on degrading acetate film, and then finally digitized, are often a challenge for human eyes, as well as for OCR. Still, OCR provides a powerful window into the vast newspaper collections on *Chronicling America*.

The project librarian ships reels, raw metadata, and a 1 TB drive in a protective case to iArchives. iArchives loads the images and data files on the drive in a single batch, and returns the drive to Vermont. The project librarian employs a quality-review process to ensure that NDNP specifications are met by checking for image quality, irregularities, and correct bibliographic data. Software tools used for quality assurance include the Oxygen XML editor, ThumbsPlus, and the NDNP's Digital Validation View (DVV). Each of these tools is described below.

The Oxygen XML editor ([www.oxygenxml.com/](http://www.oxygenxml.com/)) is used to check the XML files. Oxygen displays the XML files in a color coded, line by line format. The color-coding serves as a



Figure 1. Library of Congress Digital Viewer and Validator (DVV), Graphical User Interface (GUI).

visual marker for different elements within the XML files. When changes to the XML data are needed, this color-coding makes it significantly easier than with a black and white XML viewer.

ThumbsPlus imaging software ([www.cerious.com/thumbnails.shtml](http://www.cerious.com/thumbnails.shtml)) is used to examine the images. This software can process an entire batch and load all of the images on the batch into the viewer. Every TIFF and JPEG2000 image in the batch is examined as part of the VTDNP quality assurance. In the thumbs preview, one can easily scroll through issues and see each image. Typically, images in the preview are large enough to

spot problem pages, which can then be enlarged in the software for full sized viewing and closer inspection.

LC provides the most essential tool, the NDNF or DVV,<sup>61</sup> which is used to examine the OCR files and to check that the image headers are correct. The DVV orders the information in a manner that allows one to check OCR zoning and for correct bibliographic information. The critical function of the DVV is to validate and verify the digital signatures of the files in a batch. Validation checks to ensure that the batch data meets the structure of NDNF technical specifications, though not the specific information within that structure. Digital

signatures are then written into the batch. The presence of these digital signature files provides a ready way to quickly see that a batch has been validated. Verification is a similar process that checks the data structures, but it does not write files to the batch. Whenever data are moved, either virtually or physically, verification is used to ensure that no damage or errors occurred in transfer.

Following validation, the batch is shipped to LC on the 1 TB hard drive. LC verifies the batch and performs their quality control processes before ingesting the batch into *Chronicling America*. The process is not always straightforward. There are sometimes

problems with batches that require communication and coordination between the project librarian, the vendor, and LC so that adjustments can be made. Once the batch is ingested, LC notifies the VTDNP that its newspaper contents are available in Chronicling America.

### Outreach + Promotion = User Communities

By engaging partners and advisors from a variety of organizations, with representatives acting in distinct roles, the VTDNP simultaneously draws on input from stakeholders from a wide community and expands the connection that member organizations have to the project. Collaboration and cooperation foster responsibility, investment, and ownership by these organizations in the success of the VTDNP. Ongoing communication through the use of Basecamp project management software, an email discussion list, and regular in-person meetings or conference calls keeps partner interactions alive and vital, with involved groups of individuals in disparate locations throughout Vermont.

The project also includes a vigorous outreach and educational component. In addition to developing a project website (<http://library.uvm.edu/vtnp>), and implementing social media applications in the VTDNP blog (<http://vtdnp.wordpress.com>) and Facebook ([www.facebook.com/pages/Vermont-Digital-Newspaper-Project-VTDNP/169536416425273](http://www.facebook.com/pages/Vermont-Digital-Newspaper-Project-VTDNP/169536416425273)) project partners have given presentations and educational workshops at local, state, regional, and national venues.<sup>62</sup> The outreach component provides an ongoing avenue of communication between project partners, community members, and the general public. Besides raising awareness of the project and promoting the value and use of digitized newspapers on Chronicling America as primary resources for the

**Table 2.** Chronicling America Report from the Library of Congress: State-Specific Statistics for Vermont

Month	Total Newspaper Pages Served	Uvm Newspaper Pages Serves	Vermont Usage	
			Page Views	Visits
Jul 2011	855,701	2,254	4,926	662
Aug 2011	884,767	6,093	5,757	578
Sep 2011	1,010,576	8,024	11,074	939
Oct 2011	1,014,350	6,115	6,682	819
Nov 2011	1,086,050	12,453	8,996	746
Dec 2011	957,618	14,771	12,517	867
Jan 2012	1,186,596	15,047	15,973	1,050
Feb 2012	1,212,515	19,656	20,581	1,286
Mar 2012	1,241,476	21,707	22,928	1,323
Apr 2012	1,145,272	21,354	28,073	1,469
May 2012	881,118	12,553	17,522	1,047
Jun 2012	1,085,718	12,312	16,691	1,018
<b>TOTAL</b>	<b>12,561,757</b>	<b>152,339</b>	<b>171,720</b>	<b>11,804</b>

study of American history and culture, these outreach initiatives spark discussion and exchange of ideas that channel broader public perspectives and end user concerns into the VTDNP, which then feed into project development.

Public awareness of the VTDNP has grown parallel with the increase in Vermont titles available online at Chronicling America. Usage statistics from July 2011 to June 2012 show an increase in the number of visits and page views over the period. In July 2011, there were 9,217 pages of Vermont newspapers available, which received 662 online visits and 4,926 page views that month. The number of visits and page views show a relatively steady increase to April 2012, when there were 72,428 pages of Vermont newspapers available, which received 1,469 visits and 28,073 page views. These numbers tapered off slightly in May and June, part of a pattern that may reflect lower usage after the conclusion of the school year. The total for July 2011 to June 2012 for Vermont titles on Chronicling America is 11,804 visits, 171,720 page views, and 152,339 newspaper pages served.

As the volume of content continues to grow, it will be interesting to see how this is reflected in usage patterns.

### Outcomes

The VTDNP received NEH and LC authorization to produce nearly 130,000 unique pages of digital newspaper content during its first two-year grant. This 30 percent production increase was accorded without additional funding due to a no-cost two-month extension to the grant based on changes in the NEH calendar cycle, and the fact that our digitization vendor, iArchives, was delivering quality work at a significantly lower price (0.51/page) than was provided in our initial budget (0.95/page). Salary savings also factored in since the lengthy period that passed before hiring the project librarian (six months) and the digital support specialist (one year) gave us a slow start to production, but also permitted us to re-assess and streamline workflows to maximize production.

The decision to target monthly goals at 10,000 pages deviated from our

original work plan to produce smaller monthly batches spread over a longer period. With a project librarian, a half-time digital support specialist, and three cataloging/metadata specialists contributing a combined ten- to twelve hours per week to frame by frame inspection and metadata collation, production advanced at an efficient, high-volume rate. As the production team became familiar with the work, capacity for output was increased to levels that met, or exceeded, production goals each month.

With collaboration and encouragement from an engaged group of partner institutions, a committed Project Management Group, a dedicated production team, and knowledgeable advisers making informed selection decisions based on available titles and the historical value of those titles, the VTDNP produced just under 130,000 pages of digitized newspaper content for *Chronicling America* at the close of its Phase 1 grant in August 2012. The project has received positive community feedback while the fruits of its labor, Vermont historical newspapers on *Chronicling America*, have seen steady rise in use as content contributed by the VTDNP in its first year has grown.<sup>63</sup>

The University of Vermont was recently awarded an NEH continuation grant to re-embark with its state partners on Phase 2 of the VTDNP. In addition to digitizing newspapers in geographical areas of Vermont that we were not able to include in this first phase, Phase 2 proposes to digitize important historical newspapers in French and Italian, from Vermont's immigrant population communities, as well as some significant abolitionist and anti-slavery newspapers that were published in Vermont before the outbreak of the Civil War.<sup>64</sup> We look forward to a new round of title selections, to learning more about Vermont history and culture through our strong, historical newspaper collections, and to delivering more great content to the

LC for *Chronicling America*. The lessons learned from Phase 1 will surely inform our progress and outcomes for Phase 2.

### References and Notes

1. Bruce Cole, "The National Digital Newspaper Program, from the Chairman of the NEH," *OAH Newsletter* (May 2004): 11, accessed July 2, 2012, [www.oah.org/pubs/nl/2004may/cole.html](http://www.oah.org/pubs/nl/2004may/cole.html). For additional details, see the National Digital Newspaper Program webpages at the National Endowment for the Humanities (NEH), [www.neh.gov/divisions/preservation/national-digital-newspaper-program](http://www.neh.gov/divisions/preservation/national-digital-newspaper-program), and at the Library of Congress, [www.loc.gov/ndnp](http://www.loc.gov/ndnp).
2. Cole, "The National Digital Newspaper Program," 11.
3. Ibid.
4. Award Recipients—National Digital Newspaper Program, accessed May 17, 2012, [www.loc.gov/ndnp/awards](http://www.loc.gov/ndnp/awards).
5. "National Digital Newspaper Program," National Endowment for the Humanities, accessed September 3, 2012, [www.neh.gov/grants/preservation/national-digital-newspaper-program](http://www.neh.gov/grants/preservation/national-digital-newspaper-program).
6. The VTDNP was originally awarded funding from NEH for a two-year grant cycle, from July 1, 2010 to June 30, 2012. Due to changes in the NEH calendar, the project was notified in February 2011 of a no-cost, two-month extension for the project through August 31, 2012.
7. Cole, "The National Digital Newspaper Program," 11.
8. National Endowment for the Humanities, "U.S. Newspaper Program," accessed August 4, 2012, [www.neh.gov/us-newspaper-program](http://www.neh.gov/us-newspaper-program), and Library of Congress, "United States Newspaper Program," accessed August 6, 2012, [www.loc.gov/rr/news/usnp/usnpp.html](http://www.loc.gov/rr/news/usnp/usnpp.html).
9. "U.S. Newspaper Program," National Endowment for the Humanities, accessed Aug. 4, 2012, [www.neh.gov/us-newspaper-program](http://www.neh.gov/us-newspaper-program).
10. John J. Connell Jr., "USNP Bibliography: Newspaper and Periodical Articles, 1983–1998," accessed August 6, 2012, [www.loc.gov/rr/news/usnp/newsbibliop.html](http://www.loc.gov/rr/news/usnp/newsbibliop.html).
11. Robert Harriman, "Coordination of Cataloging Practices in the United States Newspaper Program," *Cataloging & Classification Quarterly* 6, no. 4 (Summer 1986): 15–29.
12. Robert Harriman, "The News in Review: The United States Newspaper Program," *Cataloging & Classification Quarterly* 17, nos. 3/4 (1993): 93–94.
13. Ibid., 95.
14. Ibid., 95–96.
15. Harriman, "Coordination of Cataloging Practices," 27.
16. Mark Sweeney, "The National Digital Newspaper Program: Building on a Firm Foundation," *Serials Review* 33, no. 3 (September 2007): 188.
17. Ibid., 188–89.
18. "About the Utah Digital Newspapers Program," accessed August 21, 2012, <http://digitalnewspapers.org/about/>.
19. Kenning Arlitsch, L. Yapp, and Karen Edge, "The Utah Digital Newspapers Project," *D-Lib Magazine* 9, no. 3 (Mar. 2003), accessed August 21, 2012, [www.dlib.org/dlib/march03/arlitsch/03arlitsch.html](http://www.dlib.org/dlib/march03/arlitsch/03arlitsch.html).
20. Kenning Arlitsch and John Herbert, "Microfilm, Paper, and OCR: Issues in Newspaper Digitization: The Utah Digital Newspapers Program," *Microform & Imaging Review* 33, no. 2 (June 2004): 59–67, accessed August 21, 2012, <http://content.lib.utah.edu/cdm/ref/collection/uspace/id/714>.
21. John Herbert and Karen Estlund, "Bringing the Past to the Present," *Online Magazine* 32, no. 4 (July/August 2008): 32–27, accessed August 26, 2012, <http://hdl.handle.net/1794/9914>, and John Herbert and Karen Estlund, "Creating Citizen Historians," *Western Historical Quarterly* 39, no. 3 (Aug. 2008): 333–41, accessed August 26, 2012, <http://>

- hdl.handle.net/1794/9915.
22. “First Grants Awarded in National Digital Newspaper Program,” National Endowment for the Humanities, March 28, 2005, accessed August 25, 2012, [www.neh.gov/news/press-release/2005-03-28](http://www.neh.gov/news/press-release/2005-03-28).
  23. Herbert and Estlund, “Creating Citizen Historians,” 336.
  24. Hartmut Walravens ed., *Newspapers of the World Online: U.S. and International Perspectives: Proceedings of Conferences in Salt Lake City and Seoul, 2006*, IFLA Publications 122 (Munich: Sauer, 2006). The table of contents can be viewed at [www.ifla.org/files/hq/publications/122.pdf](http://www.ifla.org/files/hq/publications/122.pdf).
  25. Helen Aguëra et al., “The U.S. National Digital Newspaper Program—Thinking Ahead, Designing Now,” in *Newspapers of the World Online: U.S. and International Perspectives: Proceedings of Conferences in Salt Lake City and Seoul, 2006*, ed. Harmut Walravens. IFLA Publications 122 (München: Sauer, 2006): 79–95. Presentation slides available at [www.loc.gov/ndnp/guidelines/docs/IFLA\\_NEHLC.pdf](http://www.loc.gov/ndnp/guidelines/docs/IFLA_NEHLC.pdf).
  26. Kenning Arlitsch, “Utah Digital Newspapers and Mountain West Digital Library”; John Herbert, “Utah Digital Newspapers”; Andrea Vanek, “The California Newspaper Project”; and J. Wendel Cox, “The U.S. National Digital Newspaper Program: University of Kentucky Libraries,” in *Newspapers of the World Online: U.S. and International Perspectives: Proceedings of Conferences in Salt Lake City and Seoul, 2006*, ed. Harmut Walravens, IFLA Publications 122 (München: Sauer, 2006): 97–134.
  27. “Publications and Presentations—National Digital Newspaper Program,” Library of Congress, accessed August 27, 2012, [www.loc.gov/ndnp/guidelines/pubs.html](http://www.loc.gov/ndnp/guidelines/pubs.html).
  28. “NDNP Extras,” Library of Congress, accessed September 30, 2012, [www.loc.gov/ndnp/extras](http://www.loc.gov/ndnp/extras).
  29. Kopana Terry, “The Digitization of Historic Newspapers on Microfilm: The Kentucky Experience,” *Microform & Imaging Review* 38, no. 2 (Spring 2009): 54–63.
  30. “Welcome to Meta | Morphosis: Microfilm-to-Digital Lecture Series,” University of Kentucky Libraries, accessed November 26, 2012, [www.uky.edu/Libraries/NDNP/metamorphosis/index.html](http://www.uky.edu/Libraries/NDNP/metamorphosis/index.html).
  31. *University of Kentucky Libraries Preservation & Digital Programs, NDNP: The Kentucky Edition* [video], accessed August 29, 2012, [www.youtube.com/watch?v=5WjJLEhC4KQ](http://www.youtube.com/watch?v=5WjJLEhC4KQ).
  32. Laura Robinson, “The Evolution of Newspaper Digitization at the Washington State Library,” *Microform & Imaging Review* 39, no. 1 (Mar. 2010): 24–27.
  33. Alyce L. Scott, “The Illinois Digital Newspaper Project,” *Microform & Imaging Review* 37, no. 4 (Dec. 2008): 163–69.
  34. “National Digital Newspaper Program: Guidelines & Resources,” Library of Congress, accessed August 28, 2012, [www.loc.gov/ndnp/guidelines](http://www.loc.gov/ndnp/guidelines).
  35. “National Digital Newspaper Program Technical Guidelines for Applicants,” Library of Congress, accessed August 28, 2012, [www.loc.gov/ndnp/guidelines/archive/NDNP\\_201214TechNotes.pdf](http://www.loc.gov/ndnp/guidelines/archive/NDNP_201214TechNotes.pdf).
  36. “About VSARA,” Vermont State Archives and Records Administration, accessed July 12, 2012, <http://vermont-archives.org/about>.
  37. *A Union List of Vermont Newspapers* (Burlington, VT: University of Vermont Libraries; Montpelier, VT: Vermont Department of Libraries, 2000).
  38. National Endowment for the Humanities, Division of Preservation and Access, *National Digital Newspaper Program*, 6–7, accessed November 27, 2012, [www.neh.gov/files/grants/ndnp-jan-17-2013.pdf](http://www.neh.gov/files/grants/ndnp-jan-17-2013.pdf). Application guidelines for 2010 and 2012 awards include similar instructions for involving an advisory board in the selection process.
  39. “National Digital Newspaper Program: Content Selection,” Library of Congress, accessed September 3, 2012, [www.loc.gov/ndnp/guidelines/selection.html](http://www.loc.gov/ndnp/guidelines/selection.html).
  40. “State Advisory Committee,” Vermont Digital Newspaper Project, accessed September 3, 2012, <http://library.uvm.edu/vtnp/advisors.html>.
  41. “Advisory Committee Briefing Book: July 1, 2010–August 31, 2012,” Vermont Digital Newspaper Project, accessed September 3, 2012, <http://library.uvm.edu/vtnp/VTDNP-BriefingBookRev20110426.pdf>; the briefing book was revised in the fall of 2012 as part of the VTDNP Phase 2 planning and preparation for title selection, accessed November 27, 2012, <http://library.uvm.edu/vtnp/VTDNP-BriefingBookRev20121121.pdf>.
  42. “Vermont History Timeline,” Vermont Division for Historic Preservation, accessed September 4, 2012, [www.historicsites.vermont.gov/html/timeline.html](http://www.historicsites.vermont.gov/html/timeline.html).
  43. “National Digital Newspaper Program: Content Selection,” Library of Congress, accessed September 3, 2012, [www.loc.gov/ndnp/guidelines/selection.html](http://www.loc.gov/ndnp/guidelines/selection.html).
  44. The Pennsylvania Digital Newspaper Project, [https://secureapps.libraries.psu.edu/content/PaNP/PaD-NP/about\\_padnp.html](https://secureapps.libraries.psu.edu/content/PaNP/PaD-NP/about_padnp.html); the VTDNP adapted a ranking form found on the PaSNP blog, accessed September 3, 2012, [www.personal.psu.edu/kkm111/blogs/pa\\_digital\\_newspaper\\_project/2008/11/scranton-titles-my-recommendation.html](http://www.personal.psu.edu/kkm111/blogs/pa_digital_newspaper_project/2008/11/scranton-titles-my-recommendation.html).
  45. Chronicling America, “About the Vermont Farmer,” accessed September 5, 2012, <http://chroniclingamerica.loc.gov/lccn/sn84023255>.
  46. Library of Congress, *National Digital Newspaper Program Technical Guidelines for Applicants*, 3, accessed July 10, 2012, [www.loc.gov/ndnp/guidelines/NDNP\\_201214TechNotes.pdf](http://www.loc.gov/ndnp/guidelines/NDNP_201214TechNotes.pdf).
  47. Ibid.

48. Ibid, 4.
49. Ibid, 3.
50. ABBYY FineReader OCR Software, accessed September 30, 2012, <http://finereader.abbyy.com>.
51. Library of Congress, *National Digital Newspaper Program Technical Guidelines for Applicants*, 3.
52. Tyler Selle, "Acetate Microfilm," *New York State Archives Publication* No. 15, revised (2003): 1, accessed July 12, 2012, [www.archives.nysed.gov/a/records/mr\\_pub15.pdf](http://www.archives.nysed.gov/a/records/mr_pub15.pdf).
53. James M. Reilly, "IPI Storage Guide for Acetate Film," Image Permanence Institute, 1993, accessed March 5, 2012, [www.imagepermanenceinstitute.org/webfm\\_send/299](http://www.imagepermanenceinstitute.org/webfm_send/299).
54. Selle, "Acetate Microfilm," 4.
55. Steve Dalton, "Microfilm and Microfiche," NEDCC Preservation Leaflet, 2007, accessed July 12, 2012, [www.nedcc.org/free-resources/preservation-leaflets/6.-reformatting/6.1-microfilm-and-microfiche](http://www.nedcc.org/free-resources/preservation-leaflets/6.-reformatting/6.1-microfilm-and-microfiche).
56. Reilly, "IPI Storage Guide," 4.
57. ProQuest holds many microfilm reels acquired in this manner, accessed July 12, 2012, [www.proquest.com/en-US/products/brands/pl\\_umi.shtml](http://www.proquest.com/en-US/products/brands/pl_umi.shtml).
58. See "USNP Preservation Microfilming Guidelines" for discussion of newspaper preservation microfilming standards, accessed September 30, 2012, [www.loc.gov/rr/news/usnp/usnpguidelinesp.html](http://www.loc.gov/rr/news/usnp/usnpguidelinesp.html).
59. Ibid.
60. For a discussion of JPEG2000 image quality, see Paolo Buonora and Franco Liberati, "A Format for Preservation of Images: A Study on JPEG 2000 File Robustness," *D-Lib Magazine* 14, no. 7/8 (July/Aug. 2008), accessed July 13, 2012, [www.dlib.org/dlib/july08/buonora/07buonora.html](http://www.dlib.org/dlib/july08/buonora/07buonora.html).
61. NDNP Technical Guidelines for Applicants, "Technical Validation of Digital Objects," 12, accessed September 5, 2012, [www.loc.gov/ndnp/guidelines/NDNP\\_201214TechNotes.pdf](http://www.loc.gov/ndnp/guidelines/NDNP_201214TechNotes.pdf).
62. "VTDNP Presentations and Tutorials," Vermont Digital Newspaper Project, accessed July 11, 2012, <http://library.uvm.edu/vtnp/presentations.html>.
63. "VTDNP in the News," Vermont Digital Newspaper Project, accessed September 7, 2012, <http://library.uvm.edu/vtnp/news.html>.
64. Vermont has the distinction of being among the first states to outlaw slavery in its state constitution in 1791. A strong abolitionist movement was in place well before the outbreak of the Civil War.

## Appendix. Vermont Digital Newspaper Project

### Phase 1. Title List (2010–2012)

Title	Publisher Location	Reel Count	Years Digitized
Vermont farmer (Newport, Vt.)	Newport, Vt.	2	870–1877
Rutland herald (Rutland, Vt. : 1823)	Rutland, Vt.	4	1836–47
Rutland County herald (Rutland, Vt. : 1848)	Rutland, Vt.	1	1848–50
Rutland herald (Rutland, Vt. : 1850)	Rutland, Vt.	1	1850–52
Rutland County herald (Rutland, Vt. : 1852)	Rutland, Vt.	1	1852–54
Burlington free press (Burlington, Vt. : 1827)	Burlington, Vt.	8	1836–65
Burlington weekly free press	Burlington, Vt.	30	1866–1920
The Rutland daily globe	Rutland, Vt.	4	1873–77
The Rutland weekly globe	Rutland, Vt.	1	1873–76
Windham County Democrat	Brattleboro, Vt.	1	1837–53
The Vermont transcript	St. Albans, Vt.	1	1864–68
Vermont daily transcript	St. Albans, Vt.	1	1868–69
Vermont phoenix (Brattleboro, Vt. : 1834)	Brattleboro, Vt.	3	1836–49
Vermont phoenix (Brattleboro, Vt. : 1855)	Brattleboro, Vt.	18	1855–22
State journal (Montpelier, Vt. : Oct. 1831)	Montpelier, Vt.	1	1836
Vermont watchman & State journal	Montpelier, Vt.	10	1836–83
Vermont watchman (Montpelier, Vt. : 1883)	Montpelier, Vt.	8	1883–1910
The People's press, and anti-masonic Democrat	Middlebury, Vt.	1	1837–38
The People's press, and Addison County Democrat	Middlebury, Vt.	0	1838–41
The Middlebury people's press	Middlebury, Vt.	0	1841–43
The Northern galaxy, and Middlebury people's press	Middlebury, Vt.	0	1843–44
The Northern galaxy	Middlebury, Vt.	1	1844–48
The Middlebury galaxy	Middlebury, Vt.	1	1848–50
Middlebury register (Middlebury, Vt. : 1850)	Middlebury, Vt.	5	1850–82
Middlebury register and Addison County journal (Middlebury, Vt. : 1883)	Middlebury, Vt.	1	1883–85
Middlebury register (Middlebury, Vt. : 1886)	Middlebury, Vt.	14	1886–1922
Spirit of the age (Woodstock, Vt. : 1840)	Woodstock, Vt.	2	1840–44
The Woodstock age	Woodstock, Vt.	1	1844–45
Spirit of the age (Woodstock, Vt. : 1845)	Woodstock, Vt.	7	1845–1913
Bennington banner (Bennington, Vt. : 1858)	Bennington, Vt.	3	1888–94
Bennington semi-weekly banner	Bennington, Vt.	0	1896, 1897, 1899
he Bennington daily banner	Bennington, Vt.	1	1877
Daily banner (Bennington, Vt.)	Bennington, Vt.	0	1891
Bennington banner and reformer	Bennington, Vt.	1	1903
The Bennington evening banner	Bennington, Vt.	8	1903–4
Caledonian (Saint Johnsbury, Vt.)	St. Johnsbury, Vt.	8	1837–67
St. Johnsbury caledonian (Saint Johnsbury, Vt. : 1867)	St. Johnsbury, Vt.	5	1867–84