

Born Digital Newspaper Preservation Workflows for the Florida Digital Newspaper Library and the Caribbean Newspaper Digital Library

Lois J. Widmer
University of Florida
Digital Services & Shared Collections
George A. Smathers Libraries
1-352-273-2916
lwidmer@ufl.edu

Laurie N. Taylor
University of Florida
Digital Services & Shared Collections
George A. Smathers Libraries
1-352-273-2902
laurien@ufl.edu

Mark V. Sullivan
University of Florida
Digital Development & Web Services
George A. Smathers Libraries
1-352-273-2907
marsull@uflib.ufl.edu

ABSTRACT

Newspapers are rich information resources frequently requested by researchers and the general public for inclusion in digital libraries. Libraries have traditionally preserved newspapers in bound volumes and microfilm, and regularly digitize from those physical holdings. Indeed, the *National Digital Newspaper Program (NDNP)* is a partnership between the National Endowment for the Humanities and the Library of Congress create an online, searchable database of U.S. newspapers with digitized historic pages. Efforts by digital libraries to include newspapers have focused on the digitization of historic content. Many newspapers are now created digitally, being “born digital.” Libraries are actively investigating how to directly capture and preserve born digital newspaper issues instead of digitizing from print or microfilm.

This paper serves as a case study of how born digital ingest workflows can support newspaper preservation and online access with the *Florida Digital Newspaper Library (FDNL)* and *Caribbean Newspaper Digital Library (CNDL)* as examples. FDNL and CNDL were two of the first digital newspaper libraries to move to born digital ingest (UC Riverside, 2011; Zarndt, 2011). The paper addresses born digital workflows, selection and collection criteria, and digital preservation for contemporary newspapers. The paper explains new problems faced by large digital libraries with current newspapers and new solutions.

Categories and Subject Descriptors

H.3.7 [Information Storage and Retrieval]: Digital Libraries – collection, dissemination, standards, systems issues, user issues.

General Terms

Documentation, Standardization, Verification.

Keywords

Digital library, born digital, preservation, newspapers, metadata.

1. INTRODUCTION

The *Florida Digital Newspaper Library (FDNL)* and *Caribbean Newspaper Digital Library (CNDL)* programs evolved from the University of Florida’ (UF) long history of collecting Florida and Caribbean newspapers in print and microfilm for preservation and access. FDNL currently has over 1.28 million pages and CNDL has nearly half a million pages of current and historic newspapers available online and in a preservation repository. New issues are added online and to the repository on an ongoing basis with current newspapers both ingested from original born digital files and digitized from print editions when no digital version exists.

FDNL and CNDL began as digital library programs in 2005 when UF ceased microfilming and began digitizing newspapers. In 2005, UF’s microfilming program reported through the Digital Library Center (DLC) where digitization and born digital ingest was underway for many projects and programs with all resulting materials openly available online in the UF Digital Collections and digitally preserved in the Florida Digital Archive.

2. PUBLISHER PERMISSIONS

In order to convert from microfilming to digitization, UF requested permissions from the publishers for all of the newspapers that had been microfilmed to allow their newspapers to be digitized, shared freely and openly online, and stored and migrated to different formats as needed for long-term digital preservation. The majority of the newspapers granted permissions. Publishers not granting permissions were primarily those affiliated with larger corporations with microfilming sales programs or other preservation and access monetization programs in place. The initial workflow was that the newspapers continued to be received in print as they had been for microfilming and were instead digitized.

3. SELECTION AND COLLECTION DEVELOPMENT FOR PRESERVATION

The selection and collection development criteria for the newspaper digital libraries were shaped by the historic priorities focused on preservation using microfilm. The preservation priorities focused on small, rural newspapers in Florida and specific titles from across the Caribbean. The program sought to include at least one newspaper for every county in Florida and at least one newspaper for every country and territory in the Caribbean. The Caribbean has many colonial and other territories

that have changed over time. A simple representation of countries or colonial groups would be grossly insufficient to preserve the news and voices of the region. For instance, Martinique and Guadeloupe are overseas departments of France, Saint-Martin is an overseas collectivity of France, and in 2010 Curaçao and Sint Maarten joined Aruba as countries within the Kingdom of the Netherlands. The news for each area captures and shapes history. For example, the US Virgin Islands were part of the Danish West Indies before becoming a US territory in 1917. While the US Virgin Islands became a territory in 1917, citizenship was only granted in 1927 and was done largely thanks to the establishment of the free press by D. Hamilton Jackson and his political writings in *The Herald*, the free press newspaper. Additionally, the Florida and Caribbean digital library programs sought to include newspapers that spoke to and for otherwise unrepresented communities. For instance, several historical African-American newspapers are included. These newspapers began when newspaper publishing was segregated and so they contain news and represent a community voice that is not covered by other newspapers. For researchers, access to the full accounting of history from multiple perspectives, especially voices silenced or removed in other sources, is of critical importance and high research value.

With the selection and collection development criteria for preservation in place, other factors also dictated which newspapers could be included.

Newspapers included in FDNL and CNDL are limited to those where the publishers granted permissions for inclusion. The vast majority of newspapers from the prior microfilming programs granted permissions; however, several did not and so could not be included. The situation for local newspapers is dramatically different from that for large, major publishers. Where major newspapers are facing a dwindling market share, local papers are thriving. The FDNL and CNDL programs frequently see existing papers add new titles, for nearby regions or on specific topics in the same region, and those are often added to the digital libraries. Occasionally, local newspapers do cease publication. In those instances, a replacement title is sought to ensure continued preservation of the local newspapers for the affected area or community.

When FDNL and CNDL began, the top priority was to provide coverage for newspaper issues no longer being microfilmed. Thus, the digitization priorities are for all selected titles from 2005 through current.

4. INITIAL DIGITIZATION

Unlike microfilming, where the majority of the work required was in the image capture, digitizing newspapers required creating metadata for the dates for each issue, quality control for all scanned images, processing for optical character recognition, verification of online loading, and verification of processing into the digital archive for long-term digital preservation.

In 2005, UF was awarded a grant from the Florida Library Services and Technology Act (LSTA) program, for the proposal "Rewiring Florida's News: from Microfilm to Digital." The grant supported the purchase of two CopiBook scanners, which could scan a full, single broadsheet newspaper page in one capture. The CopiBooks and existing UF infrastructure enabled the shift from newspaper microfilming to digitization. The grant proposal explicitly stated the need for ongoing sustainability for a Florida

newspaper digitization program, with the grant funding a portion of the initial technical infrastructure. The grant proposal also stated that there was already a known need to digitize earlier years of the selected titles and to include additional newspaper titles. Indeed, from the time the grant was submitted to the submission of the *Mid-year Report*, the project's selected current titles grew with the addition of historic titles and years from 54 to 103. As noted in the *Final Report* submitted at the completion of the grant in 2006 only 27 current titles along with 171 historic titles had loaded online and been archived for digital preservation. The Caribbean newspaper digitization was also supported through the development of the infrastructure for the Florida newspapers as well as through collaborative grants submitted by UF and partners and awarded from the Department of Education's Technological Innovation and Cooperation for Foreign Information Access (TICFIA) grant program for the *Digital Library of the Caribbean* in 2006 and the *Caribbean Newspaper Digital Library* in 2009.

For ongoing program sustainability, UF sought grant and donor funding as well as collaborative partnerships with shared resource contributions for specific projects. Additionally, UF began a process of ongoing analysis to find and implement workflow efficiencies. The overall goal was to ensure that the total cost of operating the digital newspaper libraries, including all production workflows, would be less than the costs for operating the prior microfilming program. The full program sustainability requires that the costs be controllable and predictable, that production workflows could benefit from efficiencies to reduce costs further, and that the digital newspaper libraries would benefit from maximized return on investment in terms of patron benefits and reduction of costs for other areas (e.g.; interlibrary loan costs to send out newspapers on microfilm would be reduced and eventually removed with online access to the materials).

5. BORN DIGITAL WORKFLOWS

In 2008, UF sought to make the process more efficient through born digital ingest instead of digitization from print materials. UF contacted the publishers regarding the availability of born digital files. The vast majority of the publishers responded that they were creating issues as born digital files, although several still created the newspapers using paste up. UF requested that the publishers send the born digital files when available. In November 2008, UF began receiving and ingesting born digital files for the newspapers in addition to digitizing from print.

Establishing a born digital ingest was essential for UF to ensure that FDNL and CNDL could remain sustainable with ongoing resource limitations. Because preservation was a core priority, UF also needed to ensure the validity of the born digital files and any related files or concerns that would best enable long-term digital preservation.

To begin establishing the new, born digital workflow, UF queried the publishers regarding the types of born digital files they create, systems used in the creation of those files, and available metadata. In *Preserving News in the Digital Environment: Mapping the Newspaper Industry in Transition* released in 2011, the Center for Research Libraries explains the rich metadata created and managed by news organizations in the production of newspapers. Ideally, this metadata should be included in a born digital newspaper workflow for a digital library. Simple, print-ready files like PDFs do not generally include any of this rich metadata. In

discussion with the publishers, UF learned that the small publishers in FDNL and CNDL did not have the same types of systems in place as the large publishers and they were not creating the same sort of extensive metadata. In fact, the publishers created only PDFs or similar files as the master digital files. By 2008, UF's DLC had extensive experience in ingesting, processing, and preserving PDF and similar files through the larger UF Digital Collections and the Institutional Repository where all theses and dissertations were ingested and preserved from PDF files.

Because UF's DLC had established workflows for processing PDF and similar files, the DLC next needed to acquire the files from the publishers. Different publishers could support different modes of transfer: FTP, emailed for papers where the file sizes were small enough, mailed on external hard drives, mailed on DVDs, files placed on the publisher website for harvest, and files available for harvest through electronic edition subscriptions. Each of these modes needed to be supported by the workflow. Additionally, the existing workflow for digitization of print newspapers needed to continue in parallel to support the newspapers made in paste up where no digital version existed.

Because most of the publishers are small operations, staff and other changes affect all aspects of their operations and file transfers could be forgotten or delayed. All of the new born digital workflows required that the DLC establish a schedule to ensure that DLC staff was aware if files were not received on a timely basis in order to then contact publishers. The new workflow for FTP transfers from publishers required that the DLC set up FTP accounts for the publishers and establish a schedule to check FTP folders and ingest all transferred files. Frequent communication was required to establish the schedule with publishers and support their use of FTP. The workflow for files emailed from the publisher is straightforward; however, only a few publishers have files small enough to email so it is a less used workflow. For the workflow, publishers email the files to the digital newspaper library coordinator who ingests the files.

Postal mail supports the workflows for mailing files on external hard drives and on DVDs. The DLC has many external hard drives specifically available for transferring files from partners in shared digital library programs. The workflow for transferring files using the external hard drives is that the DLC packages one of the available hard drives and mails it to the partner or newspaper publisher. The drive is logged in an internal tracking system with the hardware ID, which then references the drive information, the recipient, and the date sent. Publishers then load files to the drive and return it to the DLC. The DLC connects the returned hard drive to a quarantined machine, to avoid any potential viruses that may have been accidentally loaded, verifies that the drive is clean and copies the files into the ingest queue. After all files are moved, the drive is erased and returned to the available pool for use. The workflow for sending DVDs is a simpler version, with publishers saving files to the DVDs, which are then mailed to the DLC. The DLC copies the files into the ingest queue and discards the DVDs, unless the publisher has asked for them to be returned in which case the DLC mails the DVDs to the publisher. Initially, both of these methods were heavily used. This was in part because many publishers had born digital files from before November 2008. In cases where those issues had not already been digitized from print, the publishers transferred the born digital files to hasten the processing, loading, and archiving of their newspaper issues.

Web harvest workflows are in place to harvest files directly from publisher websites or, in some cases, to harvest files from publisher files through electronic edition subscription websites. A number of publishers load the files for new issues directly to their websites. The publishers do this on a regular basis, most often replacing the prior loaded files with the files for the newest issue and with no more than the files for a single issue or several of the most recent issues. The DLC schedule is critical to keep up with the harvest process for these newspapers to ensure that the files are harvested before they are removed. A number of publishers load their files to an electronic edition subscription service, sometimes related to their printer and print distribution. Some of the publishers have given gift subscriptions to the DLC, while other electronic edition subscriptions are funded through the UF Libraries acquisitions budget and endowments as allocated by the collection managers. With the access information from the gift or paid subscription, the DLC logs into the electronic editions and harvests the files according to the established schedule. The web harvests are the most reliable workflow because they are already part of the normal workflows for the publishers and so an extra step and extra work is not needed, unlike with FTP, email, and mail transfers. Because this is the most reliable, it is the best workflow for the DLC. As the best workflow for both publishers and the DLC, based on current experience, when a new publisher begins to support electronic edition subscriptions, the DLC works with the publisher to convert existing workflows to web harvest.

Creating and supporting the additional workflows represented an increased workload. However, this increased workload was simultaneous with a reduced workload for inventory control, imaging, and image correction of printed newspaper issues. The overall result is a reduction in workload. Initially, this reduction was not as dramatic because the DLC was still receiving print copies of the newspapers. The DLC was inventorying and holding the print issues until after the born digital files were received and ingested. If the born digital files were not received, the DLC would process from the printed issues. In 2011, this workflow was changed to ensure the print issues are only being received for the newspapers that are being digitized from print and missing born digital files are requested from publishers with no attempts to locate print copies for backfilling.

6. WORKFLOW ANALYSIS AND EFFICIENCIES

In searching for other opportunities to optimize the workflow, the process for enhancing newspaper metadata was examined and amended. In 2005 when the digital newspaper libraries began, the initial newspaper workflow followed the workflows already established for books. The workflow for books included creating table of contents information for all chapters, sections, and the like. The workflow for newspapers initially included creating a table of contents for the newspaper sections: main, sports, local, lifestyle, etc. After usability testing showed that most users were not using the table of contents, or were using it only for books, the newspaper workflow was examined. In the course of that examination, the DLC noted that many of the newspapers are brief, being under 30 pages, and all are full text searchable across the entirety of the digital newspaper library or within each issue. Given that users already had support through full text searching and a variety of page image views, and users were not using the table of contents for the newspapers, the workflow was amended to end labeling of the newspaper sections. This change reduced

the overall workload and hastened the time from initial ingest to loading online and processing into the digital archive.

With the efficiencies from born digital ingest and metadata enhancement workflow changes, one of the more demanding newspaper workflow components for both print digitization and born digital ingest was the handling of syndicated content. From 2005- March 2011, UF reviewed all newspapers for FDNL and CNDL and applied a “blur” to redact syndicated content and added a note above the blurred image that syndicated content had been removed. This process was accomplished through an automated image actions using Adobe Photoshop, with the actions designed by the Operations Manager who is an expert Adobe Photoshop user and so the work did not require additional programming expertise, where student workers conducted page-by-page reviews, selected the syndicated content area, and clicked to apply the action, which would simultaneously apply the blur and text notation. Then, the accuracy of the blurred content was reviewed within the quality control process to ensure all syndicated content had been blurred and that non-syndicated content had not been blurred. While the redaction or blurring process was relatively straightforward, the workload required student workers. Student workers rotate frequently, increasing the time needed for training and the time needed for training with frequent errors with new students.

In reviewing the process, the removal of syndicated content was found to be based on prior risk management concerns and not legal requirements. In March 2011, after review of best practices for other digital newspaper programs and discussion with the Association of Research Libraries, UF ceased the process of redacting syndicated content.

7. DIGITAL PRESERVATION

In order to support the serial hierarchy needs for newspapers within an integrated and cross-searchable digital library/asset management system, UF developed the SobekCM software. SobekCM is an integrated digital library/asset management system that supports the online user access capabilities as well as internal workflow supports. UF developed SobekCM for a variety of needs including integrated workflow management for multi-institutional collaborative projects and projects with both digitization and born digital workflows. SobekCM was developed with digital preservation as a primary concern and so it manages the local archives, where all files online are served through redundant servers which are also backed up to tape and stored in redundant offsite locations and where all files are also preserved in the Florida Digital Archive (FDA).

The FDA technical design, procedures, and policies are based on OAIS, the Open Archival Information System Reference Model (ISO 14721:2003) and on ongoing work to define and certify trusted digital repositories. For every file in each digital object (as specified in the archival information package, or AIP, created for each SIP), two master copies are written and stored on active hard drives. One copy is stored at a data center in Gainesville, Florida and one copy is stored at a data center in Tallahassee, Florida. These data centers are under the control of the State of Florida and are not private or separate institutions. The two master copies are treated as a single file by the repository software application underlying the FDA, which is named DAITSS for Dark Archive in the Sunshine State. Because the two master copies are treated as a single file by DAITSS, when any action is performed on a

file, it must be successfully performed on both master copies to be considered complete. In addition to the two master copies, traditional backup copies on tape are maintained both in Gainesville, Tallahassee, and Atlanta. While the proximity of these disaster sites is not ideal at this time, alternate or additional sites are in planning outside of the southeastern region of the US.

SobekCM tracks and maintains information about the digital preservation and archival processing for all digital objects. This information is displayed within the SobekCM online system under "Work History". The SobekCM "Work History" tracking includes the "History" which lists the workflow name (for the name of the archive and the process; e.g.; FDA ingest), date the workflow occurred, and location/notes (e.g.; the FDA IEID). Also under "Work History" is a field entitled "Archives" which lists all of the archived files including: filename, file size, last write date, and archived date.

SobekCM also includes tools for preparing files directly for submission to FDA, with or without loading to an online system. These functions are supported by the SobekCM METS Editor, which is in use by State University Libraries in Florida for preparation and submission of materials to FDA (FDA, 2011). The FDA preparation process creates the Submission Ingest Package (SIP) file with the metadata and in the format for submission to FDA, including: MD5 checksum numbers, file format and version information, and administrative and bibliographic metadata.

7. NEW PROBLEMS: EVERYTHING THAT'S FIT TO PRINT

In November 2010, major improvements were made for search engine optimization resulting in all materials in FDNL and CNDL being well crawled and indexed by major search engines. The ease of findability resulted in greatly increased overall usage as well as a number of patron requests to remove or suppress news stories of arrests, foreclosures, and graduations that appear when they conducted online searches for their names. The patrons were concerned because a simple web search with their names returned these stories first or on the first page of results from searches using major search engines. With the Florida housing market being particularly impacted by the financial downturn, UF received requests from Florida citizens requesting that stories of their home foreclosures be hidden from searches lest they impact employment opportunities. UF received a flurry of these requests immediately after the search engine optimization.

In *The longtail of news: To unpublish or not to unpublish*, Kathy English explains the new phenomenon resulting from online news archives and the request to remove content from the archives, with the removal requests resulting in a status of “unpublishing” news (2009). Similarly, newspaper archives in libraries have also faced requests to unpublish content as with the lawsuit, which was dismissed, wherein a Cornell Alumnus sued to remove a story of his arrest from the library archives of *The Cornell Daily Sun* newspaper (Stratford, 2009). Unpublishing as the actual removal of content from an archive is counter to the mission of archives and to both FDNL and CNDL. However, some support needed to be in place so that news stories in newspapers in FDNL and CNDL could not be found through commercial web searches which present the stories in a decontextualized manner as though they exist without the benefit of subsequent stories and context, which could cause negative impacts for various individuals. In

searching for guidance, UF located the *Oakland Archive Policy: Recommendations for managing removal requests and preserving archival integrity (of electronic documents)*. The *Oakland Archive Policy* seeks to protect archives and archival integrity while also supporting a productive method for responding to removal requests. In using the *Oakland Archive Policy* as a model, UF developed a procedure for handling removal requests. The procedure is that when a request to withdraw a news story is received, UF adds the newspaper issue with that story to be listed as “disallow” in the robots.txt directive, which issues the commands for search engine robots. As external search engines re-crawl and re-index the site, the newspaper issue and all stories in that issue cease to be included or shown in the search engine indexing and search results. This takes variable amounts of time based on the operation of the search engine robots. To hasten the process, UF also uses the Google Webmaster tools to request immediate removal of the link. The “disallow” using robots.txt and the removal request using Google’s webmaster tools is a temporary procedure by the UF Libraries when any requests are received to suppress items from indexing by external search engines. This temporary procedure is applied for all requests. This procedure is temporarily in place while the UF Libraries develop official policies and procedures. Once the new policies and procedures are in place, UF’s DLC will use its list of accommodated removal requests to notify the affected parties of any consequences resulting from the modifications to the policies and procedures.

8. CONCLUSION

As shown through the recent NEH award for the *Chronicles in Preservation Project* and the Donald W. Reynolds Journalism Institute’s *Newspaper Archive Summit* and subsequent whitepaper (2011), newspaper preservation is a critical concern at this time. This paper provides an overview of two newspaper digital libraries that leveraged existing infrastructure for selection, collection, and preservation from a microfilming program to establish the robust infrastructure needed for newspaper digitization for access and preservation. The robust infrastructure was created through permissions agreements with publishers, digitization workflows for analog materials, born digital ingest workflows for digital materials, and constant workflow re-evaluation for sustainable processing and for responding to new problems in an age of concerns regarding unpublishing and archives.

9. ACKNOWLEDGMENTS

Our thanks to the UF Digital Collections and Digital Library of the Caribbean for sharing and providing full, free, and open access to the grant proposals, grant reports, and other documentation referenced throughout this document.

10. REFERENCES

- [1] Center for Research Libraries. 2011. *Preserving News in the Digital Environment: Mapping the Newspaper Industry in Transition - A Report from the Center for Research Libraries, April 27, 2011*. Chicago, IL. http://www.crl.edu/sites/default/files/attachments/pages/LCreport_final.pdf.
- [2] Digital Library of the Caribbean. 2011. *Caribbean Newspaper Digital Library (CNDL)*. <http://dloc.com/cndl>.
- [3] Donald W. Reynolds Journalism Institute. 2011. *Newspaper Archive Summit: Rescuing orphaned and digital content*. University of Missouri. April 11-12, 2011. <http://www.rjionline.org/events/newspaper-archive-summit>
- [4] Educopia Institute (host for the MetaArchive Cooperative); San Diego Supercomputer Center; and the libraries of University of North Texas, Penn State, Virginia Tech, University of Utah, Georgia Tech, Boston College, and Clemson University. 2011. *Chronicles in Preservation Project Wiki - NEH Grant Funded Project*. http://metaarchive.org/neh/index.php/Main_Page.
- [5] English, K. 2009. *The longtail of news: To unpublish or not to unpublish*. Online Journalism Credibility Projects, Associate Press Media Editors. http://www.apme.com/resource/resmgr/online_journalism_credibility/long_tail_report.pdf.
- [6] Florida Center for Library Automation (FCLA). 2011. *Florida Digital Archive*. FCLA. Gainesville, FL. <http://fclaweb.fcla.edu/FDA>.
- [7] Florida Digital Archive (FDA). 2011. “METS Editor client for creating FDA packages.” FCLA. Gainesville, FL. <http://fclaweb.fcla.edu/content/mets-editor-client-creating-fda-packages>.
- [8] Florida International University. 2006. *Digital Library of the Caribbean (dLOC) Technological Innovation and Cooperation for Foreign Information Access (TICFIA), US Department of Education Grant Proposal*. <http://dloc.com/FI04071904/>.
- [9] Florida International University. 2009. *Caribbean Newspaper Digital Library (CNDL) Technological Innovation and Cooperation for Foreign Information Access (TICFIA), US Department of Education Grant Proposal*. <http://dloc.com/UF00091464/>.
- [10] McCargar, Victoria. 2011. *A mandate to preserve: Assessing the Inaugural Newspaper Archive Summit*. Donald W. Reynolds Journalism Institute; University of Missouri. <http://rjionline.org/sites/default/files/archives-whitepaper1.pdf>.
- [11] National Endowment for the Humanities and the Library of Congress. 2011. *National Digital Newspaper Program (NDNP)*. <http://www.loc.gov/ndnp/>.
- [12] School of Information Management and Systems, U.C. Berkeley. 2002. *Oakland Archive Policy: Recommendations for managing removal requests and preserving archival integrity (of electronic documents)*. <http://www2.sims.berkeley.edu/research/conferences/aps/removal-policy.html>.
- [13] Stratford, M. 2009. Judge Dismisses Libel Suit Against Cornell. January 23, 2009. *Cornell Daily Sun*. Ithaca, NY. <http://cornellsun.com/section/news/content/2009/01/23/judge-dismisses-libel-suit-against-cornell>.
- [14] University of California Riverside. 2011. California Weekly Newspapers to be Preserved Online. June 21, 2011. *UC Riverside Newsroom*. Riverside, CA. <http://newsroom.ucr.edu/2667>.
- [15] University of Florida. 2011. *Florida Digital Newspaper Library (FDNL)*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/newspapers>.

- [16] University of Florida. 2006. *Rewiring Florida's News: from Microfilm to Digital – LSTA Grant Final Report for Federal Fiscal Year 2005 Projects*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/UF00082123/00003>.
- [17] University of Florida. 2006. *Rewiring Florida's News: from Microfilm to Digital – LSTA Grant Mid-year Report for Federal Fiscal Year 2005 Projects*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/UF00082123/00002>.
- [18] University of Florida. 2005. *Rewiring Florida's News: from Microfilm to Digital - LSTA Grant Proposal for Federal Fiscal Year 2005*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/UF00082123/00001>.
- [19] University of Florida. 2011. *SobekCM: Digital Content Management System*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/sobekcm>.
- [20] University of Florida. 2011. *Preservation Systems for the UF Digital Collections (UFDC)*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu/sobekcm/preservation>.
- [21] University of Florida. 2011. *UF Digital Collections (UFDC)*. University of Florida. Gainesville, FL. <http://ufdc.ufl.edu>.
- [22] Zarndt, F. 2011. Digitization: Successful Projects and the Challenge of Born-Digital Newspaper Archives. *Newspaper Archive Summit: Rescuing orphaned and digital content*. University of Missouri. April 11, 2011. <http://www.rjionline.org/events/newspaper-archive-summit-recorded-sessions>.