2017 IFLA International News Media Conference
Collecting, Preserving, and Transforming the News – for Research and the Public
27-28 April 2017
Landsbókasafn Íslands – Háskólabókasafn (The National and University Library of Iceland)
Reykjavík, Iceland

Three Approaches to Newspaper Digitization at the University of North Carolina at Chapel Hill University Library

Barbara Ilie
North Carolina Collection, University Library, University of North Carolina at Chapel Hill, Chapel Hill, USA.
E-mail address: ilie@unc.edu

Todd Bowser
North Carolina Collection, University Library, University of North Carolina at Chapel Hill, Chapel Hill, USA.
E-mail address: tbowser@live.unc.edu

_____

Abstract:

Digitization has created a sea change in the way library patrons access and use archival and other historical material. Newspaper digitization is no exception and user demand is high for the ease of access and usability that digitized newspapers afford. To meet this need, the University Library at the University of North Carolina at Chapel Hill has approached newspaper digitization in three ways.

The North Carolina Historic Newspapers Project, a partnership between the University Library, the Library of Congress, the National Endowment for the Humanities, and the State Archives of North Carolina, places historic North Carolina newspapers on the Library of Congress' Chronicling America newspaper website. The North Carolina Digital Heritage Center, a statewide digitization and digital publishing program and partnership housed in the North Carolina Collection of the University Library, digitizes newspapers as a part of their program and makes them available on their award-winning DigitalNC site. Finally, a partnership between Newspapers.com and the University Library has made over 3 million pages of historic North Carolina newspapers available to its patrons, the State Archives of North Carolina and its Western Office, and the State Library of North Carolina.

This paper will discuss the successes, obstacles, and future work of each of these approaches to newspaper digitization. It will also explore the context in which this work was made

possible, recognizing the large and varied historical efforts within the State of North Carolina and the University Library to collect, preserve and make newspapers available.

Keywords: newspaper digitization, partnerships, collaboration, newspaper collecting, newspaper preservation

_____

INTRODUCTION
The North Carolina Digital Heritage Center, the North Carolina Historic Newspapers Project and a partnership with Newspapers.com have created three newspaper digitization streams at the University of North Carolina at Chapel Hill University Library. These projects have a few important things in common. One is that they rely heavily on previous work done in the North Carolina Newspaper Project and its participation in the United States Newspaper Program. Another is the pervading influence of the National Digital Newspaper Program on all newspaper digitization work at the University Library. And finally, each stream operates within a limited budget environment in which outside funding makes newspaper digitization possible. The beginning of this paper is a brief description of the North Carolina Newspaper Project. Following is some background on newspaper digitization in the State and at the University Library. The remainder of the paper describes the processes, lessons learned and future directions for each newspaper digitization stream.

NORTH CAROLINA NEWSPAPER PROJECT (1959-1998)
The North Carolina Newspaper Project (NCNP) was a cooperative effort between the State Library of North Carolina and the North Carolina Division of Archives and History. The planning phase included a statewide survey conducted by Duke University Library that sought previously unidentified newspaper titles throughout the state. Gaps in the catalog of newspaper titles had accumulated for decades. The Technical Services Section of the State Library collected catalog listings and physical holdings data that became part of a national database and WorldCat. The Division of Archives and History oversaw a concurrent preservation effort, creating microfilm copies of deteriorating historic newspapers. Selection of titles began with the materials thought most at risk – newspapers published prior to 1870[Wegner].

Newspapers were collected from public and private collections across the state and microfilmed at the State Archives of North Carolina. For many newspapers, complete runs of issues were not locally available. Continuity of titles occasionally required project staff to seek missing issues outside the state. Titles were sometimes completed with one or a few issues at a time, shot at the State Archives from rare copies and spliced into reels where gaps existed. In 1961, the State Archives was able to establish the Microfilm Services Center, to shoot and process all of its microfilm in-house. The documentation part of the project yielded the *Union List of North Carolina Newspapers,* as well as the first edition of *North Carolina Newspapers on Microfilm,* in 1962[Daniels]. That title continues today as the *Guide to Newspapers on Microfilm in the North Carolina State Archives,* which is a key reference text for the projects discussed below.

Well into the 1990s, project staff members continued to seek copies of historic newspapers for cataloging and filming. NCNP worked to add titles to OCLC, and to create local records for known North Carolina titles using the existing OCLC union list. From 1989-1990, Duke University Library performed project planning to support the NCNP's entry into the United States Newspaper Program. The project was absorbed into the United States Newspaper Program in 1992, receiving a three-year grant from the National Endowment for the

Humanities Office of Preservation. This support brought resources for a field survey of all 100 North Carolina counties and expansion of the collecting and cataloging effort, and for an exhaustive quality inspection of existing film holdings.

The North Carolina Newspaper Project completed its NEH grant-funded microfilming as part of the United States Newspaper Program in 1998, capping a forty-year collecting effort to preserve more than 3,000,000 pages of newspapers from hundreds of titles, spanning more than two centuries of state and local history[Wegner].

NORTH CAROLINA NEWSPAPER DIGITIZATION PROJECT (2007-2009)
The North Carolina Newspaper Digitization Project was an exploratory newspaper digitization project at the State Archives of North Carolina. It was funded by a Library Services and Technology Act (LSTA) grant. The project sought to digitize some of the state's earliest newspapers and make them available online. This was meant to serve three purposes that would inform future newspaper digitization efforts: to determine the amount of work involved in such a digitization effort, to establish best practices for outsourcing digitization work to vendors, and to create a series of K-12 lesson plans using the digitized materials. OCLC scanned the images from microfilm, and now the State Archives' CONTENTdm-based system hosts more than 23,000 pages on a site with basic search functionality. This program ceased in 2009 after the LSTA grant expired and the State Archives was unable to source additional funding.

A FOUNDATION FOR PROGRAMMATIC DIGITIZATION AT THE UNIVERSITY LIBRARY
The technological, administrative and human capital for programmatic digitization was built incrementally at the University Library. Digitization projects of varying scope and scale slowly established a base of equipment and knowledge that would benefit later projects. Beginning in 2006, there were plans underway to create a long-term digitization project in the Southern Historical Collection. These plans were implemented in 2007 with support from the Andrew W. Mellon Foundation through its funding of the Extending the Reach of Southern Sources Grant. The benefit of the technical resources gathered for those efforts (equipment as well as human expertise) carried over into numerous other projects at the University Library, including the projects discussed here.

NEWSPAPER DIGITIZATION AT THE UNIVERSITY LIBRARY – FINANCIAL ENVIRONMENT
The impact of the 2007-2009 economic recession on the University and the University Library has been significant and ongoing. In the fiscal year 2011-2012, the University's budget was cut by 17.9 percent. In the same year, the University Library sustained the largest of a series of cuts that continue today. For five of the last eight years, the University Library has experienced a net cut in its budget across three areas: salaries and wages, operations, and library materials. As a consequence, every aspect of work at the University Library is examined through the lenses of efficiency and cost cutting. This budget environment requires efforts to find outside funding and/or partner with other organizations. These arrangements have played an essential role in shaping newspaper digitization at the University Library.

AN EARLY EXPERIMENT IN NEWSPAPER DIGITIZATION AT THE UNIVERSITY LIBRARY – INTERNET ARCHIVE
The University Library worked with the Internet Archive in its earliest experimentation with newspaper digitization. In 2008 the Internet Archive attempted to demonstrate its capacity in

digitizing newspapers from microfilm by digitizing and making available two years of the University's student-run newspaper, the *Daily Tar Heel*. Budget constraints prevented this collaboration from continuing into additional titles, just as they prevented the University Library's purchase of equipment for in-house digitization from microfilm.

DIGITALNC – THE UNIVERSITY LIBRARY'S FIRST PROGRAMMATIC NEWSPAPER DIGITIZATION EFFORT (2010 - PRESENT)
About
The North Carolina Digital Heritage Center (NCDHC), which launched in 2010, made the University Library's first step in programmatic newspaper digitization. It works with cultural institutions throughout the state to make their cultural heritage materials available online through in-house and outsourced digitization and digital publication on its DigitalNC website. DigitalNC hosts a variety of material online including yearbooks, newspapers, scrapbooks, letters, records, other printed materials, and moving images. Newspapers are among their most popular materials. DigitalNC began newspaper digitization in 2011. The newspaper section of its website currently hosts 611,185 newspaper pages and continues to grow.

Staffing
DigitalNC currently employs three full-time staff, two graduate assistants, and one undergraduate assistant. Staff includes a program coordinator, digital projects librarian and digital projects programmer, as well as a director. The program coordinator sets priorities, defines services and promotes the program throughout the state. The digital project librarian oversees metadata creation, digitization, digital publication and provides quality control, working closely with the program's partners. The digital projects programmer provides computer programming support for all aspects of the program. The programmer also works in a group sponsored by the National Endowment for the Humanities and the Library of Congress to coordinate states' efforts in updating ChronAm, the open source delivery software created by the Library of Congress for the National Digital Newspaper Program.

Funding
The majority of DigitalNC's funding comes through Library Services and Technology Act grants. A small portion of DigitalNC's support comes from the University Library. The Library's support provides the space for the program and pays the program coordinator's salary. It also provides cataloging support and digital storage facilities and maintenance. Digitization of newspapers from microfilm is performed by vendors funded through the LSTA funding. The current funding for DigitalNC allows for the digitization of about 44,000 pages of microfilmed newspapers annually. This is based on a $15,000 annual request at $0.34 a page. Sometimes funds can be moved around within the program to digitize additional pages. Future funding is uncertain, as the elimination of the Institute of Museum and Library Services has been discussed in plans for the federal budget. While IMLS funding may be secure for the moment, there is a need for those involved in such projects to be prepared to respond to budget changes.

Scope
DigitalNC does not have a set date range of selection for newspaper digitization. Many early titles were digitized as part of the North Carolina Newspaper Digitization Project at the State Archives of North Carolina, which was finished in 2009. DigitalNC does not digitize newspapers that are freely available elsewhere. In its first year, DigitalNC's newspaper

collection included issues spanning the years 1824 through 1964. Current holdings date through 2016.

Selection

Newspapers are nominated by library staff and other partner cultural heritage institutions across North Carolina. The volume of nominated titles greatly exceeds digitization capacity. DigitalNC adheres to a predetermined set of selection criteria. The constraints of time and resources make it necessary to select a short list of newspapers from a large range of available papers. The selection process favors newspapers that document underrepresented regions and communities. Selection also favors public demand, and titles for which microfilm is available from the State Archives of North Carolina. Titles proposed by partners new to DigitalNC receive priority consideration. These partner groups must be willing to provide the materials for digitization if needed, and support the project in their outreach efforts. There may be a temporal or topical focus for annual selection cycles, which varies from year to year. For 2015 selections came mostly from the 20th century, specifically newspapers covering the Great Depression, World War I, World War II and the civil rights movement.

Source Materials

DigitalNC digitizes newspapers from microfilm at the State Archives of North Carolina, the University Library, and other institutions in North Carolina. A variety of other materials serve as alternatives when microfilm is unavailable. This may include originals provided from partners' collections, or PDF files created by publishers. Titles on paper are generally limited to student publications and a number of company newsletters and newspapers that have been produced across the state. The variety of accepted materials often presents a challenge, as partners or other organizations may submit microfilm or other materials in a condition less than ideal for digitization. Even in cases where microfilm is available there may be issues missing from a selected run. DigitalNC may then rely on cultural institutions or the University Library to find paper originals or other alternatives in order to maximize quality and continuity of the newspaper run.

Cataloging

Many of the titles in the growing DigitalNC collection require the creation of new MARC records. This comes up more frequently with some of the non-commercial publications that exist in a grey area with regard to classification, such as student papers and newsletters from organizations and firms across the state. A University Library cataloger creates and works on MARC records monthly, ensuring that records meet standards established by the Cooperative Online Serials Program (CONSER) of the Program of Cooperative Cataloging Program at the Library of Congress. The cataloger creates a record in WorldCat for the digital version of the newspaper. When DigitalNC made the move from ContentDM to ChronAm in 2014-2015, there was a significant amount of cataloging work to be done. Unlike ChronAm, ContentDM did not require that the newspapers have CONSER-compliant MARC records, so many records were created or edited to allow ingestion into the new system.

Copyright

DigitalNC digitizes newspapers that are in the public domain, out of copyright or titles for which permission has been obtained from the publisher. Many of the library and institutional partners with which it works have close ties to local publishers and have had success in securing copyright. A standardized form is provided to content owners by the partner that proposes each title, through which the owners may give permission and may specify the type of copyright statement that is applied. Legal counsel at the University approved of this

procedure to obtain permissions for use. Copyright statements are held for every digital record in the holdings, and the staff will honor takedown requests for contested materials.

Metadata
Metadata was formatted to National Digital Newspaper Program (NDNP) specifications from the beginning, with local modifications made as needed by DigitalNC. This metadata framework was compatible with ContentDM (the delivery software that DigitalNC previously used to serve newspaper pages) and is required by ChronAm, which is now used to deliver newspaper content.

Tools/Process
Technical equipment for in-house digitization includes a Zeutschel scanner, a Phase One overhead scanner for oversize or fragile newspapers, an Excel web app, an XML editor, the ABBYY Recognition Server software, and an application developed in-house called Paperboy. A local instance of the ChronAm software is used with modifications in response to issues such as security requirements at the University.

Digitization from newspaper originals is a part of the regular production queue at the University Library's Digital Production Center. The seed of page-level metadata creation occurs in Paperboy, which operates in tandem with the scanning process, allowing data entry through a simple interface. Student assistants fill out a set of fields (page number, volume, date, edition, issue) as each physical page is scanned. The ABBYY Recognition Server software generates OCR and XML files as it creates derivative PDF and JP2 files. The Paperboy application formats title, volume, edition, and issue numbers for easy delivery into a spreadsheet that will also include LCCN and publisher name (pulled from MARC records), and physical location information. A combination of manual editing and automated processes are involved in checking, formatting, and rechecking files and their structure. DigitalNC does not keep TIFF files, and performs a post-process to the METS files to make the JP2 the primary image format.

DigitalNC has also recently worked with two born-digital newspapers and the work proved challenging. Files arrived with disorganized directory structures and inconsistent image sizes, requiring additional work that was unanticipated. In 2016 DigitalNC successfully published a born-digital newspaper. Another title, in addition to file organization and image size issues, had inflexible rights restrictions, and work on it was tabled.

Vendors
Digitization of newspapers on microfilm is performed by a vendor. DigitalNC had an unsuccessful experience with its initial vendor. It then worked successfully with another vendor that had also done the digitization work for the North Carolina Historic Newspapers project. After this vendor decided to focus its business elsewhere and drop ChronAm clients nationwide, DigitalNC is now with its third vendor.

Storage
Images and metadata for DigitalNC's newspaper section are hosted from a local data storage array. A portion of the NCDHC's grant pays for backup storage on Amazon S3. A local server is used as temporary residence for files during in-house production. A decision was made to discard TIFFs for community papers based on storage costs and maintenance considerations. Student newspaper TIFFs resulting from on-site labor are kept on external hard drives and given to partner institutions for storage.

Content Delivery
Initially, newspapers were served up via ContentDM. The demand for better navigability as well as a concrete set of standards fueled the move to the NDNP's openly available ChronAm software. An in-house web server with a local instance of ChronAm now serves the content to users. When the papers were moved from ContentDM to ChronAm, there was a fair amount of data massaging. One challenge identified in this process is a feature of the ChronAm software that displays images only in black and white. This is not a problem for NDNP awardees producing digital pages of historic papers from microfilm, but for more recent titles digitized by DigitalNC or born-digital titles, the color content appears in black and white unless the user downloads the JP2 file. However, Center staff members are petitioning that the color library ChronAm utilizes be updated to accommodate color pages. To simplify the user experience, newspaper titles on DigitalNC's instance of ChronAm can be grouped together. For example, *The Polk County News, Polk County News,* and *Polk County News and the Tryon Bee* are consolidated into a single listing, *The Polk County News*. The ContentDM platform is still used to deliver DigitalNC's other (non-newspaper) digital content.

Data
DigitalNC is able to make bulk data available (for text mining, text analysis or other purposes) upon request, and anticipates access to be an easy process due to NDNP organizational standards. However, to date there have been no requests for bulk data.

Community
DigitalNC's collaborative approach to digitization has made them a guiding light for other programs in the region. They have been through some of the challenges that other programs now face, such as moving from one delivery system to another (from ContentDM to ChronAm), and building NDNP-compliant metadata in their in-house production. Individuals and groups seeking to create their own digital collections have turned to DigitalNC for guidance, including vendor recommendations or paths to finding project funding. DigitalNC promotes new content and shares details of their processes on their blog. The Paperboy software (under the name "newspaper-scanning-assistant"), workflow details and scripts, and additional technical guidance are all available via the NCDHC GitHub page.

Future
Though their institutional backing is secure for the moment, DigitalNC continues to explore alternative ways to sustain their collections. They will continue to face the challenge of a limited budget for newspaper digitization coupled with a high demand for the service. Efforts to improve the workflow for born-digital newspapers will be ongoing. This may require increased participation of publishers and partners who submit material for digitization. The strength of their partnerships may be their security in an uncertain future.

NORTH CAROLINA HISTORIC NEWSPAPERS – THE NATIONAL DIGITAL NEWSPAPER PROGRAM IN NORTH CAROLINA (2012 – PRESENT)

About
The National Digital Newspaper Program (NDNP), begun in 2006, is a collaborative program between the Library of Congress, the National Endowment for the Humanities, and U.S states and territories - 44 awardees to date. There are now upwards of 11,000,000 historic American newspapers pages on its Chronicling America newspaper site. The University Library

received a NDNP award in 2012, and is partnered with The State Archives of North Carolina. The award is now in its third cycle, with over 200,000 historic North Carolina newspaper pages digitized and appearing on Chronicling America so far.

Staffing

Staff includes a principal investigator, a project librarian, a University Library cataloger and graduate assistants. The principal investigator oversees all grant activities. The project librarian oversees the day-to-day work of the project, supervising one graduate assistant. A portion of a University Library cataloger's time is dedicated to the project. The project was initially funded for a full-time project librarian with two graduate student assistants, but was able to move the project librarian to 0.75 Full Time Equivalent and move from two graduate assistants to one after realizing workflow efficiencies.

Funding

The NDNP provides funding for the project librarian's salary, microfilm duplication, microfilm digitization through a vendor, outreach, and travel to the annual meeting of NDNP awardees. The University Library commits a portion of the time of the principal investigator and the cataloger to the project, and funds the graduate assistants. Additionally, the University Library provides infrastructure and the physical workspace.

Scope

The goal of the NDNP is to digitize newspapers of historic importance from all fifty U.S. states and all U.S. territories, and make them freely available online. North Carolina Historic Newspapers performs this effort for the State of North Carolina. Titles span the entire history of newspaper publishing in America from the first colonial titles in 1690 through 1963. By design, the NDNP builds upon the work of the United States Newspaper Program, which created the catalog records for and microfilm of many of the titles now made available on Chronicling America.

Selection

NDNP awardees select an advisory board – a cross-section of librarians, historians, archivists and media specialists – to propose and discuss titles in committee. The selection process strives toward a balance of information value – political, economic, cultural significance across an equitable geographic representation of the state, as outlined by the official NDNP selection guidelines. Technical analysis of the existing microfilm aims toward a standard of media quality (from the original text, to the quality of the microfilmed image, to the photographic reduction ratio) that allows for a high-quality digital version. Within each title, longer and continuous chronological runs are preferred over short runs and scattered titles, but significance of a title along with the other criteria may justify an effort for a title with limited availability. Titles are selected if only if they are not freely available elsewhere.

Source Materials

All microfilm is second-generation duplicate silver negative microfilm that is provided by a project partner, The State Archives of North Carolina. The project has also used a limited amount of positive film to fill in a run where negative copies were unavailable.

Cataloging

A University Library cataloger creates and updates MARC records as needed to ensure that they are compliant with CONSER standards, and adds them to WorldCat's CONSER database per project guidelines. The records are pulled into the local catalog, with a link that

goes directly to the digital content. The cataloger provides cataloging guidance throughout the process of selecting newspaper titles for digitization and the digitization phase. A cataloger at the Library of Congress also provides guidance.

Copyright
In the University Library's first two award cycles, NDNP guidelines called for the digitization of newspapers published on or before December 31, 1922, which are free of copyright by default. In 2016, the NDNP expanded its date range to include newspapers published through 1963 that are in the public domain in the United States because they were published without copyright notice or because the copyright was never registered or renewed. Works published in the U.S. during this time period entered the public domain if steps were not taken to renew ownership. There is now an evolving process employed by awardees to ascertain the copyright protections that might apply to titles selected in the new phase. Measures are now taken to determine ownership of titles and copyright, primarily through inspection of records from the United States Copyright Office. Project staff examine records from the range of years where copyright renewal for a given title would have taken place. The process of checking title copyright has added time to the title selection process. A second layer of embedded titles – syndicated material within the newspapers whose copyright belonged to a separate publisher – presents an additional challenge. Another is the lack of a comprehensive listing of the copyright records. There are scanned copies of the records for most years available via the Internet Archive, and others are available via Google Books. OCR allows text searching in most of these volumes with a fair degree of reliability. Inspecting copyright records for renewals in varying formats compounds the time and effort required for copyright research.

Metadata
North Carolina Historic Newspapers follows the well-defined and documented metadata standards of the NDNP. The NDNP provides guidance on metadata at the title, issue/edition and page levels. A certain number of metadata elements are mandatory. Optional elements are for use at the discretion of awardees, to meet any needs they may have locally. North Carolina Historic Newspapers uses a handful of optional metadata elements, including volume and issue number. For North Carolina Historic Newspapers metadata, the project provides the vendor with reel and title metadata and the vendor provides issue and page metadata.

Tools/Process
Reels and metadata are sent to the digitization vendor in batches. Batches are approximately 10,000 pages each. The digitized pages are then made available on a vendor-hosted instance of the ChronAm software for review. Review capabilities on this instance are hindered by limited bandwidth and system speed. This review mechanism will be fully tested when the project enters full production mode later this year. If inadequate, solutions will be sought with the vendor. After review on the vendor's instance of ChronAm, the batch is sent to the University Library for an in-house quality review process which includes the Library of Congress' Digital Viewer and Validator (DVV). Other tools used in the in-house quality review process include the command line and XML editing and image viewing software.

A previous vendor had a comprehensive in-house production environment that project staff accessed remotely. Project staff used the tool to perform robust quality review before batches were shipped to the University Library.

Once corrections within a batch are made in-house, the batch is sent to the Library of Congress for acceptance and ingestion into Chronicling America. Batches are submitted monthly to the Library of Congress.

Vendors
North Carolina Historic Newspapers had the same vendor for its first two phases. This vendor dropped it and other NDNP clients in 2016. There were some marked changes in the level of support the vendor was able to provide after it did some restructuring during this time. A new vendor has been selected for the current phase. It is experienced with NDNP awardees, and also works with DigitalNC.

Storage
The Library of Congress provides secure, long-term storage of the batches as a part of the NDNP. North Carolina Historic Newspapers also keeps the scanned newspaper pages and some of their associated files. Initially, the University Library planned on keeping the batches in their entirety, including the TIFF images, but it soon became apparent that the space the TIFF files required was in excess of the University Library's storage and maintenance capacity. The decision was made to strip the TIFF files from the batches to save space. The Library of Congress also provides long-term storage for the duplicate microfilm reels used in digitization.

Content Delivery
Newspaper pages are made available through the Library of Congress' Chronicling America newspaper website, which runs on the ChronAm software. North Carolina Historic Newspapers does not make digitized newspapers available locally.

Data
Bulk data is available for download from Chronicling America through its publicly available API. Numerous uses of this data are being discovered. In 2015, the National Endowment for the Humanities announced the Chronicling America: Historic American Newspapers Data Challenge, inviting research utilizing the project data at academic/professional and high school levels. They announced six prize winners in June 2016.

Community
One strength of the NDNP is the network that it creates. NDNP awardees routinely share information and ask questions on a listserv, and gather in person at the NDNP Annual Meeting. Members of this community learn from one another, and in turn, they share what they have learned with their home institutions and other projects. For example, in developing the North Carolina Digital Heritage Center's newspaper digitization program and before the Newspapers.com partnership was established, University Library staff looked to the NDNP for its insight on metadata and image standards. NDNP is regarded as a leader in the field of newspaper digitization.

Future
North Carolina Historic Newspapers activities will cease at the end of the current NDNP grant. However, the design of the national program is such that these digital assets will have a permanent home at the Library of Congress and be permanently and publicly available on Chronicling America.

NEWSPAPERS.COM PARTNERSHIP (2013-2015)

About
In 2012, Newspapers.com approached University Library staff and proposed a partnership in which the company would digitize the University Library's collection of North Carolina newspapers on microfilm. There was initial caution at the University Library. Staffing concerns and allowing microfilm to leave the University Library and the consequent disruption to patron use were some of the reasons for this reluctance. After further discussion, an agreement between the University and Newspapers.com was announced in 2013. The first papers were made available online on Newspapers.com in early 2014. By the end of the project in 2015 over 3 million pages were digitized.

Staffing
A portion of a full-time librarian's time was dedicated to directing the project, overseeing all aspects of its management. A part-time graduate student prepared spreadsheets of reel and title metadata, and shipped the reels to Newspapers.com.

Funding
Under the agreement, Newspapers.com fully funded the project. This included paying a portion of a staff member's salary and the full salary of a part-time graduate student. The University Library provided space to box the reels of microfilm, office space and use of office equipment. Other staff providing support to the project included the University Library's Scholarly Communications Officer and Research and Instructional Services staff. It was estimated that it would have cost the University Library more than $1,500,000 to digitize and make the pages available online, in addition to expenses for project management, staffing, digital storage and hosting.

Scope/Selection
Under the agreement, all microfilmed newspapers in the North Carolina Collection's reading room published through 1922 would be digitized. This included commercially produced microfilm. There were negotiations after the completion of the initial project to digitize material published after 1922, assuming that the parties could take appropriate care to address copyright issues, but an agreement was not reached. An exception was made for the *Daily Tar Heel*, the University's student newspaper, which dates back to 1893 and continues today. Permission for use in this case was a simpler process, and Newspapers.com now hosts a 100-year range of the *Daily Tar Heel*.

Source Materials
All digitization was done from reels in the North Carolina Collection reading room. This presented a concern in terms of valuation of the microfilm reels, were they to be lost or damaged in transit. The users of the North Carolina Collection reading room needed consideration as well, since their use of the space and the materials (microfilm reels temporarily absent from the collection) was disrupted. This impacted different users in different ways. Though a handful remained devoted to the familiar microfilm format, most were enthusiastic about accessing the newspapers on a new platform and the search capabilities of the new system.

Cataloging
There was no cataloging work to prepare for digitization as it did not figure into Newspapers.com's work model. Links to the digitized newspapers were later added into the

University Library's catalog records. Before this step, some library users did not know the newspapers were available in digitized form.

Copyright
There were no copyright considerations, as newspapers published on or before December 31, 1922 are in the public domain in the United States. The title range would finally be limited to those years in the public domain, because of copyright concerns and also because of Newspapers.com's business needs. The *Daily Tar Heel* is the one exception, because the rights issues for that title were straightforward. There are also contractual limits on what the University Library may do with the digitized files that resulted from this project. Issues of North Carolina newspapers from after 1923 are available on Newspapers.com through agreements made independently of the University Library.

Tools/Process/Metadata
Microfilm metadata was compiled and sent in a spreadsheet containing title, date, reel and LCCN information, where available. The metadata was designed to meet ChronAm interface requirements. The project's graduate assistant prepared the spreadsheets and boxed the reels for shipment. This process created a few challenges to the order and organization of microfilm in the reading room. Entire runs of titles left all at once from the reading room, but came back sporadically, sometimes causing confusion among staff and users. There was no in-house quality review process at the University Library, as the finished product was going on to the Newspapers.com website. All quality review was performed by Newspapers.com personnel.

Vendors
Newspapers.com performed all of the digitization work in this partnership.

Storage
Copies of the digitized newspapers and associated files were delivered to the University Library from Newspapers.com on a number of external hard drives. Due to the volume of data, the University Library approached Information Technology Services at the University, and they agreed to store it for the University Library.

Content Delivery
The newspapers began to be available on Newspapers.com in 2014 and are freely available to users at the University, the State Archives of North Carolina and its Western Office, and at the State Library of North Carolina. Outside of these locations, access to the newspapers requires a paid subscription to Newspapers.com.

Print and digital copies can be made for educational purposes and digital projects within the University community. These new forms of access were an immediate boon. Students who had previously needed to access historic newspaper content for a project would often arrive at the reading room seemingly all at once to obtain and view microfilm. The Newspapers.com partnership allows students to view these digitized pages from any place within the University's network and has eliminated this research bottleneck in the microfilm room. All visitors to the search room appreciate the ease with which they can search within the digitized newspapers.

While the digitized pages are a tremendous resource, there are some potential difficulties. There is no agreement in place for technical support for the University Library's and the

other institutions' access. And as with all outsourced hosting arrangements, questions of long term availability are ever present.

Data
The University Library has possession of the JP2, PDF and XML files of over 3.3 million newspaper pages. The University's Digital Innovation Lab and Odum Institute have pledged support to University members interested in using the data for textual analysis projects and historical data mining. There has been some use of the underlying data in instruction at the University. The agreement with Newspapers.com allows parties outside the University to use the files for textual analysis and data mining but limits other bulk use of the files. University Library staff are interested in measuring the impact of Newspapers.com use by patrons. At present, there is no way to access usage statistics from Newspapers.com's platform.

Community
In planning for the partnership, staff tapped into the existing NDNP community for technical specifications and metadata guidelines.

Future
As Newspapers.com's exclusive rights to this content gradually expire, the University Library plans to selectively republish titles as staffing and budget allow. There was hope to partner with other state entities to republish the newspapers in total, but the agreement as written precludes such partnerships. The agreement allows the University Library to sublicense use of the files, but only on a limited basis. When the agreement was undertaken, the University Library did not anticipate the advantages of such a partnership. Hosting this amount of digitized pages was not technically feasible for the University at the time. Discussions are underway to selectively republish the *Daily Tar Heel*. Beyond Newspapers.com's exclusive ownership period for the digitized files (three years from the publishing date for each file), the University may publish files on its own website and in its own databases, and make them available for non-commercial purposes. The last exclusive periods for the digital files will lapse on March 3, 2018. The University Library is currently exploring options for further dissemination in the state.

Conclusion
In this paper we described the process of newspaper digitization in three different streams at the University of North Carolina at Chapel Hill University Library. For each stream we presented details such as source material, copyright, storage, content delivery, and tools employed with the hope that our experiences could assist others in their own newspaper digitization efforts. The University Library is now faced with the perennial issues of access and preservation. We cannot pretend this work will live forever without a great deal of effort and ingenuity. Those who have put their time and effort into building these collections must now pool their knowledge and resources in order to sustain them.

References

Daniels, Dennis F. "The North Carolina Newspaper Project: Preserving the Printed Word." Archives & History Centennial Celebration. North Carolina Office of Archives & History. 2004 https://wayback.archive-it.org/194/20080815183405/http://www.ah.dcr.state.nc.us/centennial/features/newspaper.htm

Wegner, Ansley H. *History For All The People: 100 Years of Public History in North Carolina.* Raleigh: The North Carolina Office of Archives and History, 2003.