
2016 Satellite meeting - *News, new roles & preservation advocacy: moving libraries into action*
10 – 11 August 2016
Lexington, Kentucky USA, USA

Microfilm to Born-Digital for Current Newspapers: A Case Study from the University of Oregon Libraries

Sheila Rabun

Digital Project Manager and Interim Director of the Digital Scholarship Center, University of Oregon Libraries, Eugene, Oregon, USA
srabun@uoregon.edu



Copyright © 2016 by **Sheila Rabun**. This work is made available under the terms of the Creative Commons Attribution 4.0 Unported License:
<https://creativecommons.org/licenses/by/4.0/>

Abstract:

The University of Oregon (UO) Libraries have been microfilming and preserving Oregon's newspapers as a public service since the 1950s. However, microfilm production has become increasingly unsustainable due to the rising costs associated with maintaining microfilming equipment and supplies, and researchers increasingly prefer digital access to newspapers as opposed to microfilm access. In order to mitigate costs and meet user needs, the UO Libraries has transitioned from microfilm to digital to provide preservation and access services for current Oregon newspapers. This case study traces the project from start to finish, outlining motivating factors for making the transition, stakeholder communication strategies, initial project plan and workflow, expected and unexpected challenges, current and forthcoming solutions, improved workflows, next steps, and lessons learned.

Keywords: microfilm production, born-digital newspapers, born-digital newspaper workflow, chronam software, University of Oregon Libraries

Introduction

The University of Oregon (UO) Libraries have been microfilming and preserving Oregon's newspapers as a public service since the 1950s. By providing access to Oregon's history, these microfilms have been an invaluable resource for libraries, individuals, and organizations throughout the state, providing the foundation for the UO Libraries' [Oregon Digital Newspaper Program](#) (ODNP) and [Historic Oregon Newspapers](#) online. Over the past several years, microfilm production has become increasingly unsustainable for the UO Libraries, due to the rising costs associated with maintaining microfilming equipment and supplies. In addition, we started receiving more and more inquiries from other libraries and historical societies subscribing to our microfilm services, asking for digital access to newspapers as opposed to microfilm access. In order to mitigate costs and meet user needs, the UO Libraries decided to

investigate transitioning from microfilm to digital to provide preservation and access services for current Oregon newspapers.



Figure 1. The UO Libraries' newspaper microfilm production and resulting collection has provided the primary means of public access to Oregon's newspapers in the 20th century.

In July of 2014, we began to explore our options. We looked to our colleagues at the [University of North Texas \(UNT\) Libraries](#) and the [University of Kentucky Libraries](#), two of the largest statewide library programs providing digital access to current newspapers, for advice on making the transition. Like the UO Libraries' ODNP, the programs in Texas and Kentucky are both housed and maintained by the university libraries and also have historic newspaper digitization programs. Both programs had eliminated the possibility of providing subscription or fee-based access due to cost-benefit analyses and dedication to open information. However, we were also facing some unique challenges that differentiated our program from UNT and Kentucky. In Texas and Kentucky, the respective newspaper publisher associations require all publishers to send their born-digital newspaper issues to a clipping service to have the public notices extracted and archived, which essentially provides a newspaper aggregation service for each state. Both UNT and Kentucky were able to then receive all of their current digital newspaper content from one source, saving time and effort in coordinating with newspaper publishers to transfer their files to each library. In Oregon, newspaper publishers do not send their PDFs to a clipping service, so there is no singular supplier of current born-digital newspaper content in the state. We were tasked with figuring out the best way to work with individual publishers in order to receive their content in a digital format.

Also, unlike Kentucky and UNT, we use a local instance of the [chronam software](#) to provide online access to our digital newspapers at Historic Oregon Newspapers online. Chronam was originally developed specifically for the Library of Congress' [Chronicling America](#) website, a product of the National Digital Newspaper Program (NDNP), sponsored by the National Endowment for the Humanities (NEH). The software was specifically built to accommodate the [NDNP technical specifications](#), so several very specific assumptions were built into the

software. For example, Chronicling America allows users to search newspapers from across the United States, whereas all of the newspapers on the Historic Oregon Newspapers site are from one state. Structured batches are required in order to ingest digital newspaper content into the chronam software. Batches must conform to the NDNP standards, with a batch folder containing a batch XML file listing the contents of the batch, and LCCN folders, which typically contain “reel” folders, identifying the microfilm reel from which the content was scanned. For issues scanned from print, a “print” folder can be used in place of the “reel” folder. Issue date folders are used to house the individual pages of an issue. Each issue folder contains an XML file for issue-level metadata, as well as a TIFF, JP2, PDF, and METS/ALTO XML file for each page. Additionally, Chronicling America only contains digitized newspapers that were published before 1922 and in the public domain, so notions of copyright and embargo periods are not considered in the software. No other institution had yet attempted to use the chronam platform for born-digital current newspapers.



Figure 2. Historic Oregon Newspapers online, hosted and maintained by the UO Libraries as part of the Oregon Digital Newspaper Program (ODNP), runs on a local instance of the chronam software.

Despite these new challenges, we decided to move forward with a plan to stop microfilming newspapers, starting with those published in 2015, in favor of a fully digital receipt, processing, and delivery system for preserving and providing online access to current newspaper content. Thanks to advice from our colleagues, we knew that early and frequent communication with all of our stakeholders would be a crucial step in making a smooth transition. We would definitely need to keep our 56 microfilm subscribers in the loop, as they were paying to receive copies of the microfilm produced by the UO Libraries. Equally important was our communication with individual newspaper publishers responsible for the 100 publications that we had been microfilming, explaining how the new digital program would work. We sent our

first correspondence announcing the change to stakeholders in October 2014, with subsequent updates sent in January, June, and October of 2015, and April of 2016. Our letters to both microfilm subscribers and newspaper publishers can be found online at <http://odnp.uoregon.edu/current-newspapers/documentation/>.

Pilot Phase

As of January 2015, we stopped receiving print newspapers for microfilming, and started accepting born-digital PDFs from publishers. For the pilot phase of the project, we set up a Google Drive folder for each newspaper title and instructed publishers to use their Google accounts to login and upload their issue PDFs to their assigned folders. At this point, we were just interested in obtaining sample content to determine average file sizes in order to develop an appropriate cost model, so we did not provide any specific instructions on file naming or number of pages per file. Only a few of the publishers interested in participating did not have Google accounts, so we created accounts for them. Using Google Drive was an easy solution for most of the participating publishers, although a few requested a more automated solution for file transfer. For some very small publishers, Google Drive presented technology barriers or time requirements that hindered their ability and/or willingness to participate. However, we received files from 16 different newspapers almost immediately, and with the extensive assistance of student workers, we started an inventory, assessing the average size of each file, taking account of the variation in newspapers between 8 and over 100 pages, and number of weeklies as opposed to dailies. Our student workers entered sample metadata into spreadsheets, so we could get a ballpark estimate for student time needed to input metadata for each issue.

Using this initial data, we determined that we would be able to subsidize the cost of intake, processing, hosting, and preserving born-digital newspaper content at a lower cost than running the microfilm operation. We would not be storing TIFF files, but would store PDF/A files for preservation, thus saving server space. We would have student assistants perform image processing, metadata entry, and quality review under the direction of our Digital Production Manager. Using our existing Historic Oregon Newspapers website, we would follow as closely as possible to the National Digital Newspaper Program (NDNP) technical specifications that are required for ingesting content into the chronam software. We would create scripts to automate creation of METS XML files for issue and batch level metadata, and METS/ALTO XML files for page level OCR data. Staffing was planned to include a product owner/manager (two hours/week), workflow supervisor (six hours/week), two or three processing student assistants (40 hours/week total), one or two quality review student assistants (20 hours/week total), a systems administrator (20 hours upfront with on-going maintenance integrated into existing maintenance of Historic Oregon Newspapers), and programming support. We originally estimated eight weeks of development time needed, but this expanded due to unexpected programming needs that unfolded over the course of the project.

We required each publisher to submit a signed [license agreement](#) in order to fully participate in the new digital program. Early on, we decided that we would not be storing newspaper files from publishers who were unwilling to have their content made available on Historic Oregon Newspapers online. We provided an option of either a three month embargo before making issues available online, or no embargo, as well as four options for copyright statements:

- Free Access License - all other permissions retained (<http://www.europeana.eu/rights/rr-f/>)
- Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>)

- Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>)
- Creative Commons Attribution Non-Commercial No Derivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

As of June 2016, we have received signed license agreements for 27 newspaper titles out of the 100 we once microfilmed. Several publishers specifically cited their plans to have their own digital repository of back issues as a way to maintain revenue, as justification for not participating in our program. With limited resources and significant staffing turnover during the course of the transition, we have not yet been able to dedicate time for outreach and targeting newspapers for participation.

Initial Workflow and Unexpected Challenges

While participating publishers were actively submitting their PDFs to our Google Drive, we started to refine the workflow for processing PDFs, derivatives, OCR data and metadata to create batches for ingest into our local instance of chronam. Our programming team devised a “[PDF to text](#)” script to extract the embedded text data from each PDF file into a passable METS/ALTO XML file for each page, with tags for text blocks, lines, and words, each including coordinates to represent location on the page/image. Output might not strictly meet ALTO specs, but it works for ingest and search functionality in chronam.

Using the NDNP metadata specifications as a guide, we decided to include only the very basic metadata: LCCN (which would give us the title and place of publication), date, volume, issue number, and edition number, as well as indicate missing pages, mis-dated issues, and embargo information. Embargo information is not part of the NDNP specifications and thus not considered in the chronam software, which posed a challenge for incorporating embargo periods into our workflow. Our first approach to metadata entry was to have students manually download each issue from Google Drive into appropriately-titled LCCN folders, separate pages if necessary, and rename each page using all of the essential metadata. For example, the file naming convention was as follows: LCCN_yyyymmdded_vol_iss_sequ. A file named sn83008376_2015010201_030_001_0001 would be page one from the January 2, 2015 first edition of the *Daily Astorian*, Volume 30, Issue 1.

We then had students manually create a JSON file to go with each issue, containing information about missing pages, mis-dated issues, and embargo periods, as these elements are not consistent across newspaper issues and would thus be difficult to represent as part of the file name. Our programmers wrote scripts to transfer the metadata contained in the file name and JSON files into METS XML files for each issue. While the scripts worked, progress was hindered when we realized that the file name and JSON method was problematic. The file names were unwieldy and often fraught with typos and inconsistencies, and the JSON files were burdensome and often erroneous. Time spent on quality review and manual data revisions quickly grew beyond the resources that we had allocated for the project.

Additionally, within several months it became apparent that Google Drive was not going to be sustainable as a file transfer platform. Not only were PDF submissions taking up space on the individual publishers’ Drives, we also could not delete the files that they had uploaded due to ownership issues. There was no workable solution with Google Drive that would streamline the uploading and management of the publishers’ content. In addition, we received complaints from various sources about the publishers’ lack of understanding about what we

were doing. We had sent out multiple letters explaining the situation and the benefits to the new workflows and digital content, but our messages were not getting through in some cases. Additionally, former microfilm subscribers whose local newspapers had not yet sent in a license agreement or were not going to be participating in the digital program started to panic – what were they going to do now that we were no longer microfilming?

To address stakeholder concerns, we dedicated additional time to clarifying the information about our transition and related documentation online, including an FAQ section for publishers and subscribers. We volunteered to attend an Oregon Newspaper Publishers Association meeting to field publisher questions in person and rectify any misconceptions about our activities. For subscribers, we were able to provide some information on other possible microfilm vendors, but we had not yet done the research to feel comfortable making explicit recommendations for any one vendor. We also suggested that subscribers reach out to their local newspaper(s) and encourage them to participate, since ultimately, the decision to participate in the digital program resided with each publisher.

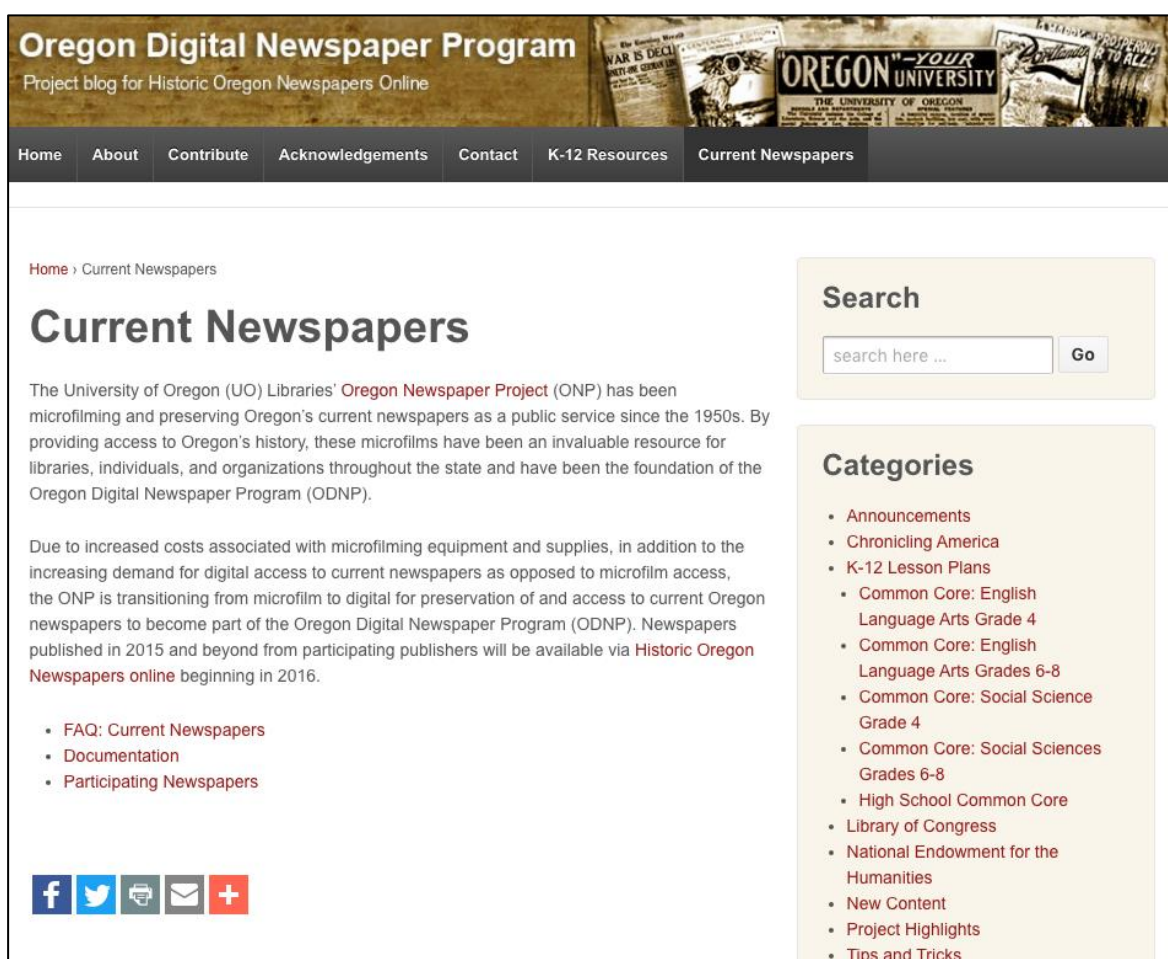


Figure 3. General information about the UO Libraries' transition from microfilm to digital for current newspapers can be found online at <http://odnp.uoregon.edu/current-newspapers/>.

Workflow Improvements

Upon looking to improve upon our initial workflow, first we set up a Secure File Transfer Protocol (SFTP) system to replace Google Drive, which required generating user names and passwords for each participating newspaper and creating instructions for publishers to install

and use FileZilla to transfer their PDFs to us. In order to automate file retrieval on our end, we required publishers to upload the files for each issue within a folder labeled with the issue date in YYYY-MM-DD format. In order to streamline our processes and ensure consistency, we set a date for removing permissions on the Google Drive folders and solely accepting PDFs via SFTP. We gave publishers a three month notice to allow sufficient time to adjust their workflows, and we sent reminders leading up to the deadline. The majority of participating publishers have made the switch, and the rest will need to do so in order to participate. For some publishers, the SFTP has presented technical challenges, but others have been able to automate their uploads, making it easier on them.

To address problems with metadata entry and quality review, we created the “[PDF to chronam administrative web application](#),” or “admin app.” The app software controls data entry to mitigate human error as much as possible. In light of time and resource constraints, the app was built to consider only the most basic data entry and review needs. Users are assigned roles of either issue curator or issue reviewer. Issue curators are required to enter metadata for Issue Date (separate from Issue Date as Labeled, should there be a misprint), Volume Number, Issue Number, Edition Order, and Edition Label, as well as Page Numbers (as printed). Our initial attempts to notate missing pages proved error-prone and resulted in hours of re-work for file naming and page sequence labels, and it became clear that including this metadata in our new workflow would take more time and effort than we could afford. In an effort to reach a minimum viable product under increasing time constraints, we decided to temporarily give up on trying to include information about missing pages, with hopes that page labels for present pages would provide users with good enough information as to potential pages that might be missing from an issue. Date information is pulled from folder names, since the files for each issue are contained within a folder labeled with YYYY-MM-DD. Title information is pulled from each title’s Library of Congress Control Number (LCCN), since issues for each newspaper are contained within folders labeled with LCCN. Date, title, and issue metadata are presented as part of the review queue process within the app, once metadata has been entered. Issue reviewers double check issue metadata and page labels, and then either approve or reject each issue, depending on metadata accuracy. Rejected issues are cycled back into the data entry queue. Metadata information is saved in a JSON file for each issue, which is automatically converted to the required issue-level METS XML file format upon batch creation. Ultimately we hope to do away with JSON in favor of storing metadata in a database for best practice.

Additionally, the admin app contains embargo information for each LCCN, so embargoed issues are held for a specified number of days before being added to a batch. The app also contains rights statements per LCCN, and inserts a Rights field into each issue-level METS XML file. Our programming team was able to hack our local chronam batch loader to read the Rights field, and manipulate the Historic Oregon Newspapers site user interface to present a copyright statement at the bottom of the screen when a user views a current newspaper on the website. (The Rights field is specific to our local instance and is not part of the base chronam software.) While the basic functionality of the app is in place, enhancements will likely need to be made in the future to improve application speed and functionality.

ODNP Admin
Newspaper Titles
MARC Org Codes
Issues
Users
Logged in as srabun

Return to Issues
Enter Metadata

Type a page label **exactly as it's printed**, or type zero ("0") if a page has no printed label. Press [ENTER] to save and move to the next page.

Page Label
3

Image 3 of 32
Jump to top
Jump to bottom


February 19, 2016
CapitalPress.com
3

Oregon wolf delisting bill draws questions in Senate committee

Bill would ratify removal of wolves from endangered species list

By MATEUSZ PERKOWSKI
Capital Press

SALFEM — A bill to ratify



Oregon farm wins radish seed dispute

Favorable decision may affect other growers in lawsuit with bank

By MATEUSZ PERKOWSKI
Capital Press

A favorable ruling won by an Oregon farm in a legal dispute

Hawman Farms of Hermiston, Ore., agreed to grow roughly \$180,000 worth of seed and has a "possessory lien" on the crop that takes precedence over the security interest held by the bank.

Under a possessory lien, a party can retain possession of property until a debt has been paid.

Northwest Bank claimed the contract between Hawman

Figure 4. The PDF to chronam admin app allows users to enter basic newspaper issue metadata and page labels.

Reject
Review Page Labels


Image 1 of 20
Jump to top
Jump to bottom

GW CLARK IS BACK TO BULLDOGGING – PAGE B1

The Blue Mountain EAGLE

Grant County's newspaper since 1868

WEDNESDAY, APRIL 20, 2016 • No. 16 • 20 PAGES • \$1.00 www.MyEagleNews.com



Publication	The Blue Mountain eagle.
Issue Date as Labeled	2016-04-20
Issue Date	2016-04-20
Volume Number	0
Issue Number	NO. 16
Edition Number	1
Edition Label	

Figure 5. Quality review of issue-level metadata and page labels is performed in the admin app.

Our continued evaluation of the entire workflow has produced the following processes as of June 2016:

Current Newspapers Workflow (<https://github.com/uoregon-libraries/pdf-to-chronam>):

- 1) Publishers upload PDFs (within folders labeled YYYY-MM-DD for each issue) via SFTP regularly
- 2) Folders moved from SFTP to local server weekly (automated)
- 3) Original files saved to master backup location (automated)
- 4) Initial file processing (automated)
 - a. PDFs within each YYYY-MM-DD folder are merged to mitigate multi-page files
 - b. PDFs from each issue are split into one page per file
 - c. Files are converted to PDF/A
- 5) Derivatives created and named to match PDFs (automated)
 - a. OCR data from PDF is extracted and converted to METS/ALTO-like XML
 - b. JP2 created
- 6) Initial file review and renaming (manual)
 - a. Files are reviewed for order accuracy and re-ordered if needed
 - b. Files are renamed in sequence order (0001, 0002, 0003, 0004, etc.)
- 7) Issues queued for metadata entry in the PDF to chronam admin application (automated)
- 8) Data entry assistants input issue and page level metadata via web application (manual)
 - a. Edit Metadata:
 - i. Issue Date As Labeled
 - ii. Issue Date
 - iii. Volume Number
 - iv. Issue Number
 - v. Edition Number
 - vi. Edition Label
 - b. Label Pages
 - i. Input printed page number
 - c. Queue for review
- 9) Quality review assistants review metadata for accuracy (manual)
 - a. Check metadata
 - b. Check page numbers
 - c. If approved, move to step 10
 - d. If rejected, issue moves back to step 7
- 10) Batch creation in accordance with NDNP batch structure (with exception of TIFF files and reel-level metadata) (automated)
 - a. Issue-level METS XML created
 - b. LCCN folders containing issue-level folders are moved into batch folders
 - c. Batches are named
 - d. Batch-level METS XML created
- 11) Final Quality review (manual)
 - a. Batch is ingested into development instance of website
 - b. Lead quality review staff views batch online
 - i. Every issue date is checked for accuracy, using the chronam “Calendar View” feature

- ii. Search functionality tested
 - iii. General review of image and metadata
 - c. If problems are found, batch is removed from the admin tool and backup files are used to correct and re-create the batch. Depending on the problem, reprocessing may start at step 5 (broken derivatives), 6 (mis-ordered pages), or 7 (metadata needs to be re-entered). If the problems are limited to a single issue, the rest of the issues can be re-batched fairly quickly, although it is currently a semi-manual and tedious process
 - d. If the batch is approved, move to step 12.
- 12) Batch ingest (automated)
- a. Batch is ingested into production website
 - b. Batch is deleted from development website

Metadata Fields Used

(see [NDNP technical specifications](#) for more detail:)

- Award Year
- Awardee Name
- Original Source Repository
- Original Source Repository Code
- Digital Responsible Institution
- Digital Responsible Institution Code
- Batch Name
- LCCN
- Title
- Volume Number
- Edition Order
- Edition Label
- Issue Number
- Issue Date
- Issue Present Indicator (only for present issues)
- Issue Date As Labeled
- Page Sequence Number
- Page Number
- Page Physical Description
- Page Present Indicator (only for present pages at this time)
- + Rights Statement (not part of the NDNP specification)

Next Steps

Despite all of the challenges and revisions we have encountered, we were able to successfully ingest over 20,000 pages of current newspaper content into our chronam instance by early 2016. We are still refining our workflow and working towards a finalized, seamless, ongoing process. Ideally with continued troubleshooting, we will deploy full automation for downloading publishers' SFTP submissions, processing derivatives, ingesting into and out of the metadata app, creating batches, and finalizing ingest. At present, this process is semi-automated with scripts existing in different platforms, requiring more human intervention than desired, and needing to be unified to fully assemble a singular, dependable system. Forthcoming workflow enhancements, if time allows, will likely include automated notices sent to publishers confirming issue receipt, and alerts to problems, e.g., missing issues or pages. Furthermore, the PDF to chronam admin app needs ultimately to be revised as there have been

many bugs that were undetected due to time constraints and other projects that required development resources. We are currently using Adobe Bridge to re-order and re-name pages, which has presented additional challenges and serves as yet another potential point for error in the workflow. We are hoping to ultimately include functionality for indicating missing pages, as well as page re-ordering and file naming, within the app. We would also like to improve system functionality for purging and re-creating batches if and when errors are found in the late stages of quality review, since this process is currently semi-manual and tedious. Once the full process is improved and streamlined, newspaper digitization and processing will take very little staff time, as a result of the scripting and automation work that the project programmers have initially provided.



Figure 10. Despite both expected and unforeseen challenges, born-digital newspapers from 2015 and beyond are now available for keyword searching and browsing at Historic Oregon Newspapers online!

Once full automation is in place to collect the rich content and metadata that we are providing, we hope to enhance discoverability by linking digital newspaper content in our library catalog records. We know that aggregating and providing access to these specialized publications is not enough without sharing and promoting these resources in larger collection platforms. The detailed and precise data that we are collecting and extracting will allow for easier data sharing and discoverability for these unique, primary source materials.

We are hoping to finish microfilming the backlog of papers published through December 2014 (estimated completion date of December 2016) and work through backlog of Google Drive submissions that we received before switching to the SFTP system. As previously mentioned, we are still looking into potential microfilm vendors if future needs arise for

current or future partners, so we have not been able to assist subscribers with this task as of yet. We have learned that thorough and frequent communication with the stakeholders is vital. In the future, it would be beneficial to do more outreach to get more publishers involved. Once our entire process is perfected and harmonized, we strive to create new partnerships for ODNP and continue adding new content.

Conclusion

The process of transitioning from microfilm to digital for current newspapers evidently has not been one without roadblocks, surprises, and even anticipated challenges. However, the importance of building a solid framework for collecting, preserving, and providing open access to current newspapers in the digital age has made this project worthwhile. Not only are newspapers valued by researchers across disciplines, they also preserve day-to-day happenings for future reference. Being the “first drafts of history,” newspapers are crucial to the historical record, as they reflect our cultural heritage in the form of primary source documents.

The work that we have done on this project—creating workflows, the metadata entry app, quality review process, and METS/ALTO creation—has been crucial to not only current born digital newspapers, but also for our in-house workflow for digitization of historic newspapers both from print (large format planetary scanner, Library Services and Technology Act (LSTA) grant, 2015-2016) and microfilm (microfilm scanner, LSTA grant, 2011), especially since our former newspaper digitization vendor is no longer taking projects.

The adjustments that we have made to our local instance of chronam have informed the [Open Online Newspaper Initiative](#) (Open ONI) community software development efforts as use cases have emerged, such as the need for rights statements. Goals of the Open ONI project include adapting the chronam software for use by state and local institutions; making installation and usage of the software as easy as possible; and building compatibility with interoperability standards such as the [Portland Common Data Model](#) (PCDM) for compound object functionality, and the [International Image Interoperability Framework](#) (IIIF) for improved access and utility. Future IIIF functionality could include the ability to allow OCR text corrections, commenting, describing, and otherwise tagging digital images. Discovery of articles and images within newspapers would definitely improve with the use of image tags and descriptions made available through the IIIF Presentation API.

Lessons Learned

1. Allow as much time as possible to for workflow creation, including testing and troubleshooting of small pre-ingest batches, and solidify workflow preferably before approaching publishers
2. Reach out to stakeholders (primarily publishers and microfilm subscribers) early and often to ensure clear communication of project goals, expectations, and next steps
3. Compile a list of potential alternative microfilm production vendors before reaching out to former microfilm subscribers
4. Avoid switching file submission systems once the project is underway; set up a permanent file transfer system ahead of time, and test, test, test!
5. Dedicate ample staff time for fielding questions from publishers and subscribers
6. If possible, avoid undertaking other major projects at the same time; do not underestimate the necessity of development time when creating/augmenting new software/systems

Workflow Charts

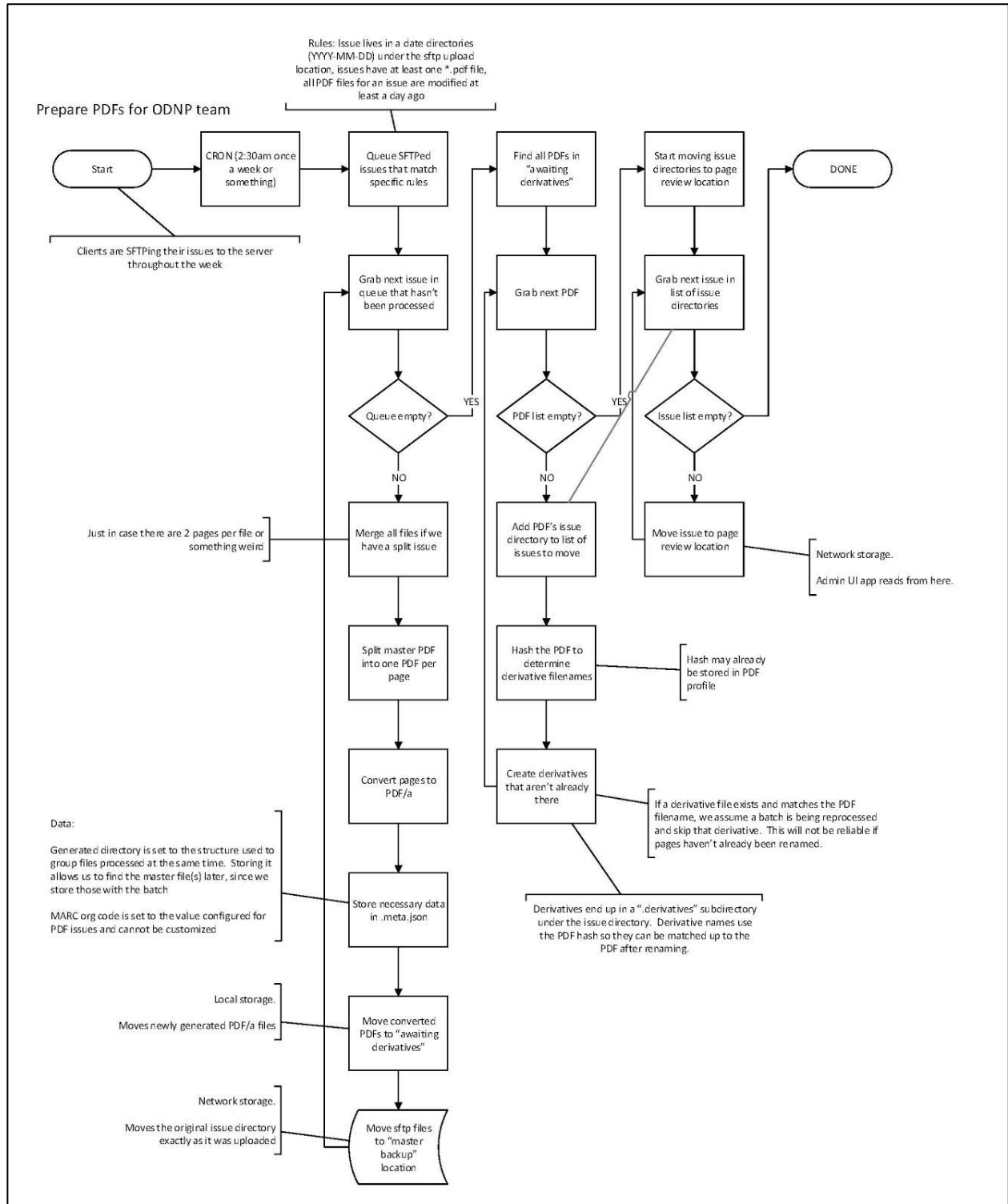


Figure 6. File receipt and preparation, steps 1-5 (Echols, 2016)

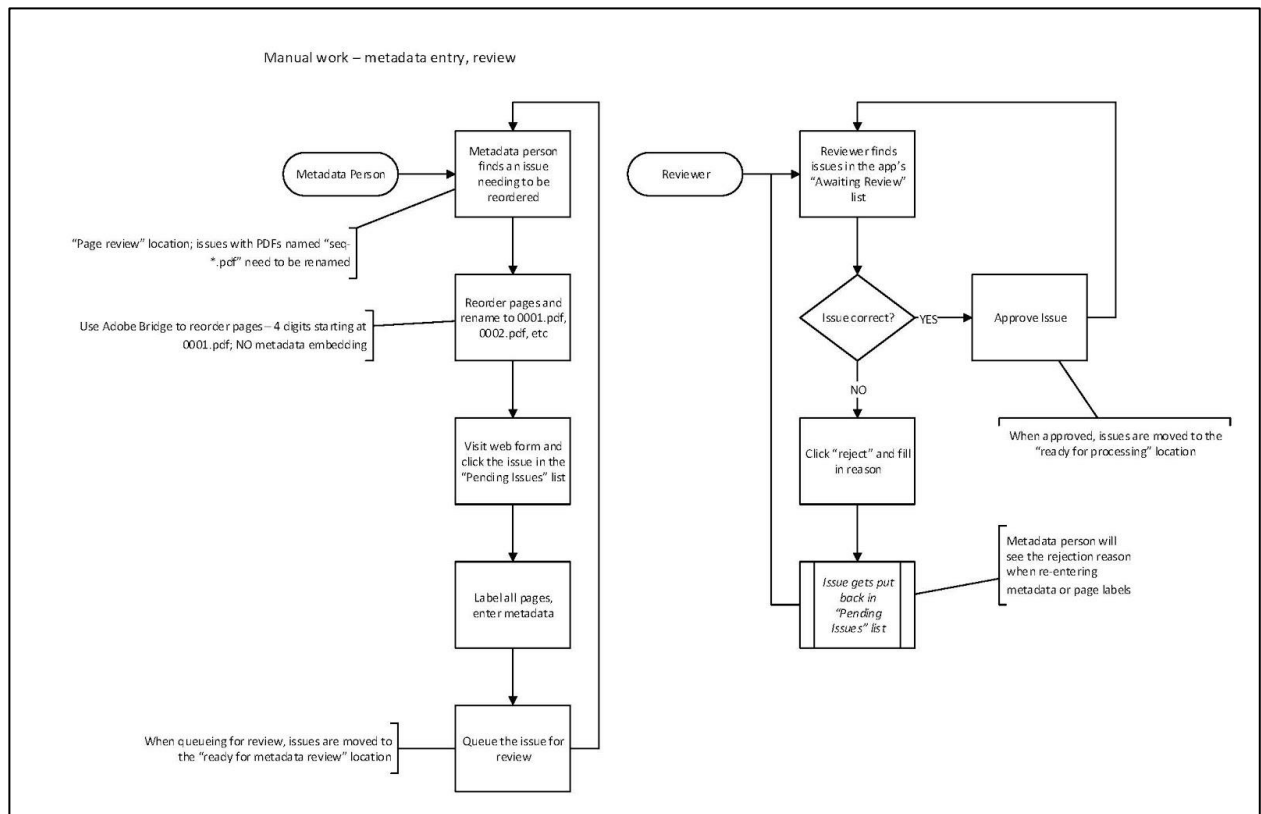


Figure 7. Metadata Entry and Quality Review, steps 6-9 (Echols, 2016)

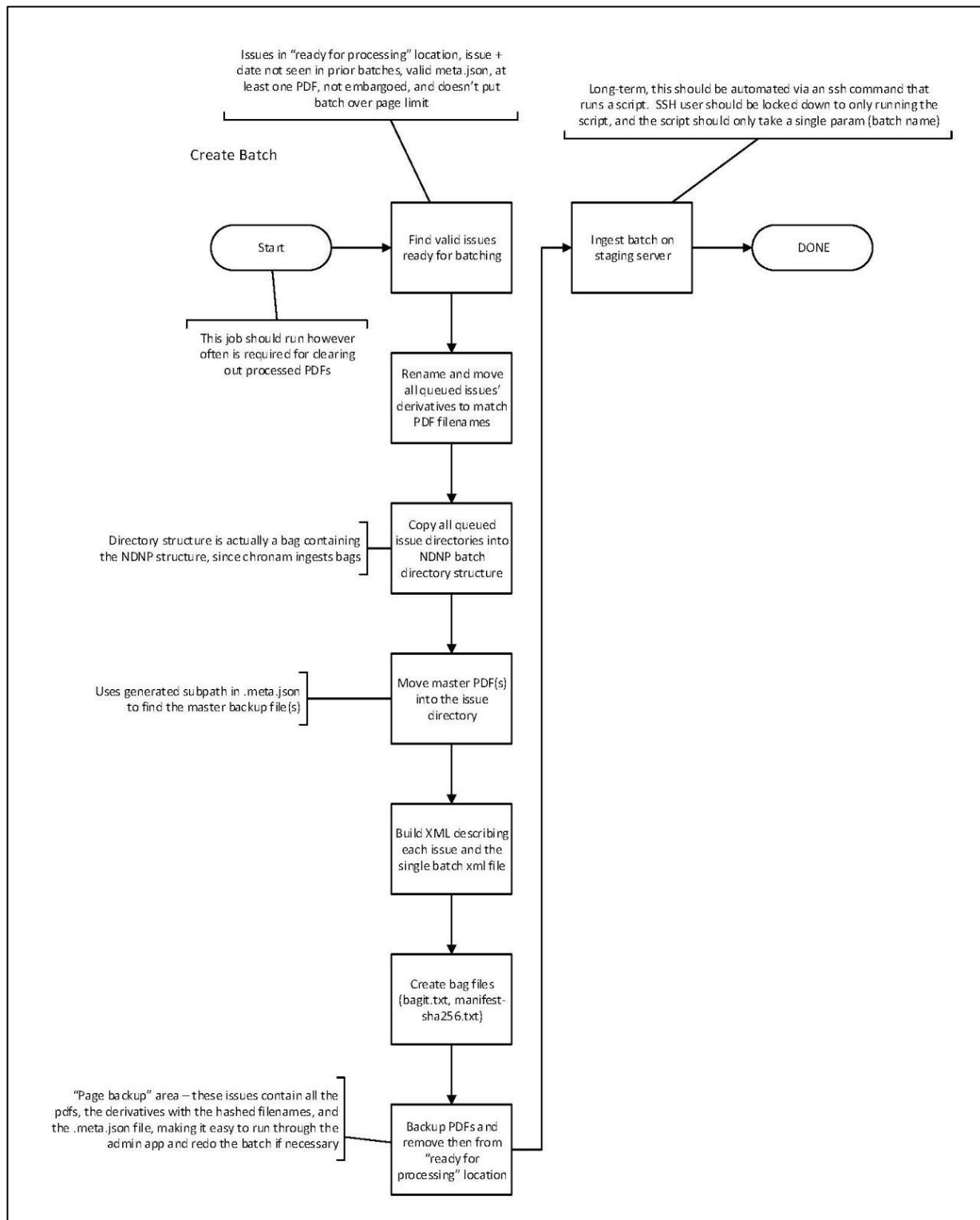


Figure 8. Batch creation, step 10 (Echols, 2016)

Verify and Ingest

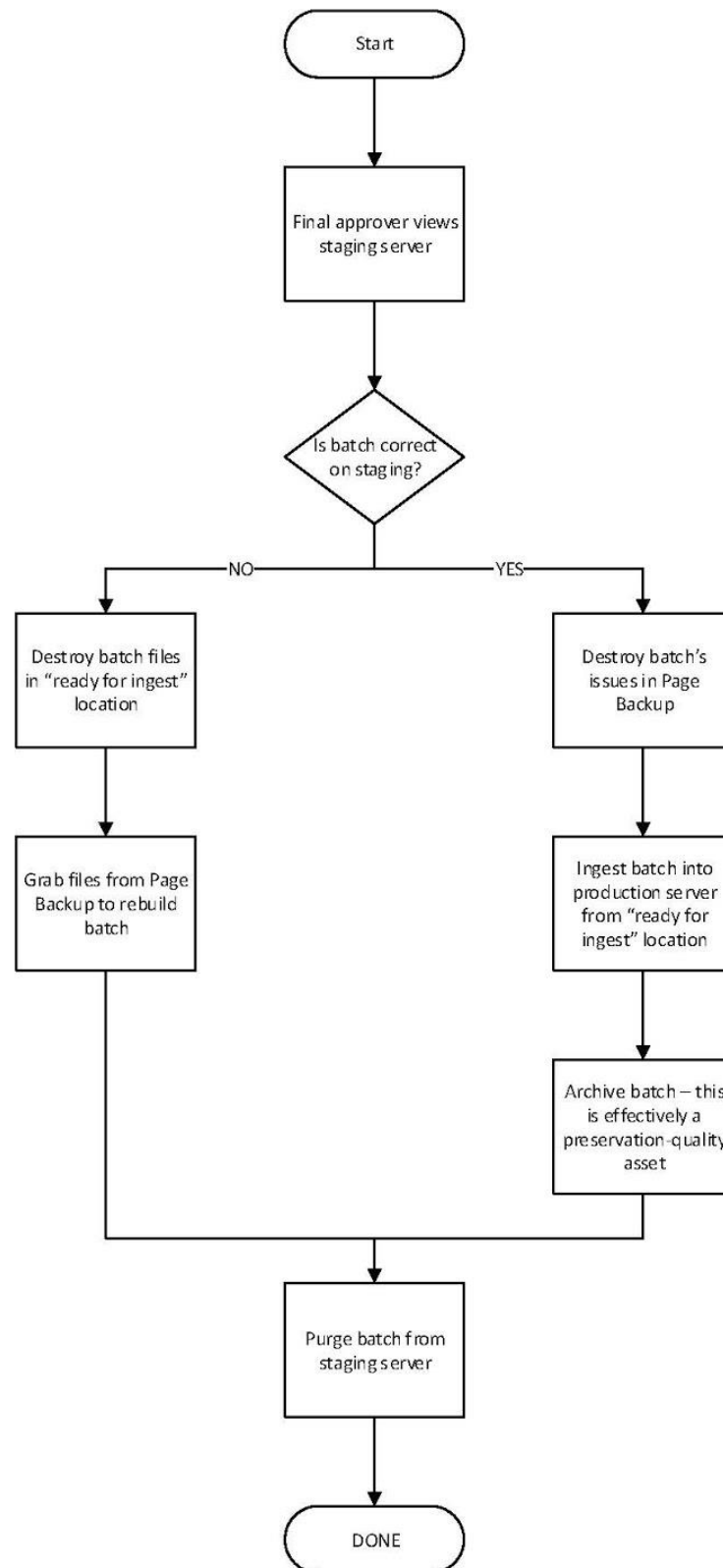


Figure 9. Final batch verification and ingest, steps 11-12 (Echols, 2016)

Acknowledgements

University of Oregon (UO) Libraries Project Team:

- Karen Estlund, former Head, Digital Scholarship Center
- Sheila Rabun, Digital Project Manager and Interim Head, Digital Scholarship Center
- Duncan Barth, Assistant Director, Library Systems
- Jeremy Echols, Programmer, Library Systems
- Linda Sato, Programmer, Digital Scholarship Center
- Heidi Scheidl, Microfilm Production Coordinator, Image Services Center/Digital Scholarship Center
- Amanda Schmutz, former Microfilm Production Coordinator, Image Services Center
- Julia Simic, Metadata and Digital Production Librarian, Digital Scholarship Center
- Randy Sullivan, Digital Production Manager, Digital Scholarship Center

UO Libraries Student Assistants:

- Erin Choi
- Darienne Christiansen-Miller
- Andrea Cueva
- Janette J. Halemeier
- Jaylee L. Jordon
- Alesha Major
- Keelan Sullivan

UO Libraries Leadership:

- Adriene Lim, Dean of Libraries and Philip H. Knight Chair
- Andrew Bonamici, Associate Dean for Media and Digital Strategies
- Deborah Carver, former Dean of Libraries and Philip H. Knight Chair

Special Thank You:

- Mary Molinaro, former Director of Preservation and Digital Programs at the University of Kentucky Libraries, and former Associate Dean for Library Technologies; University of Kentucky
- Sarah E. Seymore, Digital Metadata Technician, University of Oregon Libraries' Digital Scholarship Center
- National Digital Newspaper Program
- Participating Oregon newspaper publishers

References

Chronicling America. (n.d.) Retrieved June 30, 2016, from <http://chroniclingamerica.loc.gov>

Current Newspapers. (n.d.). Retrieved June 30, 2016 from <http://odnp.uoregon.edu/current-newspapers/>

Documentation. (n.d.). Retrieved June 30, 2016, from <http://odnp.uoregon.edu/current-newspapers/documentation/>

Duraspace/pcdm. (n.d.). Retrieved June 30, 2016 from <https://github.com/duraspace/pcdm/wiki>

Echols, Jeremy. (2016). *Workflow charts*. Unpublished documents

Historic Oregon Newspapers. (n.d.) Retrieved June 30, 2016, from <http://oregonnews.uoregon.edu>

International Image Interoperability Framework. (n.d.) Retrieved June 30, 2016 from <http://iiif.io/>

Kentucky Digital Library. (n.d.). Retrieved June 30, 2016, from <http://kdl.kyvl.org/>

LibraryOfCongress/chronam. (n.d.). Retrieved June 30, 2016, from <https://github.com/LibraryOfCongress/chronam>

Library of Congress. (2015, Sept. 29). *The National Digital Newspaper Program (NDNP) Technical Guidelines for Applicants*. Retrieved June 30, 2016 from https://www.loc.gov/ndnp/guidelines/NDNP_201618TechNotes.pdf

Open-oni. (n.d.). Retrieved June 30, 2016, from <https://github.com/open-oni>

Oregon Digital Newspaper Program. (n.d.). Retrieved June 30, 2016, from <http://odnp.uoregon.edu/>

Texas Digital Newspaper Program - The Portal to Texas History. (n.d.). Retrieved June 30, 2016, from <http://texashistory.unt.edu/explore/collections/TDNP/>

University of Oregon Libraries. (n.d.). *Non-Exclusive License Agreement for Digital Rights*. Retrieved June 30, 2016 from http://odnp.uoregon.edu/files/2015/09/ONP_License_Agreement-2j17z6i.pdf

Uoregon-libraries/pdf-to-chronam. (n.d.). Retrieved June 30, 2016 from <https://github.com/uoregon-libraries/pdf-to-chronam>

Uoregon-libraries/pdf-to-chronam-admin. (n.d.). Retrieved June 30, 2016 from <https://github.com/uoregon-libraries/pdf-to-chronam-admin>

Uoregon-libraries/pdftotext. (n.d.). Retrieved June 30, 2016, from <https://github.com/uoregon-libraries/pdftotext>